ITS INSTITUTE
Intelligent Transportation Systems

Places and Networks:

The Changing Landscape of Transportation and Technology
Final Summary Report of the STAR–TEA 21 Project

Final Report

Prepared by:

Principal Investigators:

Richard S. Bolan
Lee W. Munnich Jr.
Kevin J. Krizek
Thomas A. Horan
David Levinson

Other Contributors:

Mark Berndt
Richard Burkhard
Frank Douma
Chandler Duncan
Susan L. Handy
Andrew Johnson
Ugur Kaplancali
Ramachandra Karamalaputi
Adam Kokotovich
James Lehnhoff
Yi Li
Michael Marich
Denise McCabe
Pavithra Parthasarathi
Benjamin Schooley
Colbey Sullivan
Ryan Wilson
Peng Xu
Bhanu M. Yerra
Lei Zhang

State and Local Policy Program
Humphrey Institute of Public Affairs
University of Minnesota

CTS 08-15

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.
CTS 08-15		

4. Title and Subtitle	5. Report Date
Places and Networks: The Changing Landscape of Transportation and Technology Final Summary Report of the STAR–TEA 21 Project	December 2008
	6.
7. Author(s) Mark Berndt, Richard S. Bolan, Richard Burkhard, Frank Douma, Chandler Duncan, Susan L. Handy, Thomas A. Horan, Andrew Johnson, Ugur Kaplancali, Ramachandra Karamalaputi, Adam Kokotovich, Kevin J. Krizek, James Lehnhoff, David Levinson, Yi Li, Michael Marich, Denise McCabe, Lee W. Munnich, Jr., Pavithra Parthasarathi, Benjamin Schooley, Colbey Sullivan, Ryan Wilson, Peng Xu, Bhanu M. Yerra, Lei Zhang	8. Performing Organization Report No.
9. Performing Organization Name and Address State and Local Policy Program Humphrey Institute of Public Affairs University of Minnesota 301 19th Avenue S Minneapolis, Minnesota 55455	10. Project/Task/Work Unit No. 2009041, 2009042, 2009043, 2009045, 2009046, 2009047 11. Contract (C) or Grant (G) No.
12. Sponsoring Organization Name and Address Intelligent Transportation Systems Institute University of Minnesota 200 Transportation and Safety Building 511 Washington Avenue SE Minneapolis, Minnesota 55455	13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code

15. Supplementary Notes
http://www.cts.umn.edu/Publications/ResearchReports/

16. Abstract (Limit: 200 words)

Over the past six years, researchers from the University of Minnesota have studied the many ways in which transportation and technology intersect.

Our work has explored these intersections from many perspectives, from ways intelligent transportation systems can help police, ambulance, and other public safety providers communicate more accurately and save lives, to the use of agent-based modeling to predict how high-technology workers will influence city form--and therefore, transportation needs--through their choices about work and home location. Two other areas of study are whether and how the Internet will replace travel demand and the potential loss of privacy related to advanced transportation technologies and the public policy issues surrounding privacy.

17. Document Analysis/Descriptors	18. Availability Statement
STAR, TEA-21, privacy, transportation technology, wireless communications, emergency response, urban development	No restrictions. Document available from: National Technical Information Services, Springfield, Virginia 22161

19. Security Class (this report)	20. Security Class (this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	463	

Places and Networks:

The Changing Landscape of Transportation and Technology Final Summary Report of the STAR–TEA 21 Project

Final Report

Prepared by:

Principal Investigators:

Richard S.Bolan
Lee W.Munnich Jr.
Kevin J. Krizek
Thomas A. Horan
David Levinson

Other Contributors:

Mark Berndt
Richard Burkhard
Frank Douma
Chandler Duncan
Susan L. Handy
Andrew Johnson
Ugur Kaplancali
Ramachandra Karamalaputi
Adam Kokotovich
James Lehnhoff
Yi Li
Michael Marich
Denise McCabe
Pavithra Parthasarathi
Benjamin Schooley
Colbey Sullivan
Ryan Wilson
Peng Xu
Bhanu M. Yerra
Lei Zhang

State and Local Policy Program
Humphrey Institute of Public Affairs
University of Minnesota

December 2008

Published by:

Intelligent Transportation Systems Institute
Center for Transportation Studies
University of Minnesota
200 Transportation and Safety Building
511 Washington Ave SE
Minneapolis, Minnesota 55455

Acknowledgements

The project participants wish to acknowledge those who made this research possible. The study was funded by the Sustainable Technologies Applied Research (STAR) TEA-21 Project and the Intelligent Transportation Systems (ITS) Institute, a program of the University of Minnesota's Center for Transportation Studies (CTS). Financial support was provided by the United States Department of Transportation Research and Innovative Technologies Administration (RITA).

The project participants would like to thank Representative James Oberstar for his interest in this program and also acknowledge the outstanding efforts of the staff of the Humphrey Institute and Center for Transportation Studies for their work in ensuring the successful completion and promotion of this work.

Table of Contents

CHAPTER 5

Lee Munnich: Industry Clusters and ITS

CHAPTER 6

Lee Munnich: Privacy Issues of ITS

CHAPTER 7

David Levinson: Travel Demand Modeling

List of Tables

List of Figures

Appendix B Maps

Executive Summary

This final report on work done under the Sustainable Transportation Applied Research (STAR) – TEA 21 grant includes research on the many ways in which transportation and technology intersect over the past six years.

As modeling techniques have become more sophisticated they have led researchers to explore how decisions of individual travelers affect larger systems. This research explores the intersections of transportation and technology from many perspectives. There is enormous potential for new technologies' contributions to society: the ability to reduce congestion, to create smarter roads that help people travel more efficiently and in a more environmentally sound way, and to respond to emergencies faster and more efficiently in order to save lives.

This research highlights other issues. How do we pay for the technologies and who pays? How do public entities work with private innovators who create new technologies? Must we sacrifice privacy and freedom as we implement technological advances? Our research has convinced each of us of how complex and layered the intersections of transportation and technology are. Transportation involves millions of decision-makers.

Transportation and technology encompass a vast array of linkages and hierarchies, public and private, transportation-oriented and technological. Policymakers need to anticipate these individual agents in the process of developing policy.

Richard Bolan studied spatial patterns of information workers. He used agent-based modeling to predict how high-technology workers influence city form—and therefore, transportation needs—through their choices about work and home location.

The study sheds some light on the impact of information technology on the metropolitan community in spatial detail, and examines how such technological development might influence land development policies, transportation policies, environmental policies, and other social policies at the local and metropolitan scale. Bolan analyzed six U.S. metropolitan areas (Atlanta, Austin, Denver, Houston, Phoenix and Minneapolis-St. Paul) using the U.S. Census Bureau's Transportation Planning Package (CTPP) for 1990 and 2000.

Appendix B includes maps indicating the weighted mean center and the standard deviational ellipse patterns of both residential and workplace locations for workers in all six metropolitan areas.

Kevin Krizek studied travel behavior impacts of information and communications technologies (ICT). This research explores whether—and how—the internet will replace travel demand. Krizek investigated dimensions of ICT use at the household level for non-work travel. A household survey on ICT use and substitution investigated the substitution aspect of household travel behavior. The results help understand what degree ICT will substitute for physical travel.

Krizek also mapped the terrain of existing work to date related to ICT and household travel. This research helps serve as a blueprint to help conceptualize future research and detailed lines of inquiry relating ICT and household travel.

The affects of eliminating physical travel that is traditionally associated with activities such as work, shopping, and entertainment is included in Krizek's research. He examines the extent to which households use the Internet for such activities may depend on spatial attributes of the communities where they are located. This section answers questions related to the propensity of residents in metropolitan areas to use the Internet for financial transactions and how such use is affected by spatial attributes of retail, bank accessibility, and traffic congestion. Krizek also studied the degree to which at-home ICT use and out-of-home store travel changed from 1995 to 2003 for the selected purposes of shopping and banking.

Thomas Horan studied how intelligent transportation systems can help police, ambulance, and other public safety providers communicate more accurately and save. His research extended information-sharing dimensions to explore information sharing relative to service performance. It utilized a time-critical information services (TCIS) conceptual framework as an analytical lens. A case studies approach was employed to examine the exchange of performance-related information in a key time information critical service: a county-wide emergency medical services (EMS) system

Horan investigated the development of a software-based ontology within the context of a rural wireless emergency management (EMS) system. The ontology was developed by integrating concepts and findings from in-depth field reviews in Minnesota into an ontological software originating out of bioinformatics. Another case study in rural Minnesota investigated technology, organizational and policy dimensions of Emergency Management Services (EMS). Horan explored the nature of interorganizational dynamics in this setting and to set forth an architecture for measuring and enhancing performance. Key technology concerns were intertwined with organizational aspects.

Horan also studied the role of information systems in enhancing end-to-end performance of rural Emergency Medical Services (EMS) systems. This study used interviews and data analysis of the EMS process: mayday call, routing and dispatch, response, and treatment. Another case study focused on the relationship between the operational Mayday system and the behavior of emergency responders and participating organizations.

Lee Munnich, Jr. examined the ways that intelligent transportation systems that help travelers move more safely, quickly, and efficiently and the policy issues surrounding the potential loss of privacy that may accompany some of these technologies. Munnich studied how ITS is affecting rural economies, by conducting interviews with representatives from firms in the recreational transportation equipment industry and the wood products cluster in northwest Minnesota. The primary goal of this research was to analyze ITS use from the industry cluster perspective.

Munnich also looks at the how major current and future ITS technologies may affect privacy, and recommends a framework to address privacy issues when working with ITS. The study looked at a teen driving fatality reduction project, the Vehicle-Infrastructure Integration project, and Minneapolis's Stop on Red program to examine the ways privacy concerns are addressed.

David Levinson demonstrates how increased computing power has made agent-based modeling a tool to help planners understand how individual travel decisions add up to large-scale traffic flows. He analyzes the relationships between network supply and travel demand, and describes a road development and degeneration mechanism microscopically at the link level.

Levinson also examines the growth of a highway network based on the present and historical conditions of the network, traffic demand, demographic characteristics, project costs, and budget. The effects of expanding a link on its upstream and downstream neighbors, as well as on parallel links, are also considered.

Levinson' research concentrates on the dynamics of the orientation of major roads in a network to understand the basic properties of transportation networks. A model was developed to capture the dynamics that leads to a hierarchical arrangement of roads for a given network structure and land use distribution. He developed an agent-based travel demand model, where travel demand emerges from the interactions of three types of agents in the transportation system: node, arc, and traveler. A unique feature of the agent-based model is explored: that it explicitly models the goal, knowledge, searching behavior, and learning ability of related agents.

The STAR-TEA 21 research shows new links between transportation and technology and ways in which they influence public policy. One of the most exciting aspects of this research project has been the opportunity to explore these issues from a variety of perspectives—engineering, social policy, economics—and to share our work with those who are creating new transportation technologies and making policy decisions.

CHAPTER 1

1.1 Places and Networks: The Changing Landscape of Transportation and Technology Final Summary Report of the STAR–TEA 21 Project

1.1.1 Introduction

University of Minnesota researchers have studied the many ways in which transportation and technology intersect over the past six years. This is our final report on work done under the Sustainable Transportation Applied Research (STAR)–TEA 21 grant.

While we have been studying advances in transportation technology, advancements have continued to occur exponentially. Wireless communication has become the norm. Complex data can easily be shared across organizations, informing public policy on transportation, safety, and public policy. As modeling techniques have become more sophisticated they have led researchers to explore how decisions of individual travelers affect larger systems. This research explored the intersections of transportation and technology from many perspectives.

This research has shown the enormous potential for new technologies' contributions to society: the ability to reduce congestion, to create smarter roads that help people travel more efficiently and in a more environmentally sound way, and to respond to emergencies faster and more efficiently in order to save lives.

This research also brought other issues to the fore. How do we pay for the technologies and who pays? How do public entities work with private innovators who create new technologies? Must we sacrifice privacy and freedom as we implement technological advances? Our research has convinced each of us of how complex and layered the intersections of transportation and technology are. Transportation involves millions of decision-makers. Policymakers need to anticipate these individual agents in the process of developing policy.

Transportation and technology encompass a vast array of linkages and hierarchies, public and private, transportation-oriented and technological. A brief description of the research follows:

In Chapter 2, Richard Bolan studied *Spatial Patterns of Information Workers*. He used agent-based modeling to predict how high-technology workers influence city form—and therefore, transportation needs—through their choices about work and home location.

Bolan's study was an effort to look at the location patterns of information technology activities at the metropolitan scale. Specifically, the study focuses on the intra-metropolitan location characteristics of persons engaged in occupations that are dominated by the need for information. The study intended to shed some light on the impact of information technology on the metropolitan community in some spatial detail as well as to examine how such technological development might influence land development policies, transportation policies, environmental policies, and other social policies at the local and metropolitan scale. Bolan analyzed six U.S. metropolitan areas (Atlanta, Austin, Denver, Houston, Phoenix and Minneapolis-St. Paul) using the U.S. Census Bureau's Transportation Planning Package (CTPP) for 1990 and 2000.

Appendix B includes maps indicating the weighted mean center and the standard deviational ellipse patterns of both residential and workplace locations for workers in all six metropolitan areas.

Kevin Krizek continued his work on travel behavior impacts of information and communications technologies (ICT), exploring whether and how the internet will replace travel demand in Chapter 3.

In the section, *ICT as a Substitute for Non-work Travel: A Direct Examination*, Krizek investigated dimensions of ICT use at the household level as it applies to matters of non-work travel. His research explored (1) the pattern of substitution effect between traditional and ICT-form activities and (2) what attributes of people affect the choice of whether or not to substitute. A household survey on ICT use and substitution and investigated into the substitution aspect of household travel behavior. The results help understand the million-dollar question—to what degree will ICT substitute for physical travel.

Krizek's, *Mapping the Terrain of Information and Communications Technology (ICT) and Household Travel*, first maps the terrain of existing work to date related to ICT and household travel and identifies the predominant nature of existing study—conceptual or empirical—as well as voids in the existing knowledge base. Second, the research sheds light on emerging phenomena to help conceptualize future research by identifying and describing three dimensions by which future work should be understood: (1) the purpose of the activity (2) the effect on travel and (3) the role of sub-tasks. This research helps serve as a blueprint to help conceptualize future research and detailed lines of inquiry relating ICT and household travel.

Spatial Attributes and Patterns of Use in Household-Related Information and Communications Technology Activity, reports on how the elimination of physical travel traditionally associated with activities such as work, shopping, and entertainment, and the extent to which households use the Internet for such activities may depend on spatial attributes of the communities where they are located. This section answers questions related to the propensity of residents in metropolitan areas to use the Internet for e-commerce, e-banking, and other financial transactions and how such use is affected by spatial attributes of retail and bank accessibility and traffic congestion.

Krizek also studied *Trends of Household-Related ICT Activities*. This research focuses on household use of three categories of ICT-based activity: e-commerce, electronic banking, and other financial transactions. The aim of this study was to determine the degree to which at-home ICT use and out-of-home store travel changed from 1995 to 2003 for the selected purposes of shopping and banking.

Thomas Horan studied how intelligent transportation systems can help police, ambulance, and other public safety providers communicate more accurately and save lives in the chapter about *ITS and Emergency Medical Services Response*. His research extended information-sharing dimensions to explore information sharing relative to service performance. It utilized a time-critical information services (TCIS) conceptual framework as an analytical lens. A case study approach was employed to examine the exchange of performance-related information in a key time information critical service: a county-wide emergency medical services (EMS) system

In the section, *Devising a Web-Based Ontology for Emerging Wireless Systems: The Case of Emergency Management Systems*, Horan investigated the development of a software-based ontology within the context of a rural wireless emergency management (EMS) system. The case study investigated the utility of a new ontology-based framework for wireless emergency response in rural Minnesota. The ontology was developed by integrating concepts and findings from in-depth field reviews in Minnesota into an ontological software originating out of bioinformatics.

The next section, *Interorganizational Emergency Medical Services: Case Study of Rural Wireless Deployment and Management*, Horan continues by drawing upon complex systems theory and Interorganizational Systems (IOS) dynamics. A framework was developed for investigating technology, organizational and policy dimensions of Emergency Management Services (EMS). The case study for this investigation also took place in rural Minnesota, where a series of semi-structured interviews were conducted and supplemented by analysis of candidate EMS system evaluations. The twofold objectives of the study were to explore the nature of interorganizational dynamics in this setting and to set forth an architecture for measuring and enhancing performance. Key technology concerns were intertwined with organizational aspects.

Horan also described his study of *Performance Information Systems for Emergency Response: Field Examination and Simulation of End-To-End Rural Response Systems*, to investigated the role of information systems in enhancing end-to-end performance of rural Emergency Medical Services (EMS) systems. This study used an embedded case study approach with multiple methods and within the context of rural Minnesota. Interviews and data analysis of the EMS process identified four elements: mayday call, routing and dispatch, response, and treatment.

Finally, Horan studied *User Perspectives on the Minnesota Interorganizational Mayday Information System*. This section presents a case study of the Minnesota Mayday system, a service oriented architecture (SOA) based information system that automatically pushes select General Motors (GM) OnStar emergency data to pre-authorized emergency response and transportation stakeholders (dispatch centers, law enforcement, ambulance providers, health care facilities, traffic management centers, traveling public). The focus of this chapter is the relationship between the operational Mayday system and the behavior of emergency responders and participating organizations.

Lee Munnich, Jr. examines the ways that intelligent transportation systems help travelers move more safely, quickly, and efficiently and the policy issues surrounding the potential loss of privacy that may accompany some of these technologies. In Chapter 5, *Industry Clusters and ITS*, Munnich studies how ITS is affecting rural economies, by conducting interviews with representatives from firms in the recreational transportation equipment industry and the wood products cluster in northwest Minnesota. The primary goal of this research was to analyze ITS use from the industry cluster perspective. This chapter looks beyond the geographically constrained view of the industry cluster to assess the promise of the “value-chain” industry cluster as a construct for planning ITS architectures and operational concepts. The paper: (1) identifies and defines the value chain industry cluster as an appropriate construct for ITS planning; (2) suggests a four-step industry cluster based approach to developing integrated public and private ITS concepts of operations and architectures within the USDOT guidance for Regional ITS Architecture and; (3) offers examples of practical freight ITS issues which can effectively be addressed at the industry cluster level.

In Chapter 6, *Thinking Privacy with Intelligent Transportation Systems: Policies, Tools, and Strategies for the Transportation Professional*, Munnich looks at the how major current and future ITS technologies may affect privacy, and recommends a framework to address privacy issues when working with ITS. The study looked at a teen driving fatality reduction project, the Vehicle-Infrastructure Integration project, and Minneapolis's Stop on Red program to examine the ways privacy concerns are addressed. Although not historically a major focus of transportation professionals, privacy considerations need to become increasingly important because the public's acceptance of new ITS technologies is dependent upon the mitigation of their privacy concerns.

In the final chapter, David Levinson demonstrates how increased computing power has made agent-based modeling a tool to help planners understand how individual travel decisions add up to large-scale traffic flows. The section titled, *A Model of the Rise and Fall of Roads*, analyzes the relationships between network supply and travel demand, and describes a road development and degeneration mechanism microscopically at the link level. A simulation model of transportation network dynamics was developed, involving iterative evolution of travel demand patterns, network revenue policies, cost estimation, and investment rules.

Induced Supply A Model of Highway Network Expansion at the Microscopic Level, examines the growth of a highway network based on the present and historical conditions of the network, traffic demand, demographic characteristics, project costs, and budget. The effects of expanding a link on its upstream and downstream neighbors, as well as on parallel links, are also considered.

Induced Demand: A Microscopic Perspective, looks at the induced demand hypothesis using a disaggregate approach at the link level.

The section, *The Emergence of Hierarchy in Transportation Networks*, concentrates on the dynamics of the orientation of major roads in a network and abstractly models these dynamics to understand the basic properties of transportation networks. A model was developed to capture the dynamics that leads to a hierarchical arrangement of roads for a given network structure and land use distribution.

Levinson developed an agent-based travel demand model, described in *Agent-Based Approach to Travel Demand Modeling Exploratory Analysis*. In this model, travel demand emerges from the interactions of three types of agents in the transportation system: node, arc, and traveler. A unique feature of the agent-based model is explored: that it explicitly models the goal, knowledge, searching behavior, and learning ability of related agents.

The STAR-TEA 21 research shows new links between transportation and technology and ways in which they influence public policy. One of the most exciting aspects of this research project has been the opportunity to explore these issues from a variety of perspectives—engineering, social policy, economics—and to share our work with those who are creating new transportation technologies and making policy decisions.

CHAPTER 2

Richard Bolan: Spatial Patterns of Information Workers

2.1 Spatial Patterns of Information Workers in Six United States Metropolitan Areas

Richard S. Bolan
State and Local Policy Program
Humphrey Institute of Public Affairs
University of Minnesota
Minneapolis, MN 55455
T: (612)625-0128
E-mail: bolan001@umn.edu

Peng Xu
State and Local Policy Program
Humphrey Institute of Public Affairs
University of Minnesota
Minneapolis, MN 55455
E-mail: xuxx0135@umn.edu

2.1.1 Introduction

With the full blossoming of the information age in the 1990s, much study and research has gone into analyzing the location of high-technology industries and workers. However, these works have primarily focused on studies at the scale of the nation or large regions within the nation. This study is an effort to look at the location patterns of information technology activities at the metropolitan scale. Specifically, the study focuses on the *intra*-metropolitan location characteristics of persons engaged in occupations that are dominated by the need for information. In effect, the study is an examination of the producers and end-users of information technology—those people engaged in occupations that depend most heavily on the artifacts of the new information economy. Do information workers locate in concentrated fashion or in a highly dispersed fashion? Do they congregate in central downtown areas, in “edge city” areas, or do they disperse randomly throughout a metropolitan area? In terms of transportation demand, do information workers show different commuting patterns than non-information workers and, if so, how has this changed between 1990 and 2000? If the workplaces of information workers follow different patterns than non-information workers, does this also hold true of their residential settlement patterns?

These questions are salient to the growing attention to land use planning at local and regional scales where current patterns of land development are having an impact on environmental degradation, traffic congestion, loss of viable farmland, and social dispersion. Sprawling urban land development in the 1990s occurred in the midst of the virtual 1990s explosion of information technology (IT). Yet little is known as to what effects, if any, information technology had on the urban land market in that decade. Some have argued that IT workers need to be in close proximity, while others claim that such workers can be located anywhere—even in remote rural areas. Consequently, the present study is intended to shed some light on the impact of information technology on the metropolitan community in some spatial detail as well as to examine how such technological development might influence land development policies, transportation policies, environmental policies, and other social policies at the local and metropolitan scale.

In this study, we analyzed six U.S. metropolitan areas (Atlanta, Austin, Denver, Houston, Phoenix and Minneapolis-St. Paul) using the U.S. Census Bureau’s Transportation Planning Package (CTPP) for 1990 and 2000. Our reasoning for selecting these is provided in more detail in Section 2.1.3 but, generally, these regions had roughly similar rates of growth during the

1990s and were generally free from geographical constraints (such as an ocean, a mountain range or one of the Great Lakes). Four are southern cities and two are more northern mid-western cities.

In Section 2.1.2 we provide a brief review of background literature together with our reasoning behind our primary hypothesis. In Section 2.1.3 we define our use of the terms information technology and information workers and the rationale behind selecting the six metropolitan areas we have studied. Also in Section 2.1.3 we address the specific analytical methods employed. Section 2.1.4 provides an overview of the changes in occupational composition in the six metropolitan areas as well as the results of applying the geostatistical methods and the changes found from 1990 to 2000. Section 2.1.5 describes the results of our employment center analysis for 1990 and 2000. Section 2.1.6 provides a discussion and overall interpretation of the findings and our general conclusions as to the contribution of the study, suggested further research, and general policy implication for land use and transportation planning.

2.1.2 Background

Machlup (1962) was perhaps the first to identify the importance of information and knowledge as a driving force in economic growth and development in the early 1960s. This observation was made despite a relatively primitive technology (by today's standards) in telecommunications and computers at that time. By the 1980s, however, development of the technologies that would greatly enhance both the development of new knowledge, and its dissemination, was clearly seen as an engine of economic growth. At the same time, there was a sense that the location of economic activity no longer mattered. High-speed transmission of voice, data, and images offered a speculation that one could carry on activities at any spot in the globe. As futurist Alvin Toffler (1980) forecast, the "Information Age" would spell the "end of cities." As an illustration, there was the notion that a surgeon could operate on a rural patient without leaving his urban office. At one point in the late 1980s Camden, Maine, proclaimed itself as the "software capital" of the nation. The image was projected of young entrepreneurial software developers sitting on their schooners in Camden Harbor and developing products they could then transmit anywhere on earth. For those stuck in traffic, the new economy hinted that many would work, shop, transact banking, etc. without leaving home. The networking of computers and telecommunications media could be seen as a substitute for local ground travel, resulting in significantly reduced transportation demand and thereby lessened congestion and cleaner air.

More analytical scholars countered these extreme views and suggested reasons why the new economy has not meant the end of cities or of economic agglomeration. Indeed, information technology may well involve a resurgence of urban centers. Location still matters, but the age-old location variables—transportation costs, labor supply, etc.—may have to be supplemented by new location factors, such as proximity to knowledge centers, level of competitiveness, and effective means for diffusion of innovation (Porter, 1990). Indeed, one author suggests that the so-called "New Economy" is founded on what he calls "spatial technologies" referring to the whole complex of transportation, communications, and information technology (Coucilelis, 1996 quoted in Shen, 1999).

With the speed and flexibility of contemporary telecommunications networks, there may be situations where transportation and communication networks are in competition with each other and the communications network may be a rational substitute for the transport network.

However, Graham and Marvin (1996, 327–333) caution that the interrelations between telecommunications and transportation are not simple. They are at once both complementary and competitive. Both are more concentrated in urban areas and demonstrate considerable physical proximity. They also argue that telecommunications advances may actually stimulate travel:

> ... evidence actually points in the reverse direction to the myth of simple substitution. Three key areas of telecommunication-transport innovation currently suggest that telecommunications either generate more transportation than they substitute for, or allow rising transport demands to be accommodated and managed. (P. 331)

"Just-in-time" management is one example that can result in substantial increases in cargo trips. Wireless telematics means that the hitherto "dead" time of traveling is now available for working while traveling. Additionally, the new telecommunications networks generate entirely new forms of social and community interactions thereby augmenting travel opportunities. Finally, telecommunications can enhance the efficiency and effectiveness of transport networks, along with improved infrastructure of all kinds (energy, water supply, airlines, etc).

A growing number of studies seek to come to grips with this new awareness of communication networks (in this context is meant both the physical transportation network and the telematics network [Graham and Marvin, 1996], or what Castells [2000] terms the "space of flows"). Castells (2000), Markusen et al (1987); Glassmeier (1985), and Hackler (2000) looked at the locational aspects of what they define as "high tech" employment. Two conflicting postulates emerge conceptually. First is the idea that the technological innovations of the past two decades mean the decline of urban areas; both firms and households have become more footloose. Evolving cyberspace implies the waning importance and meaning of physical space. The contrary position is the notion that contemporary telecommunications actually strengthens central cities rather than peripheral areas and proponents of this view point to the greatly enhanced networks centered in the world cities of New York, London, Tokyo, Chicago, Los Angeles, etc. to bolster their case.

Interesting studies of the spatial patterns of World Wide Web URL addresses have reinforced the notion that urban metropolitan areas are the locus of the new information economy. Warf compiled an interesting map of global distribution of internet hosts per 100,000 persons in 1997 (2000, p. 60) showing that internet addresses have their highest density in developed countries. A number of authors have undertaken more detailed study of the locational patterns of internet addresses. Dodge and Shiode, 2000 mapped the geography of the internet in the United Kingdom. Moss and Townsend, 1997, as well as Zook, 1998, carried out similar studies in the United States. In each study, the predominance of urban areas as the physical locus of internet web pages was graphically evident. A review of the literature by Gorman (2002) observed: "communications infrastructure has disproportionately agglomerated in the largest metropolitan regions." He further concludes: "the internet is not acting as the great geographic equalizer; instead the internet is increasingly falling into a more distinct urban hierarchy."

Most of these studies primarily examine metropolitan areas in the aggregate and do not explore evolving *intrametropolitan* patterns. A recent case study by Leigh (2000) of the industrial and office market of Chicago and Atlanta suggested that so-called "edge cities" actually benefit more from the new telematics than either central downtowns or peripheral suburbs. The office market of Leigh's study is clearly a dominant end-user of new information technology but other information dependent activities may still rely on different patterns that are more efficient or effective. Manufacturing information technology hardware or other scientific hardware (such as medical technology) comes to mind as perhaps following different intrametropolitan location patterns. These activities may benefit outer suburbs. Education or health services may be information-dependent activities that might benefit central downtown areas.

In addition, examining standard industrial sectors may not provide as useful a means of analyzing the intrametropolitan implications of information technology. Recent studies have suggested that analyzing worker occupations rather than worker employers might provide a deeper understanding of economic development processes in the information age. Traditionally, studies of the impact of "high-tech" activities examined those sectors of the economy directly associated with the development, manufacturing and marketing of telecommunications products and services. However, as these products and services have been introduced into almost all sectors of the economy, firms having little to do with the "High-Tech" sector may still have persons in-house skilled in a wide variety of telecommunications services. As Markusen, et al (2001) and Fesler (2002) argue, analyzing what people *do* rather than what they *make* may well be a more productive research objective. This would seem especially true in trying to analyze the impact of information technology on the spatial form of cities. Since information technology has entered virtually all sectors of the economy, those who create, analyze, interpret, and transmit information and those who use information for decision-making have penetrated virtually all sectors. Recent studies of the petroleum industry (Rauch, 2002) and the food-processing industry (John R. Baldwin and David Sabourin, 1999; Bolan and Martin, 2003)—manufacturers of non-durable goods—bear this out. Telecommunications using computer software has permeated the retail sectors, the transportation sectors and most activities of the service sector. Inventory control, customer transactions, financial accounting, branch office coordination, marketing analyses, dispatching and scheduling, strategic planning and trend studies have become an integral part of virtually all sectors. Even local barbershops have benefited from the introduction of computer software.

Consequently, this study is an examination of the intrametropolitan spatial patterns of workers involved in occupations that enable us to identify them as Information Workers—regardless of the economic sector they may be employed in.

The primary question for this study is: are information workers—the developers and end-users of information technology—more likely to concentrate or disperse in contemporary emerging spatial patterns of metropolitan economic activity? Can these patterns be graphically and statistically described such that we can discern their influence on economic activity in general as well as on the demand for land and for transportation and communication services? It is acknowledged that answers to these questions depend a great deal on where their employers are located. However, it can also be seen that employers may alter location decisions based on the availability of communications infrastructure and well-trained information workers. More important, however, is the fact that a computer systems analyst can be found working for a food processing manufacturer, a chain of pharmacies or other retail stores, a logging company or a host of other economic activities not generally associated, in the main, with information technology.

Overall, we postulate that there are both centripetal and centrifugal forces at work on the intrametropolitan location of information workers. We, however, argue that the forces of concentration have stronger pull than the forces leading to dispersal. Even in times of economic slow-down (as has been experienced from 2001–2003), the marriage of telecommunications and computer technology continues to spawn innovation and change. Our argument is that information workers are obliged to keep their employers and their non-information workers competitive and efficient—in very broad terms. This includes improving the nature of financial, managerial, technical and operational decision-making; and improving the capacities of line workers (including their productivity which also means having concern for their education, health and mental health). In order to do this, there is a need for information workers to remain close to the key sources of innovation and to their professional associations. And while the urban area of today might be seen to have a radius of 50–60 miles, these are still difficult distances when there is need for face-to-face communication. In short, the social and economic circumstances that dictate a need for information workers to be reasonably proximate are stronger than the lure of open countryside.

2.1.3 Methods

Defining the Terms of the Study

Definition of information workers. Many prior studies attempting to capture the locational dynamics of the "New Economy" and the "Information Age" focused on certain industrial sectors. These studies were often seeking to understand the forces of emerging economic development from a regional perspective. One noted early study, in an effort to capture the "high technology" dimension of information technology, used industries possessing a relatively high proportion of scientists and engineers in their workforce (Markusen, Hall and Glasmeier, 1986). Subsequent studies also concentrated attention on industrial sectors. Recently, however, research has shifted to analyzing occupational characteristics. This has occurred because the growth dynamic of information technology has diffused into the general economy with much greater speed than could be accounted for in research activities. As one author suggests, what people "do" is becoming as important, if not more so, that what people "make."(Feser, 2001; see also Markusen, et. al, 2001) In this shift, many sectors of the economy are now given attention after being heretofore overlooked. This is particularly true of many service sector activities such as professional services, education, health, entertainment, design and financial services.

Thus, this study is concerned with the spatial patterns of individual workers who can be considered "Information Workers"—that is workers whose occupations involve them (1) in the creation and production of information; (2) in the collection, processing and interpretation of information; (3) in the transmission of information; (4) in the manufacture of information equipment; (5) in using information in training and education; and (6) in the use of information for planning and decision-making (Kurasaki and Yanagimachi, 1992). Thus, we have taken a broad view of the variety of occupations that can be considered "Information Workers."

The primary data source for the study has been the U.S. Bureau of the Census Transportation Planning Package. The 2000 version of this has followed a breakdown of occupations that fits with the Department of Labor's Standard Occupational Code. This allows us to follow the above classification of information workers with reasonable broad conformity. Unfortunately, the 1990 version had only crude generalized categories of occupations. We have, nonetheless, captured the vast majority of the concerns spelled out in the previous paragraph. This does limit somewhat our ability to analyze changes between 1990 and 2000, but it does offer an opportunity to provide a more precise portrayal in 2000.[i]

The occupations that are included in our definition of "Information Workers" include the following:

- Executive, administrative and managerial occupations
- Professional specialty occupations.
- Technicians and related support occupations.
- Administrative support occupations.

All other occupations comprise our category of "Non-information Workers." (See Appendix A for the full typology of occupations used in the CTPP.)

Selection of study areas. The analysis of *intra*-metropolitan location patterns for information workers has been carried out for six U.S. metropolitan areas. These areas were selected on the basis of a number of criteria. First and foremost, metropolitan areas were selected based upon their percentage of information workers as determined in a study by Markusen, et. al. (2001). The authors of that study used the term “I-tech occupations.” They included persons defined as scientific or engineering workers as well as systems analysts, database administrators, computer professionals and computer scientists. The study also included persons involved in engineering and architectural services, photographic equipment, plastics, medical instruments and research, testing, and evaluation services. They also added occupations devoted to financial services. With this as the basis of their definition of “I-tech” occupations, they then ranked the metropolitan areas of the United States in terms of the percentage of such workers related to total employment. We chose middle range metropolitan areas for our study, generally between 10 and 15% of the total labor force. An additional criterion was a geographic one, where we tried to select metropolitan areas that had no significant geographic barriers to development at any point of the compass (such as oceans, mountain ranges, one of the Great Lakes, etc.). The final areas chosen were:

Atlanta, GA	10.2% of total labor force “I-tech” job share,
Austin, TX	11.5%,
Denver , CO	14.5%,
Houston, TX	7.0%,
Phoenix, AZ	13.0%, and
Minneapolis-St. Paul, MN	15.3%.

Table 2.1.1 provides a comparison of the six metropolitan areas by general demographic character in 2000 using U.S. Census data. Also shown is a graph charting the trend lines of population for the six metropolitan areas for 1980, 1990 and 2000. This chart indicates roughly similar patterns of population growth, although Atlanta and Phoenix exhibited somewhat stronger growth than did Denver and the Twin Cities of Minneapolis and Saint Paul.

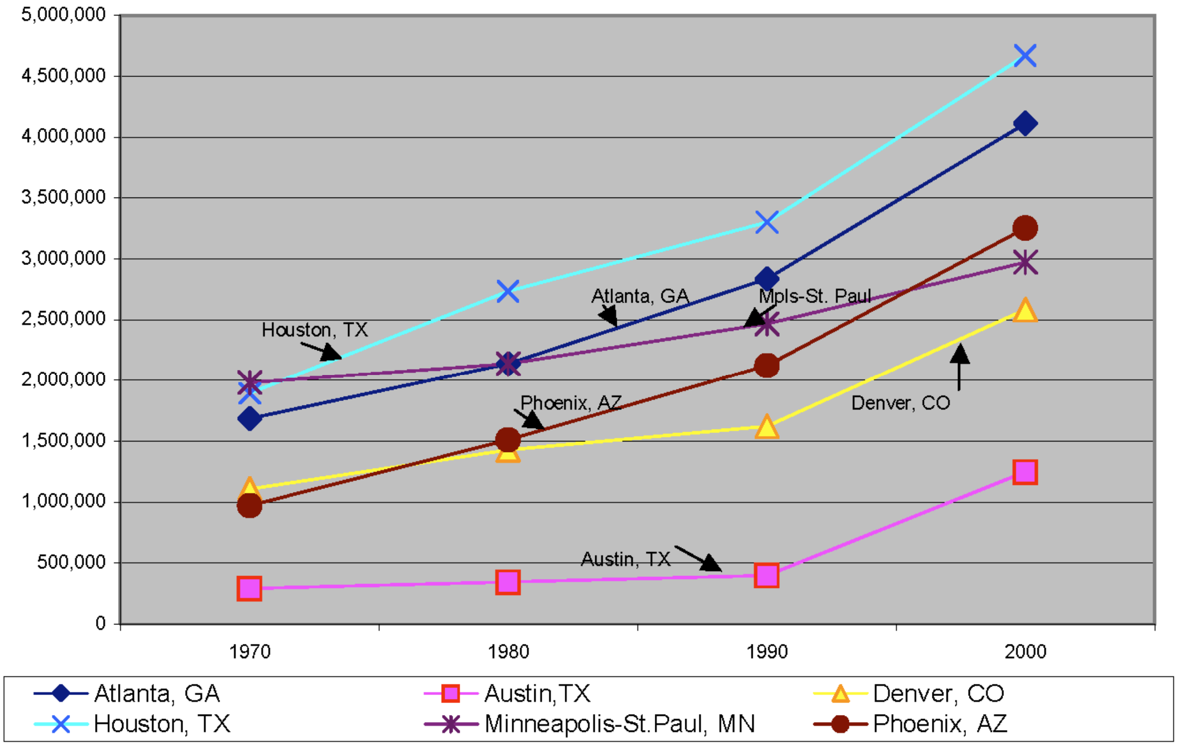


Figure 2.1.1 Trends of total population of six metropolitan areas: 1970–2000

Table 2.1.1 Basic demographic characteristics of study metropolitan areas

Metropolitan Area	Total Population, 2000	Median age (years)	% under 18 years	% 65 years and over	% 18–64	Total households	Average household size
Atlanta, GA MSA	4,112,198	32.9	26.6%	7.6%	65.8%	1,504,871	2.68
Austin, TX MSA	1,249,763	30.9	25.4%	7.3%	67.3%	471,855	2.57
Denver–Boulder,–Greeley, CO CMSA	2,581,506	33.8	25.7%	8.9%	65.4%	1,003,218	2.53
Houston, TX CMSA	4,669,571	31.9	29.0%	7.7%	63.3%	1,639,401	2.80
Minneapolis-St. Paul, MN–WI MSA	2,968,806	34.2	26.7%	9.6%	63.7%	1,136,615	2.56
Phoenix–Mesa, AZ MSA	3,251,876	33.2	26.8%	11.9%	61.3%	1,194,250	2.67
Average	2,306,408	32.8	26.7%	8.8%	64.5%	979,235	2.64

Methods

The data source for the study was almost exclusively the U.S. Bureau of the Census, Census Transportation Planning Package (CTPP) [http://www.fhwa.dot.gov/ctpp/]. The general approach has been to analyze the patterns of residential and workplace locations for both information workers and non-information workers. The analysis has been carried out through various spatial statistical analytic methods. Using software developed by Levine (2002), combined with ArcView GIS software, we were able to identify differences in the spatial patterns of both sets of workers using traffic analysis zones (TAZs) as the geographical unit of analysis. Traffic analysis zones are small areas similar in size to Census tracts.[ii]

Centrographic analysis. Assuming the centroids of the TAZs provide the basis for treating worker locations as a point distribution, we developed maps indicating the weighted mean center and the standard deviational ellipse patterns of both residential and workplace locations for both sets of workers in all six metropolitan areas. Developing these maps allowed comparisons of the locations of the weighted mean center workplace and residential patterns vis-à-vis the central downtown of each metropolitan area. The standard deviational ellipse is defined as the measure that summarizes a point pattern in terms of an ellipse rather than a circle. The ellipse has the weighted mean center as its center. The long axis represents the direction of greatest dispersion and the short axis is the direction of minimum dispersion (with the short axis always at right angles to the long axis). The more the distribution of points is equal in all directions, the more the ellipse approaches a true circle.

The standard deviational ellipse size as well as the angle of the major axis was plotted. Ellipses were drawn for both one-standard deviation (enclosing the distribution of 2/3 of the information and non-information workers) and two standard deviations (enclosing 95% of the workers). These comparisons were carried out for both 1990 and 2000, giving a rough trend pattern in the spatial distribution. These statistics provide estimates of the degree of concentration of information workers relative to non-information workers as distributed in the entire metropolitan area as well as their relationship to the central business district of the region. A significantly smaller ellipse for information workers relative to non-information workers (both in terms of length of axes and enclosed area) would indicate information workers are more concentrated.

Spatial autocorrelation analysis. Following this, we carried out an analysis of spatial autocorrelation to determine if the scattering of information workers in each urban area was purely random or if there was a probability that the spatial distribution reflected a tendency to be concentrated or to be distributed in some organized or regularized pattern.

Spatial autocorrelation analysis describes how an attribute is distributed over space. It is an indicator of the extent to which the value of a variable in one zone depends on the value of that variable in neighboring zones." (Frothingham, et al, 2000, 12). As stated by Lee and Wong:

> Spatial autocorrelation of a set of points is concerned with the degree to which points or things happening at these points are similar to other points or phenomena happening there. If significantly positive spatial autocorrelation exists in a point distribution, points with similar characteristics tend to be near each other." (2001, 78)

Thus, if the distribution of information workers is positively spatially autocorrelated, a TAZ with a high concentration of information workers would tend to be near other TAZs with high concentrations of similar workers.

The most widely used statistic measuring spatial autocorrelation is Moran's "I."

In Moran's I, the similarity of variable values is the difference between each individual value and the mean of all values for the variable in question, or:

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} c_{ij}}{s^2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}$$

and

$$c_{ij} = (x_i - \bar{x})(x_j - \bar{x})$$

thus,

$$I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{s^2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}}$$

The sample variance is given by:

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n}$$

The sample variance is given by:
The results of this calculation is compared with the expected value of Moran's "I" if "I" were to express a perfectly random distribution of the variable under study. Thus, the notation E(I) reflects the expected index value, and:

$$E(I) = \frac{-1}{n-1}$$

When:

I > E(I), the result reflects a clustered pattern where adjacent TAZs show similar characteristics;
I = E(I), the result reflects a random pattern where TAZs do not show patterns of similarity; and
I < E(I), the result reflects a dispersed or uniform pattern where adjacent TAZs show different characteristics

The different possible results are illustrated on the diagram below. The statistical significance of the results are given by the Z score:

$$Z = \frac{I - E(I)}{VAR(I)}$$

If -1.96 < Z <+1.96, then the observed distribution is not significantly different than a purely random pattern.

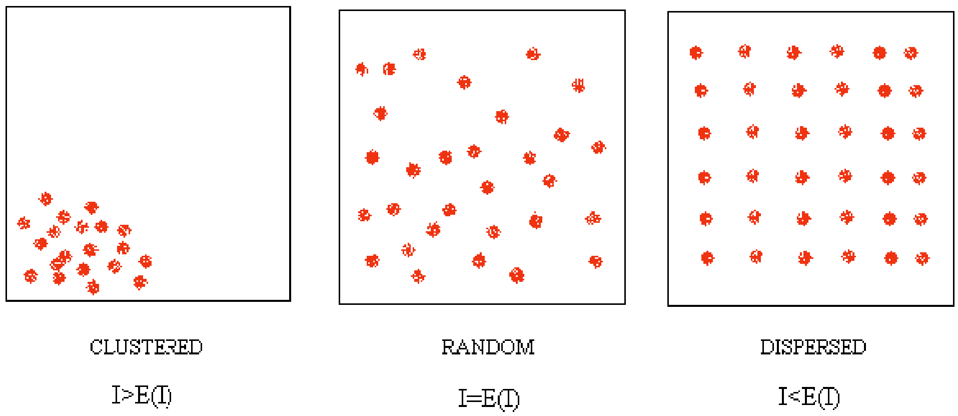


Figure 2.1.2 Spatial pattern possibilities from Moran's "I"

Workplace Locations Proximity to Their Mean Center

Another analysis was carried out to provide additional evidence of the variation in spatial distribution of information work places. Essentially, the analysis examined the relation of work places of information workers to the mean center of all work places for such workers. This was done by taking individual workplace TAZs and dividing them by the distance of their centroids to the weighted mean center of all information worker TAZs. This provided a means for a two-sample test as to whether there were significant differences between information workers and non-information workers in both their residential and workplace distributions.

This calculation is give by the following equation:

$$P = \frac{\sum_{i=1}^{n} \frac{w_i}{d_i}}{n}$$

Commuting Effort

A final statistical analysis focused on the measure of commuting effort as reported in the CTPP data. The indicator of this was the mean travel time to work for 1990 and 2000.

Identification of Major Employment Centers

As indicated above, the data set also permitted us to identify major employment centers in two distinct ways.

- For each metropolitan area, an average employment density was calculated based on total employment and total metropolitan land area (the aggregate sum of all TAZ areas as provided in the CTPP package). TAZs that had employment densities above the metropolitan average were identified as major employment centers. Where above-average density TAZs appeared contiguous to one or more similar TAZs, the contiguous areas were grouped together as a single employment center. From this, we classified the employment centers by the following typology:
    - Traditional Central Business District
    - University Center

- Edge City Center (identified as being within one mile of an interstate or other major freeway)
- Other Center (not otherwise classified).

Next, the ratio of information workers to non-information workers at each identified center was determined. Highlighted were those employment centers that had more than 60% of their workers classified as information workers, thereby providing a graphic depiction of the degree of clustering of information workplaces in or around the traditional central business district.

- We also carried out a "hot spot" analysis using the CrimeStat2 software.

"Hot spot" analysis, as its name implies, provides statistical techniques for determining concentrations of activities within a geographic area. This analysis was carried out as a check on the method outlined above. This type of analysis can be carried out for both point location data as well as zonal data. Since the CTPP works from the geographic unit of the Traffic Analysis Zone, the applied technique in this study was limited to an analytical technique known as Anselin's Local Moran's "I," also known as a local indicator of spatial association. This indicator depicts the extent to which the value of a zone is similar or different from observations of neighboring zones. It is an indicator that is valuable not in terms of the absolute value of the attribute of a zone but rather it is an indicator of relative similarity of neighboring zones. The distance of any given zone from any other zone is a factor of weighting in the calculation of the Local Moran's I—adjacent zones having greater weight for the calculation of similarity than more distant zones. In short, a traditional distance decay function is applied in the determination of the Local Moran's I.

Thus, the Local Moran's I is given by the following equation:

$$I = \frac{(Z_i - \bar{Z})}{S_z{}^2} * \sum_{j=1}^{N} \left[W_{ij} * (Z_j - \bar{Z}) \right]$$

where $\bar{Z}$ is the mean number of workers per zone over all zones,

Z_i = the number of workers in zone i,

Z_j = the value of workers for all other observations j (where j $\neq$ 1),

$S_z{}^2$ = is the variance over all zones, and

W_{ij} = the distance weight for the zones i and j. It is defined as:

$$W_{ij} = \frac{1}{d_{ij}}$$

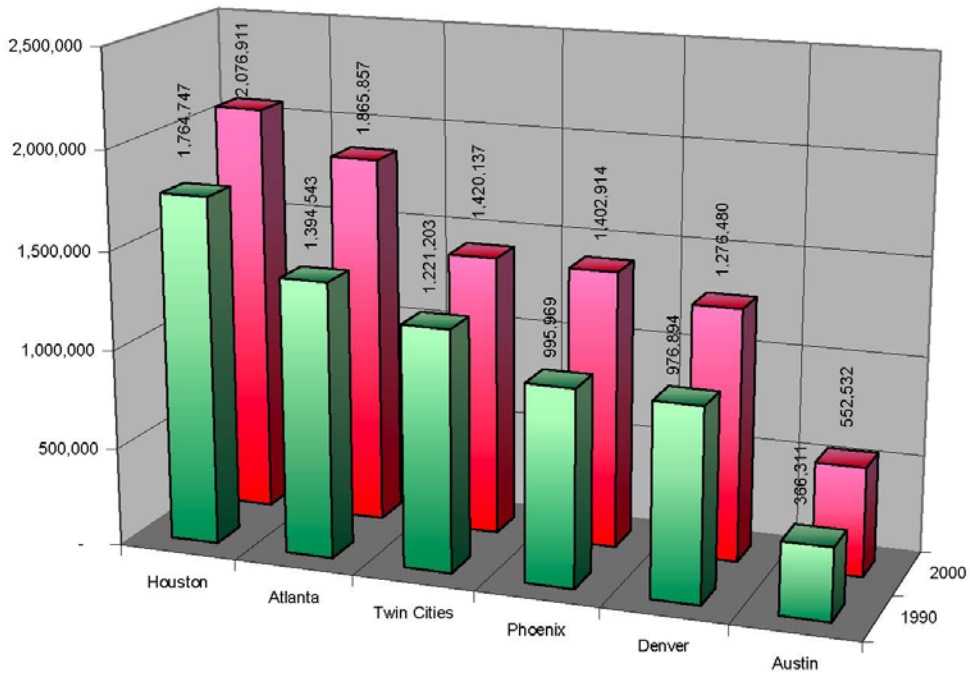
The local Moran's I indicates both "hot spots" and "cold spots"—i.e. a hot spot would be where the values of a zone are much higher than in nearby zones, while a cold spot would be where the values of a zone are much lower than in nearby zones.

2.1.4 Geostatistical Results

Overall Changes in the Occupational Composition of the Six Metropolitan Areas: 1990–2000

The decade of the 1990s generally saw the full emergence of the information economy and the data from these six metropolitan areas vividly portrays this. As discussed above, previous studies have shown that the activities involved in the information economy have tended to concentrate in metropolitan areas. In our analysis of two basic occupational categories of information workers versus non-information workers, it is apparent from the changes in occupational composition in the six metropolitan areas that these information-technology-dependent occupations were the driving force in the growth of all six economies.

Each metropolitan area saw growth in the overall number of employed workers as summarized from the Census CTPP data. In terms of overall percentage change, Austin and Phoenix advanced the fastest growth rates; Houston and Minneapolis-St. Paul had the lowest rates of growth. The table and chart below provide the data from the Census Transportation Planning Package for 1990 and 2000 by our two-way classification of information workers and non-information workers. The table and chart vividly portray the significant growth of workers in the information occupations that occurred in each of the areas. In both absolute and relative terms, the growth of information workers far outstripped the growth of other workers. This was most dramatically illustrated in the Minneapolis-St. Paul metropolitan area, where six times as many new information workers were recorded as contrasted with other workers. For the other three metropolitan areas, the growth in information workers tended to be to 2 to 2.5 times that of other workers.



	Houston	Atlanta	Twin Cities	Phoenix	Denver	Austin
1990	1,764,747	1,394,543	1,221,203	995,969	976,894	366,311
2000	2,076,911	1,865,857	1,420,137	1,402,914	1,276,480	552,532

Figure 2.1.3 Total employment in six metropolitan areas: 1990 and 2000

Source: U. S. Bureau of the Census, Census Transportation Planning Package, 1990 and 2000

Table 2.1.2 Changes in occupational distribution in six metropolitan areas: 1990 To 2000

			Absolute	**Percent**
	1990	2000	Change	Change
ATLANTA				
Information Workers	771,394	1,102,239	330,845	42.9%
Non-Information Workers	618,262	762,518	144,256	23.3%
Total Workers	1,389,656	1,864,757	475,101	34.2%
AUSTIN				
Information Workers	217,747	342,443	124,696	57.3%
Non-Information Workers	148,564	210,089	61,525	41.4%
Total Workers	366,311	552,532	186,221	50.8%
DENVER				
Information Workers	554,674	759,835	205,161	37.0%
Non-Information Workers	418,486	514,959	96,473	23.1%
Total Workers	973,160	1,274,794	301,634	31.0%
HOUSTON				
Information Workers	929,603	1,139,911	210,398	22.6%
Non-Information Workers	835,144	937,000	101,856	12.2%
Total Workers	1,764,747	2,076,911	312,164	17.7%
PHOENIX				
Information Workers	521,480	773,178	251,698	48.3%
Non-Information Workers	468,018	628,257	160,239	34.2%
Total Workers	989,498	1,401,435	411,937	41.6%
MINNEAPOLIS-ST. PAUL				
Information Workers	677,220	850,252	173,032	25.6%
Non-Information Workers	540,636	567,879	27,243	5.0%
Total Workers	1,217,856	1,418,131	200,275	16.4%

Source: U.S. Bureau of the Census, Census Transportation Planning Package, 1990 and 2000.

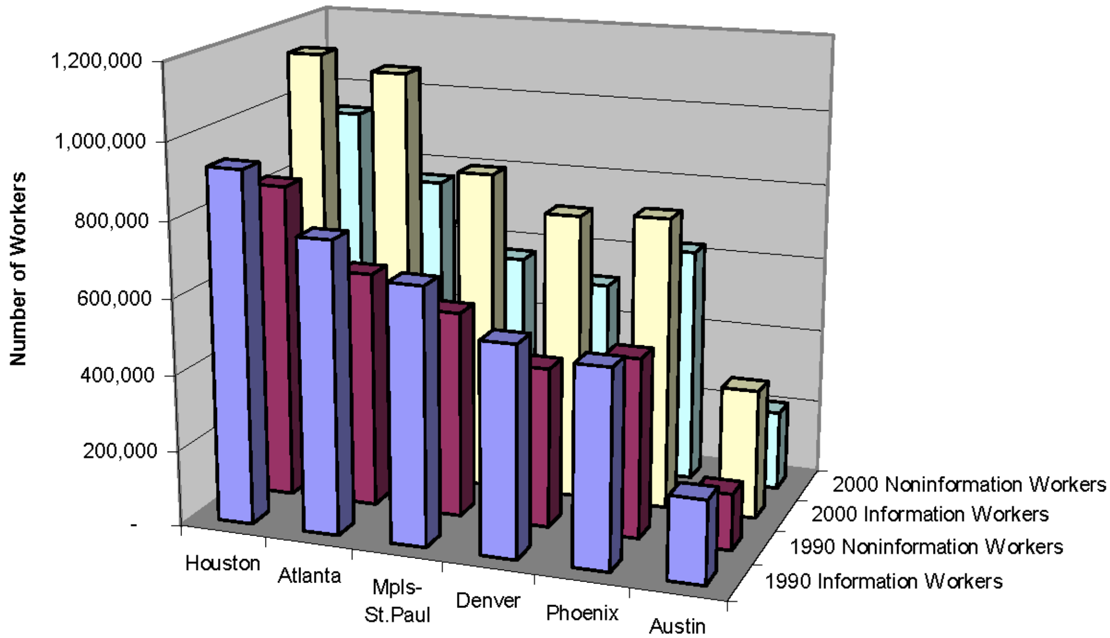


Figure 2.1.4 Changes in occupational composition in six metropolitan areas: 1990–2000

Source: U. S. Bureau of the Census, Census Transportation Planning Package, 1990 and 2000.

Information workers in the Austin area grew by the greatest relative amount at 57% followed by the Phoenix area with a 48% increase in the decade. Information workers in Atlanta grew by the greatest absolute amount overall, adding 330,000 new information workers. The Minneapolis-St. Paul area saw the greatest shift to information workers where 86% of all new workers were classified as information workers. In the Twin Cities in 1990, information workers made up 55% of all workers; in 2000 that had increased to 60% of all workers.

The aggregation of occupations into only two classes—information works and non-information workers—is not fully illustrative of the impact of information technology on these six metropolitan areas. The following charts indicate the percentage change in each based on a somewhat more refined classification—the two-digit BLS classification. These charts illustrate generally that the key drivers of change in occupations in each metropolitan area during the decade were computer and mathematical operations, business and financial operations, and management occupations. Other information dependent occupations (such as education, law, health and arts and entertainment) enjoyed modest growth somewhat in keeping with the number of non-information occupations. Occupations that were stagnant or tended to lose workers were farm related occupations, manufacturing and health support occupations. Lack of growth in these occupations was evident in all six metropolitan areas. Thus, in each metropolitan area, the occupational composition of the work force has dramatically shifted toward the activities of the information economy.

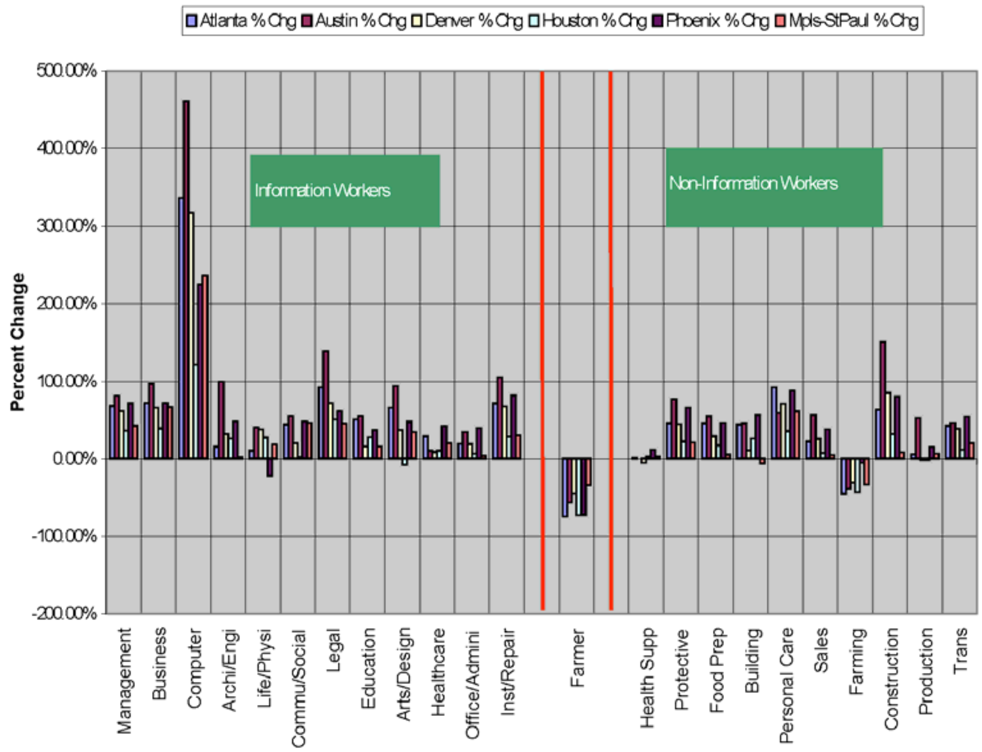


Figure 2.1.5 Percent change in all occupations in six metropolitan areas: 1990–2000

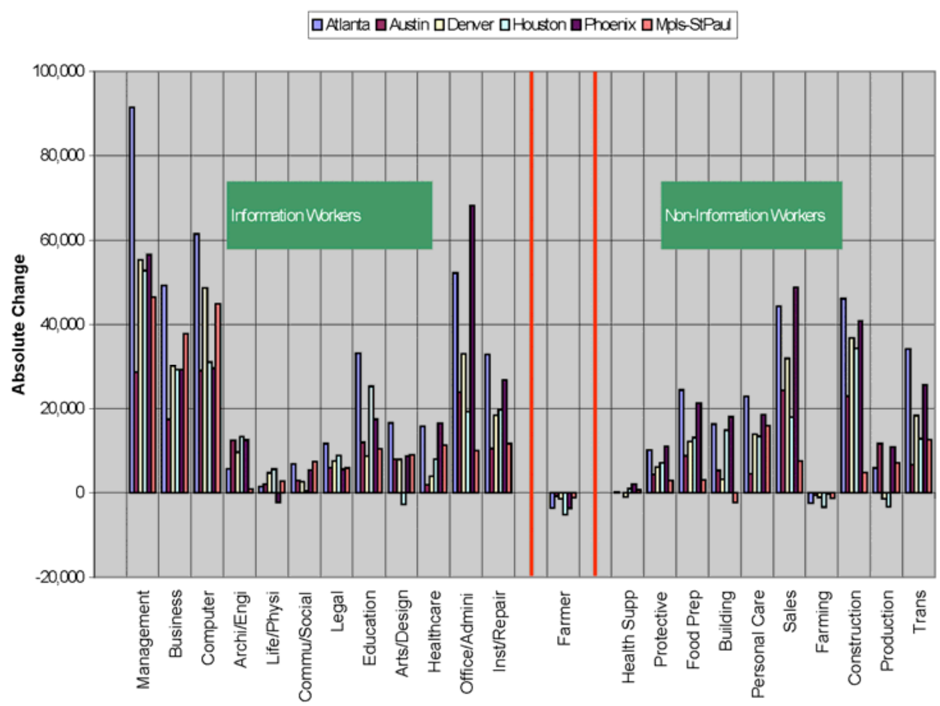


Figure 2.1.6 Numerical change in all occupations in six metropolitan areas: 1990–2000

Another important aspect of the change in these various occupations lies in their relative importance in relationship to the occupational distribution of the labor force nationally. As a calculation for this question, we used a concept similar to the use of the "location quotient" used for analyzing the importance of industrial sectors to a local economy. Thus, we determined the "location quotient" in 1990 and 2000 for each two-digit occupation, while highlighting those that seemed to show the most promising growth: management occupations, business and financial operations, and computer sciences/mathematical occupations. Thus, the general equation for determining the "location quotient" for each occupation is:

$$LQ_i = \frac{\frac{o_{li}}{\sum O_l}}{\frac{o_{ni}}{\sum O_n}}$$

Where:

LQ_i = The location quotient for Occupation "i",

o_{li} = The number of workers in occupation "i" in the local metropolitan area "l",

$\sum O_l$ = The total number of workers in all occupations in the local metropolitan area "l",

O_{ni} = The total number of workers in occupation "i" in the nation, and

$\sum On$ = The total number of workers in all occupations in the nation.

On the basis of conventional economic base theory, results that are less than one can be interpreted as signifying an activity that primarily accommodates the demand for the services of that occupation within the local metropolitan area. The ratio of workers in that activity to all workers in the local area is less than the ratio of such workers to all workers in the nation as a whole. A result greater than one suggests services of an occupation that is more than what would be expected for serving local demand alone. In other words, with such a result, the local area has more than its share of such workers and it is thereby presumed that their services have a broader market than merely local demand—in some sense, these would be services that are exported beyond the immediate metropolitan area.

The results of this analysis are shown on the two tables below. Overall, information workers hold a greater relative position to their local areas than do all such workers in the national economy. This was the case in both 1990 and in 2000. In aggregate, information workers also improved their relative position in 2000 over that in 1990. The second table depicts the changes in location quotient results for selected occupations in the decade.

In virtually all of the key information oriented occupations, these six metropolitan areas had workers that exceeded purely local demand for their services. The table shows manufacturing production workers in the last row to illustrate the relative position of such workers to information-based workers. None of the six metropolitan areas were dominated by manufacturing workers in any sense. Location quotients in each case were well below one, in contrast with almost all of the information-based occupations whose location quotients were well above one.

Table 2.1.3 Location Quotients for selected occupations in six metropolitan areas: 1990 and 2000

	Atlanta		Austin		Denver		Houston		Phoenix		Twin Cities	
	1990	2000	1990	2000	1990	2000	1990	2000	1990	2000	1990	2000
All Information Workers	**1.101**	**1.119**	**1.195**	**1.182**	**1.144**	**1.145**	**1.050**	**1.053**	**1.057**	**1.056**	**1.103**	**1.136**
Management Occupations	1.304	1.339	1.322	1.239	1.265	1.297	1.111	1.111	1.124	1.117	1.167	1.202
Business, Financial Occupations	1.292	1.402	1.305	1.375	1.251	1.360	1.103	1.165	1.121	1.154	1.163	1.453
Computer, Mathematics Occupations	1.024	1.660	1.354	2.394	1.248	2.009	1.120	1.108	1.078	1.234	1.166	1.723
Artists, designers	0.963	1.113	1.246	1.429	1.186	1.176	1.059	0.823	1.022	0.997	1.083	1.204
Manufacturing, Production Workers	0.732	0.660	0.610	0.669	0.674	0.583	0.795	0.797	0.763	0.708	0.871	0.934

Table 2.1.4 Percent change in location quotient: 1990 to 2000

	Atlanta	Austin	Denver	Houston	Phoenix	Twin Cities
All Information Workers	1.63%	-1.09%	0.11%	0.27%	-0.10%	2.96%
Management Occupations	2.68%	-6.29%	2.54%	0.03%	-0.61%	3.00%
Business, Financial Occupations	8.51%	5.37%	8.69%	5.67%	2.90%	24.89%
Computer, Mathematics Occupa	62.11%	76.86%	60.96%	-1.08%	14.46%	47.78%
Artists, designers	15.58%	14.76%	-0.83%	-22.27%	-2.39%	11.17%
Manufacturing, Production Wor	-9.84%	9.68%	-13.50%	0.25%	-7.24%	7.20%

Thus, it can be seen that, by every measure, the dynamics of growth in information-dominated occupations was the key driver of labor force expansion in each of the six metropolitan areas. Between 1990 and 2000, all six of the urban areas increased their share of information workers relative to their local economies. Four of the six increased their share of information workers relative to the national economy. The computer and mathematical occupations were the primary driving force behind the growth of all six areas, followed by business, financial and managerial occupations.[iii]

2.1.5 Geospatial Results

Centrographic Statistics

The results of the centrographic analysis are given in Table 2.1.2 The x and y coordinates of the location of the mean center were calculated for the places of residence and the places of work for information workers and non-information workers for all six metropolitan areas.[iv] Each of the six metropolitan areas has a map portfolio of four maps with a separate portrayal of information workers' workplace and residence distribution patterns as well as the workplace and residential patterns for non-information workers. These maps are presented as dot distribution maps on a base map of TAZ boundaries and major interstate highways for the year 2000. Shown on each map are the mean center for the year 2000, the Y-axis, the X-axis, the year 2000 ellipse for one standard deviation (enclosing 67% of the workers represented) and the year 2000 ellipse for two standard deviations (enclosing 95% of the workers represented). Also shown in grey are the 1990 ellipses and their X- and Y-axes. All maps are to the same scale and each dot represents 200 workers. All maps indicate a direct visual comparison of the 1990 and the 2000 ellipses.

The primary analytical information on Table 2.1.5following depicts the dimensions for each ellipse with the angle of the major axis, the length of the major Y axis and the minor X axis. The ellipse reported on the table represents one standard deviation of the dispersal patterns for each of the two classes of workers. In each of the metropolitan areas except Denver, both the Y (Major) and the X (Minor) Axes of the ellipse representing the distribution of information workers is smaller than that of the ellipse representing the distribution of non-information workers, indicating that information workers are dispersed over a smaller spatial extent than non-information workers in both the distribution of their residences and their work places. This is confirmed by the area calculation for each ellipse. For example, the elliptical area (in acres) for the work places of information workers in Atlanta is only 73% of the elliptical area for the work places of non-information workers. In the Twin Cities, this same area for information workers is similar to Atlanta at 75.3%.

Table 2.1.5 Results of centrographic analysis—six metropolitan areas: 1990 and 2000

CTPP 2000		Y Axis (mile)	Y Axis Angle	X Axis (mile)	Ellipse Acres
Place of Work					
Atlanta	Information Workers	16.32	64.27	11.95	392,071
	Noninformation Workers	18.67	64.00	13.79	517,466
Austin	Information Workers	9.92	81.62	4.47	89,238
	Noninformation Workers	11.02	82.07	4.99	110,625
Denver	Information Workers	16.24	116.68	7.96	260,033
	Noninformation Workers	16.03	112.89	9.44	304,158
Houston	Information Workers	20.39	119.99	14.25	584,288
	Noninformation Workers	22.13	113.88	15.43	686,776
Phoenix	Information Workers	13.23	148.59	8.49	225,870
	Noninformation Workers	14.55	153.76	9.26	270,824
Twin Cities	Information Worker	10.84	24.51	9.94	216,652
	Noninformation Worker	12.29	51.30	11.68	288,528
Place of Residence					
Atlanta	Information Workers	19.62	65.22	16.32	643,607
	Noninformation Workers	20.24	64.27	16.60	675,628
Austin	Information Workers	12.68	79.29	6.54	166,797
	Noninformation Workers	12.87	83.88	6.77	175,087
Denver	Information Workers	17.54	112.38	9.97	351,397
	Noninformation Workers	16.31	108.96	10.08	330,528
Houston	Information Workers	23.23	117.18	17.68	825,732
	Noninformation Workers	23.82	112.18	18.55	888,492
Phoenix	Information Workers	16.37	144.21	9.70	319,154
	Noninformation Workers	16.49	151.33	9.51	315,398
Twin Cities	Information Worker	13.71	80.09	12.57	346,542
	Noninformation Workers	14.55	83.24	13.04	381,425

CTPP 1990		Y Axis (mile)	Y Axis Angle	X Axis (mile)	Ellipse Acres
Place of Work					
Atlanta	Information Workers	12.54	68.61	10.25	258,290
	Noninformation Workers	14.19	71.43	11.80	336,577
Austin	Information Workers	6.48	84.00	3.64	47,374
	Noninformation Workers	7.01	84.03	4.36	61,456
Denver	Information Workers	14.27	118.32	7.15	205,221
	Noninformation Workers	13.82	116.16	7.61	211,431
Houston	Information Workers	20.14	121.64	15.02	608,214
	Noninformation Workers	21.21	114.51	16.57	706,822
Phoenix	Information Workers	10.90	151.04	6.63	145,350
	Noninformation Workers	11.49	152.18	6.68	154,372
Twin Cities	Information Workers	10.26	12.05	9.27	191,090
	Noninformation Workers	11.31	19.92	10.78	245,202
Place of Residence					
Atlanta	Information Workers	17.67	67.44	15.09	536,320
	Noninformation Workers	19.05	65.63	15.74	602,811
Austin	Information Workers	11.73	84.75	6.14	144,748
	Noninformation Workers	12.20	87.87	6.67	163,555
Denver	Information Workers	16.18	111.34	10.06	327,309
	Noninformation Workers	15.54	106.28	10.28	321,130
Houston	Information Workers	22.56	119.94	17.28	783,867
	Noninformation Workers	23.13	113.91	18.15	844,028
Phoenix	Information Workers	14.69	147.47	8.35	246,516
	Noninformation Workers	15.77	154.67	9.02	285,971
Twin Cities	Information Workers	12.68	71.47	11.86	302,337
	Noninformation Workers	13.78	76.90	12.59	348,880

Changes in the implied acreage of the one-standard-deviation ellipse are shown on the following graphs for information worker and non-information workers respectively.

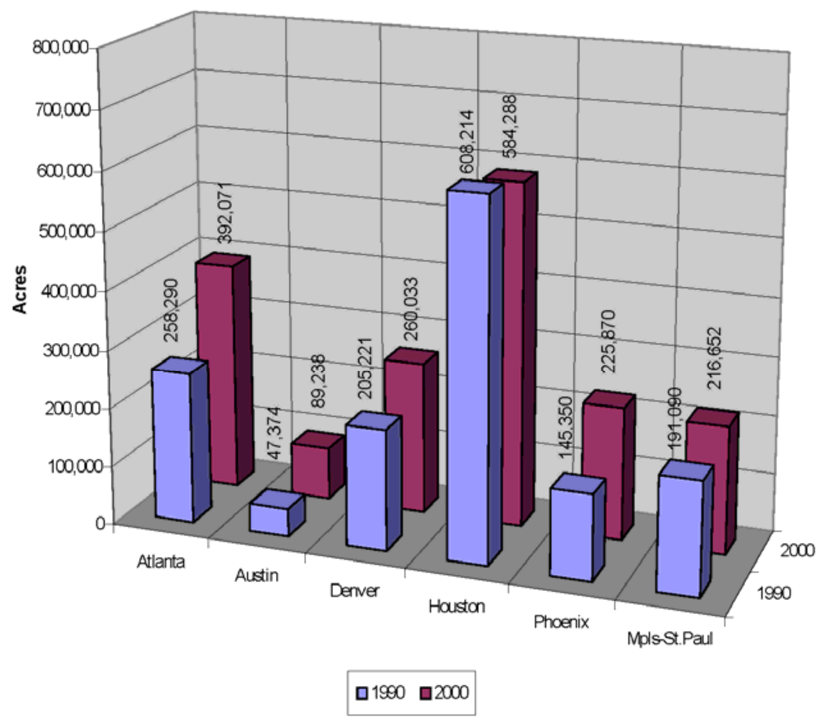


Figure 2.1.7 Information workers: acreage of workplace ellipse: 1990 and 2000

The ratios of information worker ellipse areas to non-information worker ellipse areas can be shown as follows.

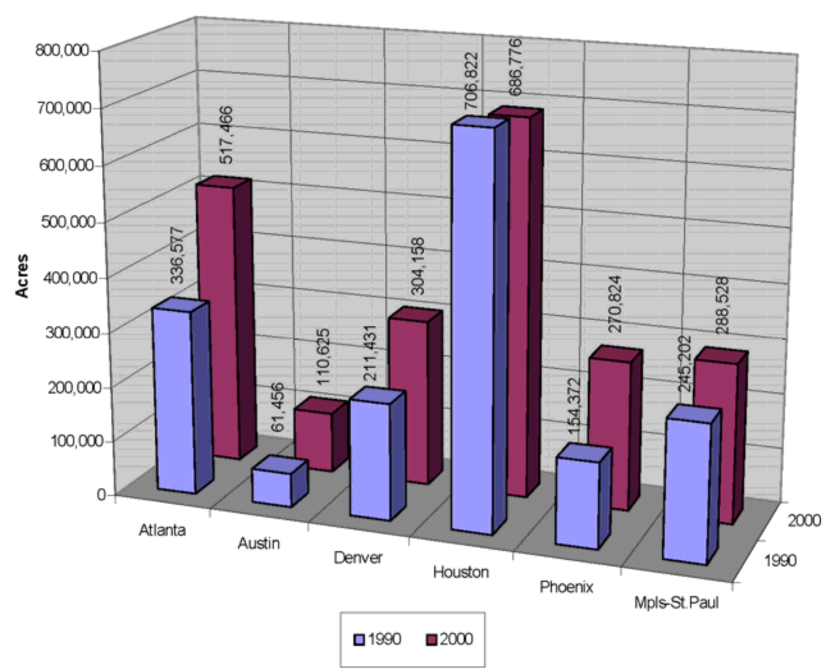


Figure 2.1.8 Non-information workers: acreage of workplace ellipse: 1990 and 2000

Table 2.1.6 Ratio of workplace ellipse* acreage of information workers to that of non-information workers

Workplace	1990	2000	Abs Chg	% Chg
Atlanta	0.767	0.758	-0.010	-1.3%
Austin	0.771	0.807	0.036	4.6%
Denver	0.971	0.855	-0.116	-11.9%
Houston	0.860	0.851	-0.010	-1.1%
Phoenix	0.942	0.834	-0.108	-11.4%
Mpls-St.Paul	0.779	0.751	-0.028	-3.6%
Mean	0.848	0.809	-0.039	-0.041

Table 2.1.7 Ratio of residential ellipse* acreage of information workers to that of non-information workers

Place of Residence	1990	2000	Abs Chg	% Chg
Atlanta	0.890	0.953	0.063	7.1%
Austin	0.885	0.953	0.068	7.6%
Denver	1.019	1.063	0.044	4.3%
Houston	0.929	0.929	0.001	0.1%
Phoenix	0.862	1.012	0.150	17.4%
Mpls-St.Paul	0.867	0.909	0.042	4.8%
Mean	0.909	0.970	0.061	6.9%

*One standard Deviation Ellipse

Analyzing these graphs and tables, it can be seen that for every 100 acres represented in the workplace ellipses for non-information workers, there were, on average, only 85 acres in the ellipses for information workers in 1990 and only 81 acres in 2000. Thus, while workplace ellipses for information workers did enlarge over the decade, their enlargement was relatively less than the spreading out of non-information worker places of employment. Information workers dispersed to some extent in the decade but not to the same extent as non-information workers, insofar as their place of work is concerned.

Table 2.1.8 Size of computer workers' ellipse to that of all information workers' ellipse: 2000

2000	All Information Workers' Ellipse	Computer Workers' Ellipse	Ratio: Computer Workers to All Information Workers
Atlanta	392,071	264,386	67.4%
Austin	89,238	68,651	76.9%
Denver	260,033	241,043	92.7%
Houston	584,288	396,473	67.9%
Phoenix	225,870	195,880	86.7%
Mpls-St.Paul	216,652	163,252	75.4%
Mean	***294,692***	***221,614***	***77.8%***

By this measure, it can be concluded that information workers, in both 1990 and 2000, were more concentrated in both their workplace location patterns and in the patterns of their residences in each of the metropolitan areas. One key exception was in the Denver-Boulder area where the residential pattern of information workers had a larger ellipse than non-information workers in both years. Another exception occurred in 2000 in Phoenix, where the residential distribution of information workers was slightly more dispersed than non-information workers—a change from 1990 where the opposite was true. In short, where information workers choose to live is slightly more concentrated than for their non-information counterparts, although the difference is considerably less evident. In fact, between 1990 and 2000 the residential distribution of information workers became almost identical to that of non-information workers.

Also of interest was the centrographic analysis for the set of workers specialized in the computer and mathematical occupations. In the map portfolio is a set of six maps (Map 25–30) depicting the centroids and ellipses for these occupations as they are distributed in each of the metropolitan areas. These maps illustrate, for the year 2000, the extent of the one standard deviation and the two standard deviation workplace ellipses for those in the computer and mathematical occupations in contrast with all information workers. The critical finding here is that for all six metropolitan areas, workers specializing in the computer and mathematical occupations were considerably more concentrated than information workers generally. This is illustrated in the following table that compares the size of the ellipse for the computer specialists as against all information workers.

Each map also provides a dot distribution by TAZs so that the concentrations of these workers stand out vividly (one dot = 50 workers). Thus, these workers in 2000 remained fairly concentrated in discrete employment districts, despite the spreading out of other information workers generally over the decade. If we can interpret these specialists as being the key innovators in information technology, it would appear that they and their places of work have not participated in the general sprawling nature of urban economic activity of the 1990s.

Spatial Autocorrelation Results

Results of the spatial autocorrelation analysis are given in the following table. Every calculation of Moran's I was found to be statistically significant, in that every pattern was distinctly different from a purely random (or "expected") pattern. All workers were somewhat more clustered in both their residential and workplace locations, although this might be expected since each metropolitan area had areas that were not developed for one reason or another (major water bodies, significant public lands, transportation [ports, railroad yards, etc.]). However, in every instance, for both 1990 and 2000, Moran's I for information workers exhibited a higher level of spatial autocorrelation than the results for non-information workers. Essentially this analysis confirms the hypothesis that information workers in 1990 were more concentrated, or clustered, in their places of employment than non-information workers. Results of this analysis is shown in the following tables.

Table 2.1.9 Results of spatial autocorrelation: Moran's "I" and expected "I" [E(I)]: 1990 and 2000

Information Workers	2000 Moran's I				1990 Moran's I			
	Moran's I	Random E(I)	Normal Z	Random Z	Moran's I	Random E(I)	Normal Z	Random Z
Place of Work								
Atlanta	0.102619	-0.000594	77.58	78.14	0.11545	-0.00107	47.11	47.41
Austin	0.132372	-0.001418	45.36	45.86	0.10039	-0.00193	27.74	28.61
Denver	0.080454	-0.000406	86.50	87.59	0.085033	-0.00081	49.09	49.67
Houston	0.082182	-0.000379	79.57	84.57	0.08966	-0.00046	72.60	75.17
Phoenix	0.081095	-0.000643	55.36	55.83	0.06965	-0.00103	26.85	27.17
Mpls-St.Paul	0.089351	-0.000833	53.55	55.29	0.09337	-0.00087	53.23	54.64
Place of Residence								
Atlanta	0.04385	-0.00059	33.40	33.44	0.09560	-0.00109	42.58	42.60
Austin	0.048098	-0.001418	16.79	16.83	0.04119	-0.00171	13.06	13.17
Denver	0.040903	-0.000406	44.19	44.22	0.027832	-0.00081	16.34	16.96
Houston	0.07755	-0.00038	75.11	75.24	0.07842	-0.00042	78.09	78.23
Phoenix	0.07747	-0.00064	52.90	52.94	0.09270	-0.00092	39.15	39.17
Mpls-St.Paul	0.02259	-0.00083	13.91	13.91	0.02906	-0.00091	16.30	16.31

Table 2.1.10 Difference of Moran's "I" from expected value of "I"

Difference from E(I) Rank Order Information Workers Place of Work	2000	1990
Austin	0.1338	0.1023
Atlanta	0.1032	0.1165
Mpls-St. Paul	0.0902	0.0942
Houston	0.0826	0.0901
Phoenix	0.0817	0.0707
Denver	0.0809	0.0858

Difference from E(I) Rank Order Information Workers Place of Residence	2000	1990
Phoenix	0.0781	0.0936
Houston	0.0779	0.0788
Austin	0.0495	0.0429
Atlanta	0.0444	0.0967
Denver	0.0413	0.0286
Mpls-St. Paul	0.0234	0.0300

Difference from E(I) Rank Order Non-Information Workers: Place of Work	2000	1990
Austin	0.1360	0.0853
Houston	0.0816	0.0617
Phoenix	0.0639	0.0637
Denver	0.0572	0.0598
Atlanta	0.0481	0.0430
Mpls-St. Paul	0.0422	0.0604

Difference from E(I) Rank Order Non-Information Workers: Place of Residence	2000	1990
Houston	0.0712	0.0685
Phoenix	0.0698	0.0733
Austin	0.0541	0.0435
Denver	0.0496	0.0286
Atlanta	0.0375	0.0682
Mpls-St. Paul	0.0244	0.0226

From these results, the differences between the actual Moran's I and the Expected I [E(I)] were calculated as shown in the Table 2.1.10.

These results indicate a reasonable approximation of a normal distribution (despite only six observations), so that a paired sample "t-test" was run comparing the differences from E(I) for information workers as against non-information workers. These results follow in Table 2.1.11:

Table 2.1.11 Difference from expected values of Moran's "I"

Paired Sample Results: Difference from Expected Values of Moran's I		
Information Workers vs. Non-Information Workers	t statistic	p statistic
Workplace: 2000	3.11	0.0265*
Workplace: 1990	3.31	0.0212*
Residence: 2000	0.47	0.6584
Residence: 1990	2.34	0.0665

*Significant at the $p<0.05$ level, 95% CI

Thus, it can be seen that, in both 1990 and in 2000, there was a significant difference in spatial autocorrelation with information workers being more concentrated than non-information workers, consonant with the findings of the previous centrographic analysis. In terms of the residential location of workers, on the other hand, results of the Moran's I calculation suggest no significant difference in spatial autocorrelation in either 1990 or 2000. Thus, for place of residence, as with the centrographic analysis, the results were less clear-cut.

For the individual cities, the smallest of the six (Austin) tended to be the most concentrated for workers in their places of work. The Austin area also showed the largest increase in concentration in the decade for both information and non-information workers. As will be shown below, the concentration of activity in Austin was associated with the one interstate highway that traverses the area. While the Denver area had information workers more concentrated than others, their information workers had a Moran's I closer to the expected value than any of the other six cities—therefore, Denver's spatial distribution was the closest to being random.

Non-information workers in Atlanta and Minneapolis-St. Paul had the most random distribution by place of work in 2000. This was also true for Atlanta in 1990. Houston and Phoenix showed stronger levels of concentration for workers' places of residence for all workers in both 1990 and 2000. Minneapolis-St. Paul had the closest approximation to a random distribution by place of residence of all workers in both 1990 and 2000.

Workplace Locations in Relation to Their Mean Center

In this analysis, two measures were developed. The first calculated the number of workers in a given TAZ and divided it by the distance from the centroid of the TAZ to the mean center of all workplaces for the given class of worker. Thus, the total number of information workers working in a TAZ was divided by the distance of that TAZ to the mean center of all information work places. From this, the mean distance of all TAZs possessing the employment of information workers from the mean center was determined. A similar calculation was carried out for non-information workers. From this, the statistical significance of the differences among information workers and non-information workers was calculated.

This calculation is given by the following equation:

$$P = \frac{\sum_{i=1}^{n} \frac{w_i}{d_i}}{n}$$

This measure is best described as the average number of workers per TAZ by proximity (P) to the mean center. Larger result quantities mean a small distance from the mean center and thus a tendency to be more clustered around the mean center. Small quantities mean that workers tend to be at a greater distance from the mean center and thus more dispersed. Small results also may indicate a smaller average number of workers per TAZ.

Table 12 provides the results of this calculation. In general, the larger the numerical result, the more clustered are workers around the mean center. In all six urban areas the results indicate that information workers tend to be more closely packed around their mean center than are non-information workers. This measure provided an opportunity to test the significance of differences between information and non-information workers. In three of the six metropolitan areas, there is high significance with a p value of less than 0.001. Only in Phoenix was the difference between information workers and non-information workers was less significant where p = 0.064.

A similar calculation was carried out for the percentage of all information workers working in a given TAZ, divided by the distance to the mean center.

$$P_{\%} = \frac{\sum_{i=1}^{n} \frac{w_i / w_t}{d_i}}{n}$$

Levels of significance were less impressive by this calculation but the results are comparable—information workers are more clustered in all six metropolitan areas by this measure. The results of this calculation are also shown on Table 2.1..

Table 2.1.12 Mean distance from workplace center: 2000 and 1990

		Number of Worker by Place of Work[1]	Percentage of Worker by Place of Work[2]		
		Mean	Sig.	Mean	Sig.
CTPP 2000					
ATLANTA	Information Workers	21.820	0.000	19.076	0.023
	Noninformation Workers	13.564		16.415	
AUSTIN	Information Workers	26.822	0.000	72.296	0.286
	Noninformation Workers	15.824		67.719	
DENVER	Information Workers	16.031	0.000	20.667	0.030
	Noninformation Workers	8.944		16.923	
HOUSTON	Information Workers	17.097	0.000	14.899	0.028
	Noninformation Workers	11.843		12.452	
PHOENIX	Information Workers	18.087	0.000	23.245	0.434
	Noninformation Workers	14.282		22.267	
TWIN CITIES	Information Workers	35.804	0.001	39.949	0.056
	Noninformation Workers	18.372		29.537	
CTPP 1999					
ATLANTA	Information Workers	28.621	0.000	37.420	0.003
	Noninformation Workers	20.463		32.594	
AUSTIN	Information Workers	20.250	0.000	93.720	0.444
	Noninformation Workers	14.419		98.426	
DENVER	Information Workers	20.713	0.000	37.784	0.023
	Noninformation Workers	13.223		32.421	
HOUSTON	Information Workers	18.431	0.000	19.682	0.002
	Noninformation Workers	12.850		15.140	
PHOENIX	Information Workers	20.825	0.000	41.349	0.234
	Noninformation Workers	17.836		39.792	
TWIN CITIES	Information Workers	29.074	0.000	41.675	0.004
	Noninformation Workers	20.250		35.404	

Source: CTPP 1990

Note:
1. Number of workers working in a TAZ divided by the direct ("as the crow flies") distance from the centroid of the TAZ to the mean center for place of work for all workers.
2. Percentage of workers working in a TAZ, relative to the total metropolitan employment, divided by the direct distance from the center of the TAZ to the mean center for all workers.

Commuting Effort Results

The Census Transportation Planning Package permits an analysis of the time involved in commuting to work on the basis of the work zone TAZ. From the overall data, an average worker density was determined for each of the six metropolitan areas. Of those TAZs with worker density above the metropolitan average, we identified those that were dominated by information workers—using the criteria of more than 60% of the employees with information occupations. TAZs that were above the metropolitan average but were not dominated by information workers are labeled "diversified" centers. TAZs with some employment while having a lower than average worker-density were considered minor areas of employment. The table below provides the results of this analysis.

Generally, TAZs dominated by information workers had longer commuting times than other types of TAZs. Where information workers dominate, mean travel times are 8% longer than those in diversified TAZs, and 38% longer than travel times in minor centers of employment. (Commute times in diversified centers are 28% longer than in minor centers.) This would seem to follow from the analysis above. Information workers are more concentrated in their workplaces, but their residential locations are equally dispersed as those of non-information workers. The residential dispersion would thereby logically imply longer commutes to more centralized job locations. For both the information dominant and the diversified centers, commute times are considerably greater than for those workers employed in minor centers.[v]

The other side of this portrayal, however, lies in the fact that commute times changed less in the decade from 1990 to 2000 for workers in the information dominant TAZs. These workers, on the average, saw their commute time increase by only 3.5 minutes whereas workers in minor TAZs had an increase of 6.6 minutes. Diversified centers experienced an increase of commute time similar to that of information dominant centers (an average of 3.8 minutes).

The data of Table 2.1.13 is also represented in the graphs on the pages following. It is interesting to note that mean commute times in Atlanta were longer than in any of the other six metropolitan areas. Generally, the largest cities in 2000 in terms of population—Atlanta, Houston and Phoenix—involved the longest commute times of the six urban areas in both 1990 and 2000.

Generally, Table 2.1.13 indicates that for TAZs above the metropolitan mean and possessing more than 60% of their workers as information workers, travel times to work tend to be higher than for those TAZs without the dominance of information workers. TAZs having sparse employment show the lowest travel times of all. Table 2.1.13 out the picture of information workers generally working in more centralized or clustered workplaces yet more dispersed in their residential locations. This results in longer times (and possibly longer travel distances). This may reflect the possibility that information workers receive greater compensation for their labor and therefore possess greater choice in residential location.

Table 2.1.13 Mean travel times of work to TAZs: 1990 and 2000

Information Worker Dominant Centers				
	1990	2000	Abs Chg	% Chg
Atlanta	30.8	35.6	4.8	15.6%
Austin	22.7	27.8	5.1	22.5%
Denver	24.0	26.5	2.5	10.4%
Houston	29.5	31.2	1.7	5.8%
Mpls-St.P	23.1	25.5	2.4	10.4%
Phoenix	25.8	30.0	3.3	12.9%
Mean	25.4	28.5	3.1	12.2%

Diversified Centers				
	1990	2000	Abs Chg	% Chg
Atlanta	28.7	33.9	5.2	18.1%
Austin	20.8	26.0	5.2	25.0%
Denver	23.2	25.4	2.2	9.5%
Houston	27.0	30.8	3.8	14.1%
Mpls-St.P	22.0	24.7	2.7	12.3%
Phoenix	23.0	26.5	3.5	15.2%
Mean	24.1	27.9	3.8	15.7%

TAZs with Employment Below Mean Worker Density				
	1990	2000	Abs Chg	% Chg
Atlanta	22.1	29.2	7.1	32.1%
Austin	18.2	26.8	8.6	47.3%
Denver	18.3	24.0	5.7	31.1%
Houston	19.8	24.8	5.0	25.3%
Mpls-St.P	17.6	22.6	5.0	28.4%
Phoenix	17.3	25.2	7.9	45.7%
Mean	18.9	25.4	6.6	35.0%

Source: CTPP 1990

Note

1. Table does not include people living in a metropolitan area and working outside of the area or workers working in a metropolitan area while living outside of the area.

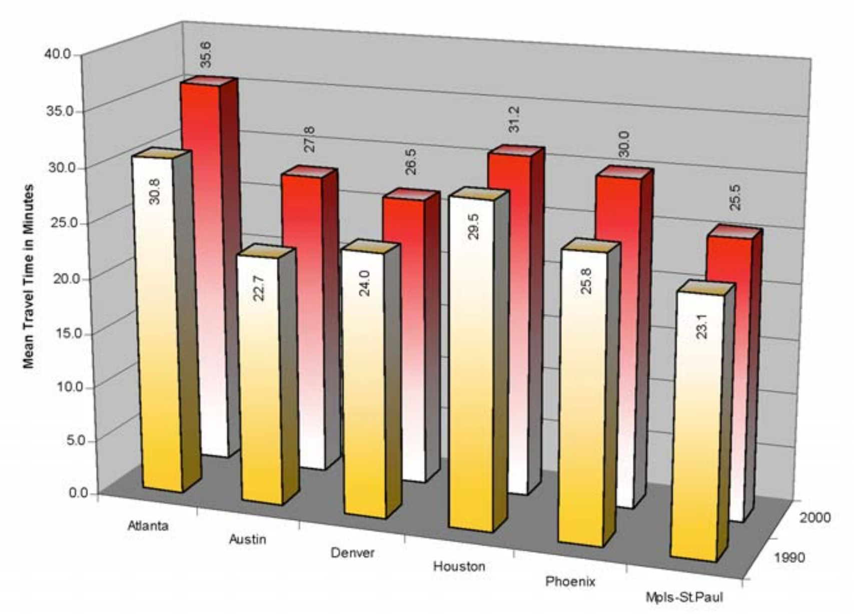


Figure 2.1.9 Change in mean travel time for information worker dominant TAZs: 1990 and 2000

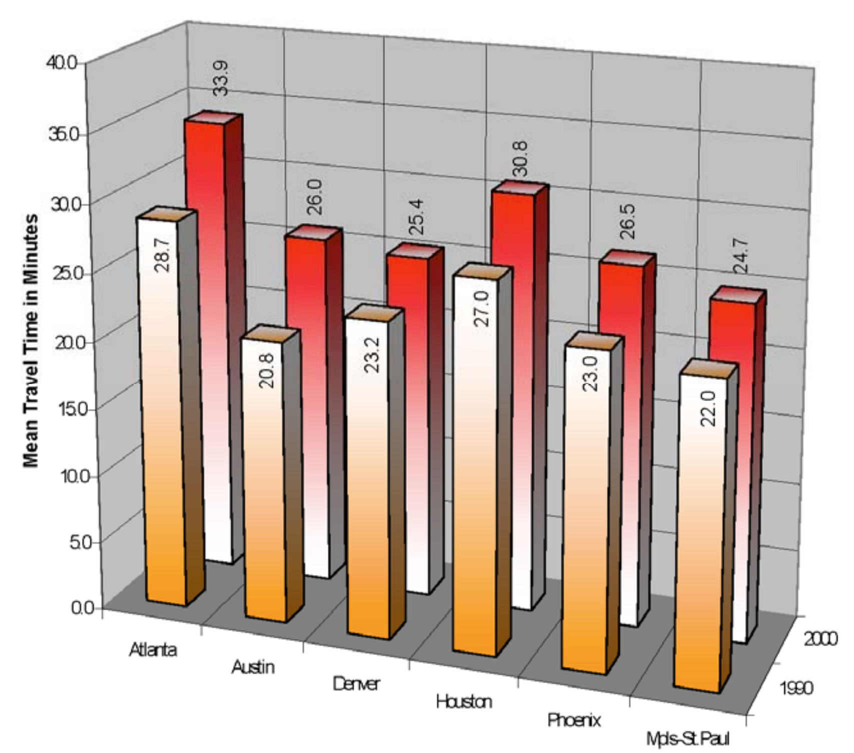


Figure 2.1.10 Mean travel time in diversified TAZs: 1990 and 2000

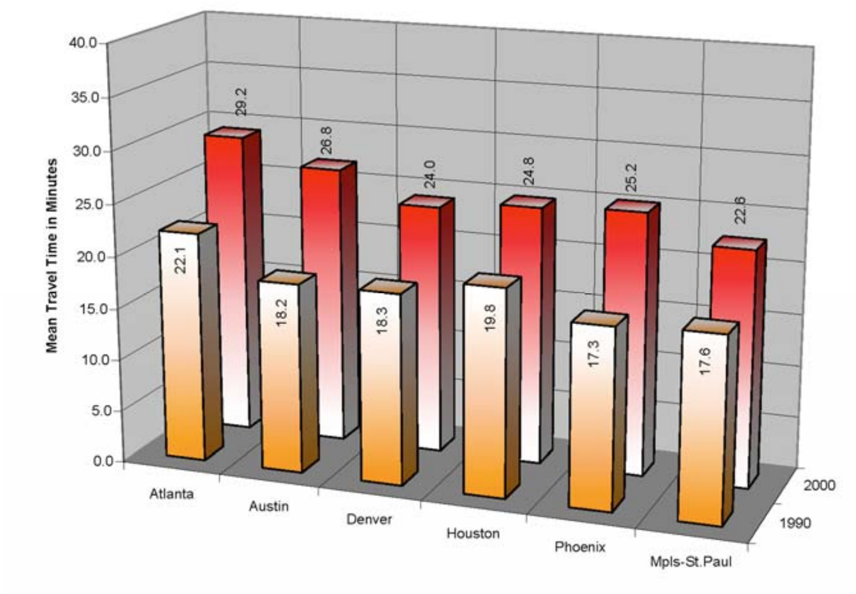


Figure 2.1.11 Mean travel time workers in low worker density TAZs: 1990 and 2000

2.1.6 Employment Center Analysis

Employment Center Identification and Classification (Maps 31–36)

The methodology for identifying major employment centers, as described in Section 2.1.3, involved determining the overall mean employment density for all workers in each of the metropolitan areas. From this the TAZs that were above the mean in employment density were identified and mapped. These formed the basis for defining TAZs as employment centers. In addition, the centers were classified from known features of the metropolitan area. Four classifications were used: (1) the central business district (or districts in the case of Denver-Boulder and Minneapolis-St. Paul), (2) "edge city" districts, (3) university districts,[vi] and (4) "other" districts. Centers were categorized as "edge city" centers if TAZs of high employment density were (1) adjacent to an interstate or other major freeway and (2) had two or more contiguous TAZs that were above average employment density. University centers were identified through various internet sources (including using maps of major universities as provided on their website) and other mapping sources. High employment density centers not meeting any of the above criteria were categorized as "Other" centers.

These centers can be viewed by referring to Maps 31 through 36, inclusive in the map portfolio. Denver and the Twin Cities had two downtown areas (Denver and Boulder; Minneapolis and Saint Paul). All six areas had "edge city" centers, as defined by their proximity to interstate highways or limited access ring roads and the presence of two or more contiguous TAZs with above average employee density. Each of the six areas had university districts. Center characteristics are shown Tables 2.1.14 and 2.1.15 following.

Table 2.1.14 Identification of centers in six metropolitan areas: 1990–2000

	No. of TAZs with Above Average Employment Density	Centers with Contiguous TAZs	No. of Central Business Districts	No. of University Centers	No. of Edge City Centers	No. of Other Centers
1990						
Atlanta	98	33	1	4	16	12
Austin	90	27	1	1	6	19
Denver	191	63	2	2	26	33
Houston	237	58	1	2	26	29
Phoenix	239	47	1	2	6	38
Mpls-St.Paul	199	71	2	1	23	45
2000						
Atlanta	200	60	1	4	28	27
Austin	109	30	1	1	10	18
Denver	458	106	2	2	39	63
Houston	258	69	1	2	23	43
Phoenix	359	48	1	2	7	38
Mpls-St.Paul	207	51	2	1	22	26
Change: 1990–2000						
Atlanta	102	27	0	0	12	15
Austin	19	3	0	0	4	-1
Denver	267	43	0	0	13	30
Houston	21	11	0	0	-3	14
Phoenix	120	1	0	0	1	0
Mpls-St.Paul	8	-20	0	0	-1	-19

From this table, it can be seen that each metropolitan area had gains in the number of TAZs with worker densities above the metropolitan mean between 1990 and 2000. Each city had gains in “edge city” centers with the exception of Houston and Minneapolis-St. Paul. Gains in “other” centers, however, were the most frequent in all cities except Austin and Minneapolis-St. Paul.

Table 2.1.15 provides further elaboration of characteristics of TAZs with worker densities above their metropolitan mean for the year 2000. This table shows the acreage involved for individual TAZs in each type of center as well as the mean density of workers per TAZ acre.

Finally, Table 2.1.15 indicates the number of TAZ's dominated by information workers (more than 60% of workers in a TAZ). Domination by information workers is discussed in the section below.

As expected, in every instance except Phoenix, worker densities in the central business districts were well above 100 workers per acre in the year 2000. The average density, excluding Phoenix, was 147 workers per acre (the average with Phoenix included was 128 workers per acre). These densities were in sharp contrast with "edge city" centers where the average density for the six metropolitan areas was 13.4 in 2000. University centers possessed densities that averaged 42 workers per acre and, thus, were mid-range between downtowns and edge cities. "Other" centers had average densities of 15 workers per acre—slightly above that of "edge city" centers.

These are characteristics, of course, that have been well known for some time. They describe the results of 20–30 year trends in the location of urban economic activity in metropolitan areas. Even in 2000, the worker densities of historic downtown areas are often ten times that of suburban employment centers, and this is born out for the six metropolitan areas under study. Despite the years of sprawl development following World War II, central business districts remain the major employment centers of each of the six regions. This is particularly true for information occupations, as will be discussed in the next section.

Maps 43 and 44 in the map portfolio (Appendix B) provide a 3-dimensional view of the patterns of worker density for each metropolitan area in 2000. As expected, Phoenix is unusual in that its central business district density of 34 workers per acre is only slightly above typical suburban densities. As well, its university centers tend to have lower densities than typical suburban densities. Houston has the tallest central spike on Map 43 as the most dense central business district with more than 200 workers per acre, followed by Atlanta with 184 workers per acre. Areas possessing two prominent central business districts show quite clearly with Denver and Boulder standing out on Map 43 and Minneapolis and St. Paul showing pronounced spikes on Map 44.

Table 2.1.15 Center characteristics of six metropolitan areas: 2000

ATLANTA **CENTERS**	**Number of TAZ's**	**Average Acres per TAZ**	**TAZ Sum Acres**	**Mean Density Workers/Acre**	**Inform. Dominant TAZs**	**% Inform. Dominant**
Central Business District(s)	26	35.1	912.9	184.1	20	76.9%
Edge Cities	93	334.3	31,089.9	14.0	44	47.3%
Other Center	76	262.5	19,951.7	22.2	31	40.8%
University	5	199.0	995.0	98.1	0	0.0%
Total in Centers	200	207.7	52,949.5			
Not in Center	1485	1,700.7	2,525,553.7			
AUSTIN **CENTERS**	**Number of TAZ's**	**Average Acres per TAZ**	**TAZ Sum Acres**	**Mean Density Workers/Acre**	**Inform. Dominant TAZs**	**% Inform. Dominant**
Central Business District(s)	19	51.1	971.2	102.4	12	63.2%
Edge Cities	25	307.7	7,692.7	11.2	12	48.0%
Other Center	62	244.4	15,153.8	13.7	31	50.0%
University	3	108.6	325.7	55.2	2	66.7%
Total in Centers	109	177.9	24,143.4			
Not in Center	597	1,584.0	945,677.0			
DENVER **CENTERS**	**Number of TAZ's**	**Average Acres per TAZ**	**TAZ Sum Acres**	**Mean Density Workers/Acre**	**Inform. Dominant TAZs**	**% Inform. Dominant**
Central Business District(s)	34	28.9	984.2	127.1	23.0	67.6%
Edge Cities	161	126.1	20309.0	15.1	83.0	51.6%
Other Center	260	108.8	28278.0	16.3	112.0	43.1%
University	3	172.1	516.3	23.0	2	66.7%
Total in Centers	458	109.0	50,087.5			
Not in Center	2004	1737.4	3,481,748.2			
HOUSTON **CENTERS**	**Number of TAZ's**	**Average Acres per TAZ**	**TAZ Sum Acres**	**Mean Density Workers/Acre**	**Inform. Dominant TAZs**	**% Inform. Dominant**
Central Business District(s)	53	19.3	1,023.5	217.8	23	43.4%
Edge Cities	65	263.2	17,105.2	18.9	34	52.3%
Other Center	130	208.8	27,145.0	15.8	51	39.2%
University	10	216.4	2,163.6	35.2	10	100.0%
Total in Centers	258	176.9	47,437.2			
Not in Center	2384	2,335.2	5,567,158.8			
PHOENIX **CENTERS**	**Number of TAZ's**	**Average Acres per TAZ**	**TAZ Sum Acres**	**Mean Density Workers/Acre**	**Inform. Dominant TAZs**	**% Inform. Dominant**
Central Business District(s)	22	96.8	2,130.0	33.8	15	68.2%
Edge Cities	90	455.0	40,954.4	7.6	40	44.4%
Other Center	244	289.2	70,556.3	11.0	93	38.1%
University	3	388.5	1,165.5	6.1	3	100.0%
Total in Centers	359	307.4	114,806.2			
Not in Center	1197	4,835.8	5,788,432.8			
MPLS - ST. PAUL **CENTERS**	**Number of TAZ's**	**Average Acres per TAZ**	**TAZ Sum Acres**	**Mean Density Workers/Acre**	**Inform. Dominant TAZs**	**% Inform. Dominant**
Central Business District(s)	34	67.5	2,294.2	104.7	29	85.3%
Edge Cities	70	322.4	22,569.2	13.3	33	47.1%
Other Center	98	236.9	23,218.5	13.4	46	46.9%
University	5	153.4	767.0	32.2	5	100.0%
Total in Centers	207	195.1	48,848.9			
Not in Center	994	1,866.1	1,854,926.8			

Table 2.1.16 Summary distribution of place of work TAZs in six metropolitan areas: 2000

	% in CBD	% in Edge City	% in Other Centers	%in University Centers
Atlanta	13.0%	46.5%	38.0%	2.5%
Austin	16.8%	25.7%	54.9%	2.7%
Denver	7.4%	35.2%	56.8%	0.7%
Houston	20.5%	25.2%	50.4%	3.9%
Phoenix	6.1%	25.1%	68.0%	0.8%
Mpls-St.Paul	16.4%	33.8%	47.3%	2.4%
Mean	**13.4%**	**31.9%**	**52.6%**	**2.2%**

Table 2.1.16 provides a summary of the distribution of workplace TAZs in each metropolitan area in 2000. Thus, on average, 13% of TAZs with employment densities above the regional mean are found in those TAZs we have defined as central business districts, 32% are found in our designated "edge cities," a bit more than 2% are found in university settings, and more than half are in those centers we designated as "other centers." These averages seem to prevail in all six areas with some notable exceptions. Houston was unusually high in central business district TAZs, while Denver and Phoenix were well below average.[vii] TAZs designated as "edge cities" are remarkably similar in the six areas. Atlanta is higher than the others, while three of the six are almost identical in their percent of the total. Thus, the mean distribution of the different designations of centers is quite similar for all six metropolitan areas in the year 2000.

Information Worker Dominant Centers (Maps 31–36)

Information Dominant Centers in 2000. In the map portfolio, Maps 31–36 display the TAZs dominated by information workers. As noted, a TAZ was defined as dominated by information workers if it had 60% or more of its workers in information occupations. Three-dimensional portrayals for employment density for TAZs dominated by information workers are shown on Maps 39 and 40.

There appear to be no distinctive patterns among the six regions, save for the central business districts of each being comprised of information dominant TAZs.

All of the urban areas saw considerable suburban development in the decade of the 1990s and the primary generalities one can gain from the experience is that they are primarily oriented toward major highways and are developed at densities between 10 and 15 workers per acre.

In Atlanta, the major ring road—Interstate 285—lies approximately 12 miles beyond the central business district and has provided a focus of development of dominant information worker TAZs. Two such concentrations lie (1) at the intersection of I-285 and the McDonald Parkway to the north and (2) at the intersection of I-285 and I-75 to the northwest. Additional dominant information worker TAZs are found adjacent to the Peachtree Industrial Boulevard to the northeast and beyond I-285. University centers include: Emory, Georgia Institute of Technology, and Georgia State University. Interestingly, while the Atlanta international airport lies to the south of the CBD, the preponderance of dominant information worker centers lie to the north.

Austin has developed in a linear fashion along a north-south axis within the parallel highways of Interstate 35 and Route 1. The University of Texas is the university center for the region and is located just north of the central business district and the state capital, thus providing a strong central anchor of information workers. A suburban center involving a number of contiguous TAZs is found at the confluence of Routes 1, 183 and 360, directly north and approximately 10 miles from the CBD. An important addition to the Austin economy in the decade was the construction of the new Dell Computer headquarters, just south of Round Rock at the confluence of I-35 and Route 1.

Three interstate highways serve the Denver region: a north-south interstate, I-25, an east-west highway—I-70, and I-76. The latter highway runs from I-80 in Nebraska and enters the Denver area from the northeast and ends with its intersection with I-70 approximately 5 miles northwest of the downtown. The region has two central areas dominated by information workers—Denver and Boulder—although Denver is significantly larger. There are two university centers adjacent to the central business districts and dominated by information workers: the University of Colorado at Boulder and at Denver. Centers dominated by information workers in this region tend to be more scattered although they are concentrated along Route 36 to the northwest between Denver and Boulder and to the southeast along I-25 between I-225 and the Route 470 ring road. One concentration lies to the west in the Lakewood area.

Houston, the largest of the six metropolitan areas studied, has an extensive major highway network with three major ring roads—an inner ring road, I-610, that ranges between 4–6 miles from the central business district, the Sam Houston toll road that circumscribes the urban area approximately 12 miles from the central business district and Route 6 that generally lies 18–20 miles from the central downtown. Radial freeways converge on the center from all major points of the compass and include Interstate highways I-10 and I-45. Despite this symmetrical pattern of freeways, the TAZs with employment densities above the mean are largely concentrated to the west of the CBD. In addition to the downtown district, TAZs with dominance in information workers are located in connection with Rice University and West University Place. Information dominant work places are also found on the western segment of the Sam Houston toll road west of Bunker Hill Village. Another notable location of information worker predominance lies to the north at the intersection of the Sam Houston toll road and I-45 with close proximity to the new George Bush Airport.

Phoenix also has a well-configured system of express highways. I-10 comes into the city from the southeast and continues west of the city. I-17 proceeds north from the center city to Flagstaff. I-10 and I-17 meet in the center with two interior loops—the first encircling the central business district and the second, lying adjacent to the east, encircles the city's major airport. Phoenix has developed at very low densities and this is also true of its employment centers. Those TAZs with employment densities above the mean tend to extend out from the central downtown along the freeways to the north and southeast with an easterly spur along Route 60 between Mesa and Gilbert. Two university centers are found at the Arizona State University main campus in Tempe and its west campus north of the city near I-17. Information workers predominate (1) in the central business district, (2) in the Tempe area and (3) at the intersection of I-17 and the Pima Freeway in the north. Overall, however, TAZs with a predominance of information occupations are well scattered and follow the general pattern of TAZs with all workers above the mean density.

The Twin Cities region, as the name implies, has the two central cities of Minneapolis and St. Paul, although Map 44 illustrates that worker densities in Minneapolis are far greater than those of St. Paul. Two major interstate highways serve the region and pass through both central business districts: I-35 and I-94. Two additional interstate freeways provide the basic ring road around the center: I-694 to the north and I 494 to the west and south. Another interstate—I-394—travels from I-94 to I-494 in the western suburbs. The major "edge city" of the region is found along the I-494 strip between the cities' major airport and the southwestern suburbs. Many historical centers can be found in close proximity to the two downtowns, principally in the Midway district lying east of the University of Minnesota, as well as in south Minneapolis and industrial districts along the Mississippi River. TAZs employment centers with a predominance of information workers can be found in the two downtown areas and along the I-494 strip, particularly in the suburbs of Edina, Eden Prairie and Plymouth. Other areas dominated by information workers occur in the Midway district, Roseville (to the northeast of downtown Minneapolis along I-35W) and in the west at the intersection of I-394 and Route 169.

Changes in Information Worker Dominant Centers: 1990–2000. As noted in earlier chapters, in the decade of the 1990s information occupations were by far the driving force of change and expansion for each of the six metropolitan areas. Our initial hypothesis suggested that while we would see a growth in the spatial distribution of information occupations, such distribution would continue to be relatively concentrated. Maps 47–52 provide a graphic depiction of the changes that took place in each of the six metropolitan areas.

Eight possible events could transpire for each TAZ geographic area. They were as follows:

1. An information dominant TAZ could remain the same in both 1990 and 2000 (these tended to be found in the central business districts of each region);
2. A TAZ above the mean density but not information dominant could also remain the same in both 1990 and 2000 (these tended to be industrial districts that did not change in the decade);
3. A TAZ having no employment or employment density below the regional mean in 1990 could become an information dominant TAZ in 2000 (these represent wholly new development in the decade);
4. A TAZ having no employment or employment density below the regional mean in 1990 could become a non-information employment center above the mean density ((these also represent wholly new development in the decade);
5. A TAZ above the mean density but not information dominant in 1990 could become and information dominant center in 2000;
6. A information dominant TAZ in 1990 could remain above the mean density in 2000 but not be an information dominant center TAZ.
7. An information dominant TAZ in 1990 could fall below the mean density in 2000 or have no employment at the end of the decade; and
8. A TAZ above the mean density but not information dominant in 1990 could fall below the mean density in 2000 or have no employment at the end of the decade.[viii]

Information Worker Centers: Extrapolative and Prosaic (Maps 53–58). Due to the variation in the growth of information workers as depicted in Section2.1.4, we carried out one further disaggregation of information workers. It will be recalled in Section2.1.3, that we laid out a broad array of functions of occupations dealing with information and information technology. Included were workers whose occupations involve them (1) in the creation and production of information; (2) in the collection, processing and interpretation of information; (3) in the transmission of information; (4) in the manufacture of information equipment; (5) in using information in training and education; and (6) in the use of information for planning and decision-making (Kurasaki and Yanagimachi, 1992). Other studies have also distinguished between "front office" and "back office" operations where the latter is devoted primarily to the prosaic tasks of data collection and entry and routine processing. "Front office" operations, in contrast, are involved in analyzing, interpreting and decision-making that impacts directly and indirectly on subsequent behavior. In addition, numerous information worker occupations are engaged in the process of designing new information technologies (both software and hardware). These workers are thereby designated as "extrapolative" information workers as distinct from "prosaic" information workers.

The occupation groups were classified as follows:

1. Extrapolative Occupations
   - Management occupations, except farmers and farm managers
   - Business and financial operations occupations:
   - Computer and mathematical occupations
   - Architecture and engineering occupations
   - Life, physical, and social science occupations
   - Community and social services occupations
   - Legal occupations
   - Education, training, and library occupations
   - Arts, design, entertainment, sports, and media occupations
   - Healthcare practitioners and technical occupations
2. Prosaic Occupations
   - Office and administrative support occupations
   - Installation, maintenance, and repair occupations

From this, we developed criteria for determining if a TAZ was more oriented toward *prosaic* information activities or if it were more oriented towards *extrapolative* activity. Generally, this was built upon our previous determination of dominant information TAZs. These areas were considered extrapolative if more than 50% of the workers held the extrapolative occupations. Thus, the overall criterion for an extrapolative TAZ was as follows: more than 60% of the workers in a TAZ were information workers and of those more than 50% were in "extrapolative" occupations. Maps 53–58 in the Map Portfolio display the resulting patterns.

In Atlanta, the central business district changed from a largely prosaic information worker center to a predominantly extrapolative center. Seven TAZs in the central district were added with extrapolative worker dominance that had not been centers in 1990. Eighteen TAZs changed from prosaic dominance to extrapolative dominance in the decade and six TAZs remained with extrapolative dominance throughout the decade. Thus, the central area experienced a significant shift in land area from a prosaic dominant center to an extrapolative dominant center. The conurbation of TAZs in the north at the intersection of I-285 and the McDonald Parkway also shifted from a prosaic grouping to an extrapolative grouping. The northwest cluster of TAZs (at the intersection of I-28 and I-75) was more of a prosaic oriented center in 1990 but shifted so that in 2000 there seemed to be roughly equal area between TAZs oriented to extrapolative dominance and those oriented to prosaic dominance.

In Austin, the central business district also shifted to TAZs possessing workers with occupations that were more extrapolative. Of 35 TAZs in the central area, extrapolative workers dominated in 29. Twelve of those had shifted from prosaic to extrapolative between 1990 and 2000. Extrapolative workers also dominated the suburban center at the confluence of Routes 1, 183 and 360. Of 14 TAZs, ten were extrapolative in character of which 6 were new in 2000. The four prosaic TAZs, however, were also new in 2000. This was a center that greatly expanded in the decade. An important new extrapolative TAZ is found in the north near Round Rock—this is the location of the headquarters of the Dell computer manufacturer that was located in Austin in the 1990s.

In 2000, the Denver area presents four distinct areas of TAZs dominated by extrapolative occupations: the two central areas of Denver and Boulder, a largely new development in the vicinity of Bloomfield on the highway between Denver and Boulder, and a conurbation in the southeast linked by I-25, I-225 and Route 470. Downtown Denver, as with Atlanta and Austin shifted from a group of TAZs that emphasized prosaic information workers to a group focused on extrapolative occupations. Of 38 TAZs making up the central area, 28 were dominated by extrapolative activity; 11 were new centers and 16 were centers that shifted from prosaic to extrapolative workers. (Four were no longer centers above the regional mean in workers, apparently displaced by other development.) In Boulder, in 2000, 18 of 33 TAZs were dominated by extrapolative occupations, four of which had shifted from prosaic since 1990. (Twelve TAZs were no longer above the regional mean in workers, reflecting possible changes in land use.) In the decade between 1990 and 2000, on the highway between Denver and Boulder (around the intersection of Routes 36 and 287, near Bloomfield) was developed a new activity center dominated by workers in the extrapolative occupations. The other major concentration of extrapolative workers lies in the southeast bounded by I-25, I-225 and route 470. This had been a center of mostly prosaic workers in 1990 with 16 of 19 TAZs shifting to extrapolative workers by 2000.

The central business district of Houston also shifted from prosaic information occupations to extrapolative. In 2000, sixteen of 21 TAZs were dominated by extrapolative occupations, up from only 3 in 1990. Twelve shifted from prosaic activites in 1990 to extrapolative in 2000. Five TAZs remained focused on prosaic occupations throughout the decade. West University Place, at the intersection of I-610 and Route 59 experienced a similar shift from prosaic to extrapolative. In 1990 only one TAZ was predominantly extrapolative; by 2000 seventeen TAZs were predominantly extrapolative. Nine of the 17 represented a shift from prosaic to extrapolative and seven represented expansion into new TAZs. Extrapolative occupations also expanded somewhat south of Rice University along Alternate Route 90 and Brays Bayou. The decade also some expansion of 1990 centers of extrapolative occupations to the west on I-10 at the western border of the region and to the southeast at I-45 and Marina Bay Drive in the vicinity of Webster. A wholly new extrapolative center emerged to the northwest on Route 249 north of Willowbrook Mall.

In Phoenix, the traditional central business district is encircled by interstates I-10 and I-17. While this area is the traditional downtown, neither in 1990 no in 2000 was it marked as an intensive center of information workers. Of a total of 38 TAZs, only 16 were represented as having more than 60% of its workers having information occupations. In 2000, nine of the information worker TAZs remained prosaic in character while six TAZs represented areas shifting from prosaic or non-information oriented to extrapolative. In summary, the Phoenix downtown area—in contrast to the other five regions studied—is marked by a greater diversity of information and non-information occupations. North of the central area, on the other hand is an area extending about five-six miles from downtown that has seen significant expansion of information workers—both extrapolative and prosaic. Twenty-three TAZs had a predominance of extrapolative workers in 2000 of which only five such TAZs existed in 1990. Six had shifted from prosaic occupations and twelve TAZs represented new developments. In the same region, 24 TAZs have a predominance of prosaic workers of which ten represent an expansion from 1990. The final notable area of expansion lies in the TAZs south of the Arizona State University Main Campus in Tempe. Apart from these concentrations, numerous TAZs represent expansions over 1990 for extrapolative occupations and are scattered throughout the region although mainly oriented to a northwest-southeast axis from Sun City to Mesa.

In the Twin Cities the downtowns of both Minneapolis and St. Paul shifted toward a dominance of workers with extrapolative occupations. For Minneapolis, in 1990 six TAZs featured a predominance of extrapolative workers and 9 had a predominance of prosaic workers. By 2000 this had shifted so that 15 TAZs were oriented toward extrapolative workers and only five toward prosaic workers. A similar pattern can be seen in St. Paul—in 1990 only three TAZs were extrapolative in character and in 2000 this had grown to nine. Another major area of growth of extrapolative occupations is found in the southwestern suburbs of Edina and Eden Prairie, influenced by I-494 and the Crosstown Expressway. This region had only one TAZ oriented toward extrapolative occupations in 1990, while in 2000 this had grown to 12. This "edge city" development is somewhat diversified with three TAZs oriented to prosaic occupations throughout the decade and four more added during the decade.

Overall, it can be seen that central business districts in the six metropolitan areas have substantially shifted toward extrapolative occupations in their workforce makeup. The only exception to this is found in Phoenix where primary growth in extrapolative workers lies in two areas—one adjacent to downtown to the north and the second south of Arizona State University. Both areas are roughly within five-six miles of the central district. Even in regions with more than one central area (Denver-Boulder, Minneapolis-St. Paul, and Houston with its downtown and West University Place), both business districts have substantially shifted territorially to districts featuring extrapolative occupations. Also notable is that new suburban development of extrapolative oriented TAZs between 1990 and 2000 tend to occur in clusters rather than randomly scattered. These new conurbations are almost entirely located in close relation to major highways—particularly on the ring roads within 10–12 miles of the central business district. Only in Atlanta do we find newly developed extrapolative TAZs beyond 18–20 miles from the center. In short, analyzing centers of employment of information worker in the decade in this fashion suggests that there is a tendency to concentrate—both in central business districts as well as in "edge city" and university districts.

Local Moran's "I"(Maps 59–64). The final analysis for examining the development of center clusters in TAZs utilized Anselin's Local Moran's I, as described in Section2.1.3. This procedure provides a local indicator of spatial association depicting the extent to which the value of a zone is similar or different from observations of neighboring zones. As noted in Section 2.1.3, the distance of any given zone from any other zone is also a factor of weighting so that a traditional distance decay function is applied in determining the Local Moran's I. It will also be recalled from the discussion in Section 2.1.3, that the key aspect of Moran's I is its variation from an expected value of I—a value that would reflect a purely random pattern.

In plotting the Local Moran's I, since its absolute value is of little significance, the maps indicate the number of standard deviations from the mean for each TAZ. A high positive standard deviation indicates a high degree of similarity with adjacent TAZs, or a "hot spot"—a location of high levels of similarity. Conversely, a high negative standard deviation would reflect a "cold spot." The variable being analyzed for the maps is worker density in each TAZ. These analyses are of interest mainly in comparison with our simpler method of determining center clusters based on the mean worker density for each region. In inspecting Maps 59–64, it is apparent that this method of analysis identifies TAZs with strong association with neighboring areas and thus tends to miss the more isolated high density TAZs. Thus, Maps 59–64 offer a visual analysis of the strongest conurbations in terms of density. Comparing them with Maps 37–42, we find the greatest congruence between the two methods in Houston, Phoenix and Minneapolis-St. Paul—the three regions that seem to most have pronounced employment districts.

Overall, utilizing this method of determining centers of employment density appears to be less comprehensive than using the prior method of working with TAZs above the regional mean density. It does appear, however, to be a method that is helpful in determining strength of association among TAZs and thereby providing a picture of density "hot spots" in each region.

2.1.7 Discussion and Conclusions

This study reflects the urban spatial dynamics of information technology and the occupations that were spawned by its expansive growth in the 1990s. It should be noted that in 1990 the full development of the Internet had not yet been realized. Its diffusion accelerated in the mid-1990s so that by the end of the decade "dot com" had become an integral part of everyday speech not only in commercial and academic circles but also in a wide variety of other activities from games to music to dating services. By 2000 e-mail was a common form of communication as was wireless telephones. In 1990, e-mail was primarily for an intelligentsia in academic research settings and governmental (primarily military) institutions. Even more than previous technological innovations, information technology in the 1990s created a host of new occupations and provided a broad array of new opportunities.

In 1990 the backbone infrastructure that is essential for information workers to engage contemporary information technology was only in beginning stages of evolution. The infrastructure available at that time was essentially concentrated in downtown centers, in university centers and in emerging office parks focused on highway and air transportation. Thus, understanding the spatial impact of information technology on information workers and their subsequent influence on urban spatial structure can begin to be seen from our analysis.

Having examined the location of workers with information occupations from a number of analytical approaches, it is apparent that these workers do appear to be more concentrated than other workers in the six metropolitan areas studied, at least insofar as the location of their work is concerned. This did not change appreciably between 1990 and 2000, although the spatial patterns became more complex. Dispersion of information-dominated workplaces did take place to some degree. However, as the analysis of workplace centers suggested, suburban job locations for information workers tended towards clustering. Central business districts in all six urban regions maintained their role in the decade as dominant centers for workers possessing information occupations. Moreover, the major shift in the central business districts was toward a concentration of what was defined as extrapolative occupations—that is occupations involved in analyzing, interpreting, designing and decision-making as distinct from prosaic “back office” administrative and technical support. Thus, central business districts further consolidated their role in innovation and wealth creation. In all six regions, central areas also had in close proximity large, research oriented urban universities and these undoubtedly played a role in enhancing this concentration. Notwithstanding this, suburban “edge city” locations also created new centers dominated by information workers. Thus, as was suggested at the outset, workers in the information occupations experienced both centripetal and centrifugal forces during the decade. In the main, however, whether in central or suburban locations, information workers tended to cluster more than workers in non-information occupations.

This was not the case in the matter of residential location, however. In 1990, information workers tended to favor residential locations that were slightly more concentrated than other, non-information workers. By 2000, however, this difference had largely disappeared and information workers were not significantly more concentrated. In short, the workplaces of information workers tend to be clustered but their choices in residential location seem to be little different than those of non-information workers. Organizations (private, public or non-profit) needing to hire significant numbers of information workers tend to make location decisions that favor clustering, but the information workers themselves behave in the residential land market the same as anyone else.

In combining workplaces that are more concentrated with residential locations that are more dispersed, commuting times to workplace TAZs for information workers were considerably higher than for workers in general, and these times lengthened during the decade. This implies that one consequence of the major advancement of information technology in the decade of the 1990s was not a reduction in the demand for transportation services but rather an increase.

In the decade between 1990 and 2000, land use planning at the local or regional scale paid scant attention to information technology. Wireless telecommunications created some pressures on local governments to adopt zoning ordinances dealing with relay antennae and the creation of the backbone infrastructure for the Internet. Some local governments did develop and carry out plans to become a “wired” city. For the most part, however, this planning was not carried out with any integral relation with land use planning. The spatial planning that did take place was often focused primarily on planning for economic development. The results of the 1990s, with information occupations driving the labor force, work locations for these occupations being clustered either in the central downtown or in suburban “edge cities,” and their residential locations being dispersed, suggests an important new look at the precepts of smart growth and efforts to alleviate traffic congestion, environmental degradation, and other problems associated with urban sprawl.

The results of this study provide a framework for further research concerning the impact of information technology on urban spatial structure. Given forecasts of future labor force with information technology occupations, and given the tendencies toward clustering as took place in the 1990s, further study of the impacts of these workers, their employers, and their technological artifacts on the commercial and industrial land markets of metropolitan regions would be important. In terms of developing such future forecasts, further study needs to address whether the diffusion rate of information technology innovation has reached a plateau that is leveling off or will it continue to grow as in the 1990s. Another important area of research would be examining the correlation of information occupations with economic sectors, such as the penetration of information occupations with the retail sector, the non-durable manufacturing sector, the energy and utility sectors, etc. Will such penetration become ubiquitous such that, over the next

10–20 years the spatial patterns of workplaces for information workers will be less unique and clustered and become more like the spatial workplace patterns of non-information workers? Follow-up studies would be important to determine whether the 1990s was a unique decade that, while witnessing a full flowering of the information economy, has played itself out and is not likely to be repeated.

In sum, for the six urban areas studied, the information economy that emerged in the 1990s spawned a new labor force driven by significant new growth in information worker occupations at the expense of other occupations. For the six areas, information workers grew at rates faster than both their local economies and the national economy. These worked out on the ground as reinforcing traditional land use patterns for clustered workspaces and contemporary popular land use patterns for the place of residence of these new workers. The study has reinforced the notion that information workers thrive best on proximity to other information workers to a significant degree but with the substantial growth and dynamic of the 1990s, proximity is a relative characteristic mediated by limits of land markets. Thus, the workplaces of information workers were more dispersed in 2000 than in 1990, but in that suburban movement, such workplaces tended to be clustered. Overall, the speculative thinking during the 1990s that information technology would change the shape, form and function of the contemporary city has not been borne out in this study. Indeed, there seems to be some reinforcement of traditional urban form.

2.1.8 References

1. Richard S. Bolan and Marcus Martin (2003).“Case Study: Information Technology In The ‘Old Economy:’ Implications For Location Decisions And Transportation Demand.” Unpubl. ms., State and Local Policy Program, Humphrey Institute, University of Minnesota.
2. John R. Baldwin and David Sabourin (1999). Innovative Activity in Canadian Food Processing Establishments: The Importance of Engineering Practices. Ottawa, Canada: Micro-Economic Analysis Division, Statistics Canada, Ms. 11F0019MPE No. 101. 19pp.
3. Manuel Castells (2000). The Rise of the Network Society, Volume 1, Second Edition. Malden, MA: Blackwell.
4. Couclelis, H. (1996). “Spatial Technologies, Geographic Information and the City,” Technical Report 96–10. National Center for Geographic Information and Analysis. University of California at Santa Barbara.
5. Martin Dodge and Narushige Shoide (2000). “Where on Earth is the Internet? An Empirical Investigation of the Geography of Internet Real Estate,” in Cities in the Telecommunications Age: The Fracturing of Geographies, edited by J.O. Wheeler, Y. Aoyama, and B. Warf. New York: Routledge..
6. Joshua M. Epstein and Robert Axtell (1996). Growing Artificial Societies: Social Science from the Bottom Up. Cambridge, MA: MIT Press with Brookings Institute.
7. Edward J. Feser (2001). “What Regions Do Rather Than Make: A Proposed Set of Knowledge-Based Occupation Clusters.” Paper presented at the Annual Meeting of the Association of Collegiate Schools of Planning, Cleveland, Ohio. November.
8. S. Graham and S. Marvin (1996). Telecommunications and the City: Electronic Spaces, Urban Places. London: Routledge.
9. Amy Glassmeier (1985). “Innovative Manufacturing Industries: Spatial Incidence in the United States,” in High Technology, Space and Society, ed. By M. Castells. Beverly Hills, CA: Sage.
10. Darrene Hackler (2000). “Industrial Location in the Information Age: An Analysis of Information-Technology-Intensive Industry,” in Cities in the Telecommunications Age: The Fracturing of Geographies, edited by J.O. Wheeler, Y. Aoyama, and B. Warf. New York: Routledge.
11. Yoshiko Kurisaki and Marumi Yanagimachi (1992). “The Impact of Inforamtion on the Economic Development of Sub-Regional Centres: A Trial Application of an “Information Activity” Index to the 43 Cities in Japan,” in The Economics of Information Networks, edited by C. Antonelli. Amsterdam: Elsevier Science Publishers.
12. Jay Lee and David W.S. Wong (2001). Statistical Analysis With ArcView GIS. New York: John Wiley and Sons, Inc.

13. Nancy Green Leigh (2000). "People Versus Place: Telecommunication and Flexibility Requirements of the CBD," in Cities in the Telecommunications Age: The Fracturing of Geographies, edited by J.O. Wheeler, Y. Aoyama, and B. Warf. New York: Routledge..

14. Sean Gorman (2002). "Where are the Web Factories: The Urban Bias of E-Business Location," Tijdschrift vor Economische en Sociale Geografic, 93:5, 522–536. Also available on the web at: http://64.227.18.27/research/web-factories/ or http://www.casa.ucl.ac.uk/cyberspace/gorman_tesg.pdf.

15. Daniel A Griffith and Carl G. Amrhein, (1991). Statistical Analysis for Geographers. Englewood Cliffs, NJ: Prentice-Hall.

16. Ned Levine and Associates (2002). CrimeStatII: A Spatial Statistics Program for the Analysis of Crime Incident Locations. Washington, DC: The National Institute of Justice.

17. A. Markusen (1987). Regions: The Economics and Politics of Territory. Totawa, NJ: Rowman and Littlefield.

18. Markusen, et al (2001). "High-Tech and I-Tech: How Metros Rank and Specialize, Unpubl. ms., Project on Regional and Industrial Economics, Humphrey Institute, University of Minnesota.

19. A. Markusen, P. Hall, and A. Glasmeier (1986). High Tech America: The What, How, Where and Why of the Sunrise Industries. Boston: Allen and Unwin.

20. Moss and Townsend, 1997

21. Michael E. Porter (1990). The Competitive Advantage of Nations, New York: The Free Press.

22. Jonathan Rauch (2001). "The New Old Economy: Oil, Computers, and the Reinvention of the Earth," The Atlantic Monthly (January), 35–49.

23. Quing Shen (1999). "Transportation, Telecommunications and the Changing Geography of Opportunity," Urban Geography, 20:4, 334–355.

24. Alvin Toffler (1980). Future Shock.

25. U.S. Bureau of the Census, Census Transportation Planning Package, 1990 and 2000. web site: www.fhwa.dot.gov/ctpp

26. Barney Warf (2000). "Compromising Positions: The Body in Cyberspace," in Cities in the Telecommunications Age: The Fracturing of Geographies, edited by J.O. Wheeler, Y. Aoyama, and B. Warf. New York: Routledge.

27. Zook, M.A. (1998). "The Web of Consumption: The Spatial Organization of the Internet Industry in the United States," paper presented at the Annual Conference of the Association of Collegiate Schools of Planning, Pasedena, CA, November 1998.

Endnotes for 2.1

Spatial Patterns of Information Workers

---

[i] We have worked out a mathematical equivalency algorithm, but this means that the 1990 data are, strictly speaking, estimates derived from this algorithm.

[ii] 2 TAZs are determined in collaboration with local metropolitan planning organizations (MPO's) and thereby differ in criteria for size and bear no relation to the overall size of any given metropolitan region.. Thus, Atlanta has 1685; Austin, 706; Denver, 2462, Houston, 2642; Phoenix, 1197; and Minneapolis-St. Paul, 1201. The average number of TAZs for the six study areas is 1650.

[iii] Computer Specialist Occupations include: Computer and Information Research Scientists, Computer Programmers, Computer Applications Software Engineers, Computer Systems Software Engineers, Computer Support Specialists, Computer Systems Analysts, Database Administrators, Network and Computer Systems Administrators, Network Systems and Data Communications Analysts and Miscellaneous Computer Specialists. Mathematical Science Occupations include: Actuaries, Mathematicians, Operations Research Analysts, Statisticians and Miscellaneous Mathematical Science Occupations.

[iv] These calculations resulted in the northings and eastings for the mean centers in each metropolitan area based on their UTM zones.

[v] Unfortunately, commute times as reported in the CTPP do not allow for persons holding more than one job. Also, the data do not distinguish between direct commuting trips (home-to-work and return) and commuting tours that may include stop-offs at schools, shops, banks, etc.

[vi] Identified were only those universities that occupied a major share of the TAZ area. Thus smaller schools or colleges occupying only a few buildings or city blocks were omitted. Hence, "university districts" are not an exhaustive or complete listing of colleges and universities in any of the metropolitan areas. Generally, only the very largest are represented.

[vii] If the highest and lowest are omitted from the mean calculation, the resulting mean remains at 13.2%.

[viii] There are of course TAZs that either had no employment or worker densities below the regional mean in both 1990 and 2000. These make up the bulk of the TAZs and are either residential areas or undeveloped areas.

CHAPTER 3

Kevin J. Krizek: Travel Behavior Impacts of Information and Communications Technologies (ICT)

3.1 ICT as a Substitute for Non-Work Travel: A Direct Examination

Kevin J. Krizek
Assistant Professor, Urban and Regional Planning
Director, Active Communities Transportation (ACT) Research Group
University of Minnesota
301 19th Ave S.
Minneapolis, MN 55455
T: (612) 625-7318
F: (612) 625-3513
E-mail: kjkrizek@umn.edu

Yi Li
Humphrey Institute of Public Affairs
University of Minnesota
301 19th Ave S.
Minneapolis, MN 55455

Susan L. Handy
Associate Professor
Department of Environmental Science and Policy
2132 Wickson Hall
University of California at Davis
Davis, CA 95616
T: 530-752-5878
F: 530-752-3350
E-mail: slhandy@ucdavis.edu

3.1.1 Introduction

Problems related to traffic congestion and auto pollution continue to mount in urbanized areas across the globe. In response, planners are devoting increased attention to alternatives to driving, including non-motorized modes such as walking and biking but also efforts that replace the act of physical travel with virtual travel through the use of Information and Communication Technology (ICT). Each time these alternatives substitute for what would normally otherwise be an automobile trip, problems like traffic congestion and auto pollution lessen a bit. Given increasing rates of use of ICT, the impact of substitution could be significant. The question is, does ICT use substitute for physical travel, particularly automobile use, and if so, how much?

ICT appears to be transforming people's everyday life in fundamental ways. From ordering appliances online to buying stocks over the telephone; paying credit card bills over the Internet to checking one's account balances at an ATM machine; chatting with friends using instant messenger to buying flowers on line. Households are completing more of their activities through a wider range of medium (namely, electronically), well beyond traditional forms. Between 2000 to 2002, there was a reported 88% increase in the number of people who have use online banking services and a 40% increase in the number of people who have made a travel purchase over the Internet[i]. In addition, the Federal Reserve Bank reported a 14% drop in the number of paper checks processed from 1995 to 2000 [2], as more and more people are paying bills and making purchases over the Internet.

Despite such dramatic increases in ICT use, vehicle travel per person continues to rise [3] and traffic congestion continues to worsen, even when considering congestion on a on a per capita level [4]. At an aggregate level, at least, these trends point to an inverse relationship, suggesting the relationship may not be as strong any many believe. A recent study showed that telecommuting has had at best a small impact on vehicle travel [5] and other anecdotal accounts suggest similar weak relationships exist for non-work travel. For a variety of reasons, people may still prefer physical travel for myriad purposes.

This paper investigates dimensions of ICT use at the household level as it applies to matters of non-work travel. We seek, in an exploratory manner, to answer two broad questions: (1) what is the pattern of substitution effect between traditional and ICT-form activities and (2) what attributes of people affect the choice of whether or not to substitute? We do so using the results of a household survey that contains several questions querying household ICT use and substitution and allows us to probe into the substitution aspect of household travel behavior. The results help understand the million-dollar question—to what degree will ICT substitute for physical travel.

Theoretical Context and Approaches from Previous Work

Salomon offered the theoretical context for this research over twenty years ago in a seminal work on ICT and travel. In it, he provided a foundation that outlined four basic types of interaction [6, 7]. These types of interaction—substitution, modification, generation, and neutrality—have been subsequently explored in several works [8–10]. The first and most anticipated interaction is the substitution of travel, generally referring to trips that are eliminated as a result of ICT use. This interaction has been the focus of most of the ICT attention among planners, and the substitution of subsistence trips, by definition, is a phenomenon inherent in the term telecommuting [11]. While the substitution hypothesis is one that holds great hope for many, the degree to which substitution occurs is considerably less than originally anticipated [8], due in part to different interpretations of what is meant by substitution.

The second interaction—modification—refers to travel that is altered by the use of ICT, primarily by a shift in the timing and routing of trips (spatial and/or temporal transformations). It may also refer to the manner in which trips are linked together (i.e., trip chaining) or even the mode of travel. The benefits of shifting even a small amount of peak hour travel to different times are mentioned as a key strategy for reducing levels of congestion. Some previous work suggests that ICT will modify travel behavior [12–14]; the outstanding issue lies in understanding whether such modification leads to more or less travel. Generation refers to any additional travel that would not have occurred but for the use of ICT. It is possible that the use of ICT stimulates demand for new activities that require travel. Finally, neutrality is used to describe instances in which ICT use does not affect travel but instead replaces other kinds of activities.

The question of the degree to which—or whether—ICT will substitute for travel remains a topic frequently broached in almost all studies on the subject. Many reports and papers make more than a passing reference to the matter [see for example, 15] in efforts to better understand its impact. The intricacies of the substitution issue, however, are myriad and complex. They are not unlike those being wrestled with for other alternatives to driving, for example, whether walking to a store replaces or is made in addition to driving to the store [16–19]. For as much attention as travel substitution receives in policy discussions, it is surprising that a clearer or more explicit methodology has not emerged to investigate its impact. Most research gently skirts the topic when it comes to estimation.

Part of the difficulty is that the relationships between substitution and total travel are both direct and indirect. Most directly, a particular instance of substitution means that the use of ICT replaces a trip (most often an automobile trip) that would otherwise be made, thereby reducing total travel. However, if the ICT-user saves time completing a purchase on the Internet and then, in turn, capitalizes on this time savings by driving to complete a different errand that s/he would not otherwise have made, the savings in travel from substitution could be partly or wholly offset by the generation of travel. In addition, the modification effect can lead to a reduction in travel even if no substitution occurs, for example, when a shopper uses the Internet to search for a closer store than she otherwise would have chosen. By examining the relationships between ICT use and total travel, either at the individual level or at an aggregate level, researchers can determine the net effect on travel of the combination of substitution, inducement, and modification, but cannot isolate the effect of each.

To isolate the substitution effect, a more direct approach is needed that examines the effects on travel of particular instances of ICT use. While several studies have used a direct approach to examine the substitution question for the work commute [20, 21], the literature offers little in addressing this concept for non-work travel. The only previous research that we are aware of that is directly applicable to the topic as it relates to non-work travel is offered by Handy and Yantis [22]. This study focused on in-home versus out-of-home versions of three non-work activities – shopping, banking, and movie watching. Respondents to a three-city household survey were asked to speculate on what they would have done the last time they used an in-home version of each activity had that option not been available. Rather than inferring substitution from changes in travel, the approach provides a direct assessment of substitution for particular instances of ICT use.

We used this work as a template for our study, modifying the survey instrument to ensure its applicability given advances in ICT since 1997 and restructuring some of the substitution questions. Recognizing that our pursuit focuses only on non-work related tips, our end result was to shed light on the following questions. For what types of activities are people most likely to substitute? How frequently do they engage in these activities? Thinking of their last transaction, would they have substituted this trip with ICT and why or why not. What factors influenced this decision? What type of person is more likely to substitute travel with ICT?

3.1.2 Survey Methodology

To examine the above dimensions, our research administered a mail-out and mail-back survey of households in the spring of 2003. The survey sample was drawn randomly from the population in three metropolitan statistical areas (MSA): Seattle, Kansas City, and Pittsburgh. Several criteria guided the selection of these cities. First, we wanted a sample of residents that were representative of relatively large US urban areas from Midwest and the East and West coasts. Next, our basic hypotheses about spatial attributes and IT availability were instrumental in stratifying our sample by balancing two other criteria: level of traffic congestion for large cities and Internet penetration. Using Texas Transportation Institute's measures of traffic congestion in US urbanized areas in 1990–1999 [4], we honed in on candidate cities representing both high and low levels of traffic congestion. Next, we balanced this criteria by controlling for IT availability in each city, as measured by rates of Internet penetration by Scarborough Research [23]. Our final selection included Seattle, Kansas City, and Pittsburgh. Seattle represented a high congestion/high technology city (for example, rating second on the congestion list and among the highest of cities in IT availability); Kansas City represented a low congestion/high technology city and Pittsburgh was selected to represent a low congestion/low technology city.

After pilot testing, we mailed the survey to 800 households in each of the three cities in May of 2003. The households were selected at random from an address database maintained by Survey Sampling International. The primary source of their database listings is telephone directories, which was supplemented with other sources, such as state identification cards. A cover letter explained the purpose of the survey and invited a household member 18 years or older to complete it. Following the Dillman [24] method, reminder postcards were mailed to the 2,400 households one week later and a the third mailing (with the complete survey) was mailed out three weeks later to each household that had not yet replied. The overall response rate is 31%, including 32% from Seattle, 30% from Kansas City, and 30% from Pittsburgh. After comparing socio-demographics with 2000 Census data for each metropolitan area, we found that the characteristics of our survey respondents differ only slightly from the overall populations in the three cities. Most notable is the fact that our survey respondents are substantially older than general population in the three cities (60–70% are above 50 years old, in contrast to 30–40% in the overall population; 36–37% are over 60 years old, in contrast to 18–28% in the overall population). Compared to census figures, other differences showed our sample to (a) have a slightly higher percentage of males, (b) be more educated, and (c) have slightly higher household incomes than the general population. On one hand, higher income levels would suggest a bias in terms of use of ICT related services. Given that more elderly populations are probably less likely to be using the Internet for services, this would likely sway the bias in the opposite direction.

Our survey queried households about their general use, level of comfort, and frequency of ICT related activities. Several groups of questions focused specifically on the degree of substitution of various activities. For ease of describing the analysis, we have divided these questions, roughly, into three different groups. Each investigates a slightly different dimension of the phenomena.

The first group asks the user about the general role of telecommunication in eliminating and affecting trips or travel in which they most frequently engage and those activities for which they are most and least willing to use ICT to substitute[ii]. The second group of questions asks the respondent to think about their last use of a particular activity and to speculate what they have done if that version had not been available[iii]. A third group aims to better understand the motivations for physical (in-store) shopping[iv]. Many of the questions asked the respondents to select from a fixed set of responses, although almost all questions had an open response category. Some questions were open ended; the advantage of the latter strategy being that participants are not influenced or constrained by a fixed set of response categories. Results from open-ended questions might confirm or might refute what we would discover from close-ended questions, or they might offer brand-new perspectives. Using three slightly different groups of questions, and subsequently three different strategies for honing on the substitution issue helps shine brighter light on one or more dimensions of the travel substitution question. None of the approaches, however, yield definitive answers on the matter. In combination with one another they help illuminate a topic often discussed and hypothesized but rarely systematically examined in the literature suggesting that analysis of this sort is more of an art than an exact science.

3.1.3 Results

ICT Impact on Select Trips

Most frequent trip from home. Our analysis begins with the first group of questions, asking the respondents to reflect on their most frequent trip and those for which they would most likely be willing to substitute with forms of ICT.

Our first step in doing so queried respondents to identify their most frequent trip, other than going to work or school. The overwhelming response to this question was the grocery store

(54%), followed in distant second by visiting friends and relatives (13%) and entertainment (11%). The next immediate question in the survey asked them if they could use telecommunications in their home to eliminate the aforementioned trip, would they choose to do so. A mere 14% said they would be willing to do so. A portion of those who would not (64%) mentioned that it was impossible to substitute this trip (e.g., transporting children to school or going to the fitness center); another 22% said they would not choose to do so, regardless.

Table 3.1.1 breaks the responses down by the seven most frequent types of trips. Our first observation is that for three of the categories (taking kids to school, going to the fitness center, and other), respondents clearly felt ICT was not a possible substitute for this activity (as judged by more than 50% of the respondents claiming so). Two categories (grocery store and entertainment), are worthy of note, although none of them appeared rank high in terms of substitution. Grocery store and entertainment enjoyed roughly the same propensity for substitution (approximately 18%); collectively, these comprise 64% of the most frequent trips. Banking was the clear front-runner in terms of substitution likelihood (27.3% responded they would be willing to do so), however, it comprised less than 2% of the most frequent trips.

In general, substitution is not perceived as possible for many of the trips that people frequently make from home. But of those trips for which ICT is deemed possible, a clear majority of the respondents appear unwilling to do so. We could calculate how many "most frequent" trips in the sample would likely be removed by ICT through using a weighted average and multiplying the number of trips denoted as most frequent by the percentage who would be willing to substitute and then summing the responses. Of the 692 "most frequent trips" identified, respondents would be willing to substitute ICT for only 100 of them (14.5%)' the and bulk of these come largely from grocery and entertainment.

Most and least willing to substitute. The above discussion is restricted because it asks the respondent to focus only on their *most* frequent trip (other than the work commute). Doing so limits the respondent to reflecting on only one type of trip. There may be any other activities, not just the most frequent, vying for substitution alternatives. We therefore asked two separate questions that broadened the scope to *all* the trips one makes from home, asking open-ended questions about which activities people would be *most* and *least* willing to substitute through the use of telecommunications.[v]

Summarizing the results to these questions, while interesting, is messy to sort though. After looking at each response, we systematically divided them into different categories (see Table 3.1.2). It is messy because the top three responses in terms of the trip *most* willing to substitute and *least* willing to substitute are alike. These include grocery shopping, non-grocery shopping, and banking. More specifically, grocery shopping ranked among the highest for each question. We are caught in a crossfire.

This suggests that people diverge on the issue of whether to substitute certain types of activities with ICT. One household's meat is another household's poison. While one person might enjoy shopping online, another person might think that the joy of in-store shopping is irreplaceable by the Internet. While one person might marvel at the convenience of being able to bank online, another person might shun at the idea of completing private financial transactions over the perceived insecure Internet. As a consequence, we get almost the same responses at the two ends of the same spectrum, suggesting important market segments. However, some activities appear to have a pattern. People are least willing to visit their friends and relatives at home with the convenience of ICT. They had rather make a trip and see their friends and relatives in person. In addition, work is among those activities people are most willing to give up.

Substitution for Banking, Shopping, and Financial Transactions

The second group of questions honed in on activities for which there is a well-known ICT substitute; these include electronic banking, shopping, and financial transactions. Our survey asked respondents to recall the last time that they used an ATM, e-shopping or the Internet to complete other financial transactions and to tell us what they would have done if that ICT option was not available. This style of question relies on respondent recall, which is troubled, in part, by people not fully remembering the detailed contextual circumstance of their decision. However, it is advantageous because it provides a real context in which they used ICT which in many ways is preferred than asking them to respond to completely hypothetical scenarios. The results are summarized in the Table 3.1.1and are displayed in one of three columns based on the projected impact the response has on overall travel.

In general, responses to these questions suggest that, without certain ICT options, more physical trips would be incurred for certain activities. Such behavior is not simple to sort through and a key question is how many more? Among those who purchase online, 78.9% would have gone to the store had they not been able to do so. Among those who use online banking, 27.6% would have visited the bank at that time if they had not had the option of banking via the web. These are direct trip savings, which supports the hypothesis that there is indeed substitution effect as an advanced form replaces the older form of the activities. When it comes to other financial transactions, our results indicate that the online version is clearly replacing use of the U.S. mail (76.4%) but this is not necessarily replacing a trip that would have otherwise been made. For this category, there is very little direct trip savings, as a mere 10.8% indicated that they would have made a physical trip for the lack of the online option. Therefore, we may say that the online version for banking is mostly substituting other in-home versions rather than reducing physical travel.

ATM use is an ICT option that presents a complex picture. Because going to an ATM also involves a physical trip, this benefit is better tallied as time-savings or enhanced convenience rather than trip reduction. The last time respondents used an ATM, nearly 50% reported that they would have visited the bank if there were no ATM option. There would be no difference in actual travel if going to an ATM involves equivalent travel as going to a bank; however, because ATMs are found at many locations other than banks, the travel distance to an ATM is likely to be shorter [22]. Another 20% reported that they would have waited till their next trip to the bank. For these people, the use of ATM represents additional banking made possible by this technology; it is more a matter of convenience than substitution. Another 28% indicated they would have cashed the check somewhere other than the bank. Under these circumstances, the substitution, if any, is not on physical bank trips.

The survey followed up on the ATM issue by asking the respondent about the overall impact and ATM has on bank visits. For those who currently use ATMs, 23.3% reported that they would have visited the bank every time they would have used ATM. This indicates direct one-to-one substitution. Slightly more than half (51.3%) said they would visit the bank fewer times than they would have used an ATM, but more often than they currently do. This shows a mixed effect of bank trip saving and increased ATM visits at the same time. For the remaining 25.4%, ATM use is not substituting bank visits for these people at all, as they would continue to visit the bank at the same frequency as now if there were no ATM option. It thus appears as if ATM availability merely increases total banking.

Modeling the Typical ICT Substituter

Our final analysis using this group of "what if" type questions aims uncover explanatory factors behind those who are more likely to substitute physical trips with ICT. We do so using a multivariate regression model operationalizing a dependant variable using a subset of the above questions, namely whether the respondent banked online or purchased online. By identifying respondents who answered affirmatively to *either* of these questions, we can distinguish individuals who may be more or less likely to substitute. The contrast group to those deemed substituters would therefore comprise both those who are not using ICT or those using ICT but only in addition to their physical trips. Using this distinction, the online activity in which substituters engage is assumed to be a direct substitution to a physical trip.

There are two considerations to keep in mind using this logic. First, those individuals coded as substituters may not *always* be substituting for the given good (as the survey questions only address the last time the person uses ICT). Second, we are asking them to reflect on what they would have done, assuming an ICT option. We acknowledge this approach fails to provide a crisp substitution question or logic, but it provides a first step in understanding this population, given the complexity of the phenomenon being modeled.

We use five different dimensions thought to explain the population of substituters. The first are socio-demographic variables including age in years, sex, education, household income, having two-plus vehicles, household structure (households without teenage children, households with teenage children), and whether the respondent is employed. The second category measures of the availability of IT related infrastructure and were gathered at both the metropolitan and the household level. The metropolitan level measures are rates of Internet penetration where, as previously described Seattle and Kansas City were deemed high and Pittsburgh was deemed low. The household measures were gleaned from the survey and captured ATM card ownership, cell phone ownership, home computer ownership, and home Internet access. The third are behavioural variables, including frequency of non-grocery shopping (where more than 5 times in the last 30 days was deemed "frequent," and otherwise is "infrequent"), and frequency of bank visits (with 1–3 times per month or more being "frequent," and otherwise is "infrequent"). The fourth are attitudinal variables which were gleaned from a dozen or so attitudinal questions answered by the respondents such as I feel comfortable using computers, technology makes my life easier, traffic makes me crazy, I think shopping in stores is a hassle, etc. The Likert responses were used as input into a K-means factor analysis procedure yielding four factors (all with an eigenvalue greater than one). These factors were subsequently titled: Pro-technology, Anti-travel, Concerned about Internet security, Outgoing/gregarious; each respondent's factor score for each factor was used as a dependent variable.

The final binomial logistic regression models were constructed after performing log likelihood tests to check whether the statistical significance of the model deteriorates when insignificant variables are eliminated from the model. The final model, presented in Table 3.1.4, suggests that the likelihood that the respondent substitutes a physical banking or shopping trip with online activities increases with seven factors. We find younger, more educated individuals with access to the Internet at home are more likely to substitute. In addition, their attitudes tend to be pro-technology and a bit introverted. Interestingly enough, however, by *not* owning a cell phone, one is more likely to substitute for travel.

The Need for Physical Travel

From the above analysis we see considerable variation in the use of ICT and different activities and in some cases, even contradiction. The reasons for engaging in ICT-related activities vary not only by use, but sometimes within the same individuals. Any discussion of substitution would be incomplete without paying attention to other reasons or justifications for going about physical travel. That is, despite being aware of ICT options for an activity, why do some people still prefer traditional forms? Our third group of questions sheds light on this matter by asking the motivations for physical (in-store) shopping. That is, what motivates people to take a physical trip instead of eliminating trip with ICT? In this analysis, we are interested in uncovering the motivating reasons for not choosing to substitute a trip for a given good/service, even though they are aware of the ICT option.

For example, we specifically asked respondents to recall the last time they made a physical trip to the bank even though they were aware of the e-banking option and why they did so (for this question, respondents could circle all reasons that apply). The most cited answers categories are, the need to deposit or check or to do other things that need to be done in person (58%), the need to do other non-banking things on the same trip (21%), and security considerations (19%). The results show that for a healthy portion of the population, traditional banking still has features more advantageous than electronic banking—namely in terms of security (particularly people's perception about it) and the diversity of services that could only be completed in person. Also, a visit to the bank might only be one convenient stop on a multi-purpose trip. The culmination of these responses suggests that at least in the foreseeable future, electronic banking is not likely to be a full substitute for traditional banking in all functions and on a large scale. We also asked the main reason people choose to buy in-store instead of online. The most cited reasons were the need to touch or try the goods (44%), the need to get the goods immediately (28%), and concern about online payment security (13%). Consistent with overall retail trends, and similar to banking, physically going to the store continues to have insurmountable advantages over online shopping, at least for now.

3.1.4 Conclusions

This work uses the results of a household in three cities seeks to understand the degree to which household ICT use substitutes, modifies or even augments household non-work travel. Consistent with findings from research focusing on telecommuting (i.e., work travel), our results suggest that the availability of ICT reduces travel somewhat. However, its substantive significance is minor and there appear to be more important determinants driving the demand for travel. To the extent individuals are using ICT for banking or shopping, its use is still pale in comparison to physical travel. This analysis shows, for example, when individuals were asked about the most frequent trip they make, the vast majority of respondents were not willing to substitute. Furthermore, many do not even consider ICT a viable option for such a trip. Some people completely dislike the idea of ICT substitution when asked about the types of trips they are most and least willing to substitute.

While we see clear-cut tendencies for certain types of activities, such as work trips (pro-ICT substitution) and meeting friends and relatives (anti-ICT substitution), people are pretty divided on their degree of willingness to substitute certain other types of activities with the use of ICT. Some are most willing to substitute their grocery or non-grocery shopping or banking trips, others are least willing to substitute these same types of activities. After aiming to model the typical substituter, we find younger, more educated individuals with access to the Internet at home are more likely to substitute. In addition, their attitudes tend to be pro-technology and a bit introverted.

Some of the specific findings discussed above can perhaps best be described by the results from one of the general and culminating questions of the survey: how does technology impact your travel. The overwhelming response to this question centers on how ICT provides convenience, efficiency, and the abundance of information. These are the key motivations using it—not necessarily reducing travel. This study further supports the notion that while many people use ICT for select activities, researchers and policy officials would be misguided to expect that ICT substitution would come on a big scale. Quite simply, ICT substitution is unlikely to take over on a large scale unless technology advancement enables the online activities to acquire most of the desirable features of traditional activities.

3.1.5 Tables and Figures

Table 3.1.1 Most frequent trip from home (other than to work/school) and whether the respondent would choose to eliminate the trip if there is an ICT substitute

	Percentage (frequency)	Willing to substitute with ICT at home (percentage)		
		Yes	No	Not possible
Grocery store	54.2 (375)	17.5	76.0	6.5
Visiting friends or relatives	12.6 (87)	9.3	76.7	14.0
Entertainment	10.7 (74)	18.9	51.4	29.7
Other	9.2 (64)	6.3	44.4	49.2
Taking kids to school or daycare	6.6 (46)	13.0	13.0	74.0
Fitness Center	5.1 (35)	0	34.3	65.7
Bank	1.6 (11)	27.3	54.5	18.2
Total	100 (692)	14.5 *		
* Weighted average				

Table 3.1.2 Open-ended responses to most and least likely to substitute

	Most likely to substitute responses (percentage)	Least likely to substitute responses (percentage)	Example responses to each open-ended question
Grocery shopping	66 (23)	94 (25)	Grocery shopping, food shopping, grocery trip, household supplies and groceries
Banking and other financial transactions	59 (21)	35 (9)	Bank/banking, deposit, check investment, pay bill, pay bill online
Non-grocery shopping	45 (16)	48 (13)	Clothes shopping, shopping mall, purchasing, book purchase, pet store or liquor store, department store, car parts and tools, comparison shopping, entertainment shopping
Visiting friends and relatives	na	60 (16)	
Work related	43 (15)	na	Work, work related, going to work, work at home, part of my job
Other	72 (25)	138 (37)	*Most*: food delivery, government services, post office, movie rental, yard work or read, video conferencing, movies, mail for billing, gas station, library/school, ticket purchasing, phone a friend, reservation, check availability, travel to doctor, other travels/visiting relatives. entertainment, church, exercise, not sure and will try anything convenient *Least*: entertainment and recreation, exercise, and eating out

Table 3.1.3 Responses to several questions querying "What would you have done last time…?"

	Having the ICT form available…		
	Reduced travel	Saved travel while still requiring other travel	Had no direct Impact on travel
Think of last time you *purchased something online*. If you had not found it online, you would have…	Gone to the store (78.9%)		Not gone to the store (21.1%)
Think of last time you used *online banking*. If no such option available, you would have…	Visited the bank (27.6%)		Do not visit the bank at that time (72.4%)
Think of last time you made an *Internet financial transaction*. If no such option available, you would have…	Made physical trip (10.8%)		Used U.S. mail (76.4%) Not completed at all (11.8%) Completed by phone (1.0%)
Think of last time you used ATM. If you had not had the option, you would have…	Visit the bank at that time (49.8%)	Waited until next trip to the bank (19.5%) Cashed check other than at bank (28.4%)	Do the transaction by phone (2.3%)
(General speaking) If no ATM is available…	Visit the bank every time you would have used ATM (23.3%);	Visit the bank fewer times than with an ATM, but more often than currently (51.3%)	Continue to visit the bank with the same frequency as now (25.4%)

Table 3.1.4 Logistic regression for ICT substitution

Independent variables	Dependent variable: Substitution with ICT, yes/no	
Socio-demographic	B	Odds Ratio
Age in years	–0.027***	.973
High education	0.407*	1.502
IT ownership		
ATM card ownership	0.541*	1.718
Cellphone ownership	–0.438*	.645
Home Internet access	2.184***	8.882
Attitudinal variables		
Pro-technology	0.926***	2.524
Gregarious/social	–0.190*	.827
Constant	–1.191	.304
Number of cases	533	
Df	7	
$\chi 2$	191.988	
–2 log-likelihood	544.051	
Nagelkerke R Square	0.404	
*significant at 0.1 level **significant at 0.05 level ***significant at 0.01 level		

3.1.6 References

1. Fox, S., *A PEW Internet Project Data Memo*. 2002, The Pew Internet & American Life Project.
2. Chatzky, J., The Check is in the Mail. Not!, in Time. 2002.
3. Handy, S.L., A. DeGarmo, and K.J. Clifton, *Understanding the Growth in Non-work VMT*. 2002, Soutwest Region University Transportation Center. p. 74.
4. Texas Transportation Institute. Urban Mobility Report, Urban Mobility Conditions. http://mobility.tamu.edu/ums/appendix_a/exhibit_a-2.pdf January 20, 2003
5. Choo, S., P.L. Mohktarian, and Salamon, Impacts of Telecommuting on Vehicle-Miles Traveled: A Nationwide Time Series Analysis. 2001, California Energy Commission.
6. Salomon, I., *Telecommunications and Travel Relationships: A Review.* Transportation Research A, 1986. **20**(3): p. 223–238.
7. Salomon, I., *Telecommunications and Travel.* Journal of Transport Economics and Policy, 1985. **19**(3): p. 219–235.
8. Mokhtarian, P.L., A Typology of Relationships Between Telecommunications and Transportation. Transportation Research A, 1990. **24**(3): p. 231–242.
9. Marker, J.T. and K. Goulias. A Framework for the Analysis of Grocery Teleshopping. in 79th Annual Transportation Research Board. 1999.
10. Hjorthol, R.J., *The Relation Between Daily Travel and Use of the Home Computer.* Transportation Research A, 2002. **36**: p. 437–452.
11. Golob, T.F., Travelbehaviour.com: Activity Approaches to Modeling the Effects of Information Technology on Personal Travel Behaviour, in Travel Behavior Research, The Leading Edge, D. Hensher, Editor. 2001, Elsevier Science/ Pergamon: Kidlington, Oxford. p. 145–184.
12. Viswanathan, K. and K.G. Goulias, Travel Behavior Implications of Information and Communications Technology in Puget Sound Region. Transportation Research Record, 2001. **1752**: p. 157–165.
13. Pendyala, R.M., K.G. Goulias, and R. Kitamura, *Impact of Telecommuting on Spatial and Temporal Patterns of Household Travel.* Transportation, 1991. **18**: p. 383–409.
14. Wells, K., et al., Telecommuting Implications for Travel Behavior: Case Studies from Minnesota. Transportation Research Record, 2001. **1752**: p. 148–156.
15. Dijst, M., ICTs and accessibility: an action space perspective on the impact of new information and communication technologies, in Travel Demand and Organization in an Evolving World, V.H.a.A.R. M. Beuthe, Editor. 2003.
16. Handy, S.L., *Urban Form and Pedestrian choices: Study of Austin Neighborhoods.* Transportation Research Record, 1996. **1552**: p. 135–144.

17. Handy, S.L. and K.J. Clifton, *Local Shopping as Strategy for Reducing Automobile Use.* Transportation Research A, 2001: p. 317–346.

18. Krizek, K.J., Neighborhood Services, Trip Purpose, and Tour-Based Travel. Transportation, 2003: p. 387–410.

19. Greenwald, M., The Road Less Traveled: New Urbanist Inducements to Travel Mode Substitution for Nonwork Trips. Journal of Planning Education and Research, 2003. **23**(1): p. 39–57.

20. Mokhtarian, P.L., A Synthetic Approach to Estimating the Impacts of Telecommuting on Travel. Urban Studies, 1998. **35**(2): p. 215–241.

21. Mokhtarian, P.L. and I. Salomon, Modeling the Choice of Telecommuting: 3. Identifying the Choice Set and Estimating Binary Choice Models for Technology-based Alternatives. Environment and Planning A, 1996. **28**: p. 1877–1894.

22. Handy, S. and T. Yantis, *The Impacts of Telecommunications Technologies on Nonwork Travel Behavior*. 1997, Southwest Region University Transportation Center: Austin, TX. p. 51.

23. Scarborough Research. Percent of Internet Penetration by Market http://www.scarborough.com/scarb2002/press/pr_internetpene_data.htm, January 20, 2003

24. Dillman, D.A., Mail and telephone surveys: the total design method. 1978, New York: Wiley.

3.2 Mapping the Terrain of Information and Communications Technology (ICT) and Household Travel

Kevin J. Krizek
Assistant Professor, Urban and Regional Planning
University of Minnesota
301 19th Ave S.
Minneapolis, MN 55455
T: (612) 625-7318
F: (612) 625-3513
E-mail: kjkrizek@umn.edu

Andrew Johnson
Oregon Department of Transportation
E-mail: ajohnson@hhh.umn.edu

3.2.1 Introduction

New forms of information and communications technology (ICT) continue to emerge as primary forces influencing people's activities and travel. More than a century ago the telephone was instrumental in establishing new patterns of both communication and shopping. Rural residents, in particular, were not able to purchase products through catalogues (e.g., Sears) and have them delivered to their outlying locations. Later in the century the energy crisis of the late 1970's prompted increased attention on reducing travel and the potential role of emerging forms of computer assisted communication. Employers have always been attracted to telecommuting as a strategy to provide employee flexibility while still maintaining the array of services offered. Similar themes, albeit somewhat modified for current settings, persist as researchers and futurists examine the influence of rapidly changing forms of information technology on society. The nascent popularity of the internet has now placed e-commerce front and center on the minds of many. Burgeoning forms of computing communication (e.g., personal data assistants, cell phones) permeate even the simplest of transactions.

But what is the impact of these new forms of information and communication on household activity? What is their potential in eliminating, reducing or modifying travel? The effect of technology, and more specifically ICT, has been studied on a number of levels in previous work. For example, Edwin Parker [1] estimated that the introduction of computers and telecommunications will change modern economics as "information" cannot be treated as a commodity. Alvin Toffler [2] later surmised that ICT improvements will in fact make cities obsolete, while Lehman-Wilzig [3] projected that telecommunications might eliminate all travel. More recently, others have offered exposes relating technology to contemporary urban life or policy agendas (e.g., Smart Growth) [4, 5]. Interests in ICT remain varied. These include (a) geographical perspectives aiming to examine loosening the relationships between spatial patterns and development, (b) psychological perspectives aiming to uncover ways in which ICT changes communication patterns, or (c) economic analysis of increasing the efficiency of freight delivery.

A perspective receiving considerable attention, and one that is at the heart of this paper, relates to the influence of ICT on household travel. Motivation for interest in this specific topic remains varied. For example, employers are interested in enhancing employee flexibility, reducing overhead costs, improving operational efficiency and minimizing time spent in travel for their employees. Commercial businesses want to provide options for customers to not travel. The most publicized interests (environmental groups, government entities) are those who look to ICT to save energy, reduce congestion, and improve air quality. A premise of such interest rests on the likelihood that ICT will reduce the need for automobile travel (either in terms of vehicle miles traveled (VMT) or number of vehicle trips). Particular attention focuses on the potential of ICT to eliminate, reduce, or temporally shift peak-period commute travel. But interestingly enough, while deployment of ICT has skyrocketed in recent years, overall household travel in terms of distance and/or trips has increased as well. This suggests that our knowledge of relationships between ICT and travel, while growing, is still tentative at best.

The work that has been pursued to date relating ICT and household travel is varied. One only needs to refer to previous reviews [see, for example, 6, 7, 8] to gain a better understanding of the central questions and range of issues associated with this general line of inquiry. But these reviews are limited in at least two contexts. First, the bulk of previous literature focuses on one aspect of ICT: affecting the work commute (hereafter telecommuting). This aspect has been the focus of most activity and development. But everyday use of ICT has migrated well beyond the bounds of the work commute. As Golob claims, "we must do better than that [focusing on telecommuting] in order to keep our field relevant to planning and policy making" [7]. Secondly, technology, and the way in which people use it, is rapidly changing. It is necessary to keep abreast of recent developments, grapple with the emerging ranges of issues, and understand current work about these developments.

In response to at least these two calls, this paper uses a literature search as the primary means of inquiry to update our knowledge of relationships between ICT and travel. It uses existing studies to help frame and update a task that has been the target of a previous "research needs assessment" presented for TRB in 2000 [9]. In doing so, this work first maps the terrain of existing work to date related to ICT and household travel. It identifies the predominant nature of existing study—conceptual or empirical—and identifies voids in the existing knowledge base. Second, the paper sheds light on emerging phenomena to help conceptualize future research. This review therefore serves to update our knowledge about a body of knowledge undergoing quick changes. The aim is that it will also serve as a springboard or a blueprint for more detailed lines of inquiry related to ICT and household travel.

Before we begin it is important to frame the issues covered in this review. The present discussion is limited to understanding the behavior of households and individuals relative to different dimensions of ICT, broadly referred to telecommunications or information technology in some circles, and ICT in others. We use the term ICT because it encompasses a variety of terms describing a wide array of communication/technological enhancements. The most commonly cited examples include Internet improvements (i.e., broadband, e-mail, increased electronic accessibility to information and opportunities, more powerful personal computers), cell phones, ATMs, and teleconferencing. ICT is about using computers to work at home and cell phones to arrange a meeting place remotely. It is about ordering gifts on line and using internet availability to enhance home schooling. It is about ordering an on-demand movie and using video-conferencing. This paper, however, is not about reviewing some of ICT's other impacts. It is less about the use of Global Positioning System (GPS), Advanced Vehicle Transportation Systems (AVTS), the general nature of Intelligent Transportation Systems (ITS), and the impact of ICT on commercial vehicle operations or business to business commerce. The focus in many of these categories is more on modifying or managing travel than replacing it. Our focus here lies in how ICT may affect the travel patterns of households and/or individuals.

3.2.2 Previous Research

Defining a Framework for Discussion

Our review uses three dimensions to map recent literature of relationships between ICT and travel. The first dimension specifies whether the study is conceptual or empirical in nature, a relatively self-explanatory pursuit. The former describes theoretical arguments or key relationships, while the latter measures changes in travel as a result of ICT. The second and third dimensions are based on, respectively: (a) the purpose of the specific activity being studied and (b) the effect that ICT has on travel. These two dimensions are described below and the literature is then mapped and discussed using this framework.

Purpose of the ICT Activity

A critical aspect relates to nature of activities that are being pursued using ICT. For example, previous study has classified eight activities in which telecommunications can be substituted for travel (telecommuting, teleconferencing, teleshopping, telebanking, tele-entertainment, tele-education, telemedicine, and telejustice) [10]. This list presents a healthy list of ICT applications; but being more than a decade also, it is one that has undergone much change in recent times. Rather than attempt to identify each of the ever-growing list of ICT applications, we find value in using a relatively broader classification scheme—one that is also useful in helping us better understand how ICT relates to a household's overall travel. The typology we adopt is based on work first introduced by Reichman [11] who defined three major classes of travel-related activities. These three classes represent:

- subsistence activities, consisting of activities that generate income, including work, business services, and sometimes schooling (travel associated with this activity is most commonly commuting),
- maintenance activities, consisting of the purchase and consumption of convenience goods or personal services needed by the individual or household, and
- leisure or discretionary activities, comprising multiple voluntary activities performed on free time, not allocated to work or maintenance activities.

This typology of activities has been employed by Pas [12, 13] and Bhat and Koppelman [14] to classify daily travel activity behavior. Such a framework also represents the cornerstone of current activity-based transportation modeling efforts (e.g., TRANSIMS) currently underway [15–24]. This classification is directly applicable to this line of inquiry as it provides a parsimonious strategy to map the demand for central purposes of ICT activity (e.g., working versus banking versus shopping) and the degree to which existing study has tended to focus on one or more of these aspects.

The Effect of ICT on Travel

Almost 20 years ago, seminal work on ICT and travel by Salomon provided a foundation that helped outlined basic types of interaction [8, 25]. These types of interaction—substitution, modification, enhancement, and neutrality—have been employed in subsequent discussion and analysis [10, 26, 27]. We employ them here as a means to better under–stand the theories and evidence of past research.

- Substitution of travel refers to the elimination of trips—trips that are no longer required as a result of ICT improvements. This interaction has been the focus of most of the ICT attention, and the substitution of subsistence trips, by definition, is a phenomenon inherent in the term telecommuting [7]. While the substitution hypothesis is one that holds great hope by many, the scale to which this is happening is estimated to be quite smaller than originally anticipated [10].
- Modification refers to travel that is likely to be altered, primarily by a shift in the timing and routing of trips (spatial and/or temporal transformations). It may also refer to the manner in which trips are linked together (i.e., trip chaining) or even the mode of travel. The benefits of shifting even a small amount of peak hour travel to different times (through telecommuting) are mentioned as a key strategy for reducing levels of congestion. The literature agrees that ICT will modify travel behavior [28–30]; the outstanding issue is understanding in what manner (e.g., more travel or less).
- Generation refers to any generation of travel that would not have occurred but for the existence of ICT. Little is known about how ICT may generate additional traffic, primarily due to the difficulty in determining causality and a lack of time-series data before and after ICT enhancement and deployment.
- Neutrality refers those instances in which ICT has no foreseeable effect on household travel behavior.

Mapping the ICT Literature

Our mapping of recent work (Table 1) reviews over 50 studies to understand the nature and orientation of existing work. We classify each study first as being primarily conceptual or empirical in nature. The second mapping dimension considers the primary nature of the specific study—its relation to subsistence, maintenance or discretionary travel. The purpose of this table and subsequent discussion is to help identify where existing research is concentrated and where future research would most usefully be targeted.

Subsistence

Subsistence activities, when viewed relative to ICT, refer to individuals working at home or other remote locations, most often with telecommunications links to a central office. Interest in telecommuting stems from the often-heralded potential of relieving peak hour traffic congestion by reducing or modifying people's work trip (on average, considered to comprise one-third of all household travel). The widespread availability of telecommunications services, combined with the use of such services, has been the principle reason this aspect of ICT has been the target for the overwhelming majority of all study of ICT use [10, 30, 33–36, 39, 40, 42, 44, 45, 47]. Additional reasons for the focus on this line of work are two-fold. First, transportation planners and modelers have long been preoccupied with the work commute (often considered the lion's share of household travel). Secondly, work related activities are typically "cleaner" to analyze than maintenance and discretionary activities because they tend to be more stable and predictable.

Table 3.2.1. Mapping the literature of travel and ICT relating to three purposes of activity

	Primarily Conceptual	Primarily Empirical
Subsistence	•Handy, 1995 [31] •Mokhtarian, 1990 [10] •Mokhtarian, 1991 [32] •Mokhtarian, 1994 [33] •Mokhtarian, 1996 [34] •Mokhtarian, 1996 [35] •Mokhtarian, 1998 [36] •Handy, 1996 [37] •Lund, 1994 [38] •Nilles, 1976 [39] •Niles, 1994 [40] •Golob, 2001 [41] •Golob, 2001 [7] •Yen, 1994 [42] •Salomon, 1986 [8] •Bernardino, 1996 [43]	•Kitamura, 1990 [44] •Pendyala, 1991 [29] •Wells, 2001 [30] •Mokhtarian, 1991 [45] •Mokhtarian et al, 1995 [46] •Mokhtarian, 1996 [34] •Mokhtarian, 1996 [35] •Kraut, 1989 [47] •Shen, 1999 [48] * •Guiliano, 1998 [49] •Hamer, 1992 [50] •Hamer, 1991 [51] •Salomon, 1991 [52] •Nilles, 1988 [53] •Nilles, 1993 [54] •Yen, 1994 [42] •Mokhtarian, 1999 [55]
Maintenance	•Salomon, 1988 [56] •Salomon, 1992 [57] •Gould, 1998 [58] •Gould, 1997 [22] •Marker, 1999 [26] •Couclelis, 2002 [59] •Mokhtarian, 2001 [60] •Golob, 2001 [41] •Golob, 2001 [7] •Underhill, 1999 [61] •Batty, 1997 [62]	•Handy, 1997 [63] •Kraut, 1989 [47] •Cairns, 1996 [64] •Kilpala, 2000 [65] •Koppelman et al, 1991 [66]
Discretionary	•Salomon, 1988 [56] •Salomon, 1985 [25] •Salomon, 1992 [57] •Underhill, 1999 [61]	•Handy, 1997 [63] •Hjorthol, 2002 [27] •Kraut, 1989 [47] •Preece, 2001 [67] •Koppelman et al, 1991 [66]

Notes:

—Some papers describe both theoretical and empirical aspects and are listed in each column.

—Some papers relate to more than one category and may be listed in multiple rows.

* This research focuses less on the impact of direct work travel and more on the impact that telecommuting has on accessibility indices. It is grouped under subsistence because of the primary focus on employment opportunities.

Despite the telecommuting piece of the ICT puzzle being subject to the most mature models and conceptual frameworks, its orientation varies considerably. For example, previous work has focused on different dimensions of the issue, such as defining telecommuting [32], forecasting telecommuting, [36, 37] understanding individual and employer attitudes toward telecommuting (e.g., flexible working arrangements, increased proliferation of self-employment) [33-35], and land use implications [48, 49, 68]. Mokhtarian [36] has attempted to assemble the substantive findings to date under a unified framework by examining current knowledge in forecasting the demand for telecommuting and the resulting transportation impacts.

The culmination of this line of work has, however, yielded somewhat conflicting findings. Some have found that telecommuters reduce their number of trips and distance traveled on telecommute days [29, 30, 54], on non-telecommute days [29, 54], or on net travel [see 36 for other reviews, 46]. Other work has uncovered anecdotal and empirical evidence suggesting travel stimulation or generation [6, 8, 10, 40], sometimes only on non-telecommute days. Some go so far as to hint that levels of traffic jams and congestion may heighten because it provides an opportunity to catch up with telephone messages and email [5].

Maintenance

The effect of ICT on maintenance travel has enjoyed perhaps the longest history of study. Researchers examined the impact of early telephone order (and delivery) businesses on the ease of rural living and the vitality of town commercial centers. Central tenets of this line of inquiry continue today with the advent of the internet and e-commerce permeating the simplest of maintenance transactions. These types of activities account for over half of all household trips (and 40% of personal miles travel), but there is a lack of knowledge about how they are affected by ICT. Some of the existing work is oriented towards the freight and delivery aspects of e-commerce [41, 64]. The bulk of the maintenance travel-related literature, however, focuses on shopping (as opposed to maintenance-type activities such as banking or paying bills).

More than a decade ago Salomon and Koppelman provided a theoretical framework to understand home shopping [56] which was later supported with an empirical investigation [66]. While its preoccupation with telephone shopping limits the applicability of such work to a contemporary setting, this work continues to enlighten current study. An initial contribution they made was to distinguish between the act of shopping (the acquisition of information) and purchasing—a distinction of growing importance related to ICT activity. A second contribution is that they articulated five primary steps in selecting a piece of merchandise: (1) entry into the market, (2) choice among alternative shopping modes, (3) information gathering, (4) evaluation of information obtained, and (5) choice between purchase, continuing to shop or exiting the market. In a contemporary expose, Couclelis [59] provided a more detailed breakdown of similar tasks (see itemized list in later section) more t to ICT purchases.

Gould [22, 58] offers an overview of the transportation implications emerging from home shopping and online commerce. Her review focused on possible changes and demands placed on delivery services, the possibility of goods with no physical delivery, and the possible growth of new retail venues. She touched on a number of the difficulties in estimating the use of home shopping, including the need for understanding the different stages of shopping, attitudinal issues and opinions of consumers, as well as how ICT travel is related to other physical travel. Marker and Goulias [26] describe a framework for understanding and estimating the use of ICT and grocery shopping. In so doing, they outline salient issues as they relate to the likely effects on traffic (e.g., substitution, consolidating loads, trip chaining), forms of delivery, and methods in modeling such activities.

To our knowledge the most applicable empirical work on maintenance travel is provided by Handy and Yantis [63]. They examined in detail the potential substitutability of three different types of activities: movie watching, shopping (non-grocery), and banking. These activities were chosen to represent the spectrum of non-work activities from entertainment to personal business with movie watching at one end (discretionary), banking at the other (maintenance), and shopping somewhere in the middle. They conducted a household survey in three different cities to explore individual use of and choices about the each of the activities. Not surprisingly, their results suggest complicated relationships between in-home activities and those requiring physical travel. For the most part, they found that out-of-home versions of movie-watching, shopping, and banking offer qualities that are not currently duplicated by the in-home versions.

Previous research on maintenance travel, while limited, has served to successfully identify salient dimensions for consideration. One issue that keeps emerging is that different goods have different potential for successful e-commerce. Gould [22] highlighted that consumers will make decisions based on the four primary costs related to shopping: the cost of the item, the cost of the time to search for the item, the time to travel, and the expense of travel. There also appears to be wide variation in how different products are perceived to be more or less convenient to purchase online. According to one survey, home electronics, computer hardware and software are the most convenient items to shop for online; clothing and apparel are the least [69]. Our responsibility as transportation behaviorists is to understand how these subtle nuances relate to household travel. Furthermore, consumers prefer using their senses for many types of shopping, for example trying on a shirt, smelling perfume, sitting in a chair. In-store shopping offers immediate gratification and social interaction [61]. Though, there seems to be consensus that significant potential remains for many other goods, most notably books, music and movies (both downloaded and delivered), travel and theatre tickets, and computer software and hardware [61, 69].

Discretionary

As evident in Table 3.2.1, the link between the use of ICT and discretionary activity is a topic that has received the least amount study. While the boundaries between maintenance and discretionary activity blur (thus the cross listing of some of the studies in Table 1), at least two related thoughts come forward with respect to discretionary activity. The first refers to the inherently social and psychological needs of humans [25]. Much of the social interaction and human contact we require is invaluable and cannot be adequately provided for via electronic means. This need has important implications in how we spend our free time. The second consideration is that few discretionary activities (e.g., theaters, sporting events, pleasure shopping) can be replicated electronically. For example, renting a movie is not a substitute for going to the theatre because the two are usually not considered to be equivalent experiences. The latter usually provides a greater degree of social interaction and a higher quality product. This notion is reinforced in the Handy and Yantis study,[63] which looked at home movie rental, television and theatre visits. Primary reasons for going to the theatre include the size of the screen, newer movies, better sound, getting out of the house and going out with friends that for many reasons cannot be duplicated without leaving home. Both of these considerations suggest that in-home and electronic versions of goods and services are not likely viable substitutes for out-of-home and physical versions. For this reason, one should expect minimal travel savings (at best) with respect to reducing the roughly one-fourth of all discretionary trips currently completed.

A final issue considers the elasticity that may be embedded within discretionary activities. Salomon suggests that the increased leisure time resulting from increased efficiency from telecommuting results in increased leisure travel [25]. (In this respect, reducing the time allotted to subsistence and maintenance activity may free up time for discretionary activity.) Although discretionary activity itself may not be becoming electronic, as ICT makes subsistence and maintenance activity more efficient, people will have more discretionary time and subsequently travel more often.

3.2.3 Issues for Further Research

Having mapped the bulk of existing literature relating ICT and travel, the second part of this review combines that learned above with recent knowledge about how ICT is employed. The aim is to articulate future lines of inquiry and topics for further research. The difficulty of this task is compounded by the rapidly evolving nature of ICT, both the manner in which it is employed and its availability. For example, online spending reached all time highs in the second quarter of 2002, up 41% from the previous year to 17.5 billion. Some predictions go so far as to suggest that online consumer spending will increase to over 70 billion this year [70] –indeed a startling growth projection. On the contrary, some skeptics point to the still miniscule share of e-commerce (a mere 0.64% of all retail sales). There remains considerable uncertainty as to how this online spending will manifest itself, particularly as it relates to travel. Past failures of electronic home shopping have been documented [71]. One need look no further than the experience of Kozmo.com (the now defunct online delivery service for sundries who closed operations after only three years) to learn of the relative uncertainty associated with this line of service.

Online users amount to an estimated 98 and 140 million people in the U.S. (between 34 and 50% of the U.S. population) [72]. This number is rising daily as Internet and other computer use becomes more deeply integrated into everyday patterns of behavior. In a similar growth phenomenon, there were over 128 million cellular phone subscribers by the end of 2001, up from 109 million in 2000 [73]. But again, the mere availability of technology does not necessarily mean it will influence travel. For example, an estimated 66 million American internet users say they are online "just for fun," including sampling music, checking sports scores or electronic gaming [74]. This category probably includes the gathering of information (e.g., reading about news events). Such use may enhance one's quality of life (by providing additional information or entertainment) but may have little role in affecting travel patterns.

The bottom line is that the research community is realizing that ICT affects different travel in different ways. A breakdown that analyzes substitution, modification, generation, or neutrality for one travel purpose is just the tip of the iceberg. We should not be surprised to see changes in one type of travel affect another type of travel. Some elements may substitute in one respect but enhance another. And, examining instances of subsistence, maintenance or discretionary travel in isolation oversimplifies issue. Past studies certainly point such issues of complexity and may identify such confounding issues as topics for future research. However, there is little writing that is available that suggests how to do it, much less a road map for framework for understanding such relationships. This part is admittedly more on the latter than the former but much needed nonetheless.

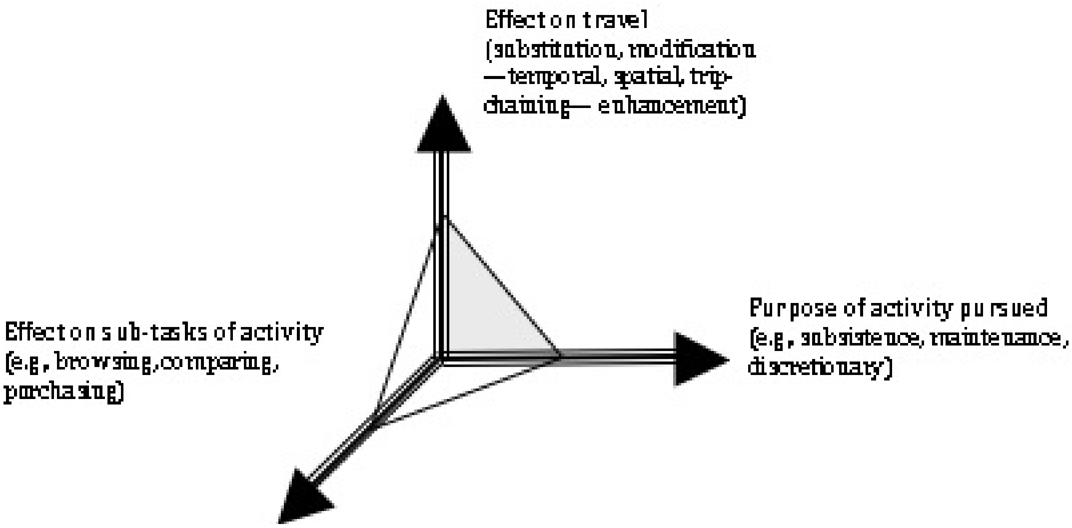


Figure 3.2.1 Dimensions for further research of ICT and household travel

To more clearly articulate the complexities that pervade this research we provide Figure 3.2.1. Our prism highlights three dimensions to represent how a comprehensive understanding of ICT and travel rests on knowledge from each leg. Absent of study of (or at least mention) one of these dimensions, a study falls short. The x-axis dimension—purpose of travel—is adequately defined in part 1 of this paper. But this dimension needs to be inextricably tied to the y-axis—the effect on travel. While described in previous literature (as well as above), the attention devoted to this topic is still relatively superficial. It is for this reason why further discussion is devoted below to the y-axis as well describing concepts introduced by the z-axis—the effect on sub-tasks of the activity.

Effect on Travel

A well recognized second dimension for inquiry relates to the effect ICT will have on travel. The principle effects—substitution, modification, generation, and neutrality—are described above. These four issues, however, comprise a limited range for this dimension. Travel is likely to be affected in ways broader and more detailed than just these four; or the travel may include more than one. It is important for us to understand the detailed behaviors that may be implied under each cover. Take for instance just one effect—modification. The three examples provided below describe some of the complexities.

Imagine the case of the teleworker, who is home during the workday and picks up groceries in the morning when the crowds are less (rather than on their typical route home from work). In this case, the substitution of the subsistence trip provides the flexibility pursue a modification (i.e., temporal shift) for a maintenance trip.

- Imagine the role that internet information would have in affecting the geographic distribution of where residents shop. Here the knowledge provided by ICT affects the geographical distribution of where the individual shopped.
- Imagine how the time freed from online grocery shopping would allow increased time to sip coffee at the café. In this example we see how the substitution of maintenance travel increased the time spent in discretionary activities.

- We quickly see the limits of any relatively simple substitution-modification-generation taxonomy. ICT may trigger changes mode of travel, the sequence of trips, the chaining of trips, and/or the time in which they are pursued. The issues are compounded when considered jointly. However, due attention is deserved.

Sub-tasks of the Activity

The third piece of the ICT-travel puzzle is one that the research community is becoming increasingly aware of. Referenced above, it relates to the different ways in which ICT enters into sub-tasks associated with any given activity. Couclelis [59] provides greater detail on the general tasks outlined in Salomon and Koppelman [56] that tend to be associated with ICT and shopping for any but the most trivial of goods. These tasks include: (a) becoming aware of need or want of a product, (b) gathering information about options, (c) searching and browsing, (d) seeking advise/expert help, (e) inspecting alternatives, (f) deciding on an item to be purchased, (g) deciding on a vendor, (h) purchasing a product, (i) tracking the status of an order, (j) getting an item to delivery point (e.g., home), (k) returning/exchanging an item, (l) seeking post-sales service. Only one of the tasks, purchasing the product, is commonly considered in analysis. It is, after all, the task in which the important transaction takes place and also the one that is instrumental to the livelihood of bricks and mortar establishments. While other tasks (for example, b, c, e, g, l) are recognized to be affected by ICT, a detailed and thorough understanding of nature of these relationships is at the heart of internet shopping and currently unavailable.

For some endeavors, the availability of online information has already changed certain customs. The process of contacting governmental agencies or individuals, searching for housing, or pricing major purchases [74] are a few examples where internet use has started to eliminate the need for travel. Take for instance the purchase of a new car, which typically involves stopping at several dealers to get the best price. It is now common practice to inspect the invoice price online prior to traveling to a single auto dealer to purchase the car. In this case "information gathering" has translated into travel substitution. But one must still be conscious of the fact that while ICT may help reduce travel associated with the acquisition of information it may not necessarily reduce travel associated with the purchase of an item. It would be interesting to identify these activities and thresholds that have been met to make such activities more mainstream.

When considering the role of ICT in completing subtasks a final matter stands out—differentiating between the population online buyers and online users. According to Jupiter Media of the 140 million online users, only 65 million (46%) are online buyers [72]. The most commonly cited reason for not purchasing online is the risk associated with relaying credit card information over the internet [69]. While this population is forecasted to increase in upcoming years, the division is an important issue for consideration because it implies a considerable population may be browsing for information.

Issues, strategies and hurdles. Much like any intervention that affects travel, the above text describes how issues relating ICT and travel are far from clear. The complexity demands continued methodological advancements and focused study. For example, many examples suggest tradeoffs exist between trip generation, mode split or other aspects of travel. The researcher is encouraged to bring methods, strategies, or techniques from other aspects of transportation research (or other fields for that matter) to tackle the problems. For example, few (if any) ICT studies closely examine these tradeoffs using econometric models designed for such purposes (e.g., simultaneous models). Frequently, such tradeoffs even fail to be mentioned. Alternatively, many of the tradeoffs may be best pursued using qualitative modes of inquiry, currently underused in transportation research. There remain seemingly endless modes of inquiry together with a burgeoning use of ICT

To help clarify examples research topics relating ICT and travel, we provide Figure 3.2.2. Each box represents an example activity for investigation; a relatively comprehensive study would most likely shed light on multiple boxes. The manner in which the activity relates to two of the dimensions is represented by its placement in the box. The lighter shading (towards the right) is used to represent the degree to which ICT may be stimulating additional travel, much to the chagrin of many transportation planners. In future pursuits, the transportation researcher is encouraged to wrestle with a few pieces of the puzzle while respecting the overall picture of the puzzle. This is indeed a difficult line to toe.

Of course, myriad difficulties will exist. And, before such difficulties can be overcome, they need to be identified. Below, we identify some of these hurdles, many of which have been introduced in previous study [25, 27, 58, 63]:

- Technologies that are widely promised for the future prove infeasible or are replaced by the next great technology before they have a chance to be adopted for mainstream use;
- The increasing familiarity of technology by children and other youth will likely have longer lasting impacts;
- There currently exists a dearth of information documenting all different purposes of travel (often just broken down by five different purposes) and modes (often failing to obtain reliable data for pedestrian trips);
- The need for uncovering and understanding the various ways in which individuals use different forms of technology (e.g., browsing versus purchasing);
- Separating subsistence from maintenance and/or discretionary travel—they are often times intricately related (for example, a decrease in subsistence travel will likely result in an increase in other travel [25]). Teasing out the effect of one is likely to tell on part of the picture;
- The interactions between different types of travel are likely to be different in various contexts. There are also methodological problems in measuring them [25];
- More extensive time use (and activity based travel) analysis is required to determine how time freed from using ICT might be used in other activities;
- The potential for cross-substitution between different types of activities (e.g., substitution of surfing the internet for watching television);
- There exists dramatic variation across households (access to educational and financial resources often determine the degree to which ICT influence household behavior).

Each box represents an example activity for investigation. The manner in which the activity relates to two of the dimensions is represented by its placement on the x and y axes (z axis not shown). Lighter shading (towards the right) represents the degree to which ICT may be stimulating additional travel.

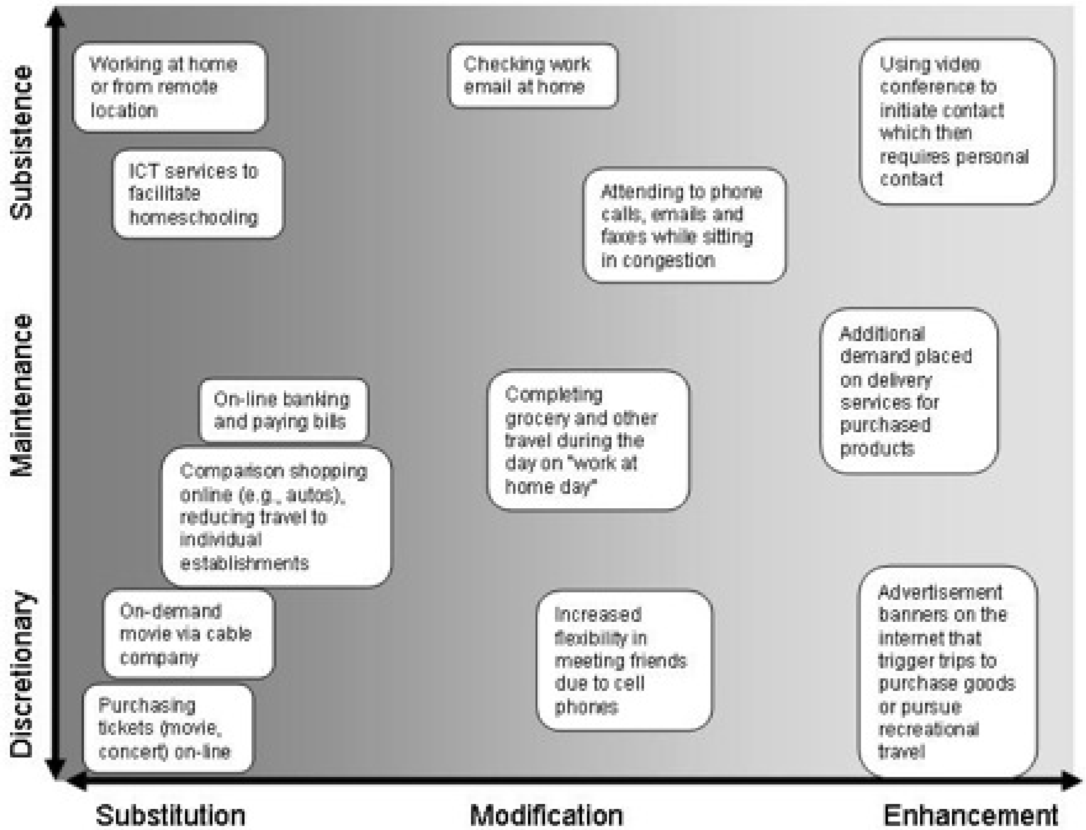


Figure 3.2.2 Examples of the complexity in which ICT affects travel

Implications for Practicing Planners and Researchers

One purpose of this work is to stimulate additional research by providing a state of the literature and concrete examples of how this work is needed. This paper goes above and beyond past work but suggesting examples of the many different ways that ICT affects travel.

Ways to expose the myriad of ways in which ICT affects household and personal travel can be better understood by pulling from other disciplines. For example, a stronger understanding of the spatial implications is especially important as the effect on efficient trip chaining and spatial-temporal components are central to understanding ICT and travel. A stronger understanding of the time and cost savings of ICT would perhaps increase the accuracy for econometric models. Finally, a deeper knowledge of the market segmentation of the different types of ICT uses would similarly enhance the ability to understand the rapid change.

Much of the ICT issue lies at the heart of land use and transportation planning. It has strong implications on crafting and revising the zoning codes in order to create more efficient land use patterns by using ICT to help off set peak period related problems. It is important to encourage development that will blend well together, i.e. blend ICT friendly and not-so-friendly developments together, blend substituters, enhancers together.

3.2.4 Summary

This paper serves two purposes to aid our understanding of how ICT and travel behavior relate.

It first provides an up-to-date review of over fifty studies on the subject by mapping and discussing existing work across three dimensions. These three dimensions are the degree to which the study is (1) primarily conceptual or empirical in nature, (2) primarily addresses subsistence, maintenance, or discretionary activities, and (3) able to comment on the hypotheses of substitution, modification, generation or neutrality. Second, it identifies and describes a framework, together with further issues to consider in future investigations relating ICT and travel. Examples are identified and there is a call for broader use of methodological advancements. The aim is that it will also serve as a springboard or a blueprint for more detailed lines of inquiry related to ICT and household travel.

The study of ICT and travel is complex and challenging—though one that provides several avenues for investigation. This task is one that has been best described by Mokhtarian [9]:

> Against the slower moving demographic and organizational changes that may be occurring naturally, the highly volatile technological environment and attendant consumer response speed up some of these natural changes and bring about entirely new ones…Thus improving our understanding of these processes is a moving target, and no sooner do we think we have made progress in answering one question then we realize that the question has been rendered obsolete or unimportant in environmental shifts.

The work spotlighted in this review relates to only one piece of the ICT puzzle—that of household travel. This is, however, an important piece and one receiving more attention. As congestion, auto-reliant travel, and centrifugal forces of development continue to rise, researchers, modelers, and policy officials are likely to demand more detailed understanding of ICT's role in tackling such phenomena. Several avenues and courses of study are needed to help uncover such phenomena; this work helps in identifying the itinerary and providing a preliminary roadmap for doing so.

3.2.5 References

1. Parker, E.B., Social Implications of Computer/Telecoms Systems. *Telecommunications Policy,* 1976. 1: p. 3–21.
2. Toffler, A., *The Third Wave.* 1980, New York: Bantam Books.
3. Lehman–Wilzig, S., Will Cities Become Obsolete? *Telecommunications Policy,* 1981. 5: p. 326–328.
4. Mitchell, W.J., Electronic Cottages, *Wired Neighborhoods and Smart Cities, in Smart Growth: Form and Consequences,*T.S. Szold and A. Carbonel, Editors. 2002, Lincoln Instuitute of Land Policy: Cambridge, MA. p. 66–81.
5. Moss, M.L. and A. Townsend, How Telecommunications Systems Are Transforming Urban Spaces, in Cities in the Telecommunication Age: The Fracturing of Geographies, J.O. Wheeler, Y. Aoyama, and B. Warf, Editors. 2000, Routledge.
6. Mokhtarian, P.L. and I. Salomon. Emerging Travel Patterns: Do Telecommunications Make a Difference. in Recent Developments.*Travel Behavior Research.* 1997. Austin, TX: Oxford.
7. Golob, T.F., Travelbehaviour.com: Activity Approaches to Modeling the Effects of Information Technology on Personal Travel Behaviour, in Travel Behavior Research, *The Leading Edge,* D. Hensher, Editor. 2001, Elsevier Science/ Pergamon: Kidlington, Oxford. p. 145–184.
8. Salomon, I., Telecommunications and Travel Relationships: A Review. *Transportation Research A,* 1986. 20(3): p. 223–238.
9. Mokhtarian, P.L., Telecommunications and Travel.*Transportation and the New Millenium.* 2000, Transportation Research Board: Washington DC.
10. Mokhtarian, P.L., A Typology of Relationships Between Telecommunications and Transportation. *Transportation Research A,* 1990. 24(3): p. 231–242.
11. Reichman, S., Travel adjustments and life styles: a behavioral approach.*Behavioral Travel-Demand Models,* P.R. Stopher and A.H. Meyburg, Editors. 1976, Lexington Books: Lexington, MA. p. 143–152.
12. Pas, E., Analytically Derived Classifications of Daily Travel-Activity Behavior: Description, Evaluation, and Interpretation. *Transportation Research Record,* 1982. 879: p. 9–15.
13. Pas, E., The effect of selected sociodemographic characteristics on daily travel-activity behavior. *Environment and Planning A,* 1984. 16: p. 571–581.
14. Bhat, C. and F. Koppelman, A Conceptual Framework of Individual Activity Program Generation. *Transportation Research,*1993. 27A(6): p. 433–446.
15. Bowman, J.L., et al. Demonstration of an Activity-Based Model System for Portland.8th World Conference on Transport Research. 1998. Antwerp, Belgium.

16. Ben-Akiva, M. and J.L. Bowman, Integration of an Activity-based Model System and a Residential Location Model. *Urban Studies,* 1998. 35(7): p. 1131–1153.

17. Bhat, C.R., J.P. Carini, and R. Misra, Modeling the Generation and Organization of Household Activity Stops. *Transportation Research Record,* 1999. 1676: p. 153–161.

18. Misra, R. and C.R. Bhat, Activity Travel Patterns of Non-Workers in the San Francisco Bay Area: Exploratory Analysis. *Transportation Research Record,* 2000. 1718: p. 43–51.

19. Jonnalagadda, N., et al., Development Of Microsimulation Activity-Based Model For San Francisco: Destination And Mode Choice Models. *Transportation Research Record,*2001. 1777: p. 25–35.

20. Kitamura, R., An Evaluation of Activity-based Travel Analysis.*Transportation,* 1988. 15: p. 9–34.

21. Bradley Research and Consulting, et al., A System of Activity-Based Models for Portland. 1998.

22. Gould, J. and T. Golob, F., Shopping without travel or travel without shopping? An investigation of electronic home shopping. *Transport Reviews,* 1997. 17: p. 355–376.

23. Ma, J. and K.G. Goulias, A dynamic analysis of person and household activity and travel patterns using data from the first two waves in the Puget Sound Transportation Panel. *Transportation,* 1997. 24(3): p. 309–331.

24. Bhat, C.R. and R. Misra, Discretionary Activity Time Allocation of Individuals Between In -Home and Out-of-Home and Weekdays and Weekends. *Transportation,* 1999. 26: p. 193–209.

25. Salomon, I., Telecommunications and Travel. *Journal of Transport Economics and Policy,* 1985. 19(3): p. 219–235.

26. Marker, J.T. and K. Goulias. A Framework for the Analysis of Grocery Teleshopping. *79th Annual Transportation Research Board.* 1999.

27. Hjorthol, R.J., The Relation Between Daily Travel and Use of the Home Computer. *Transportation Research A,* 2002. 36: p. 437–452.

28. Viswanathan, K. and K.G. Goulias, Travel Behavior Implications of Information and Communications Technology in Puget Sound Region. *Transportation Research Record,* 2001. 1752: p. 157–165.

29. Pendyala, R.M., K.G. Goulias, and R. Kitamura, Impact of Telecommuting on Spatial and Temporal Patterns of Household Travel. *Transportation,* 1991. 18: p. 383–409.

30. Wells, K., et al., Telecommuting Implications for Travel Behavior: Case Studies from Minnesota. *Transportation Research Record,* 2001. 1752: p. 148–156.

31. Handy, S.L. and P.L. Mokhtarian, Planning For Telecommuting. *Journal of the American Planning Association,* 1995. 61(1): p. 99–111.

32. Mokhtarian, P.L., Defining Telecommuting. *Transportation Research Record,* 1991. 1305: p. 273–281.

33. Mokhtarian, P.L. and I. Salomon, Modeling the Choice of Telecommuting: *Setting the Context. Environment and Planning A,* 1994. 26: p. 749–766.

34. Mokhtarian, P.L. and I. Salomon, Modeling the Choice of Telecommuting: 2. A Case of the Preferred Impossible Alternative. *Environment and Planning A,* 1996. 28: p. 1859–1876.

35. Mokhtarian, P.L. and I. Salomon, Modeling the Choice of Telecommuting: 3. Identifying the Choice Set and Estimating Binary Choice Models for Technology-based Alternatives. *Environment and Planning A,* 1996. 28: p. 1877–1894.

36. Mokhtarian, P.L., A Synthetic Approach to Estimating the Impacts of Telecommuting on Travel. *Urban Studies,* 1998. 35(2): p. 215–241.

37. Handy, S. and P. Mokhtarian, Forecasting Telecommuting: An Exploration of Methodologies and Research Needs. *Transportation,* 1996. 23: p. 163–190.

38. Lund, J.R. and P.L. Mokhtarian, Telecommuting and Residential Location: Theory and Implications for Commute Travel in Monocentric Metropolis. *Transportation Research Record,* 1994. 1463: p. 10–14.

39. Nilles, J.M., The Telecommunications-Transportation Tradeoff : Development of Policy. 1976, New York City: Wiley. 196.

40. Niles, J., Beyond Telecommuting: A New Paradigm For The Effect of Telecommunications on Travel. 1994, Department of Energy: Washington DC.

41. Golob, T.F. and A.C. Regan, Impacts of Information Technology on Personal Travel and Commercial Vehicle Operations: Research Challenges and Opportunities. *Transportation Research C,* 2001. 9: p. 87–121.

42. Yen, J. R., H.S. Mahmassani, and R. Herman, Employer Attitudes and Stated Preferences Toward Telecommuting: An Exploratory Analysis. *Transportation Research Record,* 1994. 1463: p. 15–.

43. Bernardino, A. and M. Ben-Akiva, Modeling the Process of Adoption of Telecommuting: Comprehensive Framework. *Transportation Research Record,* 1996. 1552: p. 161–170.

44. Kitamura, R., et al., Telecommuting as a Transportation Planning Measure: Initial Results of California Pilot Project. *Transportation Research Record,* 1990. 1285: p. 98–104.

45. Mokhtarian, P.L., Telecommuting and Travel: State of the Practice, State of the Art. *Transportation,*1991. 18: p. 319–342.

46. Mokhtarian, P.L., S.L. Handy, and I. Salomon, Methodological Issues in the Estimation of Travel, Energy, and Air Quality Impacts of Telecommuting. *Transportation Research A,* 1995. 29(4): p. 283–302.

47. Kraut, R.E., Telecommuting: The Trade-offs of Home Work. *Journal of Communication,* 1989. 39(3): p. 19–47.

48. Shen, Q., Transportation, Telecommunications and the Changing Geography or Opportunity. *Urban Geography,* 1999. 20: p. 334–355.

49. Guiliano, G., Information Technology, Work Patterns and Intra-metropolitan Location: A Case Study. *Urban Studies,* 1998. 35(7): p. 1077–1095.

50. Hamer, R., et al., Teleworking in the Netherlands: An Evaluation of Changes in Travel Behaviour – Further Results. *Transportation Research Record,* 1992. 1357: p. 82–89.

51. Hamer, R., E. Kroes, and H.v. Oosterroom, Teleworking in the Netherlands: An Evaluation of Changes in Travel Behaviour. *Transportation,*1991. 18: p. 365–382.

52. Salomon, I., H.N. Schneider, and J. Schofer, Is Telecommuting Cheaper Than Travel? An Examination of Interaction Costs in a Business Setting. *Transportation,*1991. 18: p. 291–318.

53. Nilles, J.M., Traffic Reduction by Telecommuting: A Status Review and Selected Bibliography. *Transportation Research,* 1988. 22A: p. 301–317.

54. Nilles, J., City of Los Angeles Telecommuting Project Final Report, . 1993, JALA International Inc.: Los Angeles. p. 89.

55. Mokhtarian, P.L. and R. Meenakshisundaram, Beyond Tele-substitution: Disaggregate Longitudinal Structural Equations Modeling of Communication Impacts. *Transportation Research C,* 1999. 7(1): p. 330–352.

56. Salomon, I. and F. Koppelman, A Framework For Studying Teleshopping Versus Store Shopping. *Transportation Research A,* 1988. 22: p. 247–255.

57. Salomon, I. and F.S. Koppelman, Teleshopping or going shopping? An Information Acquisition Perspective. *Behaviour and Information Technology,* 1992. 11(4): p. 189–198.

58. Gould, J., Driven to Shop? Role of Transportation in Future Home Shopping. *Transportation Research Record,* 1998. 1617: p. 149–156.

59. Couclelis, H., Pizza over the Internet: E-commerce, the Fragmentation of Activity, and the Tyranny of the Region. *Entrepreneurship and Regional Development,* forthcoming.

60. Mokhtarian, P.L., A. Nagurney, and J. Dong, Teleshopping versus Shopping: A Multicriteria Network Equilibrium Framework. *Mathematical and Computer Modelling,* 2001. 34: p. 783–798.

61. Underhill, P., *Why We Buy: The Science of Shopping.* 1999, New York: Simon and Schuster. 255.

62. Batty, M., The Retail Revolution. *Environment and Planning B,* 1997. 29: p. 1–2.

63. Handy, S. and T. Yantis, *The Impacts of Telecommunications Technologies on Nonwork Travel Behavior,*1997, Southwest Region University Transportation Center: Austin, TX. p. 51.

64. Cairns, S., Delivering Alternatives: Successes and Failures of Home Delivery Services for Food Shopping. *Transport Policy,* 1996. 3(4): p. 155–176.

65. Kilpala, H.K., P.N. Seneviratne, and S.M. Pekkarinen. *Electronic Grocery Shopping and Its impact on Transportation and Logistics and Special Reference to Finland,* 79th Annual

Meeting of the Transportation Research Board. 2000. Washington DC: Transportation Research Board.

66. Koppelman, F., I. Salomon, and K. Proussaloglou, Teleshopping or Store Shopping? A Choice Model for Forecasting the Use of New Telecommunications-Based Services. *Environment and Planning B,* 1991. 18(4): p. 473–489.

67. Preece, J., Sociability and Usability in Online Communities: Determining and Measuring Success. *Behaviour and Information Technology,* 2001. 20(5): p. 347–356.

68. Nilles, J.M., Telecommuting and Urban Sprawl: Mitigator or Inciter? *Transportation,*1991. 18: p. 411–432.

69. Bhatnagar, A., S. Misra, and H.R. Rao, On Risk, Convenience, and Internet Shopping Behavior. *Communications of the ACM,* 2000. 43(11).

70. Cox, B., E-commerce–Color it Green, in CyberAtlas. 2002.

71. Talarzyk, W. and R. Widing, Direct Marketing and Online Consumer Information Services: Implications and Challenges. *Journal of Direct Marketing,* 1994. 8(4): p. 6–17.

72. ePaynews, Compiled Statistics, 2002, http://www.epaynews.com/statistics/purchases.html#24.

73. Pastore, M., Wireless Voice Encroaching on Wireline Customers,*Cyber Atlas.* 2002.

74. Horrigan, J.B. and L. Rainie, Getting Serious Online, 2002, Pew Internet and American Life Project: Washington D.C. p. 22.

3.3 Spatial Attributes and Patterns of Use in Household-Related Information and Communications Technology Activity

Kevin J. Krizek
Assistant Professor, Urban and Regional Planning
Director, Active Communities Transportation (ACT) Research Group
University of Minnesota
301 19th Ave S.
Minneapolis, MN 55455
T: (612) 625-7318 F: (612) 625-3513
E-mail: kjkrizek@umn.edu

Yi Li
Humphrey Institute of Public Affairs,
University of Minnesota
301 19th Avenue S.
Minneapolis, MN 55455

Susan L. Handy
Department of Environmental Science and Policy,
University of California–Davis
2132 Wickson Hall
Davis, CA 95616

Transportation Research Record:
Journal of the Transportation Research Board, No. 1926, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 252–259.

3.3.1 Overview

Trends of Household-Related ICT Activities

This research focuses on household use of three categories of ICT-based activity: e-commerce, electronic banking, and other financial transactions. The three categories likely have distinct patterns of use and are affected by spatial attributes in different ways. Therefore, the analysis focuses on each and begins by briefly describing general trends with respect to each category.

For the purposes of this paper, e-commerce refers to the purchase of commercial goods or services over the Internet; its rise comes as little surprise to most. The Boston Consulting Group [2] estimates that online retailing in North America alone totaled $27 billion in 1999 and $45 billion in 2000. Although this dollar figure still comprises a relatively insignificant margin of total retail sales (1.7%), its amount has increased by more than 67% since 1999. Whereas most e-commerce forecasters anticipate that the shear growth in proportion of sales will likely subside, the availability of such services is likely to be of increasing impact. One need only examine trends in personal use. Among U.S. Internet users alone, those who had ever bought online has grown from 48% (about 41 million Americans) in 2000 to 61% (about 67 million Americans) in 2002—an increase of 63% [1]. It remains unclear, however, who such shoppers are in terms of demographics and computer experience; furthermore, it is unclear whether such e-commerce shopping is substituting or replacing in-store shopping.

The second category is electronic banking, defined as using automatic teller machines (ATMs), a telephone, or the Internet to engage in transactions with one's principle bank (e.g., cashing checks, deposits, transferring money, and making withdrawals). In some cases, such a banking transaction still requires a physical trip to the ATM. For this reason and because it was desirable to ensure comparability of this activity with a similar survey instrument from a previous application, it has its own category. The rise in electronic banking is interesting if for no other reason than its reported surge in growth in recent years. Looking at trends dating back to 2000, for example, the Pew Center survey estimates that online banking increased by 23 million Americans between 2000 and 2002, amounting to about 20% of the population. These proportions are generally consistent with other surveys where Gartner reported that 17% of Americans used online banking services by the end of 2002 [3] and Celent Communications reported online banking penetration at 22% in 2002 [4]. Overall, slightly less than a quarter of the population uses some form of electronic banking.

The last category distinguishes online financial transactions from electronic banking. The former is defined as completing any array of financial activities over the Internet such as buying and selling stocks, trading commodities, paying mortgages, and paying other bills. Few national surveys separate this category from online banking. It is done here primarily for two reasons. The first is that traditional financial transactions (e.g., paying utility bills or credit card statements) typically did not involve a physical trip; they were most often completed by using U.S. mail. Second, these types of transactions represent a burgeoning array of activities that previously have not been empirically researched. Available evidence indicates that this specialized ICT activity is also rising, although certainly not as rapidly as the other categories of ICT activities. According to Jupiter Research, 18.9 million U.S. households viewed and paid bills online during 2003, increasing from 12.2 million in 2002. In addition, they reported that, among those that bank online, 50% also paid their bills online in 2003 [5].

Review of Past Literature and Expected Relationships

How might spatial attributes affect the use of these categories of ICT activities? Prevailing theory suggests two contradictory hypotheses about this relationship. One theory is guided by what is commonly referred to as a variation of the innovation diffusion hypothesis [6]. In a metropolitan setting, this theory postulates that urban centers, by virtue of being at the center of innovation (intellectually or technically), exhibit a higher tendency to engage in practices or thoughts enabled by the innovation [7, 8]. Using this logic, urbanites are more likely to adopt such ICT behavior because such services (e.g., online banking) first became available in urban settings. The currently “spotty” nature of where high-speed Internet service is available throughout the United States is consistent with this theory. On the flip side, others argue that, in most metropolitan areas, different areas [i.e., central business district (CBD), first- and second-ring suburbs, and exurbs] all more or less enjoy similar levels of access to technology, thereby eliminating any noticeable differences due to geography.

A counter theory is driven by what has loosely been referred to as the “efficiency hypothesis” [9], which suggests that individuals aim to save time by combining trips [10]—or, in this case, save the additional cost associated with physical travel [11] (e.g., online shopping and banking). Such costs may derive from a variety of attributes, be it lower levels of accessibility (therefore longer travel distances) or having to fight traffic congestion in areas more urban in nature (therefore longer travel times). Virtual travel, it is argued, would more likely be used by residents in those situations to avert such costs.

Most previous work on ICT use is unable to shed direct light on these hypotheses because it has been devoid of a spatial context. Existing studies most often treat residents in metropolitan areas as a single population where spatial attributes are theorized to play little role in affecting the propensity of e-commerce or e-banking type activities. Studies that have addressed spatial attributes have often used flawed measures of those attributes; they are often measured at extremely aggregate scales (e.g., city versus rural) or are not precisely defined. Furthermore, where previous research has specifically aimed to tackle such questions, researchers have found contradictory results. For example, Farag [11] empirically tested the preceding two hypotheses with data from the Netherlands (1996–2001) and found that people living in urbanized areas were more likely to engage in e-shopping. On the other hand, however, it was found that people who do not have easy access to shops (defined as having fewer shops for nondaily goods that can be reached by car from the person's home in a certain time span) also tended to shop more online.

This paper aims to understand the degree to which residents use the Internet for e-commerce, e-banking, and other financial transactions and how such use is affected by spatial attributes such as retail and bank accessibility and traffic congestion. Controlling for a host of confounding variables, the analysis measures spatial attributes at two different scales: intermetropolitan and intrametropolitan. The intermetropolitan analysis examines aggregate use patterns among three cities, using general levels of congestion as a spatial attribute. A metropolitan area with high traffic congestion is expected to exhibit higher rates of ICT use because of the increased burden of traveling to activities. The intrametropolitan analysis uses individual households as the unit of analysis and examines relationships to nearby retail, the CBD, or banks. For these relationships, theory suggests that spatial attributes may both increase and decrease ICT use; as a result, it is difficult to posit specific hypotheses about the directions of the relationships. The hypothesized relationships examined in this paper together with the spatial units of analysis are presented in Table 3.3.1.

3.3.2 Methodology

To examine the preceding dimensions, a mail-out and mail-back survey of households was administered in the spring of 2003. The survey sample was drawn randomly from the population in three metropolitan statistical areas: Seattle, Washington; Kansas City, Kansas and Missouri; and Pittsburgh, Pennsylvania. Several criteria guided the selection of these cities. First, a sample of residents representative of relatively large U.S. urban areas from the Midwest and the East and West coasts was desired. Next, the basic hypotheses about spatial attributes and information technology (IT) availability were instrumental in stratifying the sample by balancing between two other criteria: level of traffic congestion for large cities and rates of Internet penetration. Using Institute's measures of traffic congestion in U.S. urbanized areas in 1990–1999 [12], candidate cities representing both high and low levels of traffic congestion were identified. Next, this criterion was balanced by controlling for IT availability in each city, as measured by rates of Internet penetration by Scarborough Research (2002). The final selection included Seattle, representing a high-congestion/ high-technology city (Seattle rates second on the congestion list and among the highest of cities in IT availability). Kansas City represented a low-congestion/ high-technology city, and Pittsburgh was selected to represent a low-congestion/low-technology city.

Table 3.3.1 Hypothesized relationships between ICT use and spatial attributes

		Outcome variable		
Scale of Analysis	Spatial Measure	Online Shopping	Electronic Banking	Other Online Financial Transactions
Metro Area	Higher level of traffic congestion	+	+	+
Household	Urban versus suburban residence	+/−	+/−	+/−
Household	Distance to CBD	+/−	+/−	+/−
Household	Proximity of retail activity (home)	+/−	+/−	+/−
Household	Proximity of ATM machine (work or home)	+/−	+/−	+/−

After pilot testing, the survey was mailed to 800 households in each of the three cities in May 2003. The households were selected at random from an address database maintained by Survey Sampling International. The primary source of their database listings is telephone directories, which were supplemented with other sources, such as state ID cards. A cover letter explained the purpose of the survey and invited a household member 18 years or older to complete it. Following the Dillman [13) method, reminder postcards were mailed to the 2,400 households 1 week later and a third mailing (with the complete survey) was sent out 3 weeks later to each household that had not yet replied. The overall response rate is 31%, including 32% from Seattle, 30% from Kansas City, and 30% from Pittsburgh. After comparing sociodemographics with 2000 census data for each metropolitan area, it was found that the characteristics of the survey respondents differ only slightly from the overall populations in the three cities. Most notable is the fact that the survey respondents are substantially older than the general population in the three cities (60% to 70% are more than 50 years old, in contrast to 30% to 40% in the overall population; 36% to 37% are more than 60 years old, in contrast to 18% to 28% in the overall population). Other differences showed the sample to have a slightly higher percentage of males, to be more educated, and to have slightly higher household incomes than the general population. On the one hand, higher income levels would suggest a bias in terms of use of ICT-related services. Given that older populations are probably less likely to be using the Internet for services, this would likely sway the bias in the opposite direction.

3.3.3 Results

The results of the analysis are now considered by describing relationships between ICT use and spatial attributes of urban environments. The focus is on three different ICT behaviors: online purchases, e-banking, and online financial transactions. In each case, the focus was primarily on the behavior in a binary manner, capturing whether the survey respondent performed the activity, generally within the past six months. This research is therefore less about measuring the frequency of the behavior and more about detecting its use. The following discussion is divided as follows. First, rates are described in which the three dimensions of ICT are used to glean a better understanding of general patterns. This is followed by an overview of how such patterns differ by level of IT availability and levels of congestion; this part represents the intrametropolitan analysis. Then, moving to a fuller discussion of the specific effects of various accessibilities, statistics are presented revealing penetration rates, how such rates vary by geographic area, and other descriptive information. To investigate the impact of spatial variables on e-shopping while controlling for a variety of other factors, three multivariate models were used. Doing so controls for mediating factors and therefore serves to better isolate the impact of specific measures.

General Patterns and Relative Independence of Activities

The first look at the data aims to identify general patterns of ICT use in the aggregate. This analysis provides background information about general levels of ICT among the sample and highlights the need for separate analyses for each category of ICT activity.

Examining rates of use across the three activities reveals the following. As indicated in Table 3.3.2, shopping online has the highest penetration (around 47%), with e-bankers making up a considerably smaller share of the sample (about 31%). It is also interesting to look at the degree to which the use of the different categories of ICT overlaps within individual households; 20% of the sample have completed all three activities least once; 21% reported ever completing two activities, 16% reported ever completing only one of the activities. Finally, 43% of respondents reported never having engaged in any of the three activities.

These results suggest overlap among the three activities; a fair number of individuals appear partial to one of the activities. Whereas online shopping is the most popular activity among the three, this does not necessarily suggest that all e-shoppers are also e-bankers or that they make financial transactions online (Table 3.3.3). Among those who e-bank, 23% have never shopped online and 21% have never made a financial transaction online. And among those who ever made other financial transactions online, about 20% have never shopped online. This suggests that there does not appear to be a sequential relationship between e-shopping and the other two activities. If there were a sequential relationship, then whoever had completed the other two activities would have e-shopped as well. It is therefore difficult to say that people shop online first, and then some of them go on to use online banking or do other financial transactions online. The picture of household ICT use appears to get murky very quickly. Most important for the subsequent analysis, it suggests that household use of ICT activity is complex and it is helpful to break it down to analyze the different types of ICT activities separately.

Table 3.3.2 ICT Activities in three metro areas

	Ever shopped online (% yes)	Ever use e-banking (% yes)*	Ever made other online financial transactions (% yes)	Average (% yes)	Ever completed at least two of the activities (% yes)
Overall	46.6%	31.2%	41.2%	39.7%	41.7%
Seattle	53.1%	44.2%	50.2%	49.2%	51.0%
Kansas City	50.0%	27.7%	41.5%	39.7%	42.7%
Pittsburgh	36.4%	20.7%	31.2%	29.4%	30.8%
Pearson Chi-square *p*-value	.000***	.000***	.000***		

* =p<0.1 ** = p < 0.05 *** = p < 0.01

*Among those that that *never* banked online, about one third does not have the option of online banking, while the rest choose not to use online banking even though the option is offered by their banks.

Intermetropolitan Differences

The next look at the data examines spatial dimensions by seeing how ICT use differs across each of the three metropolitan areas. Table 3.3.4 indicates that, consistent with the initial hypotheses, ICT use differs across the three metropolitan areas and appears to be correlated with ICT availability or congestion. Seattle (high technology, high congestion) appears to have the highest ICT use and Pittsburgh (low technology, low congestion) has the lowest ICT use. Seattle clearly stands out in all three activity categories (e-shopping, e-banking, other financial) with more than half the Seattle sample having experience in at least two of the three activities. In contrast, Pittsburgh is the least penetrated by ICT activities. Only 31% of the population have ever done at least two of the three activities online. Specifically, only 36% have ever shopped online, only 21% have ever banked online, and only 31% have ever made other financial transactions online. This is consistent with what would be expected given the relatively low level of IT and low level of congestion in the Pittsburgh area. Expectedly, Kansas City falls between the two with an average of 43% of the sample ever having completed at least two of the activities. About 50% have ever e-shopped, 28% have ever e-banked, and 42% have ever engaged in other financial transactions online.

Table 3.3.3 also indicates that the percentage who ever e-banked is remarkably lower than the percentages who ever completed the other two activities in both Kansas City and Pittsburgh, whereas the difference is less apparent in Seattle. In terms of e-banking, Kansas City is more like Pittsburgh than Seattle. This suggests that e-banking might be more strongly related to levels of congestion. It is plausible that, because of security concerns, people with high technology available (like those in Kansas City) are still reluctant to bank online if their traffic remains tolerable. When congestion is an issue, it appears that respondents are more likely to take advantage of e-banking. Although these endings certainly support the notion that technology and congestion appear to play a role in ICT activities, it is important not to overstate their impact. For example, it remains inconclusive whether such a coarse measure of IT penetration (i.e., provided by Scarborough Research) is accurate for more specific ICT-based services. That is, is a high-technology rating (as measured by rates of Internet penetration) synonymous with rates of ATM card ownership, cell phone ownership, home computer ownership, and home Internet access?

In response to this question, Table 3.3.4 shows each of the cities and possession of five different technologies as measured in the survey: ATM card ownership, a home computer, home Internet access, fast home Internet access, and cell phone. Seattle and Kansas City are grouped because both metropolitan areas were deemed technology friendly in the sampling stratification. As expected, those dimensions on which sample selection was based appear to show up as statistically significant. Seattle and Kansas City live up to their high-technology classification for computer- and Internet-related activities. However, the distinction is not as noticeable as one would expect. Across the remaining three services, statistically significant differences are not observed across high-technology versus low-technology cities.

Intrametropolitan Differences

Having described general relationships between each of the cities, now the relationships between ICT use and spatial attributes are described as they vary within metropolitan areas. This is done to better understand matters related to accessibility of different services (e.g., shops, ATMs). Measures representing two different levels of geography and attributes are used, each of which is derived by the hypotheses of interest and described later. The first measure aims to represent relatively coarse distinctions between areas with high levels of retail activity and shops versus those with lower levels. This dimension is captured through two measures. The first is a self-reported measure, asking respondents how far they live from the CBD. The second measure is an objective one gleaned from the census zip code business patterns data, quantifying the number of establishments in the retail trade category (Industry Code 44xxxx in the North American Industry Classification System) for the respondent's zip code. The second group primarily applies to the electronic banking category and aims to get at the degree to which there are ATMs within close proximity that are either available or used by the respondent. Both are binary variables and self-reported from the survey, which asked if there is an ATM close to the person's residence or workplace and if there is a bank close to the person's residence or workplace (i.e., within walking distance).

In Table 3.3.5, some preliminary relationships emerge. First, there appears to be a higher propensity for people living farther from the CBD to engage in ICT-based activities. Self-reported city residents are less likely to shop online than suburbanites. In addition, almost 60% of those living more than 10 mi from the CBD have shopped online (versus 40% for those living less than 5 mi from the CBD) and almost half of those beyond 10 mi have ever made financial transactions online. Second, the relationship between the availability of nearby retail and online shopping is not statistically significant, suggesting that the proximity of retail (within one's zip code) does not appear to affect one's propensity for online shopping. Examining the effects of having a neighborhood ATM produces expected results: an ATM near one's home or workplace is more likely to be used. Those respondents who do not have banks within walking distance of home or work are slightly more likely to have ever banked online (31.8%) as opposed to 30.1% among those who have banks within walking distance. However, the distinction is not statisti–cally significant.

Table 3.3.3 Relationships between ICT activities and users

Of those that …		*Ever* shopped online	*Ever* banked online	*Ever* made other financial transactions online
They have…	***Never*** shopped online	na	23%	20%
	Never banked online	48%	na	40%
	Never made other financial transactions online	30%	21%	na

Table 3.3.4 Availability of IT-related services

			Do you have…				
Sampling stratification			**…an ATM card**	**…a home computer**	**…home Internet access**	**…fast home Internet access +**	**…a cell phone**
Low Tech	Pitt	Yes	175 (74%)	163 (69%)	157 (67%)	67 (46%)	156 (65%)
Hi Tech	Seattle	Yes	198 (78%)	220 (87%)	208 (82%)	78 (40%)	167 (66%)
	KC	Yes	161 (70%)	184 (79%)	171 (73%)	72 (44%)	154 (66%)
Total within High Tech		Yes	359 (74%)	404 (83%)	379 (77%)	150 (42%)	321 (66%)
Pearson Chi-square *p*-value			.888	.000***	.002***	.426	.893

*=p<0.1 ** = p < 0.05 *** = p < 0.01

^ Seattle and Kansas City are groups because both metropolitan areas were deemed technology friendly in the sampling stratification.

+ Fast home Internet column provides percentage based on only 75%(S), 69%(KC) and 60%(P) of survey sample who have home Internet in each city. The question on Internet type branches after “Does your household have access to Internet?” Those answering no to the Internet access question are not led to answering the question regarding the type of internet they are using.

Table 3.3.5 ICT activity and intrametropolitan spatial attributes

Spatial attribute	Percent of respondents within each category of the spatial attribute	Ever shopped online (% yes)	Ever electronic banked (% yes)*	Ever made other financial transactions online (% yes)
City/Suburban	City (38.2%)	40.8%	31.4%	40.1%
	Suburban (61.8%)	50.7%	31.3%	42%
	Pearson chi-square p-value	*0.01***	*0.964*	*0.604*
Distance to CBD	<=5 miles (42.8%)	40.3%	28.4%	36.1%
	5–10 miles (25%)	44.8%	31.5%	39.3%
	>10 miles (32.2%)	58.4%	36.1%	50.9%
	Pearson chi-square p-value	*0.000****	*0.164*	*0.002****
Availability of Nearby Retail	>100 retail (32.8%)	49.6%		
	50–100 retail (33.3%)	47.1%		
	<=50 retail (33.9%)	44.3%		
	Pearson chi-square p-value	*0.506*		
ATM Availability	Within walk distance (53.5%)		37%	
	Not within walk distance (46.5%)		24.5%	
	Pearson chi-square p-value		*0.000****	
Bank Availability	Within walk distance (36%)		30.1%	
	Not within walk distance (64%)		31.8%	
	Pearson chi-square p-value		*0.631*	

*=p<0.1   ** = p < 0.05   *** = p < 0.01

The relationships thus far have been examined primarily through simple correlations of ICT use and various explanatory variables. Although some statistically significant relationships were uncovered, the bivariate analysis precludes an understanding of such phenomena in a multivariate context. Introducing control variables in the analysis and modeling such behavior at the individual level can better discern the explanatory power of the spatial attributes.

In the multivariate analysis, five categories of variables thought to contribute to higher levels of ICT were used. The first category is sociodemographic variables including age in years, sex (male = 1), education (low, high), household income (low, high), two-plus vehicles, household structure (households without teenage children, households with teenage children), and whether the respondent is employed (yes = 1). Measures in the second category, reflecting the availability of IT-related infrastructure, were gathered at the metropolitan level and at the household level, each described previously. The metropolitan level measures are rates of Internet penetration where Seattle and Kansas City were deemed high and Pittsburgh was deemed low. The household measures were gleaned from the survey and captured ATM card ownership (yes = 1), cell phone ownership (yes = 1), home computer ownership (yes = 1), and home Internet access (yes = 1). The third category comprises behavioral variables, including frequency of nongrocery shopping (more than five times in the last 30 days = frequent, otherwise = infrequent) and frequency of bank visits (one to three times per month or more = frequent, otherwise = infrequent). The fourth category captures attitudinal variables that were gleaned from a dozen or so attitudinal questions answered by the respondents; examples include "I feel comfortable using computers," "technology makes my life easier," "traffic makes me crazy," and "I think shopping in stores is a hassle." The responses on a 5-point Likert scale were used as input into a K-means factor analysis procedure yielding four factors (all with an eigenvalue greater than 1). These factors were subsequently titled protechnology, antitravel, concerned about Internet security, and outgoing/ gregarious; each respondent's score for each factor was used as a dependent variable.

The last category of measures—spatial attributes—comprises the heart of this investigation. Levels of congestion (at the metropolitan wide level) were captured using dummies of the city variable (Seattle = 1). Other measures are self-reported from the survey and include whether the respondent's home is in the city versus suburb, miles from the CBD (less than 5 mi, 5 to 10 mi, more than 10 mi), whether there was an ATM close to work or home (yes = 1, no = 0), whether there was a bank close to work or home (yes = 1, no = 0), and number of retail businesses in the immediate zip code area of the home (gathered from U.S. employment statistics). The dependent variable in three of the models was dichotomized as having ever completed the activity. However, the frequency of online purchases was examined as frequent (defined as answering two times per month and more often) versus infrequent (once per month, less than monthly, and never). The final binomial logistic regression models were constructed by performing log likelihood tests to determine whether the statistical significance of the model deteriorates when insignificant variables are eliminated from the model. The results for statistically significant variables are presented in Table 3.3.6.

The models produce interesting findings about attributes of individuals and metropolitan areas. The first observation is that only one measure of spatial attributes appears to be statistically significant—being a high-congestion city—and this was for only one of the models, the model for ever banked electronically. Thus, whereas significant spatial relationships emerged in bivariate contexts, other factors in fact explain the variation; simply put, spatial attributes, by themselves, do not appear to play a significant role in affecting levels of ICT use. In the e-shopping model, the dummy variable for low-technology city (Pittsburgh) is significant and in the expected direction. This finding indicates that households in cities with relatively low rates of Internet penetration are less likely to have ever shopped online. Similar relationships are observed in the logistic regression model for other financial transactions online. The distinction between the electronic banking model and the other two models corresponds to the descriptive findings: congestion influences only levels of electronic banking.

Table 3.3.6 Multivariate analysis of ICT activity

Independent variables	Dependent variables (logistic regression)							
	Frequency of online purchases		Ever bought online		Ever electronic banked		Ever completed Internet financial transaction other than banking	
	B	Odds Ratio	B	Odds Ratio	B	Odds Ratio	B	Odds Ratio
Sociodemographic variables								
Age in years	–.025***	.975	–.035***	.966	–.033***	.968	–.028***	.972
High education			.445**	1.560	.492**	1.636	.486**	1.626
High income	.584**	1.793						
Two vehicles or more					.572**	1.772		
Teenage children in household							573**	1.774
IT availability at household level								
ATM card ownership			.560*	1.751	.698*	2.010	.751**	2.119
Cellphone ownership							.453*	1.573
Home Internet access	2.495**	12.122	2.377***	10.773	1.786***	5.966	1.217***	3.377
Low technology			–.698***	.498			–.731***	.481
Behavioral variables								
Visit bank frequently					–.458**	.633		
Frequently shop for non-grocery items	.471*	1.602						
Attitudinal variables								
Pro-technology	.909***	2.482	.901***	2.462	.762***	2.143	.819***	2.268
Anti-travel	.244*	1.276						
Concerned about Internet security	–.467***	.627					–.285**	.752
Gregarious/social			–.266**	.766	–.192*	.825		
Spatial Attributes								
High congestion					.984***	2.675		
Constant	–3.439	.032	–.640	.527	–2.086	.124	–.981	.375
Number of cases	538		538		538		535	
Df	7		7		9		9	
χ^2	128.934		238.523		199.179		213.152	
–2 log-likelihood	424.137		505.153		509.793		526.479	
Nagelkerke R^2	.332		0.478		0.423		.439	

*significant at 0.1 level **significant at 0.05 level ***significant at 0.01 level

The results also indicate that IT availability on the household level positively affects an individual's propensity for adopting ICT activities. Understandably, individuals with ATM cards and home Internet access are more likely to engage in ICT activities. Not only is this relationship statistically robust across all four activities but the odds ratio calculations show this factor as having considerable impact. Cell phone ownership is also significant in the other financial transaction model. However, the direction of causality cannot be determined from these models. Although it is logical that having access to IT increases engagement in ICT activities, it is also possible that engaging in ICT activities leads to an increase in access to IT. For example, households that enjoy shopping online may be more likely to opt for high-speed Internet access.

Control variables were also significant in the regression results. Similar sociodemographic and attitudinal variables affect all four activities, despite some minor differences. Being young and highly educated relates positively to ICT use in all three activities. In addition, owning two or more vehicles contributes positively to adopting online banking behavior, whereas having teenage children in the household contributes positively to engaging in online financial transactions. Furthermore, people who are protechnology are more likely to adopt ICT activities, a trend that holds true for all three categories of ICT activities. Besides, people who are gregarious and who like to socialize with others are less likely to have ever used e-banking or e-shopping. The results also suggest that concern about Internet security is a significant factor that deters some people from having ever completed other financial transactions online.

3.3.4 Conclusions

The hypotheses that spatial attributes would influence the use of ICT for selected activities was not entirely borne out by this analysis. The intermetropolitan comparisons indicate that the high-technology and high-congestion city was associated with greater penetration of at least some ICT-based activities. Within metropolitan areas, residents who live in the suburbs or live further from the CBD were more likely to engage in e-shopping than residents who live in the city or close to the CBD. However, the lack of significance of these attributes in the multivariate analyses suggested the bivariate relationships are explained by other factors: high-technology status, high-congestion status, suburban location, and location relative to the CBD are all associated with other factors that more directly explain ICT use. The multivariate analyses suggest that IT availability at the household level, sociodemographic characteristics, and attitudinal factors are more important than spatial variables in explaining ICT use. If spatial attributes are not a factor in deciding to use ICT, as these results suggest, then a desire to reduce travel may not be a primary motivation for using ICT. Other studies have shown that ICT use does not always substitute for travel ([14, 15], p. 51). Nevertheless, ICT-based activities offer the option of reducing travel, and an improved understanding of the factors that do and do not influence the choice to use ICT is important for transportation planning.

3.3.5 Acknowledgements

The Telecommunications and Travel Behavior Committee sponsored publication of this paper.

3.3.6 References

1. Pew Internet and American Life Project. *America's Online Pursuits: The Changing Picture of Who Is Online and What They Do.* www. pewinternet.org/reports/toc.asp?Report=106 or 207.21.232.103/pdfs/ PIP_Online_Pursuits_Final.PDF. June 4, 2004.
2. Abdelmessih, N., M. Silverstein, and P. Stanger. Winning the Online Consumer: *The Challenge of Raised Expectations.* www.bcg.com/publications/files/Winning_Online_Consumer_Jun%2001_ofa.pdf. Sept. 14, 2001.
3. Nua Internet Surveys, *Online Banking Goes Mainstream in U.S.* www.nua.ie/surveys/index.cgi?f=VS&art_id=905358738&rel=true. Jan. 24, 2004.
4. *ClickZ Network, E-Banking Blooming, But Needs New Apps, Best Practices.* cyberatlas.internet.com/markets/finance/article/0,,5961_ 1588591,00.html. Jan. 24, 2004.
5. E-Banking, *Online Bill Paying Growth Ahead.* www.internetnews.com/stats/article.php/3112511. March 17, 2004.
6. Rogers, E. M. *Diffusion of Innovations.* Free Press, New York, 1983.
7. Hagerstrand, T. *Innovation Diffusion As a Spatial Process.* University of Chicago Press, Chicago, 1967.
8. Wolcott, P., L. Press, W. McHenry, S. Goodman, and W. Foster. A Framework for Assessing the Global Diffusion of the Internet. *Journal of the Association for Information Systems,* Vol. 2, No. 6, 2001, pp. 1–52.
9. Anderson, W. P., L. Chatterjee, and T. R. Lakshmanan. E-Commerce, Transportation and Economic Geography. Presented at STELLA Focus Group 1, Kick-off Meeting, Sienna, Italy, 2002.
10. Adler, T., and M. Ben-Akiva. A Theoretical and Empirical Model of Trip Chaining Behavior. *Transportation Research Part B,* Vol. 13, 1979, pp. 243–257.
11. Gertz, C. The Relationship Between Transportation, Urban Form and ICT: What We Know, What We Don't Know. Presented at STELLA/ STAR Focus Group 2 Meeting, National Science Foundation, Arlington, Va., 2002.
12. *Urban Mobility Report, Urban Mobility Conditions.* Texas Transportation Institute, College Station, mobility.tamu.edu/ums/appendix_a/ exhibit_a-2.pdf. Jan. 20, 2003.
13. Dillman, D. A. Mail and Telephone Surveys: *The Total Design Method.* Wiley, New York, 1978.
14. Krizek, K. J., Y. Li, and S. L. Handy. *ICT As a Substitute for Non-Work Travel: A Direct Examination.* University of Minnesota, Active Communities Transportation (ACT) Research Group, Washington, D.C., 2004.
15. Handy, S., and T. Yantis. *The Impacts of Telecommunications Technologies on Nonwork Travel Behavior.* Southwest Region University Transportation Center, Austin, Tex., 1997.

3.4 Trends in Out-of-Home and At-Home Activities

Evidence from Repeat Cross-Sectional Surveys

Ryan Wilson
Active Communities/Transportation (ACT) Research Group
University of Minnesota
301 19th Avenue South
Minneapolis, MN 55455
E-mail: wilso903@umn.edu

Kevin J. Krizek
Active Communities/Transportation (ACT) Research Group
University of Colorado
Environmental Design Building
Campus Box 314
Boulder, CO, 80309-0314

Susan L. Handy
Department of Environmental Science and Policy
University of California
2132 Wickson Hall
Davis, CA 95616

3.4.1 Introduction

Expectations remain high that information communications technology (ICT) will reduce physical travel and particularly the negative effects of automobile travel (e.g., traffic congestion and air pollution). For almost a quarter-century, government officials, travel behavior specialists, technology forecasters, and others have been monitoring how ICT might affect travel. Almost 15 years ago, a U.S. Department of Transportation report discussed the long-range implications of ICT on travel patterns, route choice, and congestion. The report identified telecommuting as a potential strategy for managing transportation demand, and other activities (including teleshopping and telebanking) as potential substitutes for auto trips [1].

Pinning down the effect of technology on travel is challenging. Almost all work to date has been cross-sectional in nature, examining behaviors at one point in time. It is helpful to consider changes over time, including trends in the use of at-home ICT activities and their relationship to travel. Prospective panel studies are ideal for examining these trends but are expensive, and so far, none have been undertaken. In the absence of panel data, cross-sectional studies can be compared, although different means of data collection as well as the rapid evolution of technology limit this approach. As a result, at-home ICT trends have been difficult to track as new questions and possibilities surface continuously. The approach offered in this paper uses repeat cross-sectional surveys—one conducted in 1995 and the other in 2003—to examine changes in ICT use and related travel behaviors.

The aim of this study is to determine the degree to which at-home ICT use and out-of-home store travel changed from 1995 to 2003 for the selected purposes of shopping and banking. First, relevant literature on the potential impacts of ICT on travel and the nature of common ICT activities provide some overall context for this exercise, and then national trends for computer ownership, Internet access, online shopping, and online banking are detailed. Next, an overview is provided of the methodological approach and statistical techniques used to examine changes in the frequency of out-of-home and at-home shopping and banking between 1995 and 2003. Finally, results are summarized, and the implications are discussed with respect to the relationship between ICT and travel behavior.

3.4.2 Previous Research

For many years, transportation agencies have expected that ICT would contribute to an effective strategy for transportation demand management. The ability of ICT to manage short-term travel demand could complement strategies and programs for mitigating congestion in the long term, 10 to 15 years [2]. ICT use could provoke numerous outcomes on personal travel; two profound possibilities are that ICT will change the type of activity engaged in (at home and out of home) as well as alter the frequency, timing, and destination of travel patterns [3]. Not without consequence, ICT activities could generate trips if they provided greater flexibility in whether, when, where, and how travel were to occur [4].

The ability of ICT to aid in transportation demand management may be limited in several ways. First, ICT trip substitution might be successful only in niche markets, such as high-income and time-constrained groups [5]. The type of product may limit trip substitution as well. Books, magazines, CDs, DVDs, software, and admission tickets are more popular online purchases than clothing, pharmaceuticals, and food [6]. Although the content of a CD does not vary across multiple vendors, consumers may be less trusting or unwilling to purchase food or prescription drugs online, preferring to inspect them in person. The percentage of multipurpose trips could also constrain the market. One study found that nearly half of shopping trips are multipurpose, implying that substituting one portion of the trip would not eliminate the trip altogether [7].

Second, the level of substitution across activities varies. One study, examining 1995 survey data, found that the degree of substitution depends on the activity and attitude. Certain out-of-home activities offer desirable qualities that at-home alternatives cannot replace, such as the social aspect of viewing a movie at a cinema with friends (3). Research in Germany found that among computer users, 74% made fewer shopping trips than nonusers, whereas the remaining 26% made more trips (8). Another study found that home shopping increases the frequency of store shopping and trip chaining (9). Still another found that online shopping reduced short automobile trips by a mere 0.31% [10). Other research found that commuters would not be willing to substitute their most frequent home-based trip [11). Collectively, these results raise significant doubts about the overall impact of ICT on travel demand.

Third, individual factors that explain trip substitution with ICT are equally unclear. Results of a study of three U.S. cities indicate that Internet availability and attitudinal factors might be as important as, or more important than, spatial attributes (e.g., trip length) in deciding whether to engage in at-home ICT activities [12). Reducing travel does not appear to be a primary motivation for using ICT. Results of a review of 65 online shopping studies indicate that the factors affecting online shopping behavior generally are Internet perceptions, vendor information, and user characteristics (e.g., sociodemographics and Internet experience). A desire to avoid trips or reduce travel is an uncommon predictor [13).

Prior research thus raises both expectations and doubts about the impact of ICT on travel. Given the rapid expansion in ICT and its use and the potential benefits of even small reductions in automobile travel, further exploration of these questions is warranted.

3.4.3 Trends from National Data

Technological improvements have increased the rates of computer ownership and Internet access worldwide. Although these changes have been most prominently realized by medium- and high-income households (14, 15), low-income households likely realize such upgrades or are able to purchase new products as technology improves and costs decrease. In this section, a snapshot of recent national computer ownership—showing Internet access, online shopping, and online banking trends—is provided for the time period analyzed. This information has two primary purposes: to document how national trends illustrate the emergence of new forms of ICT, and to provide a benchmark for determining how closely the samples reflect the general population.

Computer Ownership and Internet Access

Household computer ownership and Internet access increased steadily from 1995 to 2003. Figure 3.4.1 displays national trend data from the U.S. Census [14) and the Pew Internet & American Life Project, a nonprofit research center examining the impact of Internet on Americans [16). The gap between computer ownership and Internet access shrunk, suggesting that the Internet became more valuable and possibly more affordable to computer owners. Data from the Center for Digital Future mirrors that of Pew Internet, estimating that 65.1% of households had Internet access in 2003 compared with 46.9% in 2000 [17). The average 2003 user spent 12.5 hours per week online, up from 9.4 in 2000. These results suggest greater at-home opportunities for ICT use.

Online Shopping

Online shopping has grown at rates similar to Internet access since 2000. The Economics and Statistics Administration reports that 40.1% of Americans shopped online in 2000, compared with 52.1% in 2003 [15). Pew Internet estimates a change from 41% to 61% during the same period (16). However, the percentage of adults who purchased online remained near 45% between 2000 and 2003 [17).

The portion of total retail sales attributed to online shopping increased from 0.9% in 2000 to 2.3% in 2003 [18). As consumers realize technological innovations and become more trusting of online purchases, market share is likely to increase. At current rates, online shopping gains are likely to outpace total retail shopping over the next several years [19). The frequency of online shopping probably is increasing with sales, although these trends fail to differentiate this possibility from higher dollar purchases or simply more shopping overall.

Online Banking

Online banking trends tend to mirror those of online shopping. The Economics and Statistics Administration [15] estimates an increase of 17.4% to 27.8% in online banking between 2000 and 2003, whereas Pew Internet reports a change of 18% to 34% [16]. Another poll found that 38% of Americans used online banking in 2005 and that, despite security considerations, 81% believe it is improving overall service and will remain a banking alternative [20]. Using an ATM is an out-of-home, ICT-based banking activity that allows users to deposit checks and withdraw money from a primary bank. The convenience of ATM cards is greater than ever; a 2002 study estimates that 60% of U.S. ATMs were not at banks [21].

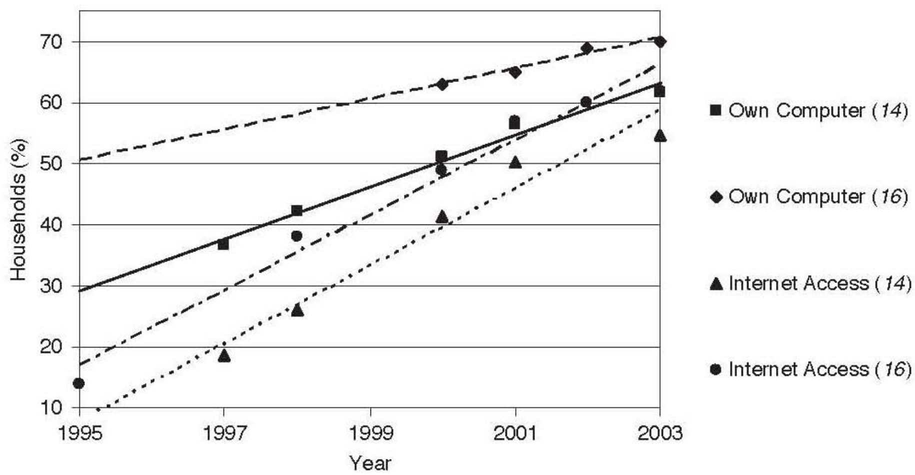


Figure 3.4.1 Percentage of American households with computers and Internet access [14, 16)]

Hypothesis

Using the data from two repeat cross-sectional surveys (conducted in 1995 and 2003), two broad research questions were posed. First, to what degree has at-home ICT use and out-of-home store travel changed from 1995 to 2003? Multiple analysis of covariance (MANCOVA) is used to determine the interaction effects, controlling for sample differences. Because congestion has worsened and technology has become more accessible over this period, higher at-home and lower out-of-home frequencies are hypothesized. Second, how much do at-home alternatives substitute for out-of-home trips, and to what degree has this amount changed between 1995 and 2003? Given travel demand expectations, greater substitution is hypothesized.

3.4.4 Data and Method

Survey Instruments

The 1995 and 2003 surveys queried ICT use; access to technology; and attitudes toward substitution, technology, and congestion. Both were administered by mail and sent randomly to 1,000 individuals in three U.S. metropolitan areas. The 1995 survey focused on San Jose, California; Austin, Texas; and Oklahoma City, Oklahoma. The 2003 survey used Seattle, Washington; Kansas City, Missouri; and Pittsburgh, Pennsylvania. The cities were chosen to test (and control) for ICT influence in areas that represent varying degrees of technology and congestion: low technology and low congestion in Oklahoma City and Pittsburgh, high technology and low congestion in Austin and Kansas City, and high technology and high congestion in San Jose and Seattle. The overall response rate was 16% in 1995 and 31% in 2003. Additional details regarding the individual instruments and select analysis results are available elsewhere [3, 11].

The only difference between the survey instruments was in the form of at-home shopping and banking: respondents reported the frequency of catalog shopping and phone banking in 1995 and Internet shopping and banking in 2000. Each represents the common at-home activities that individuals might have used as substitutes for out-of-home physical store trips at the time.

Statistical Analyses

The first research question was posed to determine whether the year of the survey (1995 or 2003) had an effect on the frequency of out-of-home and at-home shopping and banking. MANCOVA emerges as an appropriate multivariate technique to do this because it examines the difference in means of two or more dependent variables across categorical independent variables. The dependent variables should share a theoretical relationship because each measures a separate influence of the independent variables, but their outcomes should be discrete. MANCOVA differs from multiple analysis of variance (MANOVA) in that it considers the interaction effects of continuous and interval-level covariates, which act as controls for the independent variables [22]. It is more appropriate in this application because both types of independent variables exist.

Changes in the rates of at-home and out-of-home shopping and banking are defined as the dependent variables and survey year as the independent variable of interest. The multivariate F-statistic Hotelling's Trace) tests whether the independent variable survey year and each covariate has an effect on at-home and out-of-home activities. MANCOVA also generates univariate F-statistics to describe the interaction between each category of survey year (1995 and 2003) and both dependent variables.

The nine covariates control for differences in sample characteristics, attitudes, and city type. The characteristics of the sample populations (i.e., age, household income, household size, and number of household vehicles) are coded as continuous variables. City type is coded as an interval-level variable, assuming the increasing likelihood to use at-home activities in the following order: low technology–low congestion, high technology/low congestion, and high technology/high congestion. Four attitudinal questions related to technology and congestion are included. A Likert scale measures the extent to which each respondent agrees or disagrees with each statement. The scale ranges from 1 (strongly disagree) to 5 (strongly agree).

To determine whether differences between the 1995 and 2003 survey respondents are statistically significant and to answer the second broad research question, three significance tests for independent samples are used: the independent sample t-test (to compare means of two independent continuous variables), the chi-squared test (to compare means of nominal-level tabular data), and the Mann–Whitney U-test (to compare mean ranks of ordinal or higher data).

Table 3.4.1 Respondent characteristics, 1995 and 2003, Part 1

	Survey year				Comparing survey year	
	1995		2003		N = 475	N = 738
Characteristic	%	n	%	n	Statistic	
Gender					0.022[a]	
Female	42.3	195	43.0	311		
Male	57.7	264	57.0	414		
Education					−3.072[b,c]	
Less than high school	2.0	10	3.3	25		
High school	27.0	123	34.7	249		
Technical college	13.7	61	15.0	109		
College degree	37.3	169	28.3	206		
Master's or professional	17.3	77	16.0	116		
PhD	2.7	16	2.0	14		
Other	0.0	0	0.7	3		

[a]Reporting likelihood ratio from chi-squared test (2-tailed significance).
[b]Reporting Z-statistic from Mann–Whitney*u*-test (2-tailed significance)
[c]$p < 0.01$.

Results

Tables 3.4.1and 3.4.2 compare characteristics of the survey respondents, adjusting household income for inflation. The national measure of the average change in price (the consumer price index) was used to adjust for inflation: 20.01% for the entire period [23]. The means of four differences are significant. On average, the 2003 respondents are 9.4 years older, whereas the 1995 respondents have a higher education level. The 2003 respondents have a smaller average household size and own fewer vehicles per household. The cumulative effect of these differences is difficult to predict. The 2003 sample is older and slightly less educated, characteristics typical of less frequent at-home ICT users. The 2003 respondents also have fewer vehicles and persons per household, which may lead to greater disposable income.

Comparing the Sample with National Trends

Contrasting the respondents with national trends establishes benchmarks for internal and general comparisons. Of all respondents, 73% owned a computer and 43% had home Internet access in 1995 versus 78.5% and 73.9%, respectively, in 2003. Each value is above the national trends outlined earlier. In terms of at-home ICT activities, 46.6% of 2003 respondents had purchased online and 41% had banked online, both slightly above national averages. Consequently, the 1995 and 2003 respondents may have had a greater affinity toward at-home ICT alternatives and associated higher frequency of use.

Table 3.4.2 Respondent characteristics, 1995 and 2003, Part 2

	Survey year				Comparing survey year	
	1995		2003		$N = 475$	$N = 738$
Characteristic	Average	*n*	Average	*n*	Statistic	
Age	45.0	466	54.4	707	9.902[a,b]	
Household income	61,016.0	457	59,299.0	669	−0.883[a,b]	
Household size	2.6	454	2.4	723	−2.454[a,b]	
Vehicles/household	2.2	456	2.0	726	−2.637[a,c]	
Household size	2.6	454	2.4	723	−2.454[a,b]	
Vehicles/household	2.2	456	2.0	726	−2.637[a,c]	

[a]Reporting *t*-statistic from independent sample *t*-test (2-tailed significance).
[b]$p < 0.01$.
[c]$p < 0.05$.

Changing Frequency of At-Home and Out-of-Home Shopping

The second strategy uses a MANCOVA model (Table 3.4.3). The multivariate *F*-statistic for variable survey year (equal to 64.450) is significant—as were age, household income, and three attitudes—in controlling for differences in the survey year, indicating that the survey year has an effect on out-of-home and at-home shopping frequency. The overall *F*-statistic of the model is significant for both dependents. The adjusted *R*-squared is 0.141 for out-of-home shopping and 0.172 for at-home shopping.

The univariate *F*-statistics explain the interaction effect of each category of variable survey year (1995 and 2003) and the dependent variables. For out-of-home shopping, the 2003 beta parameter is 0.654 compared with base year 1995, indicating that the mean frequency of out-of-home shopping was significantly greater in 2003. Interpreting the covariate parameters, adding this value to the intercept estimates out-of-home shopping when the covariate is zero. Household income, household size, and concerns about privacy with computers have a positive effect on out-of-home shopping frequency.

The 2003 respondents also were more likely to engage in at-home shopping; the beta parameter for 2003 at-home shopping is 0.145 compared with base year 1995. At-home shopping frequency is greater in high-income households, among respondents who believe that technology helps them save time, and in high-technology/high-congestion cities. Older respondents and those who prefer to spend free time with friends spend less time shopping at home. However, the primary finding from this line of analysis is that the independent variable of interest (survey year) has an effect on out-of-home and at-home shopping frequency.

Table 3.4.3 At-home and out-of-home shopping frequency, 1995 and 2003

Variable	Statistic	Out-of-home	At-home
Survey year			
(Hotelling's trace *F* = 64.450a)			
2003 *n* = 592	Beta	0.654[a]	0.145[b]
1995 *n* = 394		0	0
	F	132.237[a]	4.904[b]
Age			
(Hotelling's trace *F* = 20.740a)	Beta	0.000	−0.014[a]
	F	0.039	40.640[a]
Household income			
(Hotelling's trace *F* = 15.309a)	Beta	0.041[b]	0.127[a]
	F	3.942[b]	28.756[a]
Household size			
(Hotelling's trace *F* = 2.8022)	Beta	0.050[b]	−0.022
	F	4.483[b]	0.657
Household vehicles			
(Hotelling's trace *F* = 1.596)	Beta	0.042	0.049
	F	1.772	1.792
City type			
(Hotelling's trace *F* = 2.843)	Beta	−0.009	0.090[b]
	F	0.071	5.401[b]
Technology helps save me time			
(Hotelling's trace *F* = 19.198a)	Beta	0.010	0.181[a]
	F	0.145	38.322[a]
I worry about my privacy with computers			
(Hotelling's trace *F* = 4.418b)	Beta	0.047[b]	−0.044
	F	4.730[b]	3.112
Traffic drives me crazy			
(Hotelling's trace *F* = 0.878)	Beta	−0.025	0.015
	F	1.238	0.345
I prefer to spend free time with other people			
(Hotelling's trace *F* = 5.398a)	Beta	−0.007	−0.107[a]
	F	0.065	10.792[a]
Intercept			
(Hotelling's trace *F* = 75. 144a)	Beta	2.799a	1.947[a]
	F	115.969[a]	50.042[a]
Corrected model	F	17.105[a]	21.488[a]
R^2 (adjusted)		0.141	0.172
[a]p < 0.01.[b]p < 0.05.			

Table 3.4.4 At-home and out-of-home banking frequency, 1995 and 2003

Variable	Statistic	Out-of-home	At-home
Survey year			
(Hotelling's trace F = 2.537)	Beta	−0.161[a]	−0.120
2003 n = 468	F	3.810[a]	1.327
00 1995 n = 269			
Age			
(Hotelling's trace F = 11.628b)	Beta	0.001	−0.018[b]
	F	0.276	22.944[b]
Household income			
(Hotelling's trace F = 4.212a)	Beta	−0.077[b]	0.046
	F	6.847[b]	1.505
Household size			
(Hotelling's trace F = 1 .824)	Beta	0.036	0.067
	F	1.165	2.531
Household vehicles			
(Hotelling's trace F = 7.879b)	Beta	0.175[b]	−0.050
	F	14.947[b]	0.746
City type			
(Hotelling's trace F = 14.152b)	Beta	−0.201[b]	0.190[b]
	F	18.313[b]	9.685[b]
Technology helps save me time			
(Hotelling's trace F = 22.915b)	Beta	−0.078[a]	0.190[b]
	F	4.454[a]	41.0856[b]
I worry about my privacy with computers			
(Hotelling's trace F = 8.452b)	Beta	0.099[b]	−0.102[b]
	F	10.015[b]	6.700[b]
Traffic drives me crazy			
(Hotelling's trace F = 1.534)	Beta	−0.051	−0.034
	F	2.422 0.682	
I prefer to spend free time with other people			
(Hotelling's trace F = 0.131)	Beta	0.007	−0.025
	F	0.027	0.234
Intercept			
(Hotelling's trace F = 58.982b)	Beta	3.062[b]	1.831[b]
	F	93.992[b]	21.253[b]
Corrected model			
	F	7.001[b]	12.650[b]
R^2 (adjusted)	0.075 0.137		

[a]$p < 0.05$.[b]$p < 0.01$.*NOTES:* Wilks' lambda *F*-statistic is for DV among IV. *F*-statistic is between subjects. Household income from 1995 adjusted for inflation.

Changing Frequency of At-Home and Out-of-Home Banking

The variable measuring survey year does not explain changes in out-of-home and at-home banking. The model in Table 3.4.4 shows an insignificant multivariate *F*-statistic equal to 2.537, although this statistic would be significant at a 0.90 confidence level. The covariates that have an effect in the model are age, household income, the number of household vehicles, one attitude, and city type. The overall *F*-statistic is significant for both dependents. The adjusted *R*-squared is 0.075 for out-of-home banking and 0.137 for at-home banking. Although the out-of-home *R*-squared is relatively low, the parameter estimates make sense and have the expected sign, suggesting a practical model with numerous unobserved effects.

The univariate *F*-statistics explain the interaction effects of the independent and dependent variables. Regarding out-of-home banking, the 2003 beta parameter (equal to −0.161) indicates that the 2003 respondents banked out of home less frequently than respondents in 1995. Interpretation of the covariates indicates that out-of-home banking frequency increases as the number of household vehicles and concerns about privacy while using computers increase. Conversely, household income, a pro-technology attitude, and a high-technology/high-congestion city have a negative effect.

The variable measuring survey year fails to explain changes in at-home banking. The 2003 beta parameter is negative, indicating that respondents in 2003 banked less often at home than respondents in 1995, but the result is not significant. At-home banking frequency decreases with age and privacy concerns while using computers. Respondents who live in a high-technology/high-congestion city and believe that computers help them save time have a positive effect on at-home banking frequency.

Changing Substitution or Inducement of Shopping and Banking Activities

Further examination of the two surveys reveals the effects of at-home ICT activities on trip substitution and inducement. The degree to which respondents substituted at-home shopping and banking for out-of-home alternatives is reported in Table 3.4.5. In 2003, 79% of at-home users said they would have visited a store had an at-home option been unavailable; this response is in sharp contrast to 20% of respondents who would have made that trip in 1995. Approximately 56% of at-home users in 1995 and 2003 reported that an at-home activity induced an in-store trip.

Unlike with shopping, respondents in 2003 were less likely to substitute at-home banking for out-of-home trips; nearly 40% of the 1995 respondents reported that at-home banking substituted for a trip compared with 27.6% in 2003. As a point of comparison, respondents have not changed in the degree of substitution of out-of-home ATM transactions for out-of-home bank trips. Roughly 56% of respondents in both 1995 and 2003 would have made a trip to the bank had their last ATM use not been possible. The relationship between using an ATM and visiting a bank—both out-of-home forms of banking—are discussed in more detail below.

Table 3.4.5 Shopping and banking substitution and inducement, 1995 and 2003

	Yes				No				Comparing 1995 and 2003
	1995		2003		1995		2003		
	%	n	%	n	%	n	%	n	Statictic[b]
Shopping									
Would have made out-of-home trip had last at-home purchase been unavailable (among at-home users)	20.2	73	79.0	271	79.8	288	21.2	73	257.824[b]
Ever made an out-of-home trip because of something seen at-home (among at-home users)	56.4	264	55.5	208	43.6	204	44.5	167	0.075
Banking									
Would have made out-of-home trip had last at-home transaction been unavailable (among at-home users)	39.5	98	27.6	64	60.5	150	72.4	168	7.675[c]
Would have made out-of-home trip had last ATM transaction been unavailable (among ATM users)	48.9	160	49.8	258	51.1	167	50.2	260	0.062

[a]Reporting likelihood ratio from chi-squared test (2-tailed significance).
[b]$p < 0.01$. [c]$p < 0.05$.

Table 3.4.6 Bank and ATM proximity to respondent home and work locations, 1995 and 2003

	Yes				No				Comparing 1995 and 2003
	1995		2003		1995		2003		
	%	*n*	%	*n*	%	*n*	%	*n*	Statictic[b]
Bank or ATM proximity									
Bank within walking distance of home	15.8	73	27.5	202	84.2	389	72.5	533	22.722[b]
Bank within walking distance of work	22.1	92	18.3	123	77.9	325	81.7	549	2.272
ATM within walking distance of home	35.4	161	40.9	291	64.6	294	59.1	421	3.537
ATM within walking distance of work	43.7	180	42.9	285	56.3	232	57.1	379	0.061

[a]Reporting likelihood ratio from chi-squared test (2-tailed significance).
[b]$p < 0.0$.

3.4.5 Discussion

The multivariate models paint a distinct picture of out-of-home and at-home shopping and banking; they suggest that activities are changing in both frequency and form of technology. The variable that measures survey year has an effect on out-of-home and at-home shopping after controlling for sample characteristics, attitudes, and city type. Survey year is notable but does not significantly explain banking frequencies. These findings are not particularly surprising, because shopping and banking patterns might be expected to evolve in terms of frequency of use and form of at-home activity.

Respondents were more likely to shop both out of home and at home in 2003 than in 1995. In terms of travel demand management, as at-home technologies improve and congestion worsens, the expectation is that at-home shopping will grow and out-of-home shopping will decline. Instead, respondents engaged in multiple forms of shopping with greater frequency. Several factors may explain this behavior. Respondents may

- Continue to prize store shopping to view a product in person, ask questions, acquire a product immediately, or engage in social activities (e.g., get out of the house);
- Chain multiple trips;
- Maximize convenience, browse, or shop for certain goods at home and continue to buy other goods out of home; and
- Take more frequent trips with shorter durations because of changing time constraints.

Respondents reported a greater degree of substitution of out-of-home shopping for at-home alternatives in 2003. Given that both out-of-home and at-home shopping increased between 1995 and 2003, this finding seems contradictory. One possible explanation is that because respondents shopped stores more frequently, they were more willing to make a store trip if the product was unavailable at home. Another possibility is that the greater frequency of at-home shopping reflected general increases in product needs and desires that respondents would satisfy with either at-home or out-of-home purchases.

The extent of store shopping inducement remains unchanged. More than one-half of respondents in 1995 and 2003 had ever made a store trip to purchase an item seen at home. This result likely reflects several behaviors. The first is price comparison, which may be conducted at home (e.g., by browsing multiple catalogs or Internet sites) to discover the best deal for their product of interest. The second is a reluctance to buy without viewing, so people may analyze products at home, then travel to a store to view the product in person. This moves toward a third explanation, which is the desire to view a product in person before purchasing.

Respondents were less likely to bank out of home in 2003 than in 1995. The decrease in at-home banking frequency between survey years is not significant, despite changes in the form of at-home banking technology. In contrast to shopping, the results indicate less banking overall. One possible explanation is online paycheck deposits and automatic bill payment, which respondents may not count toward total banking activity. Respondents in 2003 were less likely to indicate that at-home banking substituted for out-of-home physical trips. The desire to conduct certain transactions at banks (e.g., obtaining loans and cashier's checks) may continue the need for store banking.

The proximity of banks and ATMs could explain the decrease in at-home banking and limit future at-home banking growth. Table 3.4.6 displays bank and ATM proximity to respondents' homes and places of work. A greater number of respondents reported having a bank within walking distance of their home in 2003. Although the difference in ATM proximity is not significant, results indicate that roughly 40% of respondents are within walking access of an ATM at one of two key locations. The ease of accessing these locations, in addition to a seemingly increasing number of banks and ATMs at other locations—supermarkets, bookstores, and gas stations—may limit the need to engage in at-home banking and slow future growth. The various banking alternatives also may reduce the observed effect of any individual at-home banking activity on trip substitution.

The characteristics of people willing to engage in at-home activities, their attitudes, and city characteristics explain changes in the frequency and form of shopping and banking activity. The models confirm some of the expected variable effects; however, the influence is not always consistent and varies depending on the type of at-home technology. For instance, respondent age was significant in at-home models but insignificant in out-of-home models. Increasing household income has a positive effect in shopping models, a negative effect in the out-of-home banking model, and an insignificant effect in the at-home banking model. Respondents from high-technology–high-congestion cities were more likely to increase their at-home frequency in both activities but varied in out-of-home frequency. Attitudes toward technology and congestion generally had the expected signs but were not always significant. Unexpectedly, the attitude "traffic drives me crazy" was not significant in any of the models. Together, the results indicate that expectations that certain populations or specific locations will engage in at-home shopping and banking alternatives may be overstated. Instead, evolving use patterns depend of the activity (shopping or banking) and the form of the at-home activity (catalog, phone, or online).

3.4.6 Conclusion

ICT use and its potential to reduce travel have been discussed enthusiastically for many years. The longitudinal approach used herein analyzed the differences from two similar surveys to examine the evolving relationship between store travel and at-home ICT alternatives. This analysis compared typical at-home activities during two survey years: catalog shopping and phone banking in 1995, and online shopping and online banking in 2003. The variable measuring the year of the survey had a main effect in explaining changes in the frequency of at-home and out-of-home shopping but not banking.

In 2003, respondents engaged in greater amounts of at-home and out-of-home shopping. They also were more likely to substitute at-home activities for store trips; however, the overall increase in shopping seemed to override this change. Examining banking activities, a decrease in out-of-home banking and a notable, but insignificant, decrease in at-home banking were observed from 1995 to 2003. Examining multiple behaviors, the most likely explanation for the results is that with the growth in ICT use, people engage in multiple forms of shopping and banking and do so in the ways most convenient to them. ICT has expanded the number of means available for carrying out activities but has not significantly replaced the earlier means. The desire to shop in a physical store, for social or other reasons, and to touch and examine products before buying is a possible explanation, although this analysis fails to fully capture this effect. Similarly, people may be unwilling to conduct certain transactions away from banks. The proximity of banks and ATMs to home and work locations and the convenience of ATMs in other stores may not warrant online banking for many individuals.

The approach presented in this paper offers insights that pure cross-sectional studies on their own cannot. MANCOVA is a useful technique to control for differences in similar sample populations and examine longitudinal changes. The surveys were limited in scope, and this analysis was limited to the portions of the two surveys that provided sufficient and consistent data. This study echoes the difficulty in capturing the effect of evolving at-home technologies. Future studies can work toward this goal by testing the frequency of at-home ICT use before and after receiving access to the newest technologies using panel data. Notwithstanding such shortcomings, the approach presented in this paper offers a viable alternative to surmising results from a series of unrelated cross-sectional studies.

3.4.7 References

1. *Transportation Implications of Telecommuting.* U.S. Department ofTransportation, Washington, D.C., 1993. isddc.dot.gov/OLPFiles/OST/011836.pdf. Accessed Feb. 18, 2006.
2. Schrank, D., and T. Lomax. The 2005 *Urban Mobility Report.* TexasTransportation Institute, College Station, Tex., 2005.
3. Handy, S. L., and T. Yantis. *The Impacts of Telecommunications Technologieson Nonwork Travel Behavior.* SWUTC/97/721927–1F. SouthwesternRegion University Transportation Center, Austin, Tex., 1997.www.uctc.net/papers/233.pdf. Accessed Feb. 16, 2006.
4. Mokhtarian, P. L. *A Typology of Relationships Between Telecommunicationsand Transportation.* Transportation Center, University of California,Berkeley, June 1990. www.uctc.net/papers/082.pdf. AccessedFeb. 18, 2006.
5. Gould, J., and T. F. Golob. Will Electronic Home Shopping ReduceTravel? *Access: Research at the University of California TransportationCenter,* Vol. 59, No. 12, 1998, pp. 26–31.
6. Lenz, B. Will Electronic Commerce Help to Reduce Traffic in AgglomerationAreas? In *Transportation Research Record: Journal of the TransportationResearch Board,* No. 1858, Transportation Research Board of theNational Academies, Washington, D.C., 2003, pp. 39–46.
7. Gould, J., and T. F. Golob. Shopping without Travel or Travel withoutShopping: *An Investigation of Electronic Home Shopping.* Institute ofTransportation Studies, Center for Activity Systems Analysis, Universityof California, Irvine, 1997. UCI-ITS-AS-WP-97.2. repositories.cdlib.org/cgi/viewcontent.cgi?article=1046&context=itsirvine/casa. AccessedMarch 6, 2006.
8. Bhat, C. R., A. Sivakumar, and K. W. Axhausen. *An Analysis of the Impactof Information and Communication Technologies on Non-MaintenanceShopping Activities.* Institute for Transport Planning and Systems, Zurich,Switzerland, 2002. www.ivt.baum.ethz.ch/vrp/ab97.pdf. AccessedFeb. 16, 2006.
9. Ferrell, C. E. Home-Based Teleshoppers and Shopping Travel: DoTeleshoppers Travel Less? *In Transportation Research Record: Journalof the Transportation Research Board,* No. 1894, TransportationResearch Board of the National Academies, Washington, D.C., 2004,pp. 241–248.
10. Giglierano, J., and M. Roldan. *Effects of Online Shopping on VehicularTraffic.* MTI Report 01–20. Mineta Transportation Institute (MTI), SanJosé State University, San Jose, Calif., October 2001. transweb.sjsu.edu/mtiportal/research/publications/documents/01–20.pdf. AccessedJan. 3, 2007.

11. Krizek, K., S, Handy, and Y. Li. ICT as a Substitute for Non-WorkTravel: A Direct Examination. Working paper. Active Communities/Transportation (ACT) Research Group, University of Minnesota,Minneapolis, 2004.

12. Krizek, K. J., Y. Li, and S. L. Handy. Spatial Attributes and Patterns ofUse in Household-Related Information and Communications TechnologyActivity. *In Transportation Research Record: Journal of the TransportationResearch Board,* No. 1926, Transportation Research Board ofthe National Academies, Washington, D.C., 2005, pp. 252–259.

13. Cao, X., and P. L. Mokhtarian. The Intended and Actual Adoption ofOnline Purchasing: A Brief Review of Recent Literature. UCD-ITSRR-05–07. Institute of Transportation Studies, University of California,Davis, 2005.

14. Day, J. C., A. Janus, and J. Davis. *Computer and Internet Use in the UnitedStates: 2003.* U.S. Department of Commerce, Economics and StatisticsAdministration, U.S. Census Bureau, Washington, D.C., October 2005.www.census.gov/prod/2005pubs/p23-208.pdf. Accessed Feb. 16, 2006.

15. *A Nation Online: Entering the Broadband Age.* U.S. Department ofCommerce, Economics and Statistics Administration and NationalTelecommunications & Information Administration, Washington,D.C., September 2004. www.ntia.doc.gov/reports/anol/NationOnlineBroadband04.htm. Accessed Feb. 22, 2006.

16. *Latest Trends: Usage Over Time.* Pew Internet & American Life Project,Washington, D.C., 2005. www.pewinternet.org/trends.asp. AccessedJan. 30, 2006.

17. Lebo, H. The *Digital Future Report. Surveying the Digital Future: YearFour—Ten Years, Ten Trends.* USC Annenberg School Center for the DigitalFuture, University of Southern California, Los Angeles, Sept. 2004.www.digitalcenter.org/downloads/DigitalFutureReport-Year4-2004.pdf.Accessed Feb. 20, 2006.

18. *Quarterly Retail E-Commerce Sales,* 2006. U.S. Department of Commerce,Economics and Statistics Administration, U.S. Census Bureau,Washington, D.C. www.census.gov/mrts/www/ecomm.html. AccessedFeb. 17, 2006.

19. e-Data B2C: U.S. E-Commerce Growing Faster than Total RetailSpending, According to eMarketer. ITAA E-Letter, May 2005. www.itaa.org/isec/pubs/e20055–11.pdf. Accessed Feb. 15, 2006.

20. *Online Banking Gaining Worldwide Momentum.* Global Market Insite,Inc., 2005. www.gmimr.com/gmipoll/press_room_wppk_pr_10192005.phtml. Accessed Feb. 9, 2006.

21. Hayashi, F., R. Sullivan, and S. E. Weiner. *A Guide to the ATM andDebit Card Industry.* Payments System Research Department, FederalReserve Bank of Kansas City, Mo., 2003. www.kc.frb.org/FRFS/ATMpaper.pdf. Accessed March 20, 2006.

22. Garson, G. D. GLM MANOVA and MANCOVA. *Statnotes: Topics inMultivariate Analysis.* www2.chass.ncsu.edu/garson/pa765/statnote.htm.Accessed March 11, 2007.

23. *Consumer Price Indexes.* U.S. Department of Labor, Bureau of LaborStatistics, 2007. www.bls.gov/cpi/. Accessed Jan. 15, 2007.

Endnotes for 3.1

ICT as a Substitute for Non-Work Travel: A Direct Examination

[i] A total 17% of Internet users (about 14 million Americans) had ever used online banking services and 36% (about 31 million Americans) had ever made travel purchase over the Internet. In September 2002, the numbers have increased to 32% (online banking) and 50% (online travel purchase).

[ii] An example question is: Think of the last time that you…(e.g., used an ATM machine). If you had not had the option of…(e.g. using an ATM), what would you have done?

[iii] An example question is: Think of last time…(e.g. you made a trip to the bank even though you were aware of the option to bank online), why did you choose to do so? Or if you choose to go to a store instead of purchasing online, which ONE of the following is the MAIN reason?

[iv] Roughly 62% of the respondents answered these questions.

CHAPTER 4

Tom Horan: ITS and Emergency Medical Services Response

4.1 Towards End-To-End Government Performance Management

Case Study of Interorganizational Information Integration in Emergency Medical Services (EMS)

Benjamin Schooley, MIS
School of Information Science
Claremont Graduate University
Claremont, CA 91711, USA
F: +1 909/621 8564
E-mail: Ben.Schooley@cgu.edu

Thomas A. Horan, Ph.D.
School of Information Science
Claremont Graduate University
Claremont, CA 91711, USA
F: +1 909/621 8564
E-mail: Tom.Horan@cgu.edu

4.1.1 Introduction

Interorganizational systems (IOS) concepts provide a targeted means to look at the cross-organizational features of a socio-technical system (Cash & Kosynski, 1985; Williams & Fedorowicz, 2005). In its simplest form, an interorganizational system (IOS) is an information and communications technology-based system, shared by two or more independent organizations (Bakos, 1991; Johnston & Vitale, 1988). Past IOS research has focused primarily on private sector systems, some of the earliest studies discussing how IOS creates competitive advantage (see Cash & Kosynski, 1985; Guglar & Dunning, 1993; Johnston & Carrico, 1988). A more recent research effort has stressed the value of interorganizational information sharing in the public sector (e.g., Dawes, 1996; Fountain, 2001; Sawyer, Tapia, Pesheck, & Davenport, 2005;West, 2005; Zhang, Dawes, & Sarkis, 2005). E-government IOS research has involved several public sector contexts such as criminal justice (Sawyer et al., 2005; Scholl, 2005) and services to citizens (e.g., tax processing, workers compensation insurance, forest service information) (Dawes & Prefontaine, 2003; Drake, Steckler, & Koch, 2004). These and other e-government researchers have demonstrated a need to improve capabilities to share data, information and knowledge across departmental, organizational, geographic, and institutional boundaries (Rudman, Clarke, & Metzel, 2003; Sawyer et al., 2005). It has been posited that such interorganizational improvements in information sharing will improve the performance of public sector services (Dawes, 1996; Landsbergen &Wolken, 2001; Layne & Lee, 2001; Pardo, 2000). However, there has been limited attention given to the e-government service performance benefits that result across a coordinated chain of service organizations (Drake et al., 2004; West, 2005). Drake and colleagues (2004) identify the research need to further explore issues and challenges in interorganizational public services to further understand the value chain concept of information sharing within and across a public service delivery system. More specifically, an examination of what value is added or subtracted at each stage of public service delivery is needed. This study addresses the abovementioned research need within an e-government IOS domain that has received very little attention—that of the sequentially cooperative processes of government and private organizations providing emergency medical services (EMS) to citizens (Horan & Schooley, 2007). The emergency response domain introduces unique and challenging dynamics and complexities to multi-organizational information sharing, including the time-critical nature of emergency services and the need for timely information in a form that can be trusted and used by emergency responders (Arens & Rosenbloom, 2002; Dawes, Cressell, & Cahan, 2004; Sawyer, Tapia, Pesheck, & Davenport, 2004; Turoff, Chumer, Van de Walle, & Yao, 2004). Effective and timely service depends upon all participating organizations working cooperatively and utilizing information technology effectively (Mayer-Schonberger, 2003).

Few studies have addressed the service performance implications of IOS in the egovernment domain generally, and EMS specifically. Nearly a decade ago, the National Highway Transportation Safety Administration (NHTSA) (1996) identified the importance and need for research, evaluation, and development of robust data and information systems to support multi-organizational EMS. More recent studies have stated that interorganizational EMS continues to operate without a sufficient research basis to support many of its operational and information systems decisions (IOM, 2006; McLean, Maio, Spaite, & Garrison, 2002; NHTSA, 2001; Sayre, White, Brown, & McHenry, 2003). Enabling interorganizational information sharing has been identified as an important precursor to improving EMS research and system-wide services (IOM, 2006; NENA, 2001; NHTSA, 2001). Kapucu (2006) identified the need to better understand the role of communication processes, supporting information technologies, and information sharing across emergency organizations to achieve higher levels of service performance. Furthermore, a very recent Institutes of Medicine (2006) report discussed the need to further explore the socio-technical nature of and uses of interorganizational systems in EMS and how they relate to performance benefits. This study addresses this specific gap in the IOS e-government EMS literature.

Most directly, this study extends research conducted by Horan, McCabe, Burkhard, and Schooley (2005) and Horan and Schooley (2007), who present a general framework for understanding and studying a type of interorganizational public service process—one that is highly time critical. Horan and colleagues explain that future research should seek a detailed examination into how participating agencies and organizations interact with each other and with technological systems in the delivery of emergency services, including what the perceived performance benefits are for sharing information.

The research issue addressed in this investigation is: How do operational, organizational, and governance dimensions of interorganizational time-critical services influence the use of performance information from end-to-end? The concepts in this research question will be further explained below in terms of how they relate to this study.

4.1.2 Theory

The conceptual framework that guides this study is the time-critical information services (TCIS) framework (Horan et al., 2005; Horan, Marich, & Schooley, 2006) (see Figure 4.1.1). TCIS was developed to aid in the study of time-critical information services (TCIS), that is, public services that are highly time and information dependent. The researchers have developed TCIS as a heuristic that allows for a multidimensional view of "end-to-end" system performance and information sharing therein for time information critical services such as EMS. Researchers investigated prominent sociotechnical works such as Fountains (2001) technology enactment and Sussman (2002) complex, large-scale, interorganizational, open systems (CLIOS). From this investigation, TCIS was developed as a way to distinguish between different simultaneously ongoing streams of phenomena, some of which are organizational, some of which are performance-based, technological, time-dependent, etc., and frame them into an analytical lens

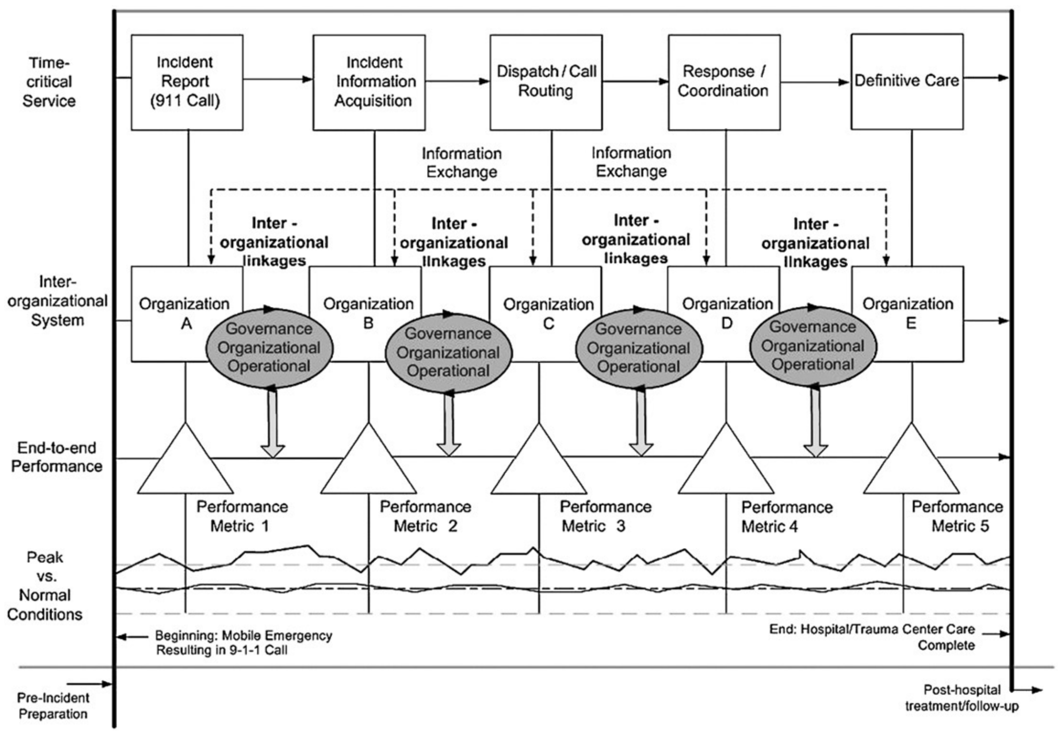


Figure 4.1.1 TCIS framework.

for interorganizational systems (IOS) analysis. short, TCIS is a conceptual framework derived from theory that is meant to simplify a complicated array of relationships, behaviors, technologies, and entities and provide utility to systems analysts. While TCIS has been refined through conferences (e.g., DG.O 2006), professional publications (e.g., Horan & Schooley, 2007), and a National Science Foundation (NSF) supported workshop on the topic (see Horan et al., 2006), the case study reported below has provided an important opportunity to further validate this approach through examination of a real-life time-critical service.

The conceptual model includes several levels of analysis for TCIS, both in regard to EMS specifically and other public services generally. These levels, shown in Figure 4.1.1, include (1) the time and information critical elements of a sequential public service process, (2) the interactions and information exchanges across multiple cooperating service organizations that include both qualitative organizational elements as well as “hard” information flow elements, (3) the end-to-end elements that consider performance metrics within and across the process flow, and (4) context variation elements such as normal versus peak conditions (in terms of service demand).

This study focuses on the interorganizational information-sharing dimensions of TCIS (Figure 4.1.1, 2nd row from the top). Their research findings and resulting framework propose a structure for understanding operational, organizational, and governance dimensions of interorganizational performance information sharing and integration to gain a deeper understanding about how information sharing influences the design and improvement of time critical public services, EMS service delivery, and information systems to support these services.

The following section of the paper further defines the “information-sharing dimensions” to be analyzed in the case study, which draws from the literature on e-government information sharing (see Table 4.1.1). Beginning with the *operational/technical* dimension of information sharing, this includes business processes, procedures, and technological resources for sharing information and data across organizational boundaries (Dawes, 1996; Sundberg & Sandberg, 2006). The "business process and procedure" aspect involves understanding what tasks are involved, who is involved, where tasks are performed, and how processes work (Curits, Kellner, & Over, 1992; Fedorowicz et al., 2005). In terms of the technical systems, Dawes and Prefontaine (2003) state that interorganizational public service performance and communication are strongly shaped by the capabilities of the technical tools. This includes understanding what computer, Internet, and communications technologies (hardware and software) are utilized to share information across organizations, how they are used, and how they influence information sharing with cooperating organizations (Dawes & Prefontaine, 2003; Scholl, 2005). These technical elements are extracted from the case studies to understand the operational dimensions of information sharing.

In a review of the literature on *organizational* models, trust has been linked to successful performance outcomes in such processes as teamwork, leadership, goal setting, performance appraisal, and more effective crisis management (Grossman, 2004; Shockley-Zalabak, Ellis, & Winograd, 2000; West, 2005). Additional organizational elements that have been found to influence performance of an IOS are effective communications, cultural differences, level of participation, power relations, and resistance to change (Allen, Colligan, Finnie, & Kern,2000; Dawes et al., 2004; Sundberg & Sandberg, 2006; Zhang et al., 2005), as well as cultural and subcultural dimensions of information sharing (Drake et al., 2004). These elements are explored in terms of how they influence interorganizational information sharing.

In addition to operational and organizational dynamics, governance issues are also explored as they affect and influence interorganizational information sharing. Fountain (2001) explains that governance structures facilitate, coordinate, and control what happens within an interorganizational collaboration episode. This includes "…roles that have been defined for participants to fill, the relationship among those roles, and regulations that govern the use of roles and relationships for participants to fill" (Holsapple & Luo, 1996). In addition to issues surrounding participant roles, this study also seeks to extract elements such as decision-making processes, rules, regulations, legal, political, and fiscal issues as posited by other researchers (Landsbergen & Wolken, 2001; Williams & Fedorowicz, 2005).

Table 4.1.1 Analytical dimensions of interorganizational information sharing

Operational dimensions
Technical systems (software and hardware)
Business processes (who, what, where, how)
Communication flows (voice and data)
Organizational dimensions
Power relations
Level of participation
Cultural, subcultural differences/similarities
Resistance to change
Trust
Governance dimensions
Participant roles
Rules and regulations
Decision-making processes
Political/legal
Fiscal

4.1.3 Performance Measures in EMS

In order to explore performance implications of interorganizational information sharing, it is important to further describe the concept of performance as it relates to this study. According to the Institute of Medicine (2001), EMS system performance could be measured from a number of perspectives including financial, safety, effectiveness, patient-centeredness, quality of care, timeliness, efficiency, and equity of the system, where each perspective overlaps with another. This study utilizes the view of a more recent IOM (2006) study where the critical performance features of the system are those that matter most to a patient: timeliness and quality of care.

Timeliness

Time has long been used in EMS to measure interorganizational system performance (IOM, 2006). It is measured by recording time stamps at specific points in service delivery. EMS service typically begins with a consumer action (placing an emergency 911 call), involves the private sector (telecommunications service provider) delivering the call, the public sector (PSAP, state police) receiving and dispatching the call, the private and/or public sector (ambulance service, fire, or police) providing first response, transport and health care services, and finally either a public or private sector hospital or trauma center delivering appropriate health care services. Time stamps are typically recorded at the time of a call, the time the call was answered, at the time resources (fire, ambulance, police, etc.) are dispatched, resources arrive to a scene, and resources depart to a hospital or trauma center, arrive to the hospital, and complete delivery.

Quality of Care

Quality care performance measurement is complex in part because of the wide range of possible types of emergency health conditions and the difficulty diagnosing them (Coffey, 1998; Spaite et al., 1995). For example, some of the most effective medical procedures for treating a heart attack victim differ significantly from the procedures for treating neck trauma patients. A significant finding from a recent Institute of Medicine (IOM, 2006) report was that quality of care performance differs across local EMS systems. In this regard, there are several ongoing health care initiatives whose primary goal is to define, improve, and standardize medical care performance measures, data collection, and analysis. For example, the Health Care Financing Administration (HCFA) has been developing core sets of measures for a number of common conditions, including acute myocardial infarction, heart failure, stroke, pneumonia, and others (Health Care Financing Administration, 2000). It is safe to say that unlike the nature of the "time" metric, the quality of care metric is more multifaceted. This case analysis explores how information is shared across organizations and how interorganizational information integration may influence the general notion of quality care EMS performance. As such, quality of care performance metrics vary from incident to incident and are embedded in the information types that have been labeled "treatment provided" and "patient condition" in the case study analysis (see Figure 4.1.5).

End-to-End Performance

The above two performance dimensions explain "what" the important patient-centered measurement domains are. This section describes "how" service is measured. For a cooperative interorganizational network such as emergency response, total performance from service initiation to resolution (from end-to-end) is essential. For example, it makes little difference for an operator to dispatch quickly if the ambulance takes a very long time to arrive and/or goes to the wrong location. The concept and term referred to as "end-to-end" has long been used by computer and social network researchers to discuss service performance across a networked event (for example, see Monge & Contractor, 1988; Provan & Milward, 2001; Wigand, 1988). We utilize this terminology and concept in our investigation of EMS service performance, from service initiation (e.g., 911 phone call), through dispatch, EMS response, and resolution (e.g., definitive health care at a hospital).

4.1.4 Methodology

This case study effort examined dimensions of interorganizational information sharing through two overlapping phases. The first phase investigated the operational and technical levels including what performance-related information is/are not collected, communicated, and/or exchanged across organizations. It examined the business processes and information flows across organizations and information systems. The types of data that were collected in phase 1 are as follows: business process documentation, performance data for the year 2005, technical information system documentation, management reports, and performance reports, inter-organizational agreements including formal and informal contracts, as well as field notes and supplemental interviews. These data were collected through field visits on location at each participating organization as well as through follow-up phone and e-mail conversations. During field visits, researchers interviewed staff and management personnel, observed both demon-strations and real-world operations, including sitting with emergency dispatch operators and riding along with fire/paramedic and ambulance crews. The organizations visited and positions of persons interviewed are outlined in Table 4.1.2. The data collection for findings reported below occurred from October 2005 to September 2006.

The phase 1 examination provided an operational understanding that allowed researchers to construct interview questions for phase 2. The second research phase examined contextual issues about the operational processes and information exchanges. Interview and round table discussion participants were selected based on their organizational roles. In particular, upper management was asked to identify persons who have extensive knowledge about and experience with interorganizational business processes and information exchanges. Semi-structured interview questions sought to understand dimensions to information sharing. The interviews also sought to understand how performance information sharing, or lack thereof, relates to performance benefits or challenges for each organizations' service and how information integration initiatives could be structured to enhance information systems and end-to-end service performance. As in the first phase, researchers took detailed field notes, and summarized observations. Issues were categorized by the three focus areas mentioned previously: operational, organizational, and governance.

4.1.5 Research Setting

San Mateo County is located between the Pacific Ocean on the west, San Francisco Bay on the east, San Francisco County on the north, Santa Cruz County on the south, and Santa Clara County on the south east. The majority of the population estimated at approximately 699,6102[i] live along a narrow corridor along the 101 freeway, with the majority of the land mass located in rural and remote mountain and coastal areas (see Figure 4.1.2for illustration).

The San Mateo County EMS Agency formed an innovative public/private partnership to provide more efficient and effective emergency medical services to its citizens throughout the county. The partnership was established in 1999 after a four-year planning process that involved nurses, physicians, paramedics, city managers, fire agencies, ambulance providers, consumers, and county staff as well as an RFP process with ambulance providers. The partnership includes the County Health Services Department's EMS office, American Medical Response (AMR) ambulance service, a Joint Powers Authority (JPA) made up of all 18 fire service agencies in the county, and the County Public Safety Communications Center. Oversight for the EMS system, both operational and medical, is provided by the EMS office of county Health Services, which holds the master contract with AMR for both ambulance service and paramedic first response. AMR has subcontracted paramedic first response service to the JPA and communications dispatch service to the County Communications Center (see Figure 4.1.3 for partnership arrangement). The 11 independent health care facilities throughout the county are not formally included in the partnership but collaborate with each of the major partners. This interorganizational arrangement was the county's first "performance-based" contract system, and it has received a number of awards from the National Council for Public– Private Partnerships, International City–County Management Award for Outstanding Partnerships, the League of Cities Helen Putnam Award for Excellence in Public Safety, the National Academies of Emergency Medicine, and the International Association of Fire Chiefs.

Table 4.1.2 San Mateo County participants

Organization	Position
California Highway Patrol	Communications center supervisor
San Mateo County EMS Agency	Administrator, medical director, project manager/data analyst, clinical coordinator, EMS-children project coordinator
American Medical Response	Paramedic supervisor, paramedic, emergency medical technician (EMT)
County Fire Joint Powers Authority	Fire chief, paramedic, JPA representative
Hospitals	Administrator, emergency physician
San Mateo Communications Center	Manager, supervisor, information systems manager
Total expert participants: 17 Roundtable discussions: 2	

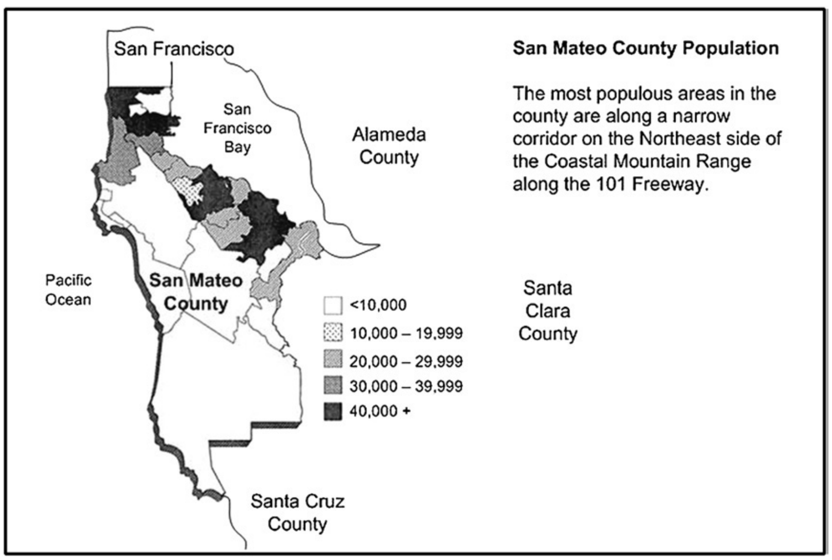


Figure 4.1.2 San Mateo County location and population

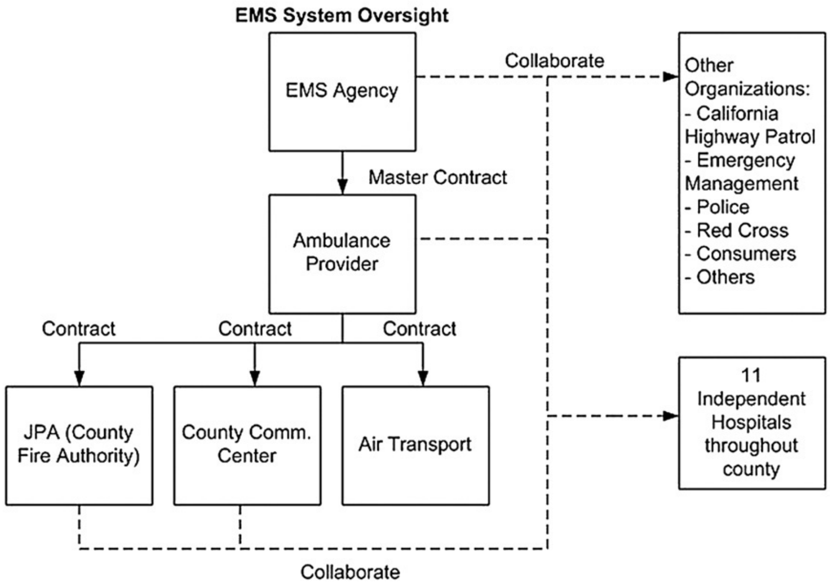


Figure 4.1.3 San Mateo County EMS partnership arrangement

Table 4.1.3 San Mateo EMS Agency contractual response times for urban, rural, and remote areas

No. of calls to county communications center—fire/EMS related	201,099
No. of recorded incidences	103,006
Patient transports to health care facilities	23,199
Urban area average response time—first responder	6:59
Rural area average response time—first responder	11:59
Remote area average response time—first responder	21:59
Urban area average response time—ambulance service	12:59
Rural area average response time—ambulance service	19:59
Remote area average response time—ambulance service	29:59

System Overview

The EMS system consists of a single consolidated dispatch center that performs all dispatch services for the 18 fire departments and ambulances within the County. The dispatch center .also provides dispatch services for the County Sheriff's department and five other local law enforcement agencies in the county. It does not dispatch for the state-run California Highway Patrol (CHP) or for the remaining 18 local law enforcement agencies in the county. However, due to the need for local agencies to reduce operational costs, several of these law enforcement agencies are in the planning stages to outsource dispatch to the County Communications Center.

While the previous system dispatched fire engines based on jurisdictional boundaries, the current system dispatches the closest fire engine and ambulance to a medical incident. The previous system provided paramedic ambulance response with a nine-minute response time standard in the more populated areas. In the current system, fire engines with paramedics on board respond within seven minutes in urban areas (with >90% compliance). Table 4.1.3 shows the response time requirements as defined in the EMS Agency master contract with the ambulance provider for urban, rural, and remote areas. All responses must be in compliance for at least 90% of all responses. Emergency ambulances continue to be staffed by paramedics but have a response time standard of 13 minutes (with >90% compliance). The cost savings to AMR for the extended response time is used to subsidize the fire departments for providing the paramedic first response service. The partnership has also standardized such functions as training, communication protocols, quality improvement, equipment (including software), supplies, and record keeping across organizations.

In terms of information systems, the EMS system features a single e-patient care record (PCR) per patient that includes select data assimilated from the Communication Center's computer-aided dispatch (CAD) system, the first responders (Fire/Paramedic crews), and the ambulance service. Ambulance crews utilize laptops with wireless connections to transmit data to the centralized PCR system server, which feeds data to anyone logged into the system, including the eleven health care facilities in the county. Though the system is accessible by each organization, the EMS Agency reports that it is primarily used by the ambulance service, with some usage by the Fire department and rare use by the health care facilities.

The EMS Agency's vision for the information system includes a single electronic patient care record (PCR) system that can share appropriate data elements with every organization in the service process, including the health care facilities. This system would utilize a suite of XML-based schemas and protocols in the transfer of its data to and from other disparate systems (law enforcement, hospitals, dispatch center, etc.). This vision is part of an ongoing initiative to integrate cross-organizational information to utilize for operational and management decision making.

This case study looks at the EMS Agency's ongoing interorganizational information integration initiative from the perspective of the operational, organizational, and governance dimensions that inhibit or prohibit interorganizational performance information exchanges. As such, and as indicated at the outset, a first step in the analysis was to look at the existing operational system in terms of business processes, information flows, and information systems across processes and organizations.

4.1.6 Assessment Service Operations

The following section provides an overview of the case study findings. Figure 4.1.4 applies the TCIS framework to the San Mateo County EMS case study. It illustrates a sequential service process involving multiple public, private, and not-for-profit organizations collecting and sharing information related to the process, the incident, and service performance. Figure 4.1.5 shows a high-level overview of the types of information collected and transmitted across organizations during an emergency incident. It shows how the EMS service process starts at a 911 phone call (or equivalent notification) and continues through a series of organizations through to delivery at a hospital. Information about an emergency incident accumulates across the process as indicated in the top portion of Figure 4.1.5 labeled "information types." The data and information instances that make up each "information type" may also change and increase in a dynamic manner as the service progresses across a series of system components and information technologies. For example, depending on the original accuracy of an incident report, location information could change or be added upon.

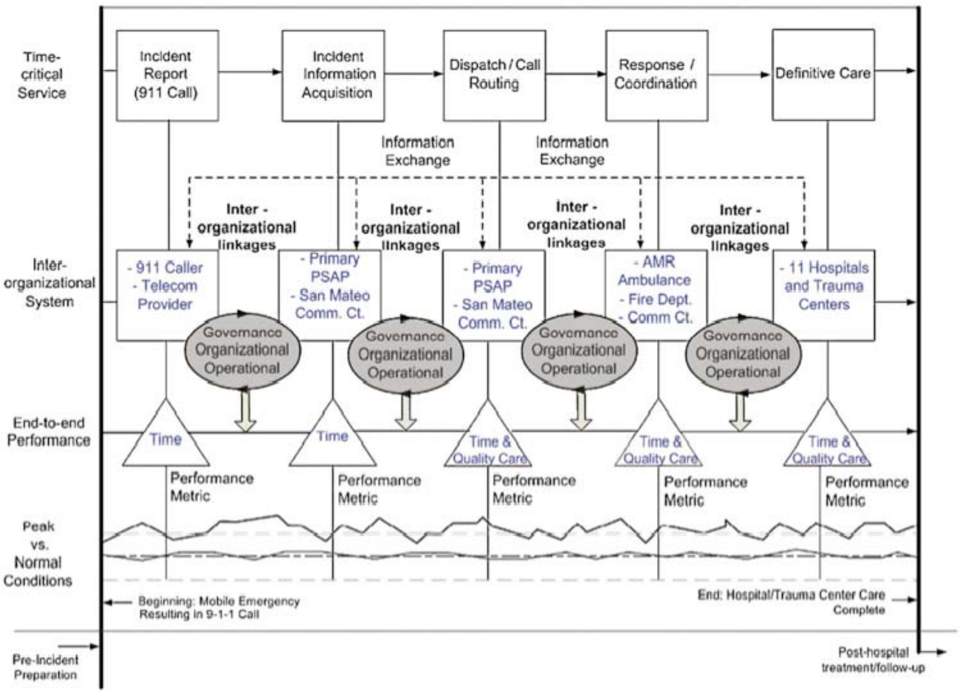
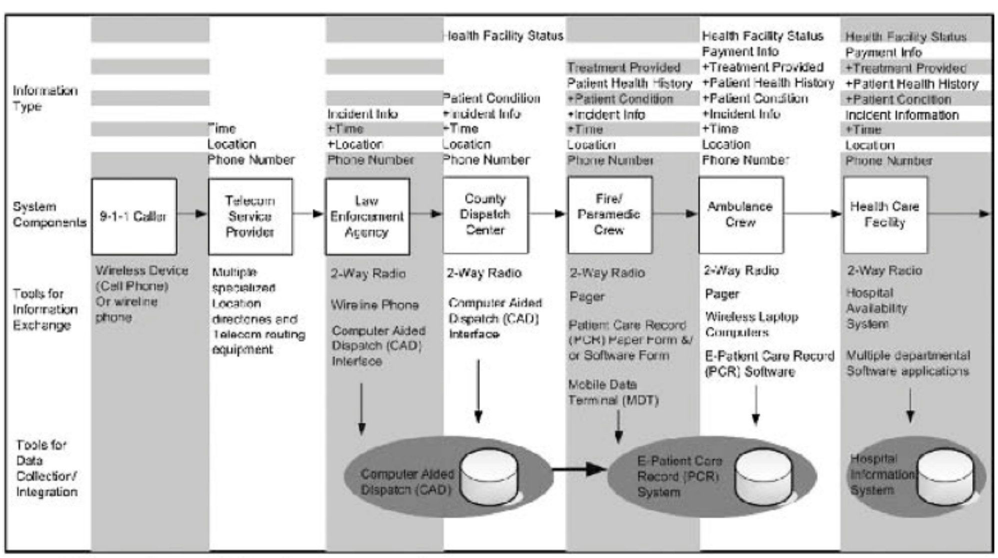


Figure 4.1.4 TCIS framework applied to San Mateo case study



+ Indicates an increase of information type across service process

Figure 4.1.5 San Mateo performance-related information exchanges across EMS system components

While the majority of information exchanges happen in a sequential manner, there are some data exchanges that occur in a nontraditional manner that is "out of sequence." For example, the hospital availability system information is sent from the hospitals to the communications center and then communicated to paramedics in the field. The diagram also illustrates the information systems used to collect, aggregate, and facilitate data exchange across organizations, such as two-way radios, wireless laptops, software interfaces, pagers, etc. Finally, the bottom of Figure 4.1.5 shows the three main systems for collecting, aggregating and analyzing data—the aforementioned PCR, CAD, and HIS systems. After these data (e.g., Figures 4.1.4 and 4.1.5) had been obtained, researchers then proceeded to investigate contextual issues and these are discussed in the next section.

4.1.7 Results

Operational, Organizational, Governance Dimensions

The following section discusses operational, organizational, and governance dimensions to interorganizational information exchange that were found through case study interviews and field observations. The section also discusses perceived performance implications of those information exchanges. The discussion below has been organized by each TCIS dimension, as well as each element of each dimension as outlined in Table 4.1.1 of this text. The case study examples illustrate the interrelated, parallel, and overlapping nature of these three dimensions across the EMS public service chain. While each example below is used to illustrate one specific dimension, it should be noted that many of the examples could also be used to demonstrate at least one other dimension. Table 4.1.4 below provides a summary overview of how the TCIS framework and Table 4.1.1 dimensions map to case study findings.

Operational Dimensions

A previous section provided a general operational description of the San Mateo EMS system. The following findings focus on operational dimensions of interorganizational information exchange. According to Table 4.1.1, these dimensions include technical systems, business processes, and communication flows that influence cross-organizational information sharing.

Technical systems. The San Mateo County EMS Agency has made significant efforts to collect and utilize incident data to manage service performance. Figure 4.1.5 illustrates the separate and disparate information systems that support EMS in the county (CAD, PCR, and HIS). An important "data type" collected across "pre-hospital" organizations (dispatch, fire, ambulance) is "time stamp" data, which enables the EMS Agency to monitor pre-hospital time-related performance. However, it was noted that very little data are aggregated (integrated) and shared across all organizations including the hospitals. Currently, a cross-organizational management information system to facilitate such a function exists in only a limited fashion. An EMS Agency representative stated:

> The missing link to making better county wide EMS decisions is to be able to tie 'pre-hospital' performance to hospital outcomes. But we don't have access to hospital data and we don't have a system that collects it.

Table 4.1.4 Summary of case study findings

TCIS dimension	Case study findings organized by Table 1 dimensions	Perceived performance implication
Operational	**Technical systems**	
	• Separate and disparate information systems across service organizations	Barrier to end-to-end integration/ performance analysis
	• Some integration between pre-hospital organizations (e.g., PCR and CAD)	Allows service segment monitoring, improves timeliness
	• Some automated data collection	Provides time benefits
	• Lack of pre-existing patient data for use in the field	Barrier to higher quality care decisions and improving timeliness
	Business processes	
	• Hospital availability data represents process change; data sent "out of sequence"	Improves quality of care decisions and transport time to hospital
	• Automated data collection allows focus on patient needs	Improves quality of medical care
	• Challenges "fitting" IT systems with dynamic EMS work flow practices	Barrier to improving quality care and timeliness
	Communications flows	
	• Many communications via voice and paper; not electronic; no data collected; new data systems not used	Barrier to assessing performance
Organizational	**Power relations**	
	• Contract holder mandates information sharing to monitor performance.	Enhances performance
	• Level of participation	
	• Each organization focuses on own individual performance as opposed to "end-to-end" focus	Barrier to end-to-end performance
	Cultural	
	• Difference between "patient" view of service and "service provider" view of service	

TCIS dimension	Case study findings organized by Table 1 dimensions	Perceived performance implication
	• Impacts how performance is perceived and measured	
	• Dynamic, time-critical work environment creates data collection challenges	Barrier to end-to-end performance analysis and improvement
	Resistance to change	
	• EMS professionals dependent and partial to voice communications	Barrier to end-to-end performance analysis
	• Technology wanted for screening 911 calls rather than for mobile health care	Impacts time and quality of care for some incidences
	Trust	
	• Paramedics perceive that some physicians trust their information more than others.	Paramedics say lack of trust degrades timeliness, physicians say it increases quality care
Governance	**Participant roles**	
	• EMS Agency role as contract holder mandates pre-hospital information sharing	Improves timeliness
	• EMS Agency role as collaborator with hospitals has limited influence on information sharing	Barrier to quality care performance improvement Rules and regulations
	• Data/information definitions change performance understanding	Appropriate definitions accurately portray performance
	Decision-making processes	
	• Challenges to enhancing decision processes with IT	Affects timeliness and quality care
	• Organizations cautious about changing processes	Protects EMS system from performance degradation
	Political/legal	
	• Privacy concerns negatively affect information sharing	Barrier to performance improvement
	• Legislation sought to mandate information sharing from hospitals	Improve performance analysis

Agency representatives explained that hospital time stamp data such as the time a patient is admitted, time medications are administered, time surgeries started and ended, and time patient is released from a hospital could be very useful for understanding how pre-hospital performance affects hospital outcomes. They discussed their belief that such an end-to-end view would enable better performance assessment and decision making, but that the current the lack of data sharing from hospitals prohibits such.

Several participants discussed the utility of information systems and data in the field. For example, fire EMT's noted how automated data collection had shortened the time required to write reports, essentially enabling them to be "on-call" for a greater portion of their work shift.

Yet, paramedics explained many difficulties related to assembling "patient history" information from friends and family members at an incident site. One paramedic explained:

> People just aren't very aware of the health information they need to know at the time of an emergency. They can't remember what they're allergic to or how to pronounce the medications they're on. They don't usually have a prescription bottle right in front of them. Family members try to dig up that information when we ask for it, but it just takes so long to do things that way.

Upon patient delivery, physicians at receiving hospitals had difficulty deciphering the prescription names written by paramedics. Participants explained that there currently exists no method to quickly assemble patient history information. Yet paramedics and physicians stated that when accurate patient information is collected, their ability to deliver appropriate medical care is greatly enhanced. One paramedic explained:

> …any information we can get about a patient—his condition, meds, history—is extremely valuable. It's a terrible situation to not know of a med that a patient is taking and then administer another med that has a negative interaction with it. You hope the worst that happens is a throw-up mess in the back of the ambulance.

Business Processes. The EMS Agency aims to improve interorganizational business processes. One specific area of focus is to automate data collection and transmission so that professionals in the field can focus more on delivering health care and less on manual processes. The hospital availability system depicted in Figure 4.1.5 is an example where data transmission does not follow the traditional service sequence. The data are sent from each of the 11 county hospitals and used by paramedics to determine which hospitals have the ability to accept new patients. One paramedic stated:

> Before the system we did a lot more talking on our radios to figure out which hospital had beds and doctors. I had times when I would show up to the ED [emergency department] and they would tell me to go to a different hospital. Granted those weren't life and death situations. But we had to deal with more driving which made us unavailable for another call, not to mention angry patients.

As alluded to above, participants noted how the process change created by information technology has created service improvements.

Participants discussed how automated data collection allows for business process changes that impact both timeliness and quality of care performance. As one emergency medical technician (EMT) described:

> I wish all of our time reporting could be automated. It may not seem like much of an effort for us to report departure and arrival [over radio], but when I know there is a bleeding child in a car, my heart is pounding and the only thing I'm interested in is getting to this kid; I just want to focus on getting there and treating her, not on reporting our time or procedures or whatever. In those really life threatening situations, we really don't record much until after transport anyways. I jot down notes on my hand [rubber glove].

In this sense, both timeliness and quality of care are impacted by how well business processes and information systems support medical care delivery.

Participant discussions revealed workflow challenges associated with the nature of health emergency work. While entering performance information is essential to reporting performance, participants felt that many of the technological tools for data collection were not suitable to workflow processes. For example, many paramedics do not use the handwritten "signature" function built into the touch-screen laptop for obtaining a patients' "consent to pay." One paramedic stated:

> The laptop is not very convenient to carry around and it's awkward to go into an emergency room, put the heavy machine on a patients lap, and ask him to sign the screen.

Many paramedics also do not bring the laptop with them to an emergency scene because of contamination issues. Participants agreed that tools designed to better "fit" the dynamic workflow environment of EMS could bring potential quality care and timeliness benefits. An EMS Agency representative confirmed:

> PDA's were a disaster when we implemented them for paramedics to use. Wireless connections were too slow and the screens were too small. But if we can find the right computer based tools for the people in the field, our ability to capture what they're doing will become easier, and our ability to improve the EMS system will be easier.

Communication Flows. Participants discussed challenges to improving the flow of information from one organization to another. One challenge discussed was how a large proportion of incident information is transmitted via voice or hand-written communications and not captured or transmitted via data systems. The EMS Agency Administrator discussed the Agency's attempts to address this issue with information technology. The Agency provided PCR system access to all county hospitals and trauma centers. However, an Agency representative stated:

> Hospital staff rarely use the system. They continue to rely on the traditional methods of receiving incoming patient reports. Basically, physicians rely on short voice 'snapshots' from paramedics in combination with paper reports.

The original purpose behind the system was to advance patient information to a physician prior to patient arrival. But as discussed previously, paramedics rarely enter patient information prior to patient arrival. In addition, physicians have no mobile means to look at the data (e.g., PDA, Tablet PC). One physician explained:

> Many IT consultants and other techies tell us that the technology exists to make our work paperless and reduce voice communications and make everything computerized. But that's not true. When a paramedic comes through the doors with a patient and describes the situation—that's valuable. We don't have voice systems that can capture those words and make them electronic. The technology to do that is still premature, despite what we're told.

Organizational Dimensions

The EMS Agency spends a significant amount of its time developing relationships with cooperating organizations at monthly Emergency Medical Care Committee meetings for the primary purpose of discussing, evaluating, and improving emergency care. The discussion below relates organizational issues to information sharing. As outlined in Table 4.1.1, organizational and Interorganizational dimensions of information sharing include trust, cultural and subcultural differences/similarities, effective communications, level of participation, power relations, and resistance to change.

Power Relations. As stated previously, time stamp data are used to analyze EMS and fire crew response times as outlined in the county EMS contract. Participants discussed how fire crews and ambulance crews monitor these segments in a "real-time" nature in order to help them meet contractual service obligations (see Table 4.1.3). One fire paramedic stated:

> We're watching the time very closely as we travel to an incident and we definitely feel pressure from our Chief to arrive before the 6-minute marker. And I'm sure he feels pressure as well.

This does not in any way imply that arriving quickly is not motivated by helping a patient. But the time element, combined with a contractual compliance marker provides a constant reminder to act quickly and avoid receiving fines for lack of compliance. An important organizational dimension to this phenomenon is the power relationship that the EMS Agency has to mandate both information sharing and performance levels for fire and ambulance organizations. And the EMS Agency Administrator unambiguously confirmed that the current performance contract system experienced improved response times over the previous countywide system.

The EMS Agency participants noted how their ability to access, aggregate, and share cross-organizational historical data grants them some degree of power to influence end-to-end cooperation, information sharing, and performance with service organizations. Participants described a monthly EMS Agency sponsored forum by which historical performance reports are shared across EMS organizations. In 2001 the EMS Agency released the first annual performance statistics report and distributed it to each emergency response organization. One element of the report displayed a list that showed response time averages for the year 2001 broken down by fire department and engine number. This was a significant source of both pride and embarrassments for the fire departments, depending on where they were located in the rankings. A fire department representative explained:

> The year after that first report, the fire department rankings completely changed. There's been kind of an ongoing competition for the best chute times since.

Participants used this example to explain how their power to share information influenced interorganizational relationships (competition), which served as a way for those organizations to focus on and improve performance. An EMS Agency representative stated:

> We [EMS Agency] want to gain greater access and control over county-wide EMS data to at least in part further influence how EMS organizations interact with each other and to help influence service performance.

Level of Participation. Upon acknowledging an understanding about "information-sharing" concepts, participants stated that organizations are not familiar with or trained to share or use data for the purpose of improving interorganizational, end-to-end performance. Rather, performance is a function of each individual organization. A communications dispatcher from the California Highway Patrol stated:

> Our job is to make sure that a medical related wireless 911 call is answered by an operator at a county PSAP. At that point, the call belongs to the county. We then dispatch our own officers and work to help them accomplish their job.

The level of participation in interorganizational information sharing across all county organizations is limited and, in general, does not include an "end-to-end" performance perspective. One EMS Agency representative stated:

> A clear understanding about service performance, an agreement from all of the organizations, is a very basic first-step to being able to improve the EMS system. The agreement has to be assembled first. Then the analysis can occur. And then changes can occur.

Cultural. Participants noted that there exists a gap between how service organizations view EMS and how customers/patients perceive the service. As an example, the EMS Agency frequently sends and receives surveys to assess customer/patient satisfaction with first response (fire) and ambulance transport services (i.e., pre-hospital services). However, many of the returned surveys contain written explanations about issues at the hospital emergency department or trauma center instead of issues relating to pre-hospital service. One EMS Agency representative explained:

> These types of survey responses represent something that is obvious and that we've been aware of for a long time. A customer's perspective of the service is that the 'pre-hospital' and hospital experience is an integrated, single event. Service evaluation is rarely conducted in such a manner.

A roundtable discussion with EMS Agency representatives revealed the general belief that a patient-centric, end-to-end evaluation would improve service performance as well as the performance perceptions of patients. But organizations have a long history of evaluating their own "leg" of a response. One ambulance provider representative stated:

> We talk about giving service in a patient focused way all the time. We believe in doing that. I don't think it's too difficult to imagine a patient's perspective. They don't really care about all the 'organizational lines,' they just want good service. But the unanswered question is, 'so what are we actually supposed to do about that?'… We are a different organization than the trauma center and the fire department. We are not one big organization and we aren't seamlessly integrated.

Researchers found cultural challenges to facilitating the input (and subsequent sharing) of patient care information. For example, the culture of the ambulance service is one where paramedics have a great deal of freedom to choose how they will work in the field, coupled with a strong sense of responsibility to provide meaningful service. Each paramedic has his/her own preferences as to how he/she uses the electronic Patient Care Record (PCR) system and there are loose organizational standards for data collection in the field. One paramedic elaborated:

> Some of us [paramedics] take the laptops with us to a scene and enter data, some start the PCR en route to the hospital, some enter the information after delivering a patient to the ED [emergency department] but before their next dispatch, and some wait all the way until the end of their work shift. Some write summary notes to shorten their reports and others, like myself, spend a lot of time writing a detailed report to cover their butts in case someone decides to sue.

Participants noted the cultural issue stems from the dynamic nature of their work. Information entry is often dictated by the differing contexts of each emergency incident including the often unknown status of a patient's health condition. One Fire department EMT stated:

> In very critical situations, we have to give complete attention to a patient. We can't have a strict protocol for how to enter data into a system. If we have a choice between entering PCR data and attending to a critically bleeding trauma patient, the choice is obvious. There was a wide range of opinions in terms of performance implications.

There was consensus that timely data entry could provide significant impacts for other professionals at different service legs, especially for physicians at a receiving hospital/trauma center. An EMS Agency representative stated:

From our perspective, it's clear that data entry and integration will allow for better performance evaluation. But if data entry causes a degradation of medical care, then the pursuit of data capture is in vain. We need to find a way to get the best of both worlds.

Resistance to Change. Participants discussed how organizations and people are often slow to change and adapt to new technologies. For example, and as shown in Figure 4.15, EMS responders are sent a text message to hand-held pagers to notify and dispatch them to an incident. First-responders push buttons on a mobile data terminal (MDT) that notifies the dispatch center that they have arrived on-scene. Wireless laptops are used by paramedics to input incident and patient information, which is automatically updated on hospital terminals for physicians to view. Despite these new data powered systems, voice over two-way radio is most often used to transmit short snapshots of much of that same information. A communications center supervisor explained

> EMS, fire, and law enforcement are all very biased towards using two-way radio communications. I don't think radios will ever be completely replaced by computers.

Upon arrival at the hospital, a paramedic provides a verbal thumbnail sketch about the patient's health condition to the receiving nurse or physician. A hospital administrator stated:

Even if the PCR has been entered, physicians rarely read it. They expect the paramedic to tell it to them in person. This illustrates the strong propensity towards traditional workflow processes, which challenges the ability to enter, aggregate, and integrate data for end-to-end performance analysis. Several participants did not feel that new technology to help them work faster and more efficiently should be a primary focus area. They discussed how many of their calls are not "time critical," not requiring a high priority ambulance response. One paramedic stated:

> Time is an important issue when it comes to trauma, cardiac arrest, stroke, and major issues. But most of our calls are not those types of issues. So often, we are dealing with people who are just feeling sick with the flu, people who don't have a car and want a ride to the hospital, people faking it to get attention, or some other minor thing. 911is really not used appropriately. People still just don't know when they're supposed to use it when not to.

While participants discussed the importance of time savings during highly time-critical circumstances, they felt that the technology focus needed to address other issues, such as the screening of calls to better differentiate between the critical and noncritical ones. But an ambulance provider representative noted:

> Paramedics aren't running to a scene like you see on TV, except when they really know that the situation is really life and death critical. Does it impact time? Sure. Does it impact a patient? Probably in some circumstances. The fact that they don't know what to expect on scene means that they make the statistical assumption that the situation is not life and death.

Trust. Organizational issues included the lack of trust that often exists between individuals from cooperating organizations. EMS Agency participants stated that in order for information to affect system performance, information must be trusted, which includes trusting the persons delivering the information. A communications center supervisor provided an example:

> Most of the time that CHP [California Highway Patrol] transfers an emergency call to us, the incident location is wrong. They even have e-911 data [location data] from cell phones. But we usually know that it is going to take us longer to respond to an emergency call from CHP and that we are going to have to use our resources to find the correct location.

In this case, the communications center does not trust information coming from the CHP and there are response time implications. Paramedics described how some hospital physicians appreciate paramedic comments more than others. Physicians at the more prestigious health care facilities tend to be less interested in paramedic comments while the less prestigious facilities spend more time discussing the patient's condition with the paramedic. One paramedic explained:

> It's a matter of physician's not trusting the information that we [paramedics] give them. Doctor's have way too much pride to believe everything we have to say.

Performance implications of cross-organizational distrust varied. Paramedics argued that trusting paramedic information could enhance service performance in highly time-critical situations. However, a physician argued that doctors have far more education, knowledge, and experience and that their disposition to question paramedic diagnoses/impressions actually helps to ensure better quality of care.

Governance Dimensions

The EMS Agency holds significant power in its relationship with the ambulance provider and the first-responder organizations due to the county EMS contract and therefore has the ability to initiate and implement integration initiatives. But participants noted that its governing power is far more limited in respect to the health care facilities in the county. In this regard, there are separate governance structures between two very critical links in the EMS process. The following discussion relates governance dimensions of information sharing including participant roles, legal definitions, policies, and rules and regulations.

Participant roles. EMS Agency representatives discussed how governing interorganizational relationships was a significant challenge at the beginning of the county EMS contract. In particular, it was challenging to define each organization's role in terms of information sharing. To help overcome this challenge, the EMS Agency holds monthly committee meetings with representatives from many county emergency organizations. A recent meeting consisted of representatives from over 16 separate organizations. An EMS Agency representative stated:

> Trying to define who shares what information and when has been an ongoing challenge due to the number of organizations involved and the logistics of coordinating between them.

Yet the function of the master EMS contract is to define roles and responsibilities. The contract states that first responder crews and ambulances must respond to an incident within the specified time frame for at least 90% of incidences (see Table 4.1.3 above). The contract also mandates the collection and transfer of "time stamp" data to the EMS Agency. The Agency Administrator spoke of an instance where ambulance response times were not in compliance with the contractual agreement. She stated:

> When we saw the report and response times were not in compliance we immediately did some research. Our inquiry uncovered that AMR had actually reduced the number of 'on-duty' ambulances. We fined AMR according to contract, they put another ambulance in the field, and response times went back to the appropriate levels the following quarter.

Participants used this example to illustrate their participant role—as an oversight organization whose goal is to manage and improve the end-to-end service performance.

They have been able to effectively leverage their role to influence interorganizational information sharing and improve performance, which they continue to do with every contract renewal (which is up for renewal in 2007). While the master EMS contract serves as a defining document for structuring the roles of pre-hospital organizations, the Agency's role in regards to county hospitals is far less defined. An EMS Agency representative explained

> We've been working with the hospitals for a long time and we just don't have much of an ability to get data from them. We basically have to try and get them to voluntarily share information with us. I've been looking for ways to get local and county governments to play a stronger legislative role, but right now our oversight role is rather limited.

The EMS Agency's governing role rests with their ability to develop relationships of trust with hospitals as opposed to a set of formal rules and regulations that mandate information sharing at the county level. The Agency Medical Director stated:

A stronger role, including funding and additional access to information, would allow us to better assess EMS performance and provide better emergency medical care guidance for the county.

Rules and Regulations. Participants discussed how performance results are also influenced by how performance information is defined, or the rules and definitions that describe the information. For example, for the County Communications Center, a "call" is not a "call" until it is answered. And according to the EMS Agency master EMS contract, "response time" starts when a dispatch message is sent to a crew and ends when the emergency vehicle parks at the closest practical point to the emergency scene. An EMS Agency Analyst explained

> This time segment measurement stops when a vehicle parks. It does not include any additional travel to the scene-up to the fourth floor of a building, across a football field, or inside a crumpled car that we can't open.

Participants explained how these rules and definitions are paramount when sharing information across organizations, creating service level agreements (SLA's), when comparing performance information from one system to another, or when comparing performance over long periods of time. The Agency Analyst explained

> One definition change could at first glance look as though performance improves or degrades when in reality it was only the definition that changed.

The EMS Agency typically takes very careful measures to define terms when drafting new or renewed EMS contracts so as to ensure performance improvement rather than degradation.

Decision-making processes Participants discussed how interorganizational decision-making processes need to be understood and integrated into technical information-sharing systems. An example was provided in relation to the aforementioned hospital availability system. A "divert" code is used to signal long service wait times, a lack of hospital beds, or an insufficient number of hospital staff. The system has significant performance implications as it may signal longer ambulance travel times for a patient, but could also direct a patient to the most appropriate surgeon at a distant facility. A hospital administrator explained a flaw in the system:

> Divert codes are not always accurate… It's pretty common for a physician with a specialty, for example a neural surgeon, to be "on-call" for more than one hospital. But whoever made the [hospital availability] system assumed that a surgeon can only be on call for one hospital at one time. So one out of the two hospitals gets a "divert" code, even though the surgeon could be there within minutes. This is misleading for the transport teams. The EMS system decision-making process is that the first of the two hospitals to request assistance from the on-call doctor would win her services. These rules were not accurately portrayed in the technical system. Participants discussed how these types of failures in interorganizational information sharing have obvious performance implications in terms of timeliness and quality of care. Participants noted that EMS organizations tend to take a significant amount of caution before adapting or changing decision processes, largely due to IT experiences similar to the one mentioned above. An EMS Agency representative explained

We are very cautious and usually pretty slow to change at least in part because of the types of risks involved. We can't have computer systems that cause confusion or otherwise don't improve performance.

Political/legal. Another governance issue relates to political and legal concerns with data privacy. The EMS Agency performs monthly quality improvement exercises by conducting in-depth analyses on a few randomly chosen emergency incidences. An EMS Agency representative explained

> We randomly select a few EMS incidences and present them as case studies at monthly Emergency Medical Care Committee Meetings. We are dealing with actual emergencies, so we have to take care to eliminate any personally identifiable information. These case studies are extremely useful because we look at the emergency all the way through from the Communications Center to the trauma center.

The case studies are presented and analyzed at the meeting by attendees from all EMS organizations. However, the hospitals are concerned that if quality service analysis were conducted on a large scale (for most or all incidences), the probability of violating HIPAA (data privacy) regulations would increase, and would subsequently increase their risk of political and/or legal action. From the perspective of one hospital Administrator:

> The HIPAA, privacy, socio-political aspects really act as a significant deterrent for health care providers to release any patient information to anyone. Partially because HIPAA can be confusing and was not well constructed in the first place. But also because of the bad publicity and litigation that could accompany violations.

The primary performance implications of not having hospital outcome data are that pre-hospital data cannot be tied to hospital outcomes, making end-to-end performance analysis incomplete. In a related sense, policy and legal dimensions to information exchange have been pursued by the EMS Agency. Trauma centers are mandated to share hospital admissions data at the state level but not at the county level. The EMS Agency has made failed attempts to access these hospital admissions data sets from county hospitals. However, one EMS Agency representative stated:

> The trauma centers don't have much incentive to give us that data. We're trying to make the case that a combined pre-hospital and hospital data set will help improve their services. It's not so much that they want to withhold data, but more of a situation where it's a low priority for them. Basically, they'll give us the data if they are mandated to do so.

The EMS Agency seeks legislation to obtain the data before it is integrated and aggregated into the state data set. They believe the performance implications will be that end-to-end performance analysis will be made possible if granted access to the data. The Administrator explained:

> My colleague [the County EMS Administrator] in San Diego pursued legislation to get that data and I am told that it worked. She says it has been very beneficial.

4.1.8 Discussion

Returning to the research questions posed at the beginning of the paper, the case study findings indicate that information systems are used to manage the performance of individual organizations and their respective separate segments of an emergency incident. Information systems are also used in a partial manner to capture and analyze time stamp data across multiple organizations, which in this case includes primarily the "pre-hospital" organizations. An information system to integrate and manage end-to-end data, including data from the hospitals and trauma centers, is a goal yet to be realized. Information systems are used in an even more limited fashion to capture and analyze "quality of care" data across service organizations to assess and manage end-to-end quality of care performance. This San Mateo case exemplifies this phenomenon with its multiple disparate information systems (CAD, PCR, HIS), which integrate on a very limited basis. Quality of care data analyses can be constructed on a case-by-case basis as demonstrated at monthly Emergency Medical Care Committee meetings. This is promising evidence that end-to-end quality of care performance management could occur on a larger scale in the future. Participants noted the performance value achieved through current integration across PCR and CAD systems. For example, paramedics are able to view limited demographic and medical emergency "type" data prior to arriving on scene, which allows them to make important medical preparations (e.g., bring the right equipment). These disparate examples and participant responses indicate that a management information system that acts as a "facilitator" to link information systems and organizations could be a valuable enabler of endto- end performance analysis and improvement. But the case study analyzed here indicates that there is much work to be done before such a goal is realized.

Issues and challenges to interorganizational information integration range across operational, organizational, and governance dimensions. In this regard, and in response to the research question, these three dimensions influence to what degree information is shared or IT is deployed across organizations from end-to-end. In terms of the *operational* dimensions, the dynamic, complex, time-critical, and multivariable nature of emergency and trauma care work creates a number of technology usability issues and challenges, which cause emergency professionals to take a conservative approach to new information and technologies to support EMS. For example, inputting data is the last thing on a paramedics mind in a life and death situation. To avoid the risk of service degradation, liability, and negative patient health consequences, emergency professionals explain the need to be cautious in the application of new IT to support emergency medical work processes.

In terms of *organizational* dimensions, the case study provides support for the thesis that interorganizational alignment including a shared set of goals and cooperative agreements can facilitate information sharing. The San Mateo example illustrates the value of working cooperatively at monthly meetings to create a shared set of goals across organizations to enhancing information sharing. In contrast, issues of trust and cultural differences create interorganizational information sharing and technology gaps. The apparent "pre-hospital" vs. "hospital" information-sharing chasm illustrates this phenomenon. In terms of governance dimensions, case study examples illustrate that clear lines of interorganizational authority and accountability tend to enhance information sharing and technology. The San Mateo performance-based contract defined clear authoritative boundaries and information sharing was enforced. In contrast, unclear lines of authority and accountability tend to limit information sharing and technology—such was the case between the EMS Agency and the 11 county hospitals.

In sum, the case study highlights how information integration holds the potential to significantly affect timeliness and quality of care performance for end-to-end e-government EMS services, especially where careful and deliberate IT enabled interorganizational business process changes occur. The hospital availability system illustrates the impacts of this phenomenon where information sharing allowed paramedics to transport patients to an available and appropriate hospital, essentially improving both timeliness and care delivery to patients.

Methods for Evaluating Interorganizational Information Integration (III) Initiatives

It is perhaps best to sum up by returning to where the article began—the overall conceptual TCIS framework. Figure 4.1.4shows the aforementioned time-critical information services (TCIS) framework as it is instantiated in the San Mateo case study. The use of TCIS illustrated the need to understand interorganizational workflows, partnership arrangements, organizational structures, governance structures, and organizational policies as a valuable step to improving end-to-end service performance. That is, this case study illustrates how a heuristic model derived from socio-technical frameworks can be used to understand a range of issues when considering an interorganizational information integration initiative.

4.1.9 Conclusion

As exemplified by this special edition, there is a pressing need to understand and improve interorganizational information integration for e-government. The contribution of this case study (and related concepts) is to raise the issue of a certain type of information—performance information—and the integration therein. The research aims to contribute to a nascent literature of e-governmental performance (Pardo, 2000; Stowers, 2004; West, 2005). But it does so by not asking about the isolated performance of new systems, but rather, how can new systems be implemented in such a manner so as to identify and enhance the performance of the entire end-to-end system. More specifically, it is aimed to contribute to the development of conceptual and empirically based principles and guidelines that can improve understanding about the role of information systems in integrating performance data in the management and governance of time-critical interorganizational public services. While there has been much discussion in business industry literature about business intelligence (BI) and business performance management (BPM) (Frolich & Ariyachandra, 2006; Miranda, 2004), this paper illustrates potential benefits to government performance management (GPM). From an applied perspective, the goal is to provide EMS organizations with tools to aid in the development of EMS information systems to enhance service performance, reduce disability consequences, and save lives.

Additional research is needed. During case study interviews, it became apparent that “quality of care” measurement is challenging to address. It includes numerous variables—and measurement differs across EMS systems. There is a need to better understand how to measure quality of care and how information systems can best extract such data in highly complex, dynamic, time-critical health care delivery environments.

To address this need, future studies should take at least three distinct approaches. One line of research would utilize qualitative methods to assemble and construct in-depth user profiles of paramedics and physicians using IT (e.g., PCR) systems in the field. The case study interviews revealed the challenges in entering, using and sharing information when operating under time information critical circumstances, such as during traumatic events. This effort would focus on extracting detailed context of health care work practices and processes to inform the design and development of more appropriate IT systems for these environments. Another research focus would investigate how an architectural view could be utilized in understanding end-to-end performance. This would entail the construction of "quality of care" performance architecture for end-to-end time critical information services. It would draw upon multiple existing system architectures such as the National Intelligent Transportation Systems (ITS) architecture, the NextGeneration 911 Architecture that is currently under development, and Comcare's E-Safety architecture. It would then seek to validate the architecture through multiple case studies and expert workshops with academics and EMS practitioners. A final approach would utilize performance data from case study entities to perform a series of computer-based simulations to observe EMS operations under normal and extreme (service load) conditions. The focus of the simulations would be to explore how information at varying points in service delivery could impact end-to-end performance. Each of these research areas address the growing importance and need to focus on the role of interorganizational information integration in improving end-to-end public sector services.

4.1.10 References

1. Allen, D., References, Colligan, D., Finnie, A.,& Kern, T. (2000). Trust, power, and interorganizational informationsystems: The case of the electronic trading community translease. *Information Systems Journal,* 10, 21–40.
2. Arens, Y., & Rosenbloom, P. (2002, March 1). Responding to the unexpected. Report of the Workshop. New York,NY.
3. Bakos, J. Y. (1991). Information links and electronic marketplaces: The role of inter-organizational information systems in vertical markets. *Journal of Management Information Systems,* 8(2), 31–52.
4. Cash, J. I. J., & Kosynski, B. R. (1985). IS redraws competitive boundaries. *Harvard Business Review*, 63(2), 134–147.
5. Coffey, D. (1998). Self-organization, complexity, and chaos: The new biology of medicine. *Nature Medicine,* 4(8),882–885.
6. Curits, B., Kellner, M., & Over, J. (1992). Process modeling. Communications of the ACM, 35(9), 75–90.
7. Dawes, S. (1996). Interagency information sharing: Expected benefits, manageable risks. *Journal of PolicyAnalysis and Management,* 15, 377–394.
8. Dawes, S., Cressell, A., & Cahan, B. (2004). Learning from crisis: Lessons in human and information infrastructurefrom the World Trade Center response. *Social Science Computer Review,* 22(1), 52–66.
9. Dawes, S., & Prefontaine, L. (2003). Understanding new models of collaboration for delivering governmentservices. *Communications of the ACM,* 46(1), 40–43.
10. Drake, D., Steckler, N., & Koch, M. (2004). Information sharing in and across government agencies. *Social Science*
11. Computer Review, 22(1), 67–84.
12. Fedorowicz, J., Gelinas, U., Grogan, J., Howard, M., Markus, L., Usoff, C., & Vidgen, R. (2005). Modelingphysical barriers to interorganizational system implementation success. Working Paper from Bentley CollegeDepartment of Information Systems.
13. Fountain, J. (2001). Building the virtual state*: Information technology and institutional change.* Washington, DC:Brookings Institution Press.
14. Frolich, M., & Ariyachandra, T. (2006). Business performance management: One truth. *Information SystemsManagement,* 23(1), 41–48.
15. Grossman, M. (2004). The role of trust and collaboration in the Internet-enabled supply chain. *Journal of AmericanAcademy of Business,* 5(1/2), 391–396.
16. Guglar, P., & Dunning, J. (1993). Technology based cross-border alliances. In R. Culpan (Ed.), *Multinational strategic alliances.* Binghamton, NY: Howarth Press, Inc.

17. Health Care Financing Administration. (2000). Quality of Care-PRO Priorities. National Clinical Topics (Task 1).Online. Available at http://www.hcfa.gov/quality/11a.htm

18. Holsapple, C., & Luo, W. (1996). A framework for studying computer support of organizational infrastructure.*Information and Management,* 31(1), 13–24.

19. Horan, T., & Schooley, B. (2007). Time-critical information services. *Communications of the ACM* (Accepted andforthcoming).

20. Horan, T., Marich, M., & Schooley, B. (2006). Time-critical information services: Analysis and workshop findingson technology, organizational, and policy dimensions to emergency response and related E-governmentalservices. Proceedings of the 2006 International Conference on Digital Government Research, ACMInternational Conference Proceeding Series, Vol. 151, June, 2006.

21. Horan, T., McCabe, D., Burkhard, R., & Schooley, B. (2005). Performance information systems for emergencyresponse: Field examination and simulation of end-to-end rural response systems. *Journal of HomelandSecurity and Emergency Management,* 2(1), 4.

22. Institute of Medicine (IOM). (2001). Crossing the quality chasm: *A new health system for the 21st century.*Washington, DC: National Academy Press.

23. Institute of Medicine (IOM). (2006). *Emergency medical services: At the crossroads.* Washington, DC: NationalAcademy Press.

24. Johnston, H. R., & Carrico, S. R. (1988). Developing capabilities to use information strategically. *MIS Quarterly,*12, 37–47.

25. Johnston, H. R., & Vitale, M. R. (1988). Creating competitive advantage with interorganizational information. *MISQuarterly,* 12(2), 153–164.

26. Kapucu, N. (2006). Interagency communication networks during emergencies: Boundary spanners in multiagencycoordination. *American Review of Public Administration,* 36(2), 207.

27. Landsbergen, D., & Wolken, G. (2001). Realizing the promise: Government information systems and the fourthgeneration of information technology. *Public Administration Review,* 61, 206–218.

28. Layne, K., & Lee, J. (2001). Developing fully functional E-government: A four stage model. *GovernmentInformation Quarterly,* 18(2), 122–136.

29. Mayer-Schonberger (2003). Emergency communications: The quest for interoperability in the United States andEurope. In A. Howitt, & R. Pangi (Eds.), *Countering terrorism: Dimensions of preparedness.* Cambridge, MA:MIT Press.

30. McLean, S., Maio, R., Spaite, D., & Garrison, H. (2002). Emergency medical services outcomes research:Evaluating the effectiveness of prehospital care. *Prehospital Emergency Care,* 6(2), 172–179.

31. Miranda, S. (2004, March/April). Beyond BI: benefiting from corporate performance management solutions.*Financial,* 58–61.

32. Monge, P. R., & Contractor, N. (1988). Communication networks: Measurement techniques. In C. H. Tardy (Ed.),*A handbook for the study of human communication* (pp. 107–138). Norwood, NJ: Ablex.

33. National Emergency Number Association (NENA). (2001). Report card to the Nation: The effectiveness,accessibility and future of America's 911 service. Columbus, Ohio: The RCN Commission and NationalEmergency Number Association. Retrieved on Sept. 15, 2005 from http://www.nena.org/Initiatives/RCN/RCN_FINAL_2–24.PDF

34. National Highway Transportation Safety Administration (NHTSA). (1996). *Emergency medical services agendafor the future.* Washington, DC: DOT.

35. National Highway Transportation Safety Administration (NHTSA). (2001). *National EMS research agenda.*Washington, DC: DOT.

36. Pardo, T. (2000). Realizing the promise of digital government: It's more than building a web site. Center fortechnology in government. Albany, NY: University at Albany, SUNY.

37. Provan, K. G., & Milward, H. B. (2001). Do networks really work? A framework for evaluating public-sectororganizational networks. *Public Administration Review,* 61(4), 414–423.

38. Rudman, W., Clarke, R., & Metzel, J. (2003, July). Emergency responders: Drastically underfunded, dangerouslyunprepared. Report of an Independent Task Force sponsored by the Council on Foreign Relations. Retrieved onSept. 15, 2005 from: http://www.cfr.org/content/publications/attachments/Responders_TF.pdf

39. Sawyer, S., Reagor, S., Tyworth, M., & Thomas, J. (2005). From response to foresight: Managing knowledgeand integrated criminal justice. In S. Newell & R. Galliers (Eds.), Proceedings of the 2005 OrganizationalLearning and knowledge Capabilities Conference, Cambridge, MA, 17–19 March, 2005, Bentley College.

40. Sawyer, S., Tapia, A., Pesheck, L., & Davenport, J. (2004). Mobility and the first responder. *Communications of theACM,* 47(3), 61–64.

41. Sayre, M., White, L., Brown, L., & McHenry, S. (2003). National EMS research agenda: Proceedings of theimplementation symposium. ˆ 10(10), 1100–1108.

42. Scholl, J. (2005). Interoperability in e-Government: More than just smart middleware. Paper presented at the 38^{th} Hawaiian Conference of System Sciences (HICSS38), January 2005.

43. Shockley-Zalabak, P., Ellis, K., & Winograd, G. (2000). Organizational trust: What it means, why it matters. *Organization Development Journal,* 18(4), 35–48.

44. Spaite, D., Criss, E., Valenzuela, T., & Guisto, J. (1995). Emergency medical service systems research: Problems ofthe past, challenges of the future. *Annals of Emergency Medicine,* 26(2), 146-152.

45. Stowers, G. N. L. (2004). *Measuring the performance of E-government.* Washington, DC: IBM Center for theBusiness of Government.

46. Sundberg, H., & Sandberg, K. (2006). Towards e-government: A survey of problems in organisational processes.*Business Process Management Journal,* 12(2), 149–162.

47. Sussman, J. (2002). Representing the transportation/environmental system in Mexico City as a CLIOS. Paperpresented at the 5th Annual U.S.-Mexico Workshop on Air Quality, Ixtapan de la Sal, Mexico.

48. Turoff, M., Chumer, M., Van de Walle, B., & Yao, X. (2004). The Design of a Dynamic Emergency ResponseManagement Information System (DERMIS). *Journal of Information Technology Theory and Application (*accepted 12/25/03 and forthcoming).

49. West, D. M. (2005). Digital government: *Technology and public sector performance.* Princeton, NJ: PrincetonUniversity Press.

50. Wigand, R. T. (1988). Communication network analysis: History and overview. In G. Goldhaber & G. Barnett (Ed.), *Handbook of organizational communication* (pp. 319–359). New York: Ablex.

51. Williams, C., & Fedorowicz, J. (2005). A framework for analyzing cross-boundary e-government projects: TheCapWin example. DG.O 2005, 313-314.

52. Zhang, J., Dawes, S., & Sarkis, J. (2005). Exploring stakeholders' expectations of the benefits and barriers ofe-government knowledge sharing. *Journal of Enterprise Information Management,* 18(5/6), 548–567.

4.2 Devising a Web-Based Ontology for Emerging Wireless Systems

The Case of Emergency Management Systems

Thomas A. Horan, Ph.D.
School of Information Science
Claremont Graduate University
Claremont, CA 91711
E-mail: Tom.Horan@cgu.edu

Ugur Kaplancali, MSMIS
School of Information Science
Claremont Graduate University
Claremont, CA 91711
E-mail: Ugur.Kaplancali@cgu.edu

Benjamin Schooley, MIS
School of Information Science
Claremont Graduate University
Claremont, CA 91711
E-mail: Ben.Schooley@cgu.edu

4.2.1 Introduction

Ontologies are of increasing interest to Information Science researchers and professionals (McGinnius, 2002). This interest stems from both their conceptual use of organizing information and their practical use in communicating about system characteristics (Juriska et al. 1999). Many ontological frameworks have already been developed by academic disciplines such as computer science and bio-informatics and applied to a broad variety of businesses from high-tech industries to agricultural sectors (Noy et al. 2000). Within the field of IS, attention to "ontology driven information systems" is now on the rise because the concept of ontology promises a framework for communicating among architectures and domain areas.

In general, ontology refers to explicit specification of a conceptualization. Ontology development and use of supporting tools offer an opportunity to utilize a unifying framework that embodies objects and concepts, their definitions and relationships between them (Gruber 1993). Ontologies also make representative content available for knowledge sharing-a set of "consistent vocabularies and world representations necessary for clear communication within knowledge domains" (Leroy et al. 1999). Three main uses of ontology are for communication, for computational inference and for reuse (and organization) of knowledge (Gruninger and Lee, 2002).

This study is motivated by all three uses of ontology in the course of developing a software-based ontology-driven system to tackle the complexity of emerging wireless EMS. This research takes an inductive approach to ontology system development and applies it within a framework to clarify the domain's (wireless EMS) structure of knowledge. Moreover, it aims to contribute to the use of ontologies for determining and achieving high-quality data relative to system attributes and functioning. As Wand and Wang (1996) note, "to design information systems that deliver high-quality data, the notion of data quality must be understood. An ontologically based approach to defining data may be the ticket of success in real world systems (Wand and Wang, 1996, p.88)."

Emergency Management Systems

Wireless EMS in rural areas is the specific domain of our research. There is increasing pressure to use this emerging system (EMS Wireless Mayday) for medical emergencies, yet little is known about its functionality and performance dimensions. Conditions driving the problem include rapid growth of cellular phone use for mayday, strong policy interest in "first-responder" Mayday as a consequence of September 11, and policy regulations toward enhanced 911 (e-911) capabilities throughout the US. Statistics about system growth document this rise: Wireless 911 calls have grown from 22,000 per day in 1991 to 155,000 per day in 2001, and in many regions wireless calls represent the major form of emergency notification (CTIA, 2002). In short, the mobile (cellular) phone has become the de facto safety lifeline, particularly for mobile travelers and especially in rural areas.

While there are several policy, market, and technological pressures leading to emergency management system growth, the full system is quite dynamic and still unfolding—hence, it is not very well understood. Drawing upon the findings from preliminary fieldwork, the authors have identified technical, organizational, and policy dimensions to wireless EMS systems. In this paper, this framework is advanced by specifying its details within the parameters of a web-based ontology. The process of creating this ontology is the subject of this paper.

4.2.2 Method

Research Approach

The ontology development methodology used falls into the category of *Inductive Approach* (Holsapple and Joshi, 2002). Development techniques ascribing to this approach require observing, examining, and analyzing a specific case in the domain in a non-static fashion. As mentioned by Holsapple and Joshi (2002, p. 44) "The resulting ontological characterization for a specific case is applied to other cases in the same domain." The *Inductive approach* to ontology design fits perfectly to our purpose of using, developing, and reusing the ontology in spiral fashion as to validate the conceptual framework.

This inductive approach is particularly appropriate because of the emergent nature of wireless EMS services—that is, the system is growing rapidly and very dynamically due to a number of market and technology considerations. As advanced by Markus, Majchrzak, and Gasser (2002) such a context lends itself to a design theory approach whereby the system is captured at a point in time, whilst its eventual functioning may be undetermined.

Knowing this, it is our intention to focus on wireless emergency response management implementation with particular attention to "on the ground" technical and "non-technical" dimensions. In this case, the "on the ground dimension" was the context of rural deployment in Minnesota. Minnesota has a distinctive approach on delivering emergency services to rural areas as compared to other states. The wireless EMS is not limited to e-911 infrastructure where Public Service Answering Points (PSAPs) play the key role. Nine centers were established called TOCCs (Transportation Operation Communications Center), in different counties to aid emergency response agencies and incident management dispatches. Our fieldwork involved in-depth field interviews and site visits in Minnesota while concurrently developing subsequent versions of the ontology. The findings and concepts from these field reviews were then integrated into the ontological software that includes its knowledgebase populated with collected data.

EMS Ontology Development Method

The ontological-development task is to analyze this framework and the cases it was derived from using a platform-independent ontological software product. This process is summarized in Figure 4.2.1and includes the following steps:

1. Initial Interviews and Field Visits (Case study 1 input to develop socio-technical framework for EMS)
2. Develop the first version of the ontology.
3. Apply Ontology to second case study, including gather performance metrics and analyzing data using project software/knowledge base.
4. Develop new versions of the ontology by revising the data collected.
5. Publish ontological knowledgebase in the web for online collaboration with other EMS/e-911 efforts.

This analysis also addresses how to measure and communicate EMS performance by using an ontology and knowledgebase designed for collaborative analysis and communication. Thus the availability of solid wireless EMS measurements and metrics for measuring performance is important. In order to define and classify quantitative performance metrics previous EMS data collected by TOCCs and EMS agencies will be used. Examining such data could tell us:
How have response times changed over time?
Where are the weak link(s) in the process?
How will TOCCs deal with new technology?
How scalable is the system in terms of capacity to manage rapid growth?
What are the appropriate performance metrics and how can these be captured in an ontological framework?

These steps aim to execute a preliminary round of knowledge acquisition effort for the domain of EMS in rural areas. At the end of the ontology and knowledgebase development process our goal is to create the domain specific knowledge for rural Minnesota's EMS/e-911. It can then be shared with other EMS/e-911 initiatives in other states, which should further evolve and enrich the ontology. To accomplish this research objective, Protégé 2000 was adopted as a tool for devising the EMS ontology.

Adoption of Protégé 2000

The EMS ontology was developed using Protégé knowledge acquisition software. Protégé 2000 is developed by Stanford University's Medical Informatics Group as an ontology editor and knowledgebase editor (Grosso, et al. 1999). It is a java-based, platform-independent tool for developing ontologies and knowledgebases. Protégé 2000 fit our objectives for the following reasons:
It is platform independent.
It has a user friendly GUI.
Simple implementation (almost no code writing is required).
Ontologies developed by Protégé 2000 can be published in the Web.
Instances for classes can be stored and retrieved easily.
It is scalable.
Adequate technical support is provided online.
Plug-ins are available for a variety of purposes.

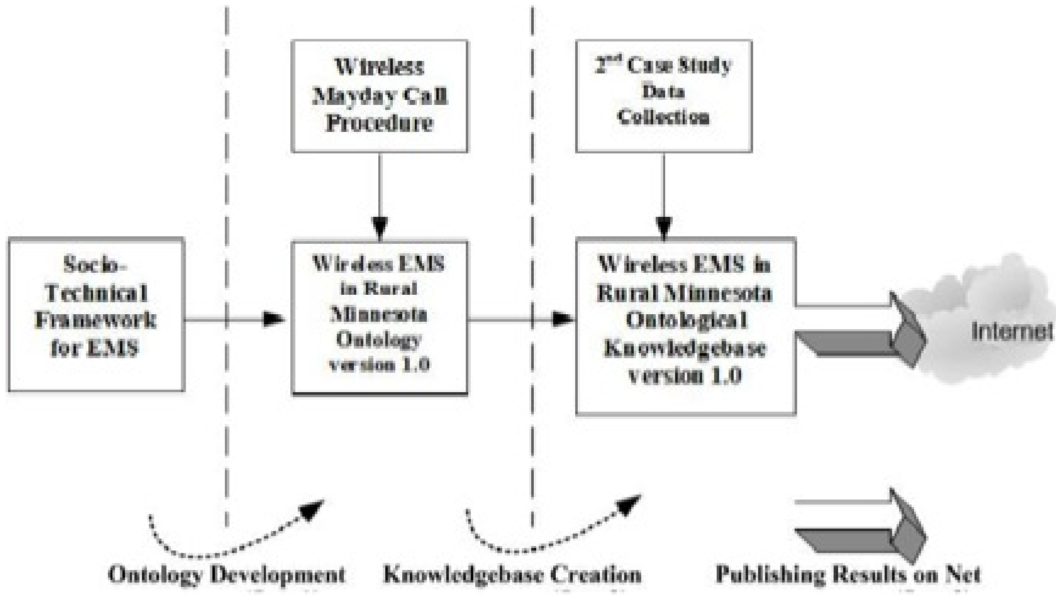


Figure 4.2.1 Research approach to wireless EMS and ontology

It has been proven in a research environment for several years.
The software is under continuous development and improvement.

In short, Protégé 2000 is generally well equipped to portray and organize the ontology for EMS services in a visually oriented and structured manner. Moreover, web publishing of the outcome in the form of an ontology and knowledgebase will increase the accessibility to the domain knowledge. Protégé capabilities in this regard would allow researchers to browse EMS ontologies and knowledge bases rather than scanning hundreds of pages of technical consultancy papers and documents to quickly find and navigate domain specific knowledge.

4.2.3 Findings to Date

To date, we have completed the case study for use in devising the ontology, developed a version 1.0 of the EMS ontology using Protégé 2000, and are currently in the process of organizing the second case study to apply the ontology and create a knowledgebase. The findings below are based on these results to date.

A Socio-technical Framework for EMS

Figure 4.2.2 provides a high-level overview of EMS systems in rural Minnesota as a result of field visits and interviews. The framework helps to define the EMS system along several key strata: organizations, technology, and policy. A brief summary of each layer of the framework follows.
Organizations—The framework illustrates some of the public and private organizations involved in the Minnesota EMS and the general interorganizational relationships between these organizations.
Technology—The top layer of the framework illustrates some of the essential networks and communications technologies used by Minnesota EMS organizations to carry out their individual and interorganizational functions.

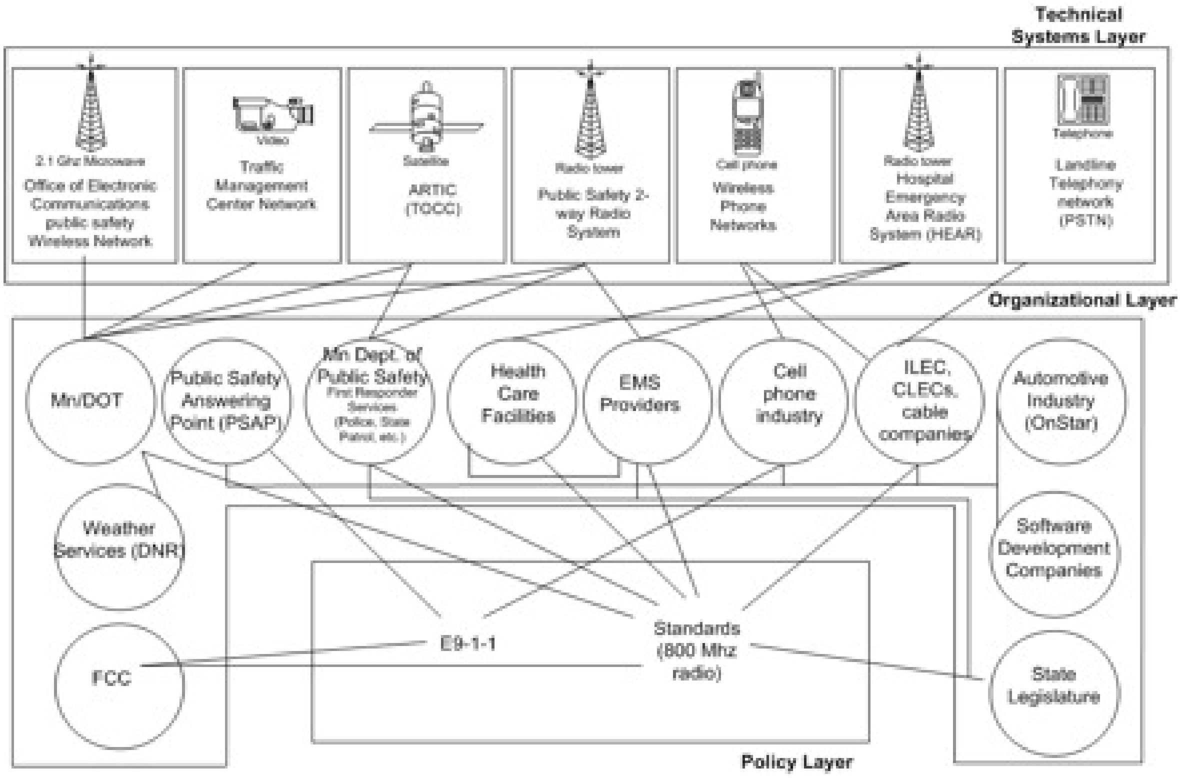


Figure 4.2.2 EMS in rural Minnesota

Policy—In order for EMS interorganizational relationships (i.e. partnerships, joint ventures, etc…) to succeed, policies need to be developed that facilitate the interorganizational use of new and existing communications technologies. The overarching EMS technology-related policies currently under development in the state are illustrated (e-911, 800 Mhz radio).

Wireless Emergency Call Routing

While this general system architecture is useful in defining system strata, for the purposes of developing the ontology, it was necessary to translate the overall EMS system architectures into a process that traces the information flows across the EMS system. Figure 4.2.3 below shows the wireless mayday call routing procedure in rural Minnesota designed for use in the preliminary design phase of the ontology. Essentially, the information flow is charted, from the originating emergency call, to the e-911 center (PSAP) and out to various emergency service providers. This procedure will be explained thoroughly in following pages.

EMS Ontology (Version 1.0)

The socio-technical framework described above represents an architectural blueprint for version 1.0 of EMS ontology. All or some of the components and their relationships in this architecture can be translated into an ontological framework. The first step in developing the ontology is to define classes. As a general proposition, these classes are based on the framework and process identified in Figures 4.2.2 and 4.2.3. The superclasses are determined first followed by subclasses. Five super-classes were defined and these are: (1) Incident Report, (2) Incident Information Acquisition, (3) Routing or Dispatch, (4) Response and Coordination, and (5) Data Management. Figure 4.2.4 provides a tree diagram representation of the ontology developed for

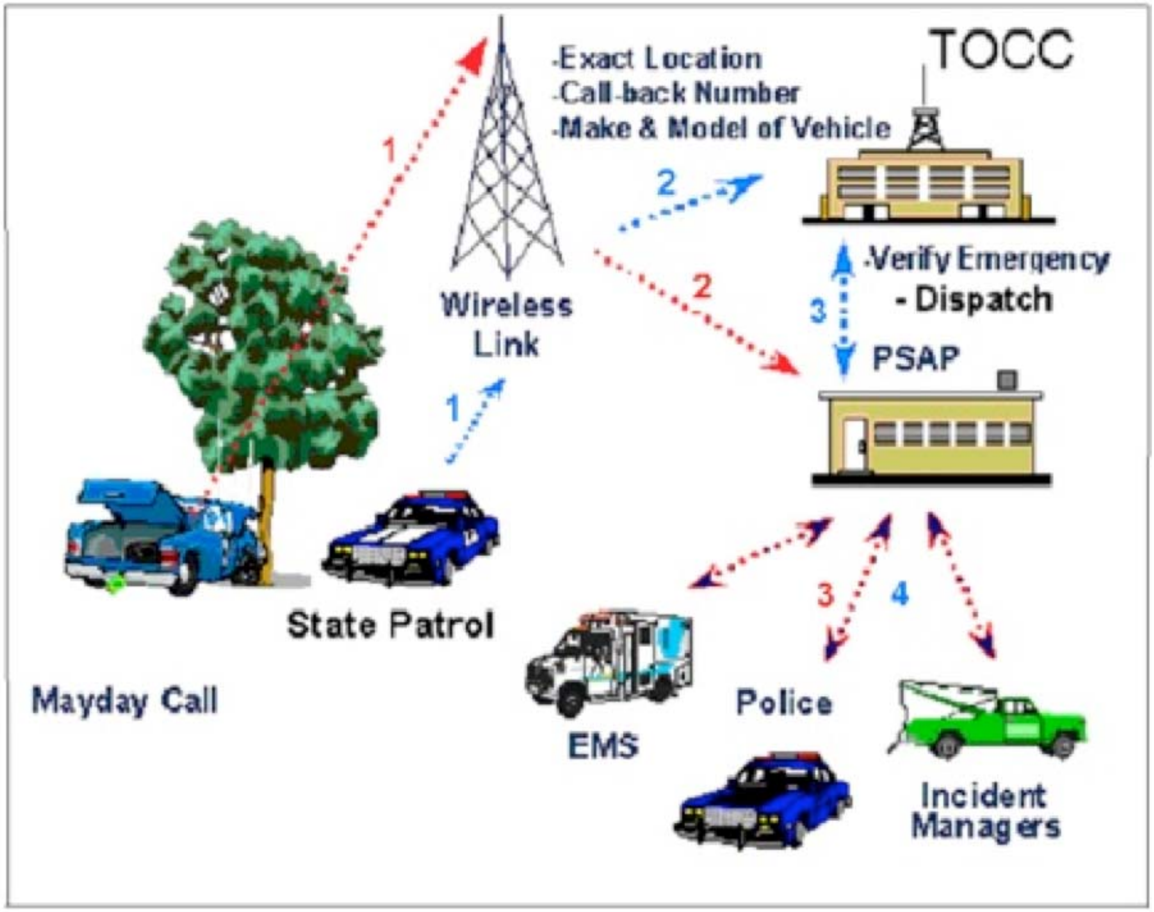


Figure 4.2.3 Wireless mayday call routing procedure in rural Minnesota

Wireless EMS in Rural Minnesota. The number of superclasses may increase as the case study evolves. Their current representation in Protégé system can be seen in Figure 4.2.5.

At the current stage of the development all superclasses have few subclasses. However, more subclasses can be added to the superclasses through updated versions of the ontology since the emergency management in rural Minnesota is an emerging system.

Many super/subclasses are created as a manifestation from the wireless mayday call procedure described above. For example, the subclasses called "911 Calls", "Automatic Crash Notification", and "Radio Communication" under the *Incident_Report* superclass are part of the wireless mayday call routing procedure as well as the technical systems layer of the Socio-technical framework. On the other hand, there are subclasses that are not directly related to both Socio-technical framework and mayday call routing procedure. One of them is "Benchmark Systems" under the *Data_Management* superclass. This subclass plays an important role in knowledgebase development with crucial nationwide information attached for comparison. That is, it establishes the basis for communicating performance, another key goal for the EMS ontology.

A second step in the Protégé knowledge acquisition system is to define the *Slots* with their value type and cardinality rules. This is an important process, as the instances recorded later must comply with these rules. Some documentation in the form of text can also be included for

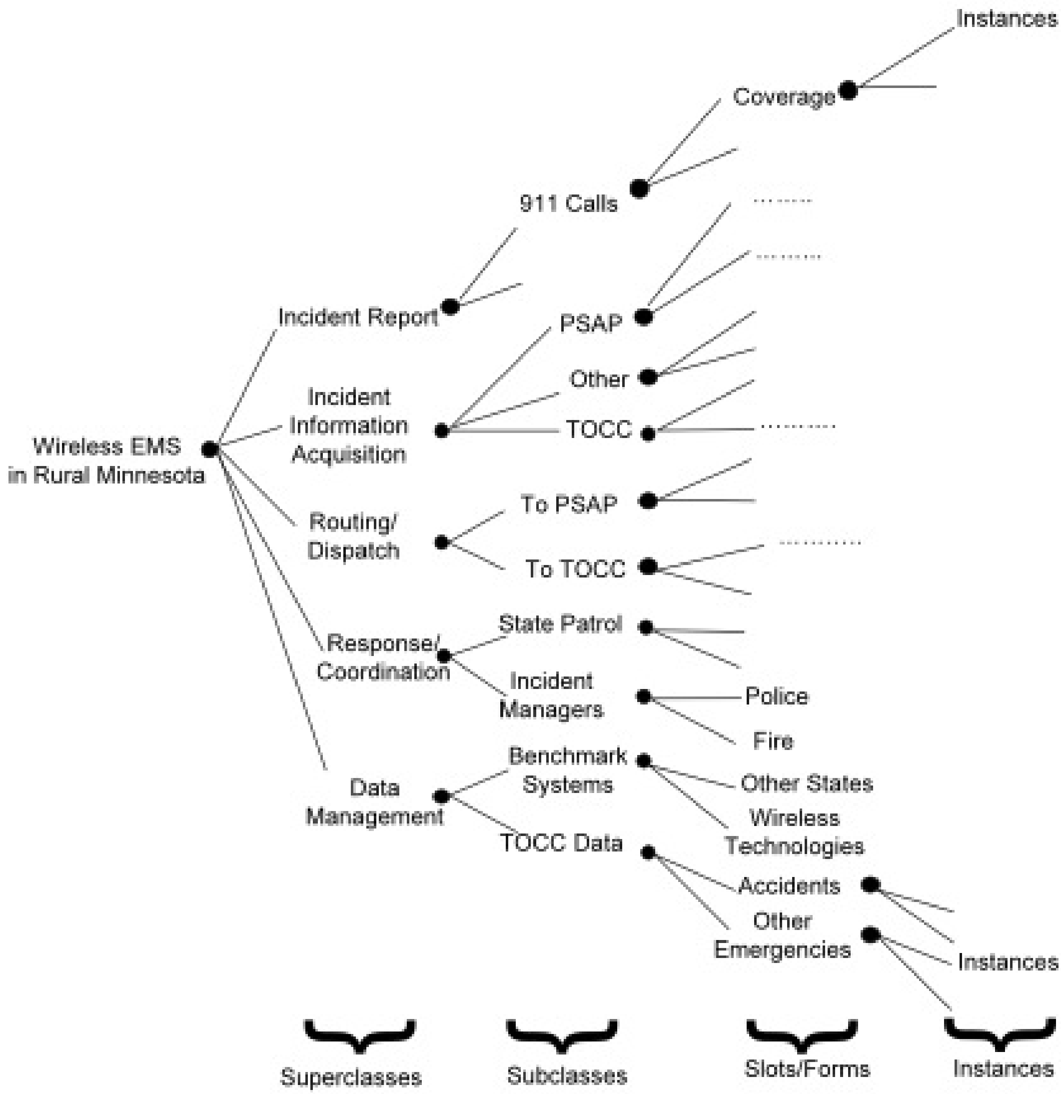


Figure 4.2.4 Tree diagram of wireless EMS ontology for rural Minnesota

each slot. After defining all the *Slots* for each subclass, Protégé automatically generates *Forms* for acquiring data as *Instances*.

4.2.4 Issues for Further Research

During the early summer 2003, the second case study will be conducted and the ontology applied to data and information obtained. This ontology will be presented in a web-enabled format and communicated to case study participants. Following completion of the second case study, a series of performance simulations will be conducted using ARENA business software. Findings from this analysis will be integrated into the ontology framework as well.

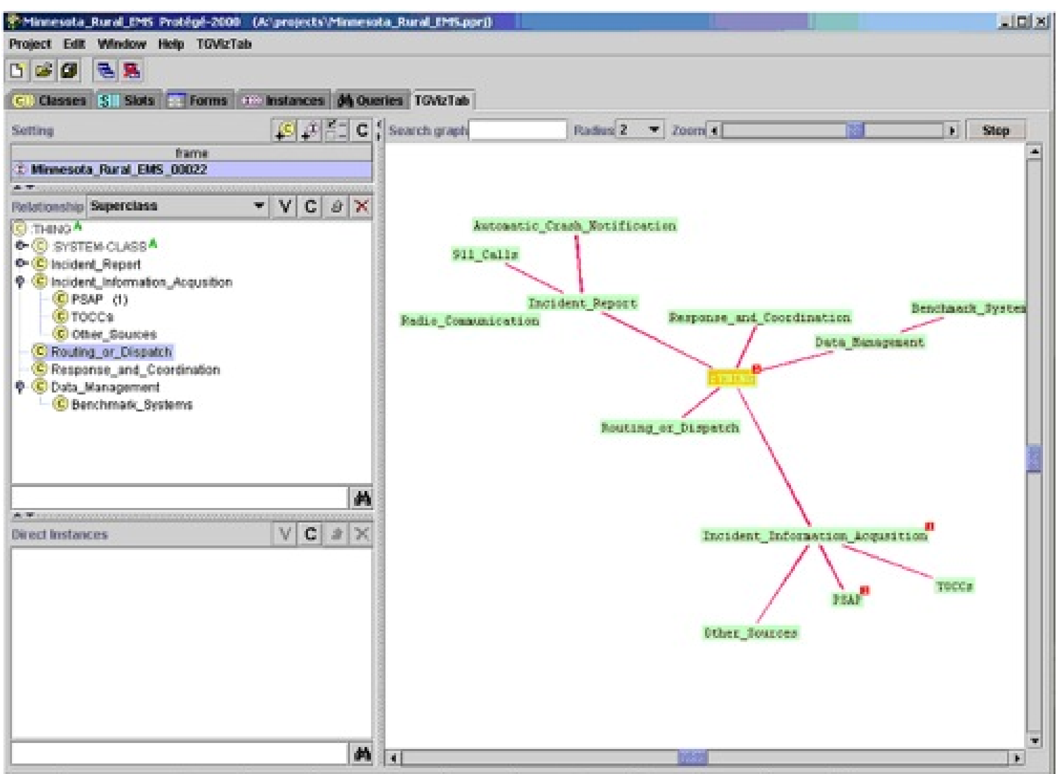
Local stakeholders were briefed on the ontology development in late 2002, and expressed support for its use; however, several steps remain to integrate the system into a paper-dominated means of characterizing and reporting on system performance. The ultimate goal of this research is to create a new means to articulate, measure, and drive development of emergent systems, such as wireless EMS. In this sense, the ontology provides a graphically-oriented evaluation element to system analysis, thereby facilitating knowledge sharing among system providers. That is, ontologies (including performance information) could be made available on the web (and through online research report channels) such that the data could be available for benchmarking, research investigations, and technology/knowledge transfer.

Indeed, from a research perspective, the EMS ontology being developed is aimed to contribute to the increasingly active area of ontological-driven research (see Mayr, 2002; Weber, 2002). The approach in this EMS research is particularly suited to enhancing the fields understanding of how a robust ontology can be devised and shared using a platform independent software program (Protégé), and done so in a manner that stresses data reporting on systems performance. It is worth noting that domain areas such as bioinformatics have moved aggressively ahead in using ontologies in this manner (though the nature of “performance” data can be quite different) (see Grosso, et. al, 1999). Expanding this dynamic use of ontologies is a worthy pursuit for the field of IS, and one that will have benefit to many domain areas as well as advancement of the field in general.

4.2.5 Acknowledgements

The authors gratefully acknowledge the support provided by the U.S. Department of Transportation, the Minnesota Department of Transportation, and the ITS Institute at the Center for Transportation Studies, University of Minnesota. Our research would not have been possible without their research and grant support. Important intellectual support for our research came from the Humphrey Institute, especially Lee Munnich and Frank Douma. An early draft of these findings was presented at the Annual Meeting of the Transportation Research Board, Washington, D.C., 2003.

Figure 4.2.5 Wireless EMS ontology for rural Minnesota (Protégé 2000)



4.2.6 References

1. Cellular Telecommunications and Internet Association (CTIA). Wireless 911 Call Statistics, available at http://www.wow-com.com/industry/stats/e911/. 2002.
2. Grosso, W.E., Eriksson, H., Fergerson, R.W., Gennari, J.H., Tu, S.W., and Musen, M.A., “Knowledge Modeling at the Millennium (The Design and Evolution of Protégé 2000),” in: 12th Workshop on Knowledge Acquisition, Modeling, and Management, Banff, Alberta, 1999.
3. Gruber, T. “A Translation Approach to Portable Ontology Specifications,” *Knowledge Acquisition* (5:2), 1993, pp. 199–220.
4. Gruninger, M., and Lee, J. “Ontology Applications and Design,” *Communications of the ACM (*45:2), 2002, pp. 39–41.
5. Holsapple, C., and Joshi, K. “A collaborative approach to ontology design,” *Communications of the ACM,* Vol. 45, No. 2 (February 2002), pp. 42–47.
6. Jurisica, I. J. and Mylopoulos, E. Yu. “Using Ontologies for Knowledge Management: A Computational Perspective,” in *Proceedings of theAnnual Conference of the American Society for Information Science,* Washington, DC, 1999, pp. 482–496.
7. Leroy, G., Tolle, K. M. and Chen, H. "Customizable and Ontology-Enhanced Medical Information Retrieval Interfaces," presented at the IMIA WG6 Triennial Conference on Natural Language and Medical Concept Representation, Phoenix, Arizona, December 1999.
8. Markus, M.L., Majchrzak, A., and Gasser, L. “A design theory for systems that support emergent knowledge processes,” *MIS Quarterly* (26:3), 2002, pp. 179–212.
9. McGuinness, D. L. "Ontologies Come of Age," in *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential,* D. Fensel, J.Hendler, H. Lieberman, and W. Wahlster (eds.), MIT Press, 2002.
10. Noy, N.F., W. Grosso, W., and Musen, M. “Knowledge-Acquisition Interfaces for Domain Experts: An Empirical Evaluation of Protege-2000,” Twelfth International Conference on Software Engineering and Knowledge Engineering (SEKE2000), Chicago, IL, 2000.
11. Wand, Y., and Wang, R. “Anchoring Data Quality Dimensions in Ontological Foundations,”*Communications of the ACM*, 39(11), 1996, pp.86–95
12. Weber, R. “Conceptual Modeling and Ontology: Possibilities and Pitfalls,” I ER 2003, LNKS 2305, 2002. pp.1–2.

4.3 Interorganizational Emergency Medical Services

Case Study of Rural Wireless Deployment and Management

Thomas A. Horan, Ph.D.
Associate Professor, School of
Information Science and
Director, Claremont Information and
Technology Institute
Claremont Graduate University
E-mail: Tom.Horan@cgu.edu

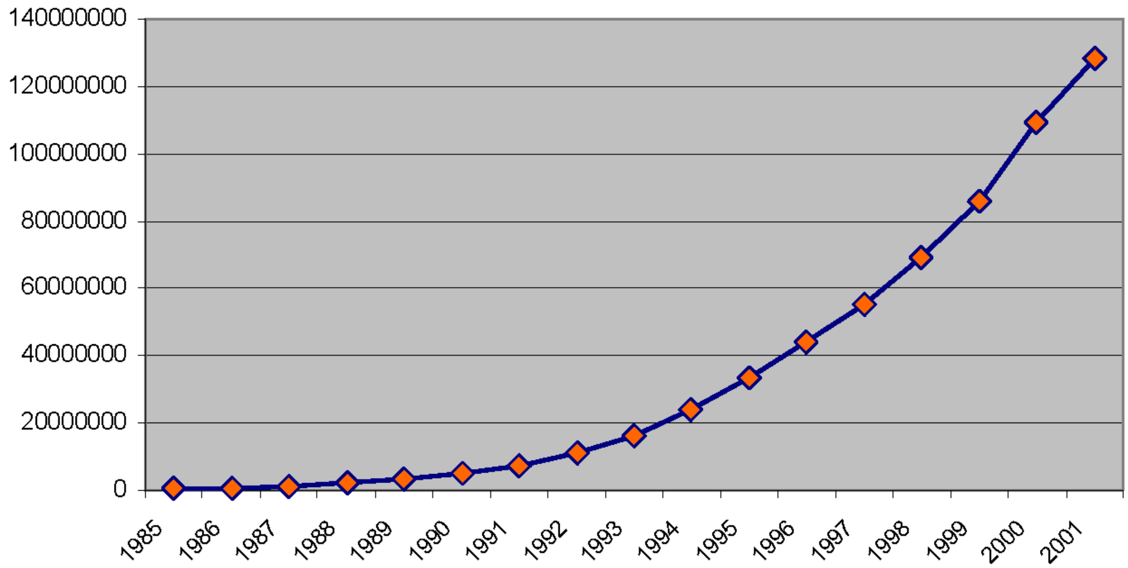
Benjamin Schooley, MIS
Research Associate and
School of Information Science
Claremont Information and Technology Institute
Claremont Graduate University
E-mail: Ben.Schooley@cgu.edu

4.3.1 Introduction

The United States emergency 911 system, established over 30 years ago, was originally conceived to be a wireline telephone means to call for help during emergencies. However, the rapid rise of wireless systems and their subsequent use for 911 calls is straining this infrastructure (Jackson, 2002; NENA, 2001). There are more than 120 million wireless users making about 155,000 emergency calls a day across the United States (see Figures 4.3.1 and 4.3.2). Indeed, mobile phones have become an important means to delivering emergency response and saving lives. A wireless 911 phone call can shave valuable minutes from the time otherwise required for a caller (or motorist aid) to find a conventional phone to access emergency medical services (Tavana, Mahmassani, & Haas, 1999).

This steady increase in private sector wireless subscribership and resulting mobile 911 use has created a need to better understand the implications of this rapidly growing system on the entire Emergency Management Services (EMS) system. While wireless-driven EMS services were generally not envisioned, it has emerged as a major "safety net" for millions of mobile users and travelers. Service providers, public agencies, and technology providers must therefore grapple with how to evolve this critical public safety and health care system (Folts, 2002; Jackson, 2002; NENA, 2001).

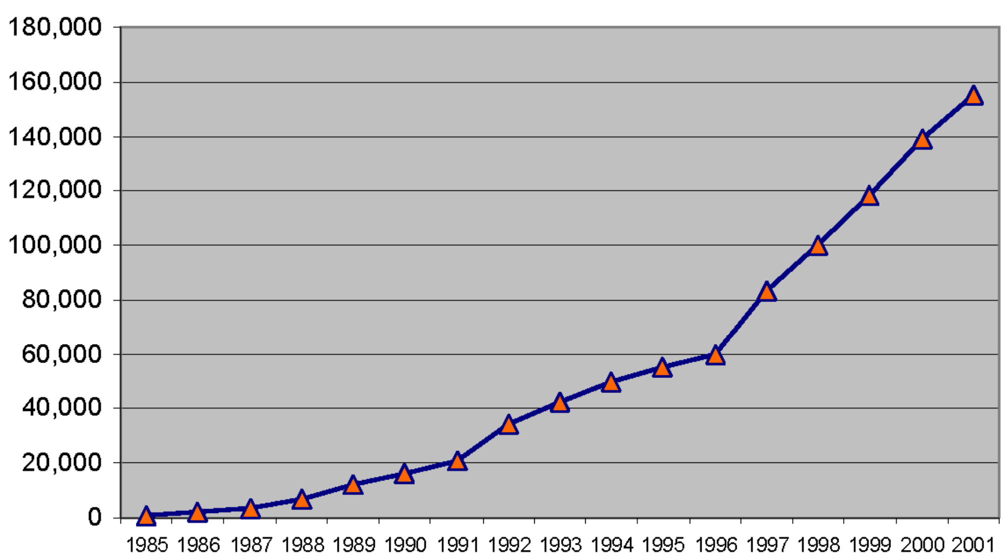
One illustration of the dynamic challenges confronting EMS providers is the U.S. Federal Communications Commissions (FCC) mandatory requirement for wireless communications services (cellular and PCS) to provide automatic location identification of a wireless 911 (hereafter referred to as "e-911") phone call to an appropriate Public Safety Answering Point (PSAP) (FCC, 2001). This requirement has grown out of a technological limitation of early mobile emergency communications, namely that the location of a caller was not identifiable. Both the private carriers and public agencies are currently working to accomplish this difficult requirement. Although the technical requirements for building these systems have been thoroughly outlined, the execution of the service has materialized slowly (Christie et al., 2002; Zhoa, 2002). One reason for this is the difficulty involved in committing to one of several viable technology alternatives to provide e-911. For example, one e-911 technology choice is a satellite-based system, which places a GPS-equivalent chip in mobile phones along with location readers at the receiving point. In addition, terrestrial systems offer several alternatives (e.g. time of arrival, time difference of arrival, angle of arrival) (see Christie et al., 2002; Zhoa, 2002).



Note. From Cellular Telecommunications Industry Association (CTIA) web site retrieved July 10, 2002 from http://www.wow-com.com/industry/stats/e911/.

Figure 4.3.1 Wireless phone subscriber growth in the U.S.

With the current rate of technological change, selecting the one best solution, or combination of solutions, for long term system investment is a difficult and daunting task for system administrators and designers (Proietti, 2002).



Note. From Cellular Telecommunications Industry Association (CTIA) web site retrieved July 10, 2002 from http://www.wow-com.com/industry/stats/e911/.

Figure 4.3.2 Estimated number of wireless emergency calls per day in the U.S.

In addition to the technology selection challenge inherent in mobile EMS services, deploying "end-to-end" e-911 systems, from the originating caller, to the PSAP, to a variety of 911 service providers, will require new technologies and nontraditional partnerships. In particular, wireless carriers, emergency dispatch center administrators (e.g. PSAPs), law enforcement, fire and EMS officials, automotive companies, consumers, technology vendors, and state and local political leaders will need to cooperate to deliver an integrated set of e-911 services (Jackson, 2002; Lambert, 2000; Potts, 2000). These partnerships are particularly critical in rural areas. According to the U.S. Department of Transportation, more than 59% of fatal automobile crashes in 2002 occurred on rural roads (National Center for Statistics and Analysis, 2003). The Minnesota Department of Transportation (Mn/DOT) reports that only 30 percent of miles driven within the state are on rural roads, yet 70 percent of fatal crashes occur on them (Short Elliot Hendrickson Inc., 2000). In addition, 50 percent of rural traffic deaths occur before arrival at a hospital. Appropriate medical care during the "golden hour" immediately after injuries is critical to reducing the odds of lethal or disability consequences. Crash victims are often disoriented or unconscious and cannot call for help or assist in their rescue and therefore rely heavily upon coordinated actions from medical, fire, state patrol, telecommunications and other entities (Lambert, 2000).

The present study focuses on the evolving Emergency Medical Services (EMS) system in rural areas of Minnesota. Minnesota provides a useful case study location for three reasons. First, rural wireless telecommunications infrastructure (coverage), and telecommunications infrastructure in general, is not as developed as in larger metropolitan areas (FCC, 2002; Parker, 2000). This offers the researchers an opportunity to explore specific barriers that may not exist in metropolitan areas. Second, Minnesota has been aggressively pursuing Intelligent Transportation Systems (ITS) initiatives for several years and thus offers a test bed for various EMS research projects, some of which have been conducted very recently (Nookala, 1998; Nookala, Gardner, & Bland, 1998). Third, rural areas of the state are rapidly moving towards implementation of wireless EMS and are therefore in the process of developing partnerships with stakeholder organizations (Nookala et al., 1998; U.S.DOT, 2002). This provides researchers with an opportunity to explore local implementation of effective EMS from interorganizational perspectives.

To understand the means of operating effective EMS infrastructures in Minnesota, this paper will utilize a theory-based case study approach. As Yin (1988) has noted, case studies can be particularly useful in exploring the holistic interactions of a phenomenon of interest, including the extent to which theory-based dynamics are operative in the setting. In the case of EMS, the theory that is drawn upon is work in the area of complex systems, and within this context, Interorganizational Systems (IOS). In terms of bounding the system under study, EMS is presented as a subsystem of the Intelligent Transportation System (ITS), the latter (ITS) being a well-defined U.S. architecture for deploying technologies in surface transportation.

Specifically, the study will utilize the National ITS Architecture as a "lens" to analyze and explore EMS infrastructures in rural Minnesota from technology, institutional and policy aspects. Relevant principles from complex systems and interorganizational linkages are raised, leading to the specific research objectives of the study. Using this theoretical and conceptual orientation, results from the case study interviews and reports are summarized. Based on this empirical material, the conceptual architecture for rural EMS dynamics is presented, including its implications for measuring and enhancing system performance as well as future research on IOS operations within complex dynamic systems.

4.3.2 Background

Complex Systems

Allen et al (2001, p. 2) offer the following as a starting point to the subject of complex systems: A system is complex when it is composed of many "components and interconnections, interactions, or interdependencies that are difficult to describe, understand, predict, manage, design, or change." The definition points out two general features of complex systems: the myriad interconnections between parts of a system, and the dynamic nature of these interconnections (Allen et al., 2001; Maier & Rechtin, 2000; Moses, 2002; Sussman, 2002).

Looking beyond this introductory definition, some underlying concerns when studying complex systems is the self-organizing and adaptive nature of the system. By self-organizing we mean the "dynamics of systems that arise endogenously and spontaneously from their structure (Sterman, 2000, p. 22)," such as the spontaneous growth of 911 demand from wireless devices. By adaptive we mean the "capabilities and decision rules of the agents in complex systems that change over time (Sterman, 2000, p. 22)," such as the evolving partnerships between public and private sector organizations in delivering new e-911 services. These two system characteristics have also been described by the term "emergent properties" – properties or behaviors of a system that emerge over time or space, including those that arise in response to behavior of other systems and environments (Allen et al., 2001).

Considerable research has been done to better understand complex systems. Several researchers have studied how to understand, engineer, implement and manage implementation of complex emerging systems and related technologies (Bansler, Damsgaard, Scheepers, Havn, & Thommesen, 2000; Bhattacherjee, 1998; Shaw, 2002). There has, however, been relatively scant attention to its application to physical infrastructure systems (Zimmerman & Cusker, 2001). One notable contribution is from Sussman (2002) and colleagues who speak of "Complex, Large-Scale, Integrated, Open Systems" (CLIOS). These researchers are actively testing several complexity-related concepts to large-scale infrastructures such as transportation. A significant aspect of this analysis is examining the nesting of technological systems within institutional processes and linkages. As Dodder and Sussman (2002) note:

> We therefore have "nested complexity" when the physical system is being "managed" by a complex organizational and policymaking system. However, while we make a distinction between the physical system and policy system–which captures the primary stakeholders as well as the policymaking and other decision-making institutions – we also need to explicitly represent the connections between the physical and policy systems. Indeed, an important step in the CLIOS representation process is to identify and characterize these policy-physical system links. Understanding nested complexity is a necessary step in moving towards better integrating institutional and policy design with physical system design (p. 4).

To restate in the domain and terminology of this study, an infrastructure system (such as EMS) includes a technical system that is nested in a social and institutional system. The degree and nature of these linkages are complex, and in this sense complex includes dynamic, emerging, and not fully predictable elements. Moreover, their (CLIOS) approach suggests the utility of *portraying* a complex system as an important conceptual step toward understanding how the system operates and evolves. Finally, included in this portrayal is the need to examine *links* across the various institutions, including their various socio-technical permutations. This last concept can also be considered the domain of interorganizational systems (Teo, Wei, & Benbasat, 2003).

Interorganizational Systems (IOS)

Within a generally complex system, interorganizational systems (IOS) concepts provide a targeted means to look at the cross-organizational features. In the broadest sense of the term, IOS help to foster relationships between independent organizations using information technology. Cash and Konsynski (1985) define IOS as an automated information system shared by two or more companies (Cash & Kosynski, 1985). Similarly, Bakos (1991, p. 34) defines IOS as "an information system that links one or more firms to their customers or their suppliers, and facilitates the exchange of products and services." Johnston and Vitale (1988, p. 154) state that "An IOS is built around information technology, i.e. around computer and communications technology that facilitates the creation, storage, transformation, and transmission of information."

While most of the earliest studies conducted on IOS centered on the objective of private sector "competitive advantage" (Bakos, 1991; Cash, 1985; H.R. Johnston & Carrico, 1988; H.R. Johnston & Vitale, 1988; Konsynski, 1993), some of the recent studies have focused on a cooperative rather than a competitive dimension of IOS implementation between partner organizations (Kumar & van Dissel, 1996; Meier, 1995; Williams, 1997). These studies describe barriers and conflicts related to relationship building between partner organizations and their effect on functionality and/or adoption of IOS (Chatfield & Yetton, 2000; Chau & Tam, 1997; Chwelos, Benbasat, & Dexter, 2001; R. B. Johnston & Gregor, 2000). In addition, some of these studies imply that adoption and implementation of IOS are complex matters due to the difficulty of building cooperative relationships within the system (Lawrence, Hardy, & Phillips, 2002; Oliver, 1990; Premkumar & Ramamurthy, 1995). This is evidenced by a high failure rate of interorganizational relationships (partnerships, joint ventures, etc…) and their representative information systems (Cavaye & Cragg, 1995; Hart & Saunders, 1997).

Evoking the complexity concepts noted above, one of the most recent IOS studies explains that IOS adoption and continued cooperation are influenced by factors beyond intra- and inter-organizational dimensions, including the complex institutional networks in which companies are embedded (Teo et al., 2003). IOS involve the cooperation and commitment of all the participating members in working through these linkages. As such, these participants "may have complex economic and business relationships between themselves that result in a number of technical, social, political, and economic factors influencing the adoption of IOS (Premkumar & Ramamurthy, 1995, p. 303)."

IOS has also been analyzed in the context of complex infrastructure systems. Amin (2000, p. 263) has analyzed institutional-technological linkages that form the basis of electricity infrastructure and services and has observed their importance to the broader array of infrastructures. He notes, "The increasing complexity and interconnectedness of energy, telecommunications, transportation, and financial infrastructures pose new challenges for secure, reliable management and operation. No single entity has complete control of these multi-scale, distributed, highly interactive networks, or the ability to evaluate, monitor, and manage in real time."

In the case of electricity for instance, success and failure can be seen in interorganizational management of demand spikes during peak periods including the ability to prevent cascading failures, such as happened in New England in 1967 (Amin, 2000) and again in the Summer of

2003. In the case of transportation, it can be seen in the ability to manage air travel across a range of dynamically changing climate conditions. In the case of EMS, it can be seen in the ability to handle a rapidly escalating amount of wireless service calls using new technological systems.

The Present Case: EMS

This research examines the interorganizational and broader complex system that is responsible for providing EMS systems in rural areas. The case of EMS provides an interesting illustration of IOS in that public and private agencies need to interact both technically and organizationally in a time critical fashion to deliver health care services to travelers and other users of mobile communications. The "end-to-end" EMS service begins with a consumer action (placing the call), involves the private sector (the cellular service provider) delivering the call, the public sector (PSAP or state police) receiving and dispatching the call, the private sector again (this time an ambulance service) providing transport and health care services, and finally, either a public or private sector hospital to deliver additional health care services. Moreover, as is expected in complex systems, there can be emergent properties to these arrangements that result from innovative technological and institutional systems devised to accommodate this new communications medium for soliciting emergency help.

As suggested by Sussman (2000), Sterman (2000) and others (Amin, 2001; Weick & Sutcliffe, 2001), a significant first step is to chart out the domain of the subject system, including the dynamics in play in terms of both the technology and the institutional system in which it is nested. This study begins such an analysis by describing the domain of EMS, and then analyzes the dynamics through an in-depth case study analysis. As such, this research offers an approach to portraying the entire system in such a way as to provide a complete look at how the IOS linkages occur and where the weak links might occur.

4.3.3 Theory

Intelligent Transportation Systems (ITS)

A broad range of diverse technologies, known collectively as intelligent transportation systems (ITS), includes information processing, communications, control, and electronics. Sensory devices, software programs, cameras, cellular, landline, and satellite telecommunications are among many technologies included to ITS. These technologies provide the intelligent link between travelers, vehicles, and infrastructure (Francois, 2000). The integration of various advanced technologies into the transportation system has formed several subsystems within ITS and hundreds of smaller divisions within subsystems (Lockheed Martin, 1997, 1999).

Emergency Management Services (EMS) is one of the most distributed subsystems of ITS. The EMS subsystem includes many different organizations, services, and technologies. First responder services, such as law enforcement, fire, and state patrol, health care facilities, state departments of transportation, wireless and wire line telecommunications service providers, emergency response call centers and some private organizations compose the complex EMS infrastructure (Lockheed Martin, 1997, 1999).

According to the Minnesota Department of Transportation ITS program, there is a need to integrate new technologies with the existing emergency response infrastructure (To & Choudhry, 2000). In addition, there is a need to integrate organizations within the infrastructure through private-private and public-private partnerships and thus create cross-organizational synergies for institutionalizing EMS (Short Elliot Hendrickson Inc., 2000; Jackson, 2002; NENA, 2001; To & Choudhry, 2000).

National ITS Architecture

For this research, the National ITS Architecture is applied as a domain specific "lens" to observe current functionality of EMS in rural areas of Minnesota. The National ITS Architecture was developed by the U.S. Department of Transportation as a framework to define the interactions between the transportation and telecommunication domains to create and offer ITS services throughout the nation (Parsons, 2000). The purpose for applying the National ITS Architecture in this study is that the architecture was specifically designed to focus on the complex aspects of transportation (mobility) and technology, or as the designers state, to "mitigate the complexity involved in dealing with numerous complex entities (Lockheed Martin, 1997)." The approach of the architecture was to provide a high-level logical and physical delineation of ITS (of which EMS is considered a subset) using a tri-partite delineation of dimensions: transportation, communications, institutions. The National ITS Architecture includes the following three layers (see Figure 4.3.3):

1. **Transportation Layer—**This is the physical ITS infrastructure. This layer identifies key players and establishes a common terminology for existing and future ITS subsystems. The Architecture encompasses essentially (1) travelers; (2) vehicles; (3) management centers; and (4) roadside appliances.
2. **Communications Layer—**This information infrastructure connects the technological elements of the transportation layer. The Architecture carefully lays out (1) what types of information and communication support various ITS services; (2) data sharing and use by physical entities (subsystems); and (3) sets of standards to facilitate data sharing and use.
3. **Institutional Layer—**This layer determines the socioeconomic infrastructure of organizations (agencies of all governmental levels, public and private entities) and their social roles, reflecting jurisdictional boundaries. The institutional layer includes developing local policy, financing ITS, and creating partnerships to guide ITS development (Lockheed Martin, 1999, p. 1-1 and 1-2).

With regard to the present study, this framework was used for constructing interview questions, determining the scope of the document analysis, analyzing findings, and developing an architecture for effective EMS, taking into consideration the specific case of rural Minnesota.

4.3.4 Methodology

Research Objectives and Approach

There were two related research objectives of this study. The first was to identify key technical, organizational and policy dimensions to EMS within the context of rural Minnesota. The second objective was to devise a system architecture for EMS that could be used to better understand and improve EMS system performance. These objectives were accomplished through a qualitative theory-driven case study that utilized interviews and supplemental EMS related reports.

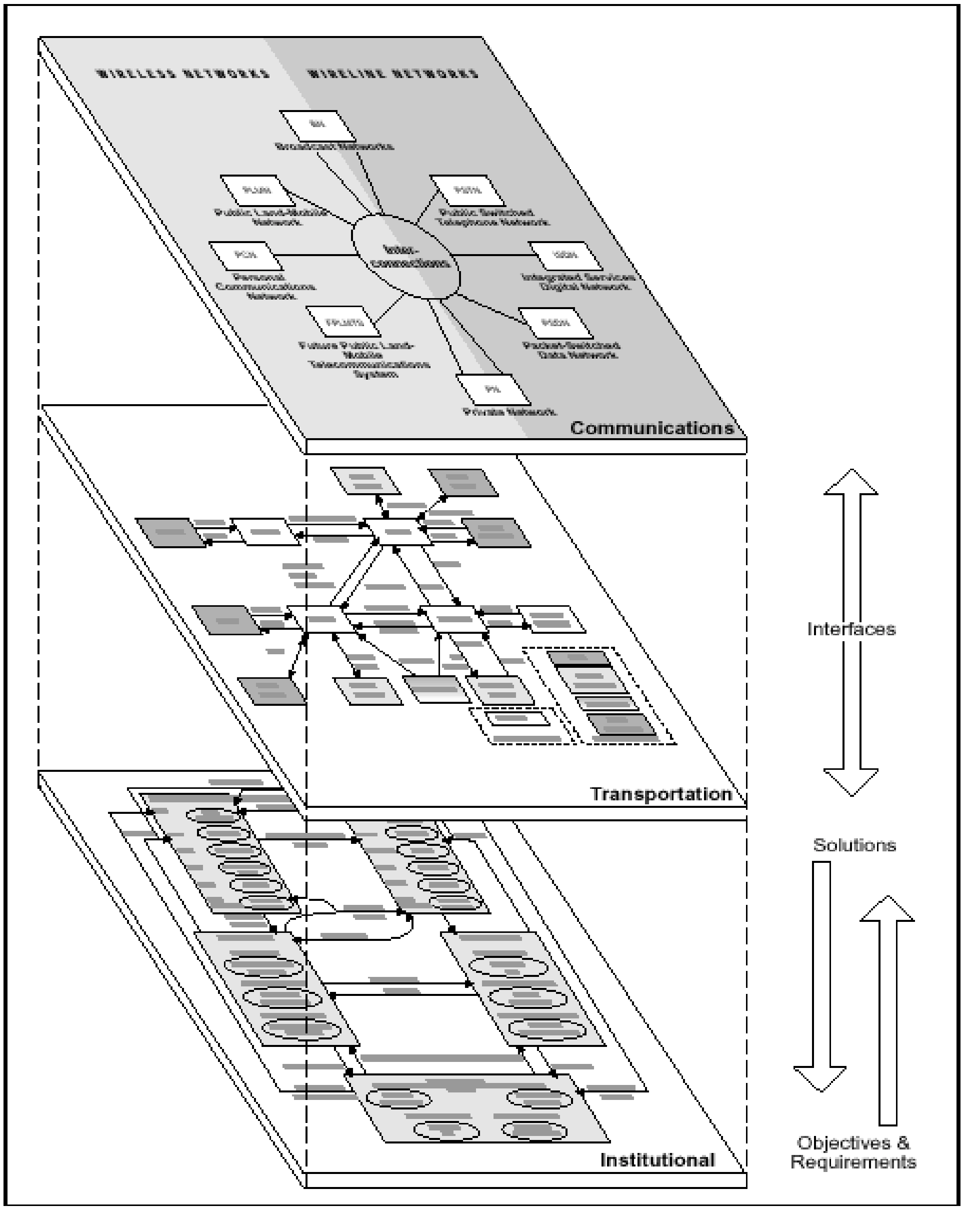
Specifically, a case study design was used to ask "What" the interorganizational challenges are to effective EMS systems. Yin (1988) states that when a "What" research question is asked and an exploratory study is contemplated, an embedded case-study design is appropriate. In addition, an embedded case study design is effective for acquiring information about a single individual, entity, or process that has many subunits of interest (GAO, 1991; Yin, 1988). This case study examines EMS systems within a single entity, the state of Minnesota. The embedded subunits in this case are the multiple EMS organizations within the rural Minnesota towns, as well as individuals engaged in local policy and program management.

Interviews

The primary empirical effort involved two rounds of semi-structured, in-depth interviews, plus a site visit with representatives of multiple organizations that participate in public-private and private-private EMS partnerships in Minnesota. Kerlinger (1986) states that semi-structured interviews are appropriate to gain in-depth exploration into ideas and relationships not initially considered. An open-ended-item technique was used to provide a framework for constructing questions to place minimal restrictions on the single or multiple interviewees within each organization (Kerlinger, 1986; Yin, 1988). Slightly under half of the total number (i.e., 6 of 13) of interviews were conducted in person and, conversely, slightly over half (i.e., 7 of 13) were conducted over the telephone. A similar set of questions were asked of all interviewees, with time devoted to technological, institutional, and policy dimensions of EMS. Written summaries were compiled for each interview.

A first round of in-depth interviews were carried out in the fall, between October 15 and November 14, 2001. Representatives from six public and private organizations in rural Minnesota provided their responses to semi-structured questions, which reflected technical, institutional, and policy aspects of wireless technology integration into EMS. Interviewees were identified through previous Minnesota ITS case research at the University of Minnesota Humphrey Institute of Public Affairs (see Kuhn & Douma, 2003). Organizations included (position/s in parenthesis): Virginia County State Patrol (Supervisor), Virginia Country Fire Department (Training Officer), Minnesota Department of Transportation Office of Electronic Communications (Director; Communications Planning Director), Duluth Economic Development Association (Duluth City Councilman), Rochester Police Department (Communications Manager; Supervisor), Mayo Medical Transport (Manager), and the City of Rochester mayor's office (Mayor). Results from the first round of interviews were documented for later synthesis. First-round interviewees suggested additional organizations and persons for second round semi-structured interviews. This method used by researchers to identify interviewees is referred to as "snowball" or "chain" sampling (Trochim, 2000).

In Winter 2002, February 7 through March 5, a second round of interviews were conducted with a range of stakeholders active in the e-911 domain. In this round, expert representatives from seven public and private organizations presented their opinions and views on institutional, technology and policy issues related to the Minnesota EMS. The organizations included (position/s in parenthesis): the Rhode Island e-911 Board (e-911 Executive Director), Public X-Y Mapping Project (Project Director), AK Associates (President), mn/DOT (ITS Program Director), The Mayo Clinic (Dispatch Center Manager), Minnesota Department of Administration (911 Product Analyst; 911 Product Manager), Metropolitan 911 Board (Executive Director). This interview group helped researchers identify a final opportunity—namely a field visit to a designated site.



Note. From "National ITS architecture documents: Communications document," by Lockheed Martin Federal Systems and Odetics ITS Division, 1997, January, Prepared for the Federal Highway Administration, U.S. Department of Transportation, p. 11.

Figure 4.3.3 National Intelligent Transportation System architecture

In Summer 2002, researchers made a site visit to Virginia, MN where the first of nine new state Transportation Operations Communications Centers (TOCC's) were implemented. The TOCC is a direct result of a Minnesota ITS test project, ARTIC, (Short Elliot Hendrickson Inc., 2000) and a newly formed EMS partnership. The organizations interviewed included (position/s in parenthesis): Minnesota State Patrol (Captain; Lieutenant; Communications Center Supervisor; Radio Dispatch Operator), Minnesota Department of Transportation (District Engineer; Maintenance Superintendent; Communications Director), Minnesota Department of

Administration (Safety Administrator), Virginia Public Safety Answering Point (PSAP) (Communications Center Supervisor). Representatives were asked similar open-ended questions about EMS policy, technology, and organizational issues within the specific context of the rural EMS deployment. Similar to the first round, written summaries were prepared containing interviewee perspectives in these three dimensions (technology, organizations, policy), as well as other issues raised during the course of the interview.

Supplemental Empirical Material

A secondary material gathering technique included the use of available information and provided a supplement to the expert interviews (GAO, 1991). A series of research and project reports relative to EMS in Minnesota and nationwide were reviewed with attention to policy, organizational, and technology issues (See Appendix C). Generally, these reports helped inform the context for the Minnesota case study and provided specific data on selected implementation trends and challenges. Specifically, there were two uses for these reports. The first use was in framing the research project to assist in understanding background issues and in constructing interview questions. The second use of these reports was as concurring sources of information on issues identified in interviews. For example, interviewees discussed challenges related to a lack of technical standards for radio communications throughout the rural areas of the state. These issues were verified in the 800 MHz Statewide Report (MN/DOT, 2001).

4.3.5 Results

Both the interviews and reports served to outline the EMS system in rural Minnesota and to identify major difficulties and achievements in establishing well-functioning and efficient EMS services. Experts discussed *technological* problems with a lack of rural wireless coverage and the need to upgrade unreliable technology to better serve the public during emergency responses. Several *organizational* issues were identified including difficulties establishing interagency cooperation, which was seen as key to successful implementation of ITS technologies. Experts discussed *policy* issues such as a lack of funding to upgrade technologies, and a lack of technical standards, or definition of standards for organizational and industry direction. Findings related to each of these dimensions (technology, organizational, policy) are elaborated upon below, with practical and research implications discussed in the subsequent and final section.

Technology Issues

EMS is comprised of many organizations and services that greatly rely upon technology to perform the vital liaison function for coordinating actions within the IOS (Boyd, Maier, & Caton, 1998; Corbin & Noyes, 2003; Pearce, 2000). At its base, wireless EMS is composed of a wireless network for communicating distress calls to a public safety answering point (PSAP), which receives the call and hands it off to responders. This latter function (recovery, handoff) involves computer-aided dispatch (CAD). While there is no doubt that the technology exists to create state-of-the-art EMS, major barriers exist to implementing and managing them (Jackson, 2002; NENA, 2001). The challenge of deploying new EMS services was seen in the design of Minnesota’s Mayday Plus demonstration (To & Choudhry, 2000). This demonstration showed that automated location devices could be used to increase the effectiveness of medical and road assistance and that collision severity notification systems were able to transmit requests for additional help and special medical instructions to emergency room or trauma center personnel at the Mayo Clinic in Rochester, Minnesota. However, while the technology is advanced and effective, the costs are currently prohibitive for a wide scale deployment. Two specific cost prohibitive technological issues identified by interviewees were in relation to wireless coverage and upgrading systems for better performance (To & Choudhry, 2000).

Coverage. Interviews confirmed that in recent years the Minnesota EMS has had to respond to a significant increase in mobile-based emergency calls. The Minnesota State Patrol answered 650,000 cellular 911 calls from 2000 to 2001 (or 1,780 daily), and numbers of cellular 911 calls have increased 15 to 20 percent every year (Jonassen, personal communication, Fall 2001; To & Choudhry, 2000). The Mayo Clinic, who serves a mostly rural territory, reported that 7080% of emergency calls are made from a cellular phone (Lyden, personal communications, Fall 2001). Many of these calls originate from rural areas of the state, where access to wireline telephone service is not always easily available. Experts noted that local emergency response specialists must rely upon their general sense of the area to locate a victim without specific directions (Gustafson, Jonassen, Lyden, personal communications, Fall 2001). Several interviewees also discussed inadequacies in developing wireless infrastructure in rural areas, including a limited coverage area and improper call routing to the correct Public Safety Answering Point (PSAP) (Beutelspacher, Pollig, & Moody, personal communications, Spring 2002). For example, the Virginia State Patrol explained that emergency wireless calls often skip over Lake Superior from Upper Michigan to the dispatch center in Virginia, Minnesota. These calls have to be redirected to the proper county and then to the city dispatcher. This creates time delays and other difficulties related to immediate response to emergency calls and the location of their related accident sites (Jonassen, personal communications, Fall 2001).

Technology Upgrades. Several interviewees noted that timely deployment of new technological systems remained one of the essential issues in creating effective interorganizational linkages for EMS in Minnesota, especially in rural areas. Interviewees explained benefits to newly implemented technologies, including enhanced efficiency among call dispatchers and state patrol officers in arriving to accident locations (Gustafsson, personal communication, Fall 2001). Although Minnesota partnerships widely deployed special software, radio, cellular, Geographic Positioning Systems (GPS), Automatic Vehicle Location (AVL) and other advanced technologies; experts concluded that the EMS infrastructure demanded additional resources (Beutelspacher, Moody, and Pollig, personal communications, Winter 2002). Systems need to be added, upgraded, and integrated with existing transportation information systems. Yet, there are two major difficulties associated with upgrading systems.

First, there is often a paucity of local funds in rural areas to do so (Gustafson, Jonassen, Hogan, Mulleneaux, personal communications, Fall 2001). For example, 40–50% of Minnesota's 119 PSAP's, the pivotal unit for dealing with 911 calls, are lagging in their upgrades to the FCC's new e-911 regulations, which has required a call-back number and location identification to accompany each 911 call into the PSAP (e.g Phase II) (Beutelspacher, Moody, and Pollig, personal communications, Winter 2002). Besides funding, a second major impediment to upgrading systems keeping up with technological demands of rapid changes in wireless EMS (Heroff, Hogan, Mulleneaux, Terry, personal communications, Fall 2001). For example, experts mentioned that as the number of emergency calls from cellular telephones has increased, wireless trunks are often busy, leaving wire line trunks available. It is difficult for decision-makers to know if they should retrofit existing wire line trunks to accept wireless calls, continue to add additional wireless trunks, and if so, how many, or upgrade their technology to include acceptance of both wireless and wire line calls without separating the trunk types (Kraus, LaBelle, personal communications, Spring 2002).

Organizational Issues

The diverse range of stakeholders interviewed reflects the myriad organizations involved in the delivery of EMS services. Interviewees were quite aware of these linkages and described several benefits to partnership creation. As noted below, interviewees explained that in order to establish a well functioning EMS using advanced technologies, coordinated relationships between individual member organizations of the system must be established. In doing so, several interviewees noted the distinct cultural differences across the organizations that make up an EMS IOS (Kujala, Maddern, Ralidak, Salo, Sheehey, personal communications, Summer 2002). Barriers to partnerships of technology sharing include lack of trust, funding, and knowledge. Other barriers include the reluctance to adopt new technology by government agencies, which reflect the concerns of the undefined role of government, the ability or inability to control technology, and service in the urban versus rural areas (Hogan, Jonassen, and Terry, personal communications, Fall 2001).

Cross-Agency Resource Sharing. Interviewees described synergies and barriers created between organizations by sharing technological resources and associated costs (Erjovec, Herbold, Jonassen, Kujala, Maddern, personal communications, Summer 2002). Interviewees stated that the sharing of resources allows organizations to benefit from economies of scale and upgrade to new technology (Hogan and Terry, personal communications, Fall 2001). Interviews with members of a recently created partnership between the Minnesota Department of Transportation (Mn/DOT) and the Minnesota State Patrol (MSP) illustrate. They stated that there was a “Tremendous positive impact in system efficiency with the integration of [partner] communications” systems (Short Elliot Hendrickson Inc., 2000). For example, communications and dispatch for the two organizations were concentrated into one communications center saving costs on equipment and personnel. In the event of a large rural highway traffic accident due to icy conditions, the communications center was able to dispatch both state patrol and transportation maintenance crews simultaneously, reducing response time for both organizations to reduce further health risks to the public (Erjovec, Herbold, Jonassen, Kujala, Maddern, Ralidak, Salo, Sheehey, personal communications, Summer 2002).

However, experts also explained problems in getting new systems adopted across partner organizations (Erjavek, Herbolt, Maddern, Ralidak, Salo, Sheehey, personal communications, Summer 2002). mn/DOT’s representatives confirmed that a partnership heavily evoked issues of financing arrangements (including technology and costs incurred from participation in the partnership) (Hogan and Terry, personal communications, Fall 2001). As a consequence, these organizations had varying levels of information interconnectivity, ranging from new and integrated CAD systems (in Rochester) to fairly elementary phone-transfer operations (in Brainerd) (McJoynt, personal communications, Winter 2002; Jonnasen, Fall 2001).

Dynamic Public-Private Partnerships. Partnerships that consist of public and private organizations encounter both barriers and synergies to creating effective EMS. A representative from the Mayo Clinic, the main private partner in the Rochester EMS partnership, described the situation when the clinic purchased its partner, Gold Cross Ambulance Service (GCAS)

(Canfield C., personal communications, Fall 2001). The Mayo Clinic made the purchase because GCAS did not adopt new technology for their vehicles fast enough and therefore could not keep pace with the Mayo Clinic's technology. In contrast with this example, the Virginia Fire Department stated that their local hospital, the main private partner in the Virginia EMS partnership, was struggling to adopt new technology for emergency response purposes (Gustafson D, personal communications, Fall 2001). These two examples demonstrate the unbalanced nature of partnerships and the uneven distribution of opportunities for member agencies.

There was no common theme in terms of the best structure, distribution of roles, responsibilities and burdens for maintaining effective EMS public-private partnerships. Perspectives on partnership roles and responsibilities varies in each organization and in some cases, this causes discrepancies in EMS performance. Several interviewees explained that this major barrier stemmed from a lack of trust between partnership organizations, especially as relates to public-private partnerships. Experts from Mn/DOT's Communications Technology Office (CTO) pointed out that government agencies could not always rely on private companies to cooperate for emergency purposes (Hogan and Terry, personal communications, Fall 2001).

For example, recognizing a strong potential in using cellular technology in emergencies, the CTO proposed the solution that the government agency receive a government mandated priority access service from wireless phone carriers during a large-scale emergency. Wireless carriers argue, however, that this action would lower consumer cell phone call completion rates and pose undesired risks to the private cellular carriers (Hogan, Terry, personal communications, Fall 2001). The CTO also suggested bypassing private cellular systems altogether for emergency response purposes and instead utilize the 800 MHz radio frequencies band for 2-way radio State Agency purposes (Hogan, Terry, personal communications, Fall 2001; Mn/DOT, 2001).

Policy Issues

The interviews (and reports) suggest that federal, state and local policy had a critical role in determining incentives and terms for EMS functionality. Retrospectively, this has involved both regulation (e.g. e-911) and funding (e.g. ITS funding). Interview and document survey results indicated a need for greater interaction between government agencies of all jurisdictions to encourage rapid implementation of advanced technological products into the emergency response infrastructure (Jackson, 2002; Kranig, Labelle, Moody, & Pollig, Winter 2002, personal communications; NENA, 2001; SRF, 2000; To & Choudhry, 2000).

A federally-sponsored demonstration entitled "Advanced Rural Transportation Information and Coordination (ARTIC)" was conducted in rural Minnesota and illustrates the successful use of a collaborative approach to establishing the EMS in northeastern counties of Minnesota (Short Elliot Hendrickson Inc., 2000). Before deployment of ARTIC, agencies independently developed duplicate record keeping, which stretched scarce resources and degraded the quality of service to the public. Under the ARTIC project, several state and local agencies created an alliance, shared resources, and implemented ITS technology to design well-managed and efficient emergency response systems that served rural roads. In consequence, the demonstration also served to underscore the influence that new policy initiatives and funding could have in shepherding collaboration in EMS service provision.

Funding. Government policy comprises not only direct participation in creating the efficient EMS but also determination of a set of incentives to establish and develop EMS's within the state. As previously mentioned, the federal government has required wireless carriers to implement e-911 service to further improve the quality of emergency services. While taking measures to comply with the federal rules, Minnesota's PSAP's have encountered funding barriers. Insufficient budget sources and lack of private investments inhibited their ability to upgrade systems and equipment to support e-911 (Beutelspacher, Moody, Pollig, personal communications, Winter 2002). Considering the successful experience of Rhode Island in gaining funds from a $0.47 telephone surcharge per customer (LaBelle, personal communications, Winter 2002), Minnesota experts supported a similar solution to the issue: allocation of upgrade costs through increasing the statewide telephone charge from 10 to 27 cents per customer (Beutelspacher, Moody, Pollig, personal communications, Winter 2002). An interim default step appears to be the heavy use of transportation funding (such as ITS funding for ARTIC) to support advances in EMS that would otherwise not be possible under current funding structures (Hogan and Terry, personal communication, Fall 2001).

Standards. The study results also demonstrated how effective policy could help overcome another major barrier to implementation: a lack of standards. Besides the previously mentioned challenges that competitive sector wireless telecommunications carriers face in choosing location identification standards, experts described additional issues. For example, one such issue is how the new e-911 systems will perform relative to established standards for existing PSAPS. The Metropolitan 911 Board is analyzing the technology design that receives the high volume of calls per accident to understand if traditional standards will remain appropriate in the new communication systems. Is one phone line per 1000 people still applicable in the region? Is 60 seconds still the average length of time of a 911 call? The answers to these questions could change the technological design of PSAP's (Beutelspatcher, Moody, and Pollig, personal communications, Winter 2002).

While federal regulation is driving e-911 related location-identifying wireless services, there are emerging concerns over the nature of location information that will be provided. Experts stated that the lack of a uniform method for describing incident locations has long been a major impediment to rapid and effective emergency response in diverse metropolitan and rural areas (Hogan, Terry, personal communications, Fall 2001). When the Federal Geographic Data Committee adopted the U.S. National Grid (USNG) standard (FGDC-STD-001-2001), the problem of interoperability of location services seemingly disappeared. The USNG corrected discrepancies in map products and provided a countrywide consistent grid reference system as preferred in data applications in emergency response. However, the USNG has not been rapidly incorporated at governmental levels—including the federal e-911 regulations-- and therefore has not realized the full potential of its advantages (McNeff, personal communications, Winter 2002). In the study context, implementation of the USNG for uniform geo-addressing would significantly increase the effectiveness of GPS applications in emergency response measures.

EMS Architecture

As noted at the outset of this paper, one of the (two) research objectives was to develop an architecture for portraying the EMS. Such an architecture needs to be domain-specific to EMS. Moreover, it can be thought of as an enterprise architecture to the extent that it conveys both technology and actions relative to system goals (Ross, 2003). Such architecture is achievable, based on the characteristics and insights documented through this case study. Figure 4.3.4 provides a high-level overview of EMS systems in rural Minnesota as a result of these field visits

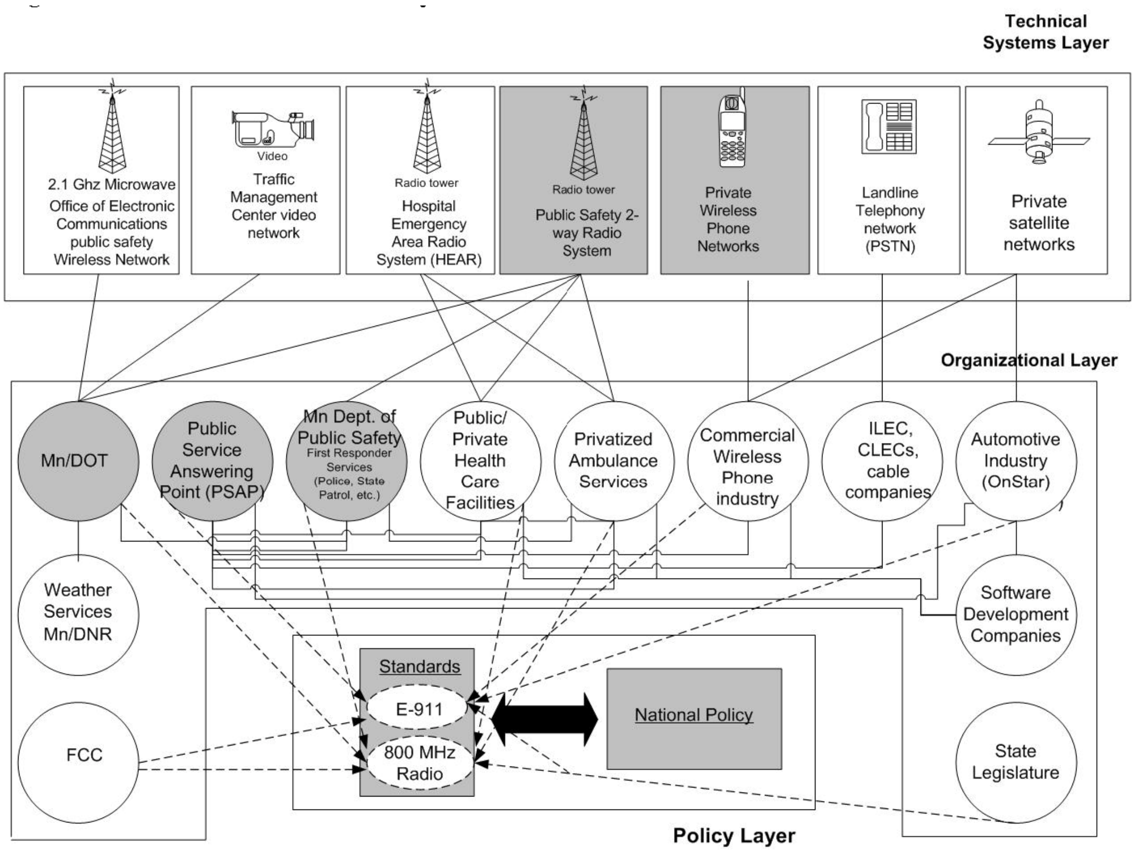


Figure 4.3.4 Architecture for EMS systems in Minnesota

and interviews. The framework helps to define the EMS system along several key strata: organizations, technology, and policy; and to identify possible critical links (shaded gray) in the overall system. A brief summary of each layer of the framework follows.

- Technology—The top layer of the framework illustrates some of the essential networks and communications technologies used by Minnesota EMS organizations to carry out their individual and interorganizational functions.
- Organizations—The framework illustrates some of the public and private organizations involved in the Minnesota EMS and the general interorganizational relationships between these organizations.
- Policy—In order for EMS interorganizational relationships (i.e. partnerships, joint ventures, etc…) to succeed, policies need to be developed that facilitate the inter-organizational use of new and existing communications technologies. The overarching EMS technology-related policies currently under development in the state are illustrated (e-911, 800 MHz radio).

4.3.6 Implications

This study has raised several policy, organizational and technological issues for EMS in rural Minnesota as well as rural areas more generally. The architecture highlights several critical areas that arose from this review and therefore have implications for future advancements. Such areas, denoted in gray in Figure 4.3.4, provide a focal point for discussing practical and research implications of the architecture for both EMS specifically and complex IOS systems generally.

Technology

The rise of wireless systems has a prominent effect on EMS. The rapid growth of wireless technologies and applications in transportation related emergencies detected by citizens and communicated via wireless systems signifies the increasing role of customer-based distributed systems in creating knowledge about service demand and delivery of services (Horan, 2002). The challenge is to devise and deploy new sets of location-specific devices (and call centers) that can in essence keep-up with emergent consumer demand. Advanced technologies further intensify the challenges of using increasingly distributed, complex, and interorganizational systems. In particular, it strains the private sector to address the wireless coverage gaps that exists in rural areas, and the public sector to develop complementary systems for communicating during emergencies as well as handling the ongoing (and growing) call volume in an efficient manner.

Issues of Interagency Cooperation

The EMS partnerships of rural Minnesota present dynamics in building interagency relations under current federal, state and local policies. Their experience revealed a set of new considerations in management of interorganizational systems. For example, Departments of Transportation do not usually consider communications infrastructure to be a core function of their agency, yet they may be in the best position to provide backbone infrastructure services, as the 800 MHz report has suggested. Mn/DOT (2001) is currently one of the major radio users in the state and has the most significant need for radio communications. However, the state will have to construct the infrastructure in order to meet their current and emerging needs. Major interorganizational barriers were generated from varying positions in terms of finance, competencies, trust, and interest in cooperative relations. One consequence has been the general lack of an “end-to-end” orientation in terms of service delivery performance. A major challenge to interorganizational system (IOS) success in EMS is finding the right mix of funding, structure, and organizational agreements across disparate organizations (e.g., PSAPS, Departments of Transportation, State Police/Patrols, Ambulance Services) that would facilitate an “end-to-end” orientation to EMS systems.

Policy

The e-911 standards have provided a set of technological, organizational, and policy challenges for many localities, including rural Minnesota. Embedded in these challenges is the broader policy challenge of finding the resources to deploy a system that may be technologically feasible but, from a financial point of view, outside the traditional level of funding available for such enterprises. Ubiquitous emergency management service is costly and raises public policy issues as to who absorbs this burden. The terrorist attacks on September 11, 2001 have only heightened general interest in domestic preparedness and provision of emergency services (Howitt, 2001). These services ultimately fall upon a range of public and private service systems similar to EMS projects in Minnesota. The EMS study demonstrates that policies on implementation, standards, and sufficient funding can create favorable climate for interagency cooperation and hasten the upgrading of systems to promote regional implementation of the ITS architecture. However, there are important policy interests in emergency management (and now, homeland security), which can impose costs on regions that they are not prepared to absorb. Consequently, policy initiative may be needed to ensure adequate EMS services in certain locales, such as in rural areas.

Research Directions

As noted in the introduction, this case study sought to examine rural EMS using complex systems as a conceptual framework and with interorganizational systems as a lens for looking at the importance of cross-institutional linkages in systems. Using an embedded case study methodology, the study found a range of technical, institutional, and policy dynamics at work. While the practical implications noted above are offered as a set of tangible directions for improving performance, from a research perspective there is a need to consider these findings within the context of complex and interorganizational systems as well as within the specific domain of (physical) infrastructure services.

The research objectives in this study reflect observations made by Sterman (2000), Sussman (1999) and others (Amin, 2001; Bansler et al., 2000; Bhattacherjee, 1998) that a significant first step in understanding a complex system is to portray the dynamics in operation, with particular attention to the range and style of interlinkages. The findings above reveal a highly heterogeneous set of linkages. For example, in comparison to the private-private form of relationships typically studied in IOS (Chwelos et al., 2001; Hart & Saunders, 1997; Hong, 2002; Johnston & Vitale, 1988; Premkumar, 2000), this case exhibits a wide range of public and private sector providers. The dynamic nature of these emergency service partnerships provides a rich canvass to understand how a relatively disparate set of partnerships can produce time critical services. Such an understanding (including pitfalls to such partnerships) adds to the inter-organizational understanding of complex systems.

Beyond a general consideration of interorganizational dynamics in complex systems lies the domain specific contribution to (physical) infrastructure services. In significant part due to advances in technology, physical infrastructures (e.g. transportation, electricity, water) have become highly complex, information-intensive, dynamic operations managed by an array of public and private sector entities (Horan, 2003). Analysis such as that offered in this research provides a conceptual lens for assessing how this infrastructure system is performing and consequently how system performance could be improved by appropriate organizational, technological and policy interventions.

The following are offered as future research directions for enhancing our understanding of both EMS systems specifically as well as complex and interorganizational systems more generally.

Better Understanding of System Performance. Building on this analysis of EMS dynamics, the next research step would be to devise a means to assess the performance implications of these dynamics. This next stage would be to draw upon this architecture to develop a set of performance analyses, such as through computer-aided systems modeling software. More specifically, advances in ontological and related modeling software (e.g., Protégé, ARENA) provide a means to meet this objective and have demonstrated applicability to systems similar to EMS (Musen, 2000). Sterman (2000, p. 898) suggests that such "improvements in graphics and animation are needed to display the global dynamics of complex models." In the context of this case study, each organization within the EMS attempts to understand its own performance. Nevertheless, each is a part of a larger system. Little is known how the entire system is performing at any given point in time. Modeling the rural Minnesota system will allow each organization to better understand how they are performing relative to the whole. From a practical perspective, such analysis would provide those agencies responsible for making decisions about the statewide system with a tool to visualize the entire dynamic system. From a research perspective, such analysis would enhance understanding of how heterogeneous organizations provide time critical services. To return to Sussman and Dodder (2002), they note, “Once the general structure of the CLIOS has been established, the next stage is to use this information to gain a better understanding of the overall system behavior, and where possible*, emergent* system behavior (p. 22, italics added).” This first research direction attends to the need to assess overall system performance; the following research direction addresses the emergent nature more directly, with particular reference to the interorganizational dimension.

Better Understanding of Inter-organizational Dynamics. Previous IOS research has typically focused on how standards-based and commonly used communications technology such as EDI can be linked to value-chain performance (Chatfield & Yetton, 2000; Chwelos et al., 2001; Hart & Saunders, 1997; Premkumar, Ramamurthy, & Nilakanta, 1994). This case study provides an example of how both public and private organizations in the rural Minnesota emergency management services system are evolving in the wake of consumer-driven pressure to improve the performance of a public service, emergency response. As such, the case provides a glimpse into the dynamics of interorganizational dimensions between public and private organizations and the challenges that these organizations face when attempting to create cooperative relationships. From a management perspective, these myriad arrangements document and extend the IOS management challenges outlined by Kumar and Crook (1999, p. 34), “With the emergence of IOS, managers across functional areas (i.e., IS, financial, marketing, etc) must understand the multidisciplinary nature of managing interorganizational systems. Managing the collaboration means managing the technology, as well as managing the individual organizations within the partnership.” Based on the present case study, this IOS management advice can be extended to consider the policy, governmental as well as private sector linkages in IOS.

Another IOS research dimension is to consider what Sterman (2000) refers to as the “organizational and social evolution” that takes place in such cooperative partnerships. One promising line of research is to identify core information content and processes and analyze the changing linkages across organizations in transmitting this information. Organizations can improve through a process of evolution, meaning that the agents that perform better can, under the right circumstances, replace those that perform worst, with certain interorganizational forms disappearing altogether (Rao & Singh, 1999). The case study identified some promising new alliances, such as between transportation and the state patrol. It would be instructive to assess the evolutionary path of these alliances, and conversely the possible decline of other forms of collaboration (possibly the PSAP). Such an analysis would help capture the emerging nature of the system over time, from the perspective of time-critical EMS information delivery.

Better Understanding of Range of EMS Systems. A common complaint about case studies is that it is difficult to generalize from one case to the entire population. However, Lincoln and Guba (1985), Yin (1988), and others (Lee, 1989) have stated that a different perspective should be taken, namely to consider the implications of the case study on the general theory (rather than general population). In this sense, the embedded case study does have generalizable attributes, in terms of the complex interorganizational dyanamics at work. Nonetheless, it would be useful to examine the robustness of these findings across rural circumstances and then compare the variations to similar enterprises in urban environments. Given the dynamic nature of the system, it would be particularly useful to examine naturally occurring "spikes" in EMS requests (such as during extreme events). This would have both practical and research significance. From a practical point of view, homeland security has become a major policy concern, and one critical aspect of this enterprise is the ability of first responders (e.g. EMS) to ramp-up efficiently (Bollwage, 2003). From a research perspective, such extreme events can provide unique opportunities to examine these systems, including the functioning of IOS coupled organizations under "normal accident" conditions, such as system overload or collapse (Perrow, 1999). By examining these naturally occurring spikes, a more comprehensive understanding of EMS systems will be achieved. This understanding will, in turn, provide a domain specific contribution to understanding how complex dynamics operate within infrastructure systems such as EMS, and how interorganizational linkages affect their success, particularly under trying circumstances.

4.3.7 Conclusion

In sum, while the proposition that we are living in a "Networked Society" is fairly straightforward and confirmable, the reality of making these networks work in a manner that is effective and affordable is a much more complicated affair. There is no doubt that the technology exists to support a fully operable, state of the art, e-911 EMS system, from the mobile caller, to the local PSAP, to the emergency response services. But the architecture developed as a result of this case study suggests that the organizational and policy dimensions exert strong influences on the nature, style, and timing of system improvements. From a management perspective, this research is aimed to help develop "mindful" EMS managers—that is those that can have an awareness of the entire system and respond to changes in an adaptive manner (Weick and Sutcliffe, 2001). From a research perspective, the aim is to provide a theory grounded model on how to integrate performance information (systems) into the analysis of infrastructures, including not only EMS, but also other complex (physical) infrastructures and thereby contribute to the growing and increasingly diverse field of complex and interorganizational systems.

4.3.8 Acknowledgements

The authors gratefully acknowledge the support provided by the U.S. Department of Transportation, the Minnesota Department of Transportation, and the ITS Institute at the Center for Transportation Studies, University of Minnesota. Our research would not have been possible without their research and grant support. Important intellectual support for our research came from the Humphrey Institute, especially Lee Munnich and Frank Douma. The authors are also indebted to the anonymous reviewers, as their comments greatly enhanced the quality of the final version. An early draft of these findings was presented at the Annual Meeting of the Transportation Research Board, Washington, D.C., 2003.

4.3.9 References

1. AEMSA. (2001). *Emergency medical services radio communications needs assessment report: For the rural aspects of the State of Minnesota*: Arrowhead Emergency Medical Services Association (AEMSA) and Minnesota State Legislature, St. Paul, MN. Retrieved June 2, 2003 from, http://www.emsrb.state.mn.us/docs/communications-21.pdf
2. Allen, T., Moses, J., Hastings, D., Lloyd, S., Little, J., McGowan, D., et al. (2001, October). *Engineering Systems Division Terms and Definitions.* Paper presented at the MIT Engineering Systems Division Internal Symposium, Cambridge, MA.
3. Amin, M. (2000). National infrastructures as complex interactive networks. In S. Weyrauch (Ed.), *Automation, Control, and Complexity: An Integrated Approach* (pp. 263–286). New York: John Wiley and Sons.
4. Amin, M. (2001, January). Toward self-healing energy infrastructure systems. *IEEE Computer Applications in Power,* 20–28.
5. Bakos, J. Y. (1991). Information links and electronic marketplaces: The role of interorganizational information systems in vertical markets. *Journal of Management Information Systems, 8*(2), 31–52.
6. Bansler, J. P., Damsgaard, J., Scheepers, R., Havn, E., & Thommesen, J. (2000). Corporate Intranet implementation: Managing emergent technologies and organizational practices. *Journal of the Association for Information Systems, 1*, 1–41.
7. Bhattacherjee, A. (1998). Management of emerging technologies: Experiences and lessons learned at US West. *Information and Management, 33*, 263–272.
8. Bollwage, E. (2003, June 24). *Testimony of J. Christian Bollwage Mayor of Elizabeth, New Jersey on Homeland Security* (Testimony before the Democratic Task Force on Homeland Security, U.S. House of Representatives): Washington D.C. Retrieved August 24, 2003, from http://www.usmayors.org/uscm/news/press_releases/documents/bollwage_062403.pdf
9. Boyd, M., Maier, M., & Caton, J. (1998). *Critical Incident Management Guidelines* (No. FTA -MA-26-7009-98-1, DOT-; VNTSC-FTA-98-5): Washington D.C.: Volpe National Transportation Systems Center and Federal Transit Administration.
10. Cash, J. I. J. (1985). Interorganizational Society: An Information Society Opportunity or Threat? *Information Society, 3*(3), 199–212.
11. Cash, J. I. J., & Kosynski, B. R. (1985). IS Redraws Competitive Boundaries. *Harvard Business Review, 63*(2), 134–147.
12. Cavaye, A. L. M., & Cragg, P. B. (1995). Factors contributing to the success of customer oriented interorganizational systems. *Journal of Strategic Information Systems, 4*, 13–30.
13. Chatfield, A. T., & Yetton, P. (2000). Strategic payoff from EDI as a function of EDI embeddedness. *Journal of Management Information Systems, 16*(4), 195–211.

14. Chau, P. Y. K., & Tam, K. Y. (1997). Factors affecting the adoption of open systems: An exploratory study. *MIS Quarterly, 21*(1), 1–18.

15. Christie, J., Fuller, R., Nichols, J., Chen, A., Hayward, R., Gromov, K., et al. (2002, April 1518). *Development and deployment of GPS wireless devices for e-911 and location based services.* Paper presented at the IEEE Position Location and Navigation Symposium, Palm Springs, CA.

16. Chwelos, P., Benbasat, I., & Dexter, A. S. (2001). Research report: Empirical test of an EDI adoption model. *Information Systems Research, 12*(3), 304–320.

17. Corbin, J., & Noyes, P. (2003). Traffic incident management planning: The case for mainstreaming. *Institute of Transportation Engineers Journal, 73*(2), 38–41.

18. Dodder, R., & Sussman, J. M. (2002, May 29–30). *The concept of a CLIOS analysis illustrated by the Mexico City case.* Paper presented at the Engineering Systems Division Internal Symposium, Cambridge, MA.

19. FCC. (2001). *Fifth report and order; First report and order: Implementation of 911 Act (*Electronic version No. WT-Docket No. 00–110 and CC Docket No. 92–105): Washington, D.C.: Federal Communications Commission.

20. FCC. (2002). Third Report: Inquiry concerning the deployment of advanced telecommunications capability to all Americans in a reasonable and timely fashion, and possible steps to accelerate such deployment pursuant to section 706 of the Telecommunications Act of 1996 (No. CC Docket 98–146). Washington, D.C.: Federal Communications Commission.

21. Folts, H. (2002, July). Standards initiatives for emergency telecommunications service (ETS). *IEEE Communications, 40,* 102–107.

22. Francois, F. (2000). Chapter 1: Introduction. In University of Minnesota, Center for Advanced Transportation Technology, ITE, USDOT, ITSA (Ed.), *Intelligent Transportation Primer*. Washington, D.C.: Institute of Transportation Engineers.

23. Hart, P., & Saunders, C. (1997). Power and trust: Critical factors in the adoption and use of electronic data interchange. *Organization Science, 8*(1), 23–38.

24. Hauer, J., Feige, J. M., & Bombach, V. (1998). *911 dispatching: Best practices review* (No. 9806): St. Paul, MN: State of Minnesota, Office of the Legislative Auditor.

25. Hong, I. B. (2002). A new framework for interorganizational systems based on the linkage of participants' roles. *Information & Management, 39*(4), 261–275.

26. Horan, T. (2002, May). Integrating the end user into infrastructure systems: Directions in the*U.S. Intelligent Transportation Systems program.* Paper presented at the Annual Meeting of the Information Resources Management Association, Seattle, WA.

27. Horan, T. (2003). Integrating the end user into insfrastructure systems: A customer centric approach to the design and function of Intelligent Transportation Systems. In G. Ditsa (Ed.), *Information Management: Support Systems & Multimedia Technology* (pp. 63–77). Hershey, PA: Idea Publishing.

28. Howitt, A. (2001, November). *Domestic preparedness.* Paper presented at the Taubman Center for State and Local Government Symposium, Harvard University, MA.

29. Jackson, A. (2002). *Recommendations for ITS technology in emergency medical services* (No. PATH Record Number 26202): Washington DC: Medical Subcommittee of the ITS America Public Safety Advisory Group.

30. Johnston, H. R., & Carrico, S. R. (1988). Developing capabilities to use information strategically. *MIS Quarterly, 12*, 37–47.

31. Johnston, H. R., & Vitale, M. R. (1988). Creating competitive advantage with interorganizational information. *MIS Quarterly, 12*(2), 153–164.

32. Johnston, R. B., & Gregor, S. (2000). A theory of industry-level activity for understanding the adoption of interorganizational systems. *European Journal of Information Systems, 9*(4), 243–266.

33. Kerlinger, F. N. (1986). *Foundations of Behavioral Research 3:e*. Chicago, IL: Holt, Rinehart and Winston.

34. Konsynski, B. R. (1993). Strategic control in the extended enterprise. *IBM Systems Journal, 32*(1), 111–131.

35. Kuhn, E., & Douma, F. (2003, Jan. 12–16). *Utilizing transportation technology to support strategic management initiatives.* Paper presented at the Academy of Sciences Transportation Research Board 82nd Annual Meeting, Washington D.C.

36. Kumar, K., & van Dissel, H. G. (1996). Sustainable collaboration: Managing conflict and cooperation in interorganizational systems. *MIS Quarterly, 20*(3), 279–298.

37. Lambert, T. (2000). Chapter 6: The role of ITS in the Emergency Management and Emergency Services community. In U. o. M. Center for Advanced Transportation Technology, ITE, USDOT, ITSA (Ed.), *Intelligent Transportation Primer*. Washington, D.C.: Institute of Transportation Engineers.

38. Lawrence, T. B., Hardy, C., & Phillips, N. (2002). Institutional effects of interorganizational collaboration: The emergence of proto-institutions. *Academy of Management Journal, 45*(1), 281–303.

39. Lee, A. S. (1989). A scientific methodology for MIS case studies. *MIS Quarterly, 13*(1), 33–50.

40. Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills: Sage Publications, Inc.

41. Lockheed Martin. (1997). *National ITS architecture documents: Communications document*: Federal Highway Administration, U.S. Department of Transportation, Washington D.C.

42. Lockheed Martin. (1999). *National ITS architecture documents: Physical architecture*: Federal Highway Administration, U.S. Department of Transportation, Washington D.C.

43. Maier, M., & Rechtin, E. (2000). *The Art of Systems Architecting, 2:e*. Boca Raton, FL: CRC Press LLC.

44. Meier, J. (1995). The importance of relationship management in establishing successful interorganizational systems. *Journal of Strategic Information Systems, 4*(2), 135–148.

45. Mn/DOT. (2001). *800 MHz team report to the 2001 Minnesota Legislature: 800 MHz statewide shared public safety radio system*: Minnesota Department of Transportation (Mn/DOT), Department of Public Safety, Department of Administration, St. Paul, MN.

46. Moses, J. (2002, May 29–30). *The anatomy of large scale systems*. Paper presented at the Engineering Systems Division Internal Symposium, Cambridge, MA.

47. Musen, M. A. (2000). Ontology-Oriented Design and Programming. In Cuena (Ed.), *Knowledge Engineering and Agent Technology* (pp. 1–20). Amsterdam: IOS Press.

48. National Center for Statistics and Analysis, NHTSA, U.S. Department of Transportation. (2003). *National Center for Statistics and Analysis Fatality Analysis Reporting System (FARS) Web-Based Encyclopedia*. Retrieved July 21, 2003, from http://www-fars.nhtsa.dot.gov

49. NENA (2001). Report card to the Nation: The effectiveness, accessibility and future of America's 911 service. Columbus, Ohio: The RCN Commission and National Emergency Number Association.

50. Nookala, M. (1998). Managing transportation and technology together. *Institute of Transportation Engineers Journal, 68*(9), 49–55.

51. Nookala, M., Gardner, W., & Bland, C. E. (1998). *Rural Minnesota: Recognizing the benefits of a statewide ITS strategic plan.* Paper presented at the 8th Annual ITS America Conference: Transportation Technology for Tomorrow, Detroit, Michigan.

52. Oliver, C. (1990). Determinants of interorganizational relationships: Integration and future directions. *Academy of Management Review, 15*, 241–265.

53. Parker, E. B. (2000). Closing the digital divide in rural America. *Telecommunications Policy, 24*(4), 281–301.

54. Parsons, R. (2000). Chapter 19: Issues in developing and implementing the National ITS Architecture. In University of Minnesota Center for Advanced Transportation Technology, ITE, USDOT, ITSA (Ed.), *Intelligent Transportation Primer*. Washington, D.C.: Institute of Transportation Engineers.

55. Pearce, V. (2000). What have we learned about freeway, incident, and emergency management and electronic toll collection? In *What Have We Learned About Intelligent Transportation Systems?* (pp. 21–44). Washington, DC: Federal Highway Administration.

56. Perrow, C. (1999). *Normal Accidents: Living with High-Risk Technologies*. Princeton: Princeton University Press.

57. Potts, J. (2000, Summer). Wireless phone calls to 911: Steps toward a more effective system. *Currents,* p. 44–49.

58. Premkumar, & Ramamurthy, K. (1995). The role of interorganizational and organizational factors on the decision mode for adoption of interorganizational systems. *Decision Sciences, 26*(3), 303–324.

59. Premkumar, Ramamurthy, K., & Nilakanta, S. (1994). Implementation of electronic data interchange: An innovation diffusion perspective. *Journal of Management Information Systems, 11*(2), 157–172.

60. Premkumar, G. P. (2000). Interorganizational systems and supply chain management: An information processing perspective. *Information Systems Management, 17*(3), 56–76.

61. Proietti, M. (2002). Carrier choices in location: The system integrator's view. *GPS World, 13*(3), 23–28.

62. Rao, H., & Singh, J. V. (1999). Types of variation in organizational populations: The speciation of new organizational forms. In J. Baum & B. McKelvey (Eds.), *Variations in Organization Science* (pp. 63–77). Thousand Oaks: SAGE Publications, Inc.

63. Ross, J. (2003, March). Creating a strategic IT architecture competency: Learning in stages. *MIS Quarterly Executive, 2,* 13–15.

64. Scientech, I. (2000). Comprehensive state communications network plan: Senior management review of network infrastructure. St. Paul, MN: Mn/DOT.

65. Shaw, N. G. (2002). Capturing the technological dimensions of IT infrastructure change: A model and empirical evidence. *Journal of the Association for Information Systems, 2*, 1–44.

66. Short Elliot Hendrickson, Inc. (2000). *Advanced Rural Transportation Information and Coordination (ARTIC) operational test evaluation report*: Minnesota Department of Transportation (Mn/DOT). Retrieved June 1, 2002, from http://www.itsdocs.fhwa.dot.gov/jpodocs/repts_te/@@801!.pdf

67. SRF Consulting Group. (2000). *TOCC Concept and Migration Plan*: MN Department of Public Safety and Department of Transportation. Retrieved May 1, 2002, from http://www.dot.state.mn.us/guidestar/pdf/toccreport00/migrationplan2000.pdf

68. Sterman, J. (2000). Business Dynamics: Systems Thinking and Modeling for a Complex World. New York: McGraw-Hill/Irwin.

69. Sussman, J. M. (2002, January 21–24). *Representing the transportation/environmental system in Mexico City as a CLIOS.* Paper presented at the 5th Annual US-Mexico Workshop on Air Quality, Ixtapan de la Sal, Mexico.

70. Tavana, H., Mahmassani, H., & Haas, C. (1999). Effectiveness of wireless phones in incident detection: Probabilistic analysis. *Transportation Research Record*(1683), 31–37.

71. Teo, H. H., Wei, K. K., & Benbasat, I. (2003). Predicting intention to adopt interorganizational linkages: An institutional perspective. *MIS Quarterly, 27*(1), 19–47.

72. To, H., & Choudhry, O. (2000). *Mayday Plus operational test evaluation report*: Castle Rock Consultants and Minnesota Department of Transportation, Office of Advanced Transportation Systems. Retrieved January 3, 2002, from http://www.dot.state.mn.us/guidestar/pdf/maydayplus/maydayplus.pdf

73. Trochim, W. M. (2000, August 2). *The Research Methods Knowledge Base, 2:e*. Retrieved March 27, 2003, from http://trochim.human.cornell.edu/kb/index.htm

74. U.S.DOT. (2000). *National Intelligent Transportation Systems Program Plan* (No. EDL #11943). Washington D.C.: U.S. Department of Transportation ITS Joint Program Office.

75. U.S.DOT. (2002). *Rural Intelligent Transportation Systems (ITS) toolbox*. Retrieved February 5, 2003, from http://www.itsdocs.fhwa.dot.gov//JPODOCS/REPTS_TE/13477.html

76. Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the Unexpected: Assuring High Performance in an Age of Complexity* (1 ed.). San Francisco: Jossey-Bass.

77. Westwood, P. S. (1998). *DIVERT evaluation report*. St. Paul, MN: Minnesota Guidestar, Department of Transportation, Federal Highway Administration.

78. Williams, T. (1997). Interorganisational information systems: Issues affecting interorganisational cooperation. *Journal of Strategic Information Systems, 6*, 231–250.

79. Yin, R. K. (1988). *Case study research: Design and methods*. Newbury Park, CA: Sage Publications, Inc.

80. Zhoa, Y. (2002, July). Standardization of mobile phone positioning for 3G systems. *IEEE Communications, 40,* 108–116.

81. Zimmerman, R., & Cusker, M. (2001, May 26). *Bringing technology to infrastructure: A white paper for a research agenda.* Paper presented at the NSF/ICIS Workshop, Bringing Information Technology to Infrastructure, Arlington, Virginia.

4.3.10 Reports Reviewed

Background Review to Frame Research Project

1. Advanced Rural Transportation Information and Coordination (ARTIC) Operational TestEvaluation Report (Short Elliot Hendrickson Inc., 2000).
2. National Intelligent Transportation Systems Program Plan (U.S.DOT, 2000).
3. Recommendations for ITS Technology in Emergency Medical Services (Jackson, 2002).
4. Transportation Operations Communications Center (TOCC): Concept and Migration Plan(SRF, 2000).

Supplemental Review to Verify Information Identified in Interviews

1. 800 MHz Statewide Report (Mn/DOT, 2001).
2. 911 Dispatching: A Best Practices Review Summary (Hauer, Feige, & Bombach, 1998). Comprehensive State Communications Network Plan: Senior Management Review of Network Infrastructure (Scientech, 2000).
3. During Incidents Vehicles Exit to Reduce Time (DIVERT) Evaluation Report (Westwood, 1998).
4. Emergency Medical Services Radio Communications Needs Assessment Report (AEMSA, 2001).
5. Mayday Plus Operational Test (To & Choudhry, 2000).

4.4 Performance Information Systems for Emergency Response

Field Examination and Simulation of End-to-End Rural Response Systems

Thomas A. Horan
Claremont Graduate University
Claremont, CA 91711
E-mail: tom.horan@cgu.edu

Richard Burkhard
Claremont Graduate University
Claremont, CA 91711
E-mail: burkha r@cob.sjsu.edu

Denise McCabe
University of Minnesota
301 19th Ave S
Minneapolis, MN 55455
E-mail: DeniseCMcCabe@netscape.net

Benjamin Schooley
Claremont Graduate University
Claremont, CA 91711
E-mail: ben.schooley@cgu.edu

4.4.1 Introduction

There has been a substantial amount of emergency response and crisis management research aimed at improving the effectiveness of the emergency response infrastructure (Davis 2002, Hale 1997, Perrow 2000). Within the field of Information Science, Turoff et al. (2004) note that a "…great deal of the literature on emergency response prior to 9/11 focuses on the response of commercial firms to emergencies or crises largely restricted to the corporate environment" (p. 3). However, since 9/11, many fields have begun to grapple with the societal dimensions of emergency response. In a recent summary of the technological platforms developed to manage emergency response, Graves (2004) finds that an array of technological tools now exist to share and visualize information and, importantly, to facilitate rapid and efficient decision-making.

Despite these existing technology tools, there is a significant gap between the promise of technology and resulting system performance. As a review of 9/11 found, "…data driven emergency response is a goal yet to be realized," (Dawes et al. 2004, p. 52). This is because emergency response is a system that lies at the interface of human systems and artificial systems (Burghard 2004). Consequently, understanding emergency medical services (EMS) system performance requires a detailed examination of how the participating agencies and entities interact with technological systems in the delivery of emergency services. Moreover, the system needs to be considered as a complete "end-to-end" process where component systems and behaviors operate within the context of interlocking multi-chained systems.

End-to-End Performance

A mobile-based end-to-end EMS system is made up of an interorganizational network of service providers. Looking from one "end" to the other "end," a medical emergency response involves multiple government agencies and nongovernment organizations, from the time an incident occurs, an emergency mobile communication (911 phone call) is made, answered by a Public Safety Answering Point (PSAP), dispatched to public agency resource (e.g., fire, police, ambulance), and treated at the scene or ambulanced to a hospital. Understanding end-to-end performance is essential to improving the timeliness and quality of service delivery, including under normal and crisis conditions.

To understand how networks of organizations can use information systems to enhance end-to-end system performance, it is important to look at the cross-organizational information flows that occur during an incident. But as Provan and Milward (2001) explain, most of the past inter-organizational network research has been, "…lacking an examination of the relationship between interorganizational network structures and activities and measures of effectiveness" (p. 414). The methods to accomplish measuring this end-to-end performance have traditionally been to assess aggregate outcomes for the population being served, and to examine the costs of the service within a given community (Provan and Milward 2001).

However, for EMS, an aggregate measure of effectiveness relating to the responsiveness of the service would need to start with the time of the incident and would then include the average elapsed time until notification of emergency authorities (911 phone call), the dispatch of resources, the arrival of services to a scene, the delivery of medical or safety related services to citizens involved in an incident, and finally definitive emergency care (Horan and Schooley 2004). Mobile EMS evaluations today do not feature such end-to-end measurements; rather they focus on individual agency performance. The research reported here aims to fill this gap in EMS performance conceptualization and measurement.

Mobile EMS in a Rural Context

The substantive context for this research project is mobile emergency response services in rural settings. The rise of mobile telecommunications has become an important means to deliver faster emergency service. In 2003 there were more than 158 million wireless users making about 198,000 emergency calls a day across the United States. A wireless 911 phone call can shave valuable minutes from the time otherwise required for a caller (or motorist aid) to find a conventional phone to access emergency medical services. This is particularly important in light of the number of fatalities that occur on rural U.S. roadways. Rural fatalities account for approximately 60 percent of all vehicle fatalities in the United States, with even higher figures in predominately rural states (e.g., Maine, 90%; Minnesota, 73%; Montana, 92%; Wisconsin, 75%; Wyoming 84%)(NCSA 2002).

Time matters greatly in responding to highway emergencies, and in rural areas the average time that lapses between a rural crash and the victim's arrival at a hospital is 52 minutes, compared to 34 minutes for an urban crash victim (NCSA 2002). Moreover, these data only address vehicle crashes. Additional fatalities and injuries result from other medical problems that occur on our highways, such as heart attacks and strokes, not to mention other medical emergencies that occur off-highway like recreational accidents. A recent study shows that all rural accident victims are over seven times more likely to die before arrival at a hospital if the emergency medical services' response time is more than 30 minutes (Grossman et al. 1997).

Precise and reliable information also matters greatly during emergencies. Rural responders are often frustrated by incomplete information about the location of the incident, while health care providers can be ill prepared to deal with the nature of the accident if they do not have advance information about the patient's condition and medical needs. Appropriate medical care immediately after injury is critical to reducing an accident victim's chances of physical disability or death, which helps to explain the need to improve system performance by reducing response times and providing more informed responses, especially in rural areas (Hale 1997, NENA 2001).

This research study investigates the role of information systems in enhancing end-to-end performance (time and information) of rural EMS systems. The article reports on empirical findings from the second phase of a multi-year study. During the first phase, an exploratory analysis was conducted on the policy, technology, and organizational challenges related to EMS deployment (Horan and Schooley 2004). The follow-on phase reported in this paper focuses on a multi-method analysis that was conducted based on an embedded case study in rural Minnesota. EMS system performance data garnered from a series of field interviews and site visits with multiple organizations allowed researchers to develop a framework (i.e., a performance ontology) and knowledgebase populated with real performance data. These data were used to perform a preliminary simulation of rural EMS systems performance under normal and crisis conditions for purposes of analysis and evaluation.

Case Study Setting: Brainerd-Baxter, Minnesota

The case study reported here is in the Central Lakes region of Minnesota, specifically the 2800 District of the Minnesota State Patrol. As noted in Figure 4.4.1, this district (herein referred to as "Brainerd-Baxter") includes seven counties: Aitkin, Cass, Crow Wing, Hubbard, Kanabec, Mille Lacs, and Morrison. Located 131 miles northwest of the Twin Cities Metropolitan Area, the region has always been a popular tourist destination and continues to draw tourists, retirees, and people who are seeking employment in the growing area. In fact, Brainerd-Baxter's population of 18,733 doubles during the summer months due to tourists, particularly on popular weekends (Brainerd Lakes Area Chamber of Commerce 2003). The seven counties are home to almost four percent of the state's population and cover ten percent of Minnesota's land. The cities of Brainerd and Baxter are adjacent to each other and they are the most populous of the region. Brainerd is also the county seat of Crow Wing County.

The principal organizational entities for dealing with rural mobile emergencies in Minnesota are the Transportation Operations and Communications Centers (TOCCs). TOCCs are a recent cooperative effort between the Minnesota Department of Transportation (Mn/DOT) and the Minnesota State Patrol (MSP) for the purposes of sharing resources like buildings and IT systems, and creating interorganizational efficiencies. TOCCs serve as regional centers throughout the state for transportation operation services including incident and emergency response, multi-agency dispatching and fleet management, interagency communications, collection and dissemination of road conditions and closures, traffic management, and potentially, integrated transit operations (Mn/DOT 2003). TOCCs are an integral part of Minnesota's enhanced 911 (e-911) infrastructure and have a working relationship with PSAPs to perform coordinated emergency response. TOCCs respond to almost all highway related emergencies and wireless 911 calls coming from motorists on highways, rural roads, and rural areas. In this study, the TOCC examined is located in Baxter, while the county PSAP is located in Brainerd.

4.4.2 Research Approach

This case study provides the setting to address three research questions. First, what are the essential end-to-end performance elements of rural EMS systems? Second, what is the composite performance picture that emerges from end-to-end performance data, including under simulated conditions? Third, what are the local contextual factors surrounding interorganizational end-to-end rural EMS system performance? Two methods were used to address these research questions. First, the research team conducted field interviews, site visits, and follow-up

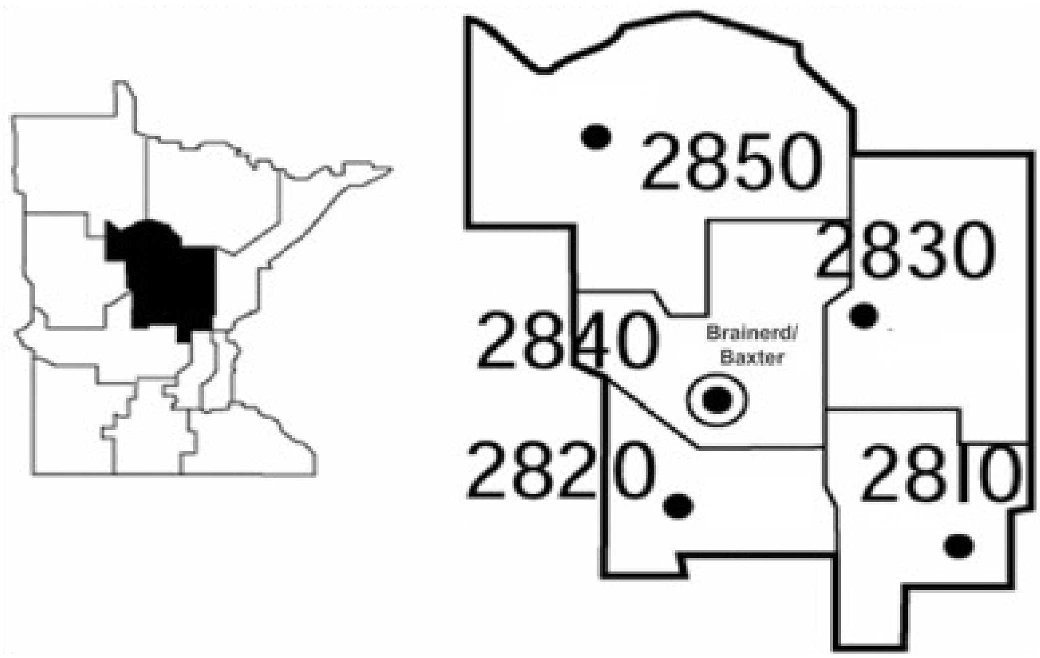


Figure 4.4.1 Minnesota State Patrol District 2800, Branerd-Baxter region

Source: Minnesota State Patrol, 2004

interviews, and gathered data. Second, the researchers created an overall data framework using case study data, and performed a simulation based on these available data.[ii]

Field Interviews and Data Gathering

The research team made four site visits and interviewed representatives at each site.[iii]During each meeting, the research team provided an overview of the study objectives and asked interviewees to comment on their organizations' roles in each step of the EMS process as well as the availability of data documenting the performance of their activities for every relevant step. The interviews provided a local perspective on the rural emergency response process from end-to-end and on interorganizational dynamics and relationships that characterize the end-to-end process. Along with the qualitative interviews, this study used data for the year 2002 from four sources: the national Fatality Analysis Reporting System (FARS), Minnesota FARS, the Baxter TOCC, and the Crow Wing County PSAP. These data were identified during the field visits and obtained through follow-up connections. In most but not all cases, the data were available in electronic formats.

Performance Data Analysis and Simulation

Once the data were obtained, they were organized into a data framework (technically, a factual performance ontology) that portrayed the end-to-end nature of the case study system for the year 2002. These data were then drawn upon to conduct a preliminary simulation using the ARENA simulation program.[iv]This simulation involved sampling the 2002 data that were specifically related to call volume, dispatch processing, and response modes, and then simulating the range of response times and emergency services wait times, or queues, that could occur under normal and disaster conditions. For the disaster scenario, the call volumes were increased to explore the subsequent impacts on overall system performance. (Additional details on simulation parameters are provided in the simulation findings section.)

4.4.3 Results

The field interviews and subsequent data provided an array of information about both the general EMS process and how it operates in the Brainerd-Baxter context. By reviewing the EMS process with local representatives, four essential elements were identified: 1) mayday call, 2) call routing, 3) dispatch, and 4) response (see Figure 4.4.2). The following sections describe these elements and report related performance information that was available about the Brainerd-Baxter setting.

Call Volumes and Dispatch Consequences

The mobile rural highway EMS process begins when a mayday call is made (Step 1) and routed to the TOCC (Step 2). In Phase I of wireless e-911 systems, when a wireless subscriber initiates a call, the closest tower picks up the signal. The wireless service provider's network has a switching center that works much like the switches on wireline calls, reading the digits and forwarding the calls accordingly. The wireless service provider must program every tower to immediately send any 911 calls to the appropriate 911 tandem. In addition, in the case of wireless e-911 Phase I, the wireless subscriber's phone number is sent along with the wireless signal. In the case of deployed wireless e-911 Phase II, the emergency call receiver is provided with the subscriber's callback number, and his/her latitudinal and longitudinal location within a few hundred feet.

In the Brainerd-Baxter region only about a third of the region is equipped with the e-911 Phase II upgrade (Minnesota Statewide 911 Program 2004). For all 911 calls, the Baxter TOCC and county PSAPs mutually decide which agency responds to particular calls. For example, in Crow Wing County a wireless call is routed either to the PSAP or the TOCC depending on the location of the originating call's cell tower. Generally, the PSAP takes calls that are in close proximity to towns whereas calls near state highways are routed to the TOCC.

In terms of performance data, the Baxter TOCC did not have a record of how many calls were received as their systematic recording begins once a dispatch is initiated (called an Initial Compliant Report or ICR). In 2002, the Baxter TOCC documented a total of 7,215 ICRs. Based on parameters and estimates provided by Baxter TOCC staff in terms of number of calls per ICR and percent mobile calls, a rough estimate of annual calls is 21,745 (see also Horan et al. 2004). TOCC representatives noted that each ICR is associated with between 1 and 20 calls.

Follow-on interviews with Baxter TOCC representatives provided additional perspective as to the impact of these mayday calls on TOCC performance. With specific regard to the rise of mobile calls, TOCC representatives noted that while they are now generally made aware of accidents sooner because of the ability of victims or witnesses to call 911 using mobile phones, they also have to deal with the "over-calling" of an accident during busy periods—that is, they can receive so many (passerby) calls that the dispatch center can be overwhelmed. A second somewhat unintended impact is that they receive "stray" wireless 911 calls that they cannot validate over the phone because of a bad connection or a 911 call with no voice communication, and as a consequence they must send a state trooper out to verify the potential incident. Unfortunately, there is no management information system that records these incidents.

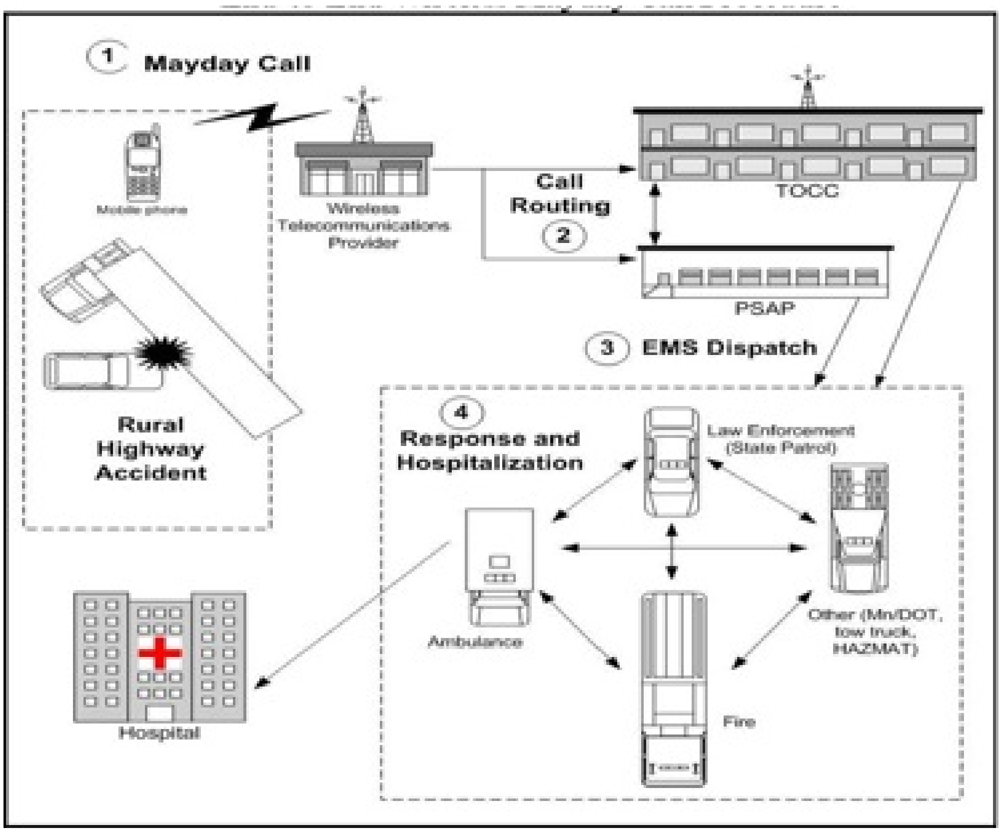


Figure 4.4.2 End-to-End wireless mayday call procedure

Response Times and Fatalities

In terms of response, the general process is that appropriate police, EMS (ambulance), fire and towing services are dispatched depending on the nature of the incident (Step 3). As the situation warrants, the victims are then transported to the appropriate medical or trauma location (Step 4).

From the perspective of the Baxter TOCC, once the dispatch is successfully made the process is in the hands of the EMS providers. While local representatives interviewed were not aware of overall response times, such parameters are available through the FARS database, a database that tracks response times for those highway accidents involving fatalities. According to the FARS data, on average 60.5 minutes passed between the time a mayday call was made and the victim(s) arrived at the hospital in the Brainerd-Baxter region compared with 33 minutes in urban Minnesota during 2002 (see Table 4.4.1).

Table 4.4.1 EMS notification and response intervals in Minnesota, 2002

Interval (in minutes)	Urban Areas		Rural Areas		MSP-D2800	
	Mean	Range	Mean	Range	Mean	Range
Accident/EMS Notification	2.1	0–28	5.8	0–275	4.9	0–61
Accident/EMS Arrival at Scene	10.3	1–73	21.3	1–285	20.8	2–94
Accident/EMS Arrival at Hospital	33.0	5–86	51.8	12–197	60.5	17–197

Source: FARS and Minnesota FARS, 2003.

The Minnesota Department of Public Safety's Office of Traffic Safety reported 3,325 crashes in the Brainerd-Baxter region in 2002 that resulted in fatalities, injuries, and property damage. Almost half of these crashes occurred in Crow Wing County. According to the Minnesota FARS data source, there were 71 crash-related fatalities in the Brainerd-Baxter region—almost 15 percent of all rural fatalities in Minnesota for that year.

As mentioned at the outset, the Brainerd-Baxter area's population surges during special events in the summer. As a consequence, the number of crashes and other incidences also increase. Figure 4.43 shows the number of ICRs, crashes, and fatalities for the following weekends in 2002: the fishing opener, Memorial Day, Fourth of July, the average number of crashes for four major race (BIR) weekends, Labor Day, the summer weekend crash average, and the crash average for all other weekends. Of note, there are 35 more crashes over the Fourth of July weekend than during an average weekend. Interviews with local representatives also revealed a strong implicit knowledge about the spikes on accidents during these periods, though there was no systematic means to summarize the nature and magnitude of the elevation.

As referenced in the findings above, an integrated database of end-to-end performance is not available for emergency managers in the Brainerd-Baxter area. The best that can be done presently is to assemble a "composite sketch" for the year 2002 based on the assorted data assembled through the case study. Figure 4.4.4 presents this data from an end-to-end perspective, that is, from the initial emergency call through response and hospitalization or fatalities.

Once this composite picture and supportive data are available, it becomes possible to link the

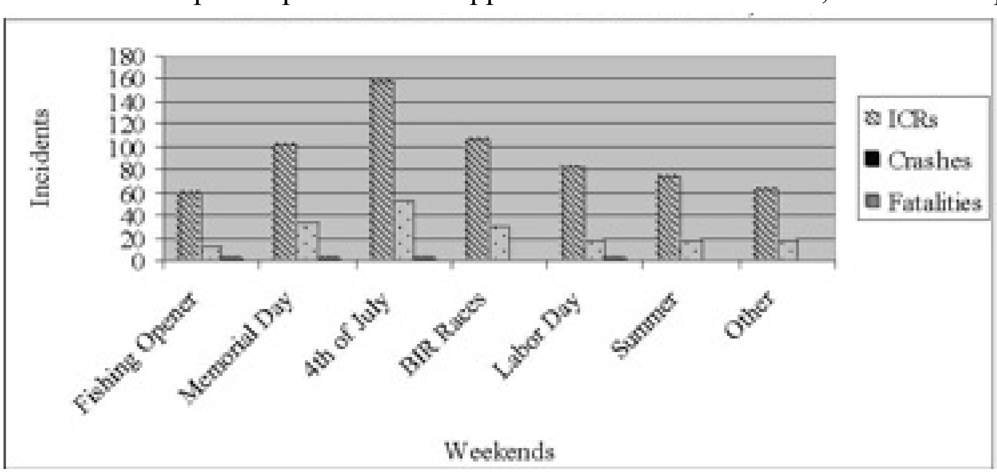


Figure 4.4.3 Brainerd-Baster weekend incidences, 2002

Source: Baxter TOCC, 2003

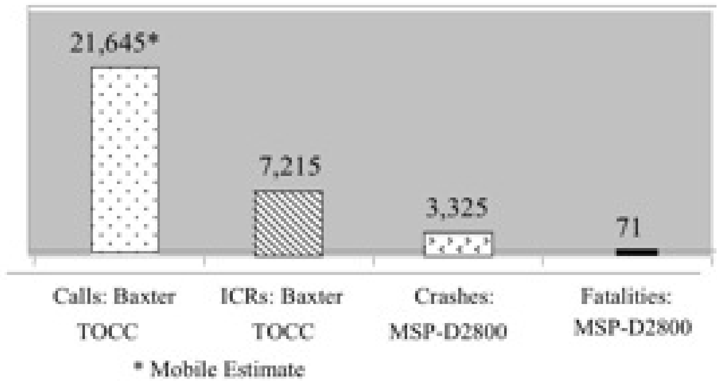


Figure 4.4.4 Factual ontology and knowledgebase data (2002)

processes together to understand how the system is performing. One tool for assessing end-to-end rural EMS systems performance is computer simulation because it allows detailed research into the performance of proposed or existing systems by modeling the behavior of key processes and entities of those systems (Kelton et al. 2004). Computer simulation also enables detailed sensitivity analyses of various scenarios such as the potential response of the EMS system to catastrophic conditions. Several major functions of the Baxter TOCC and regional PSAPs were modeled in the simulation, including call generation, TOCC-activated emergency services, and PSAP-activated emergency services.

To conduct the simulation, a judgmental sample of ICRs was extracted from the raw 2002 data beginning with 36 days sampled at 10-day intervals to limit overrepresentation of specific weekdays. In order to better capture the expected range of daily call frequencies, data from several days with expected high emergency activity (e.g., July 4 and December 31) were included in the sample. This sample yielded a range of 8 to 34 incidents per day with a mean of approximately 14 incidents per day. The distribution of call times was examined for the entire data set of 7,210 incidents, yielding an approximately uniform distribution of calls throughout the day with the exception of marked decrease in incidences from approximately 3:00 AM to 6:00 AM.

Once this basic data set was determined, two scenarios were created: one normal and one disaster. For the normal scenario, each ICR is associated with one to twenty calls, with an approximate mean of five calls per ICR, resulting in a uniform distribution with extremes of 42 and 180 minutes. This profile forms the basis of incident generation for the normal case. Calls per incident were generated based on an estimated distribution of 20 percent of incidents generating 2, 4, 6, 8, or 10 calls, yielding a median of 6 calls per incident. Each of the regional PSAPs directly dispatches ambulance, fire, city police, or tow trucks, following the same rules employed by the TOCC emergency services. During the normal scenario, events and calls are generated according to the rules in the preceding discussion of the normal. Key results of this simulation scenario run for a 14-day period and are included in Table4.4.2. A disaster scenario was also created to simulate the effects of crises situations on the EMS system, with insurgence of an additional incident every hour for a 48-hour pulse. The disaster simulation scenario repeated the normal scenario duration of 14 days.

Table 4.4.2 Representative outcomes of normal and disaster scenarios

	Normal Scenario–14days			Disaster Scenario–14days(a)		
Total Incidents Emergency Calls – TOCC Dispatcher 1 Emergency Calls – TOCC	187 883 264			228 1,029 353		
	Normal response hours)	Case unit	Activity per response unit per incident hours	Disaster response hours)	Case unit	Activity per response unit per incident hours
Response Unit	Min	Mean	Max	Min	Mean	Max
Baxter Dispatcher 1	0.02	0.08	0.16	0.02	0.08	0.16
Crow Wing County Sheriff	0.54	1.16	1.90	0.63	1.37	2.20
Baxter Tow	0.61	1.18	1.95	0.63	1.39	2.28
Mn/DOT	1.31	3.64	7.61	1.62	4.58	9.09
Morrison Fire	1.83	3.47	7.09	1.95	4.22	9.07
Morrison Ambulance	0.63	1.01	1.56	0.71	1.42	2.31
Response Unit	Normal response hours)	Case unit	Activity per response unit per incident hours	Disaster response hours)	Case unit	Activity per response unit per incident hours
Baxter Dispatcher 1	0.0	0.20	0.92	0.0	0.25	1.33
Baxter Tow	0.0	0.06	1.18	0.0	0.27	7.64
Crow Wing County Sheriff	0.0	0.11	2.19	0.0	1.90	11.35
Mn/DOT	0.0	0.15	3.54	0.0	6.18	26.26

Note (a): The Disaster Case includes a 48-hour disaster period of one additional incident per hour.

As indicated in Table 4.4.2, the normal scenario resulted in a total of 187 incidents, with 883 calls to Dispatcher 1 and 264 calls to Dispatcher 2. The disaster scenario increased the total number of incidents to 228, with a corresponding increase of calls to the two Dispatchers of 1,029 and 353. Each incoming event identified by the Dispatcher is followed by various categories of emergency response service activity, such as sheriff unit response, ambulance unit, or Mn/DOT highway response crew. Minimum, mean, and maximum durations for the activities of each EMS unit are presented in Table 4.4.2. In addition, the table also shows the minimum, mean and maximum queue durations, which are the periods that individual emergency events must wait for initiation of the assigned emergency service.

Several preliminary observations can be drawn from these results in Table 4.4.2. First, a number of system bottlenecks can be clearly identified. In the model, TOCC dispatchers spend a minimum of 1.2 minutes (0.02 hours) handling a dispatch and an average of 4.8 minutes (0.08 hours) per dispatch. The Baxter TOCC dispatcher may be able to begin dispatching the necessary services immediately but the normal scenario may take up to an hour to dispatch the needed services, due to increased event frequency, duration of dispatcher activity per event, and capacity constraints. Moreover, in the visual model it becomes clear that this dispatching event is tightly coupled with response service as demonstrated by subsequent queues in Morrison Ambulance. Conversely, the long queues in dispatching Mn/DOT vehicles are not coupled with any subsequent events.

Second, it is clear that the disaster scenario has a dramatic impact on the overall performance of the EMS system leading to extended periods of non-response or delay. The disaster scenario led to a 21 percent increase in total emergency calls over the normal scenario. However, the disaster scenario shows delays of up to 80 minutes (1.33 hours) for the actual emergency request to be logged and dispatched. Emergency service provision shows similarly extended delays in the disaster scenario with maximum wait times for towing services increasing from about an hour to over seven hours, and maximum wait times for Mn/DOT services increasing from 3.5 hours to more than a day.

Local Contextual Factors

Management information needs. The field visits and interviews focused on obtaining data relative to performance. However, it quickly became apparent that the availability of performance data varied considerably and its use was quite ad hoc.[v]While simulations can provide researchers with a tool for assessing linkages among rural EMS responses, local field managers need a dynamic information system that can monitor and guide EMS response. Brainerd-Baxter TOCC and Crow Wing County PSAP representatives expressed such a need for quality information management systems. For instance, they noted that it would be helpful for TOCC management to know the number of calls that the TOCC received over a given time period that were managed there or transferred elsewhere. In addition, Baxter TOCC officials would like to track information on call volume, calls for service, and the outcome of 911 calls. TOCC officials are in the process of upgrading to a computer aided dispatch system but the implementation of the system has taken more time than anticipated, straining the record-keeping capabilities of the current system. In this sense, there appears to be a disconnect between information systems coordination and service coordination whereby the former is lagging behind the latter. Despite this gap, local managers expressed confidence in being able to use their intuitive knowledge of what happens in the region to manage their processes successfully.

Inter-organizational cooperation. As noted at the outset, TOCCs are the result of a recent cooperative effort between the Minnesota Department of Transportation and the Minnesota State Patrol for the purposes of sharing resources (buildings, dispatch systems) and creating inter-organizational efficiencies. Within the Baxter TOCC, MSP and Mn/DOT are in constant communication with each other. Mn/DOT expects MSP to provide them with real-time information and MSP expects Mn/DOT to be proactive in maintaining adequate roads. For example, when MSP notices a problem on the road they call Mn/DOT for emergency repair. However, there is some disconnect as MSP runs on a 24/7 schedule, whereas Mn/DOT keeps conventional business hours. In order to avoid confusion during emergency response, the two organizations focus on different aspects of the emergency. According to a Mn/DOT representative, this means that MSP focuses on the time-critical response, while Mn/DOT takes responsibility for non-critical items. While Mn/DOT and MSP in the Brainerd-Baxter area work together to eliminate redundancies and share responsibilities, there is infrequent interaction between the two organizations during the actual 911dispatch and response. An example of an interaction is when Mn/DOT maintenance personnel are asked to investigate the 911scene when they are closer to it than the closest MSP officer. In general, then, the MSP (vis-à-vis Mn/DOT) leads in reporting accidents and receiving dispatches, but does not have access to an overall profile of the EMS process. Moreover, MSP representatives noted that there was not a forum for sharing EMS performance information across the full range of organizations involved—that is, not only Mn/DOT, but also emergency response and health care providers.

Policy issues. Funding and financial constraints are the significant inhibitors of enhancing EMS systems. Interview responses highlighted the budgetary constraints confronting the Brainerd-Baxter region's EMS system and these concerns were also reflected in the recent 2002 Annual Report of the Department of Administration (2003). This report found three financial threats to the Minnesota Statewide 911 program: 1) if the call volume is greater than projected, program funding will be inadequate; 2) the costs of implementing e-911 services are great; and 3) maintaining and improving 911 services is prohibitive due to increasing costs. All interviewees stressed that strategy and vision for financing TOCC operations in the long run are needed, especially considering the rate at which technology is changing. Despite an increase in call volume, the Baxter TOCC has the same number of radio operators that it had five years ago and new e-911 dispatching technologies have not been fully implemented. Finally, it should be noted that local representatives felt it was important to note the political quagmire of having many PSAPs (119—one for each county) in a state the size of Minnesota. While such a situation responds to local political interests in having local jurisdictional control over emergency dispatch and response, it is viewed in some quarters as a problematic drain on 911 resources.

4.4.4 Discussion

The case study reveals that while local EMS providers may have an intuitive understanding of how the entire system performs, there is a lack of systematic data to support, confirm or refute perceptions about overall performance. With 60 percent of the highway fatalities occurring in rural areas, there is a continuing need to find ways to improve rural emergency services, and understanding and improving the efficiency of the EMS component is a fundamental step in this direction. This research aims to serve this need as it explicates "end-to-end" performance dimensions and data elements that can be used to monitor, plan, and simulate performance. While some research strands have investigated performance measurement and evaluation in "private" interorganizational business processes for the purposes of process improvement, very little research has been conducted along these lines for public interorganizational networks such as EMS.

It is important to fill this gap in interorganizational, time-critical performance research and this area could be further investigated by expanding this study to collect additional data at a national level. The extant data and simulation analysis of the case study provided initial examination of several important features of rural wireless EMS systems and the flow of information from end to end. While the principal motivation for conducting the simulation was to conduct a "proof of concept" in terms of the ability to simulate rural EMS performance based on extant data, the results raise items for further analysis (including simulation). On this score, the next step in building an analytical model in support of EMS performance would involve calibrating the model, including normal and disaster scenarios, to help identify the nature and the range of interactions among jurisdictions.

From a systems performance perspective, dealing with crises or extreme events can be a pivotal test of overall system management capability (Horan and Sparrow 2004). Perrow (2000) describes the need to "anticipate, prepare for, respond to, and, when possible, even prevent extreme events." From such a perspective, "Extreme events emerge as a powerful focus for organizing research activities that can advance scientific knowledge and directly benefit society (p.2)."

In the post 9/11 environment, public security and safety (i.e., Homeland Security) during large-scale, catastrophic, extreme events has become a major policy concern. One critical aspect of this enterprise is the ability of first responders (e.g. EMS) to ramp-up efficiently (Arens and Rosenbloom 2002, Bolwage 2003). After all, a key component of emergency management and clear public expectation is timely and appropriate emergency services to victims, whether during extreme or normal conditions (Barbera, Macintyre, and DeAtley 2003).

Given the dynamic nature of extreme events, it is particularly useful to examine naturally occurring "spikes" in emergency (911) requests, as demonstrated in this study by examining 911 requests during holiday weekends. By examining these naturally occurring spikes, a more comprehensive understanding of the time-critical nature of emergency response systems could be achieved, including a better understanding of the role of interorganizational networks. For example, the simulation raises issues about how to manage queues in dispatching during both normal and disaster scenarios, especially for "tightly-coupled" queues such as in dispatching. From the case study, it appears that innovative approaches are needed to handle this surge without major staffing increases. One issue to consider may be remote surge assistance that could be offered to rural areas during predictable peak periods as well as on-demand for emergencies. Such an avenue could be explored both qualitatively as well as through simulation analysis. Another interesting step would be better assess and plan for long distance responses in rural areas. Technology can assist in beginning diagnosis and treatment sooner, thereby mitigating (some of) the travel time associated with far off rural accidents. But such systems need to be more fully developed and tested.

Finally, while it is analytically valuable to understand "end-to-end" performance of rural mobile EMS systems, the local contextual factors of the case study serve as a reminder that the system is operating in a human, organizational and policy context. Just as the performance systems need to be understood in detail, so too must one understand the organizational, fiscal, and policy constraints that affect the ability of localities to launch and support innovative EMS information systems. It is one thing to understand the current "end-to-end" performance, it is quite another to have the organizational, policy, and fiscal leverage to enact improvement. Performance knowledge is helpful, but in the end, it is but an analytical contribution to an important but thorny transportation safety and emergency situation. A fully effective response would undoubtedly encompass fiscal support, organizational changes, technological innovation, and the innovative use of information systems to monitor and contribute to improved response times and reduced fatalities.

4.4.5 References

1. Bollwage, E. (2003, June 24). *Testimony of J. Christian Bollwage Mayor of Elizabeth, New Jersey on Homeland Security (*Testimony before the Democratic Task Force on Homeland Security, U.S. House of Representatives): Washington
2. D.C. Retrieved August 24, 2003.
3. Brainerd Lakes Area Chamber of Commerce (2003). Retrieved February 4, 2005 from: http://www.explorebrainerdlakes.com/
4. Burghard, P. (2004). "The Combined Systems Point of View in Crises Management." In *Proceedings of the First International Workshop on Information Systems for Crisis Response and Management ISCRAM2004 (*Eds. B. Carle and B. Van de Walle), Brussels May 3–4, pp. 51–56.
5. Cellular Telecommunications Industry Association (2004). *Wireless 911 and Distress Calls*. Retrieved Oct. 22, 2004 from: http://www.ctia.org/public_policy/statistics/index.cfm/AID/216.
6. Davis, S. (2002). "Virtual Emergency Operations." *Risk Management*, 49(7), pp. 46–52.
7. Dawes, S., Cressell, A., Cahan, B. (2004). Learning from Crisis: Lessons in Human and Information Infrastructure from the World Trade Center Response. *Social Science Computer Review*, *22*(1), 52–66.
8. Department of Administration, Minnesota Statewide 911 Program. (2003). "Statewide 911 Emergency Telephone Service Program Report." St. Paul, Minnesota.
9. Graves, R. (2004). "Key Technologies for Emergency Response." *In Proceedings of the First International Workshop on Information Systems for Crisis Response and Management ISCRAM2004 (*Eds. B. Carle and B. Van de Walle), Brussels May 3–4, pp. 133–138.
10. Grossman, D.C., Kim, A., Macdonald, S.C., Klein, P., Copass, M.K., Maier, R.V. (1997). Urban-rural differences in prehospital care of major trauma. *Journal of Trauma, 42*(4), 723–729.
11. Hale, J. (1997). "A Layered Communication Architecture for the Support of Crisis Response." *Journal of Management Information Systems*, 14(1), pp. 235–255.
12. Horan, T., Kaplancali, U., Schooley, B., Burkhard, R., McCabe, D. (2004). "Performance Information Systems to Improve Rural Mobile Emergency Medical Services." In *Proceedings of the Americas Conference on Information Systems*, New York, NY, August 5–8, p. 2098–2109.
13. Horan, T. and Schooley, B. (2004). "Interorganizational Emergency Medical Services: Case Study of Rural Wireless Deployment and Management." *Information Systems Frontiers*. (Forthcoming).
14. Horan, T. and Sparrow, R. (2004), "Managing Digital Infrastructures," In Zimmerman, R. and T. Horan (eds.)*, Digital Infrastructures,* London: Routledge Press, London.

15. Kelton, W.D., Sadowski, R.P., and Sturrock, D.T. (2004). *Simulation with Arena: 3rded*. New York, NY: McGraw Hill.

16. Minnesota Department of Transportation (2003). Guidestar/TOCC Program, Retrieved November 27, 2003 from

17. http://www.dot.state.mn.us/guidestar/projects/tocc2.html.

18. Minnesota Statewide 911 Program. (2004). Retrieved September 15, 2004 from http://www.911.state.mn.us/

19. National Center for Statistics and Analysis (NCSA), U.S. Department of Transportation. (2002). *National Center for Statistics and Analysis Fatality Analysis Reporting System (FARS) Web-Based Encyclopedia*. Retrieved July 21, 2003 from http://www-fars.nhtsa.dot.gov.

20. National Emergency Number Association (NENA). (2001). *Report card to the Nation: The effectiveness, accessibility and future of America's 911 service*. Columbus, Ohio: The RCN Commission and National Emergency Number Association.

21. Perrow, C. (2000). Extreme Events: A Framework for Organizing, Integrating and Ensuring the Public Value of Research, Paper prepared for Extreme Events: Developing a Research Agenda for the 21st Century, Boulder, CO.

22. Provan, K. G., & Milward, H. B. (2001). "Do Networks Really Work? A Framework for Evaluating Public-Sector Organizational Networks." *Public Administration Review, 61*(4), 414–423.

23. Turoff, M., Chumer, M., Van de Walle, B., and Yao, X. (2004). "The Design of a Dynamic Emergency Response Management Information System (DERMIS)." *Journal of Information Technology Theory and Application*s (2004) 5:4, 1–36.

4.5 Time-Critical Services

Emergency medical services have never been more ready for the implementation of time-critical interorganizational information services for the public good.

Thomas A. Horan, Ph.D.
School of Information Science
Claremont Graduate University
Claremont, CA 91711, USA
Tom.Horan@cgu.edu
Fax: +1 909/621 8564

Benjamin Schooley, MIS
School of Information Science
Claremont Graduate University
Claremont, CA 91711, USA
Ben.Schooley@cgu.edu
Fax: +1 909/621 8564.

4.5.1 Introduction

The folks of Haleyville, Alabama (population 4,000) understood in 1968 how a communications system could help during an emergency. For on February 16 of that year, the first-ever 911 phone call was placed by Alabama Speaker of the House Rankin Fite from Haleyville's City Hall to U.S. Rep. Tom Bevill at the city's police station.[vi]

Since that day, emergency responses have become inextricably linked to the nation's telecommunication systems operationally and to the public sector organizationally. Indeed, in the almost four decades since the Haleyville call, the nation's emergency response system has evolved into a sophisticated network of dispatchers and responders and the communications system has evolved into an array of mobile and computer-aided networks. This evolution has occurred within the context of public demands for better and faster emergency medical services (EMS) to minimize the consequences of accidents and other health emergencies.

This article explores the general concept of "time-critical information services," within the specific context of EMS. While private sector-oriented information systems have focused on the critical role of information technology in achieving Just-in-Time (JIT) delivery and improved supply-chain management, our thesis is that similar attention is needed to those public sector services that are also highly time- and information-dependent. EMS represents an illustrative application domain of JIT in public services, where the adage "time is money" translates into "time is lives."

Let's begin with a fundamental question: Why is EMS a time-critical information service? In terms of time, the reason is clear: from the moment a 911 call is placed to a local public safety answering point (PSAP) and answered by an emergency operator, time is of the essence—every moment of delay can significantly reduce an accident victim's chances of survival. For example, the proportion of fatalities from car accidents in rural areas are far greater in number than in urban areas largely due to the additional time required for resources to respond to remote locations. In rural areas the average time that lapses between a crash and the victim's arrival at a hospital is 52 minutes, compared to 34 minutes for an urban crash victim.[vii]The result is that fatalities on rural roadways account for approximately 60% of all vehicle fatalities in the U.S., with even higher figures in predominately rural states (for example, Maine, 90%; Minnesota, 73%; Montana, 92%).

But the service is not just time-dependent, it is *information*-dependent as well: this information can include a range of data on the location of a caller, the nature of the call, health condition

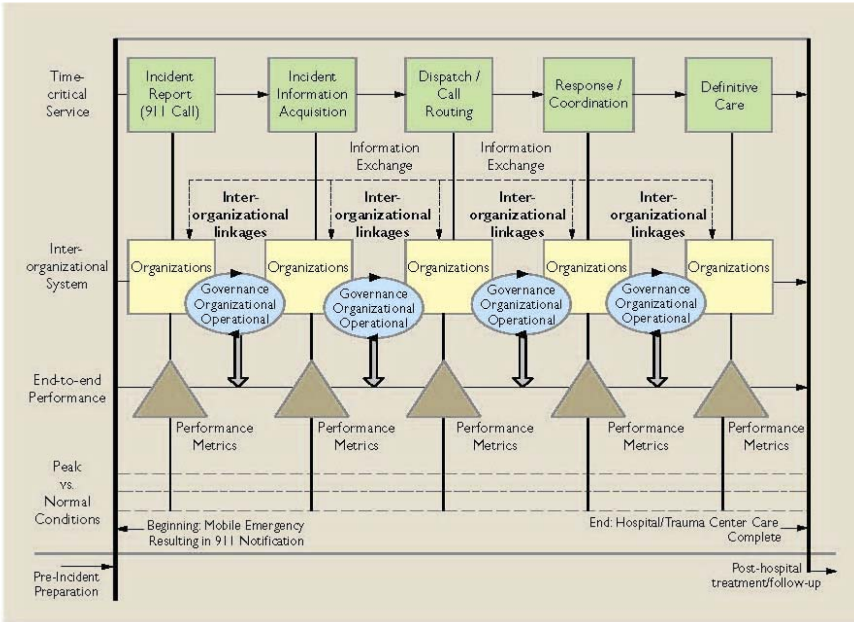


Figure 4.5.1 Time-critical information services dimensions (adopted from COMCARE).

of an injured person, or health data on biological constitution (for example, blood type) of the injured party. It can also include information used by various organizations to guide service operation and performance, such as what services have been rendered to a victim. The challenge for information and computer sciences is to devise new approaches and systems that facilitate rapid use of accurate information for time-critical use in EMS and related public services.

Over the last three years, our research team has investigated the performance of rural mobile EMS systems. This research has been conducted in multiple phases, encompassing field visits, interviews, focus groups, and performance data analysis and simulation. More recently, we have initiated a multi-tiered study (sponsored by the National Science Foundation) to refine our conceptual model and conduct empirical case studies. This analysis has led us to identify several features we think are important to time-critical information services. We expound upon these concepts here, as well as provide a summary illustration in Figure 4.5.1.

4.5.2 Time and Information Linkages

Time efficiency has been a primary objective for computer and information professionals for decades. This focus has created concepts such as JIT and business process reengineering (BPR), which have become central to private sector business operations and information technology planning. However, the issue of time is less well understood in the e-government domain, even though it can be a critical performance feature of public service. For many services, time matters in terms of days and weeks, for example, time savings possible from online tax submission. Or, in the case of determining disability or eligibility benefits, time is measured in months and years. In the case of EMS, time is measured in minutes and seconds. In short, time matters, even for government services.

As illustrated by the top row of Figure 4.5.1, a point of departure for examining EMS is the linear sequence of events. In our research, we have looked closely at the time-dependent course of a 911 call, tracking the “handoff” from one organization to another to understand the performance of the system [4]. The EMS service typically begins with a consumer action (placing the emergency call), involves the private sector (telecommunications service provider) connecting and delivering the call, the public sector (PSAP) receiving and dispatching the call, the private and/or public sector (ambulance service, fire, law enforcement) providing first response, transport and health care services, and finally, either a public, private, or not-for-profit sector hospital or trauma center delivering appropriate health care services.

New technologies provide technical platforms to enhance and alter this sequential process. For example, the rise of mobile telecommunications has become an important means to deliver faster emergency service. At mid-year 2006 there were more than 219 million wireless users making about 240,000 911 calls a day across the U.S.[viii] A wireless 911 phone call can shave valuable minutes from the time otherwise required for a caller (or motorist aid) to find a conventional phone. However, there are still multiple technical issues to resolve. For example, location information is still primarily transmitted by voice despite the large push toward advanced e-911 systems to provide geospatial data along with a mobile 911 call. As will be noted in the Virginia example here, the opportunity now exists to integrate location and other data in a manner that alters the traditional sequential handoffs by allowing the data to be available to all parties in the emergency response chain at any time in the process.

Related to the timeliness of information is the quality of the information. Even in the early days of the 911system, it was important to get accurate information about the emergency. Conveying voice information on the location and nature of a distress call proved valuable to response agencies to help prepare for the unique circumstances of each individual incident. As more information becomes digitized, a greater degree of precision in locating and responding to emergencies is possible. For example, GPS-enabled vehicles can provide dispatchers with up-to-the-second visual data about their location. However, Pettersson and colleagues [9] note how such emergency information needs to be visible "at a glance" and in a manner that facilitates collaborative problem solving by dispatchers amidst an array of available data. In this sense, information quality is about having the right information at the right time in the right format.

One of the key problems of governmental services is that information typically travels serially and sequentially, from one processing unit to the next, often with time-consuming feedback loops when incomplete or inaccurate information is detected. New technologies, such as computer-aided dispatch, can establish technical interorganizational linkages, but do not ensure a comprehensive and reliable relationship between organizations during dynamic emergency response situations [11]. For this, a socio-technical approach is needed that integrates an organizational perspective into technical solutions. For EMS, this perspective needs to emphasize the linkages across organizations.

4.5.3 Interorganizational Systems (IOS) Cooperation

A time-critical information service is typically interorganizational in nature—there is a handoff from one agency/organization to another. For private sector businesses, JIT, BPR, and related concepts have for many years been extended beyond the boundaries of a single organization. Information systems that extend across organizational boundaries have become both a complex technical undertaking as well as a challenging interorganizational phenomenon as it includes a network of information systems and organizations, across supply chains, to improve business processes with partners, vendors, suppliers, distributors, and sales channels. Over two decades of experience has demonstrated that such complex undertakings necessitate a great deal of organizational understanding and change [3].

For emergency response, new technologies can be developed and implemented to enhance organizational and interorganizational cooperation, but this does not negate the need to understand unique and varying characteristics of response organizations. Our research has examined, for example, interorganizational handoffs between state patrol agencies, departments of transportation, fire agencies, emergency medical service providers, and health care facilities [5]. This examination revealed how cultural differences between agencies can affect cooperation. For example, we found state patrol agencies to be more service-oriented and more hierarchical with their information flows than transportation departments, and this difference required ongoing attention to ensure information sharing.

A key finding at our recent expert workshop was the need to focus on the multilevel linkages between cooperating EMS organizations (see Figure 4.5.1, second row from the top). More specifically, the workshop participants noted how inter-linkages occur at the operational, organizational, and governance levels.[ix] In a subsequent case study, a representative from a northern California EMS agency told us about their use of cross-organizational management controls to monitor the performance (response times) of the subcontracted ambulance service provider. This agency's review of performance data revealed that the subcontracted ambulance service had reduced staffing levels, which had resulted in longer response times than were contractually allowed. The situation was corrected and response time performance returned to normal. In this case, cross-organizational information sharing of performance data affected management and governance decision making, which directly impacted the timeliness of emergency services.

End-to-End Perspective

For time-critical information services, total performance is essential to the operation (see Figure 4.5.1, third row from the top). It makes little difference for an operator to dispatch quickly if the ambulance takes a very long time to arrive, goes to the wrong location, or if a receiving hospital has too few physicians on duty to see an incoming patient. The critical descriptor here is "end-to-end."

Measuring performance across service processes and organizations is essential to understanding how public services are delivered to the public, the level of service (timeliness, quality) with which they are delivered, and how the service network can be improved to deliver better services under time- and information-critical circumstances.

The challenge is how to implement this "end-to-end" concept within and across emergency provider organizations. We are taking a step in this direction by creating an ontology and knowledge base to collect performance information about each step in the end-to-end process. In our early case study research in a rural region of Minnesota, we were able to collect incident data from multiple organizations and disparate information systems to account for a year (2002) of performance. We began with an estimated 21,745 mobile 911 phone calls, producing 7,215 EMS responses to 3,325 automobile crashes involving 71 deaths. We found the average ambulance response interval to be 60.5 minutes, (from EMS notification, to arrival at the scene, to arrival to a hospital), which is significantly above the state average for both urban and rural areas [4].

Our work in this area has made clear that a need exists for public agencies to first agree on how to measure performance across agencies and then build performance tracking across new or existing public agency information systems. While public sector information systems are often implemented to address separate silos of a governmental process, the end-to-end nature of time-critical information services facilitates or at least allows for information systems that can report on overall system performance. The real challenge then becomes working across very different organizational cultures (for example, departments of transportation, law enforcement, fire, and ambulance) to achieve a holistic understanding and use of performance information.

Normal and Extreme Performance

It is said that a crisis can define character; in EMS, a crisis can define performance. End-to-end performance is not only a function of system processes but also a function of exogenous occurrences such as storms, natural disasters, or terrorist attacks. Research focused solely on a system's normal behavior cannot fully characterize the full range of possible system dynamics. Extreme events provide unique opportunities to examine systems, including how cooperative organizations function when exogenous variables cause a range of system-wide "stress" issues, including increased service demand that can lead to system overload or even collapse.

That is, extreme events can be a pivotal test of overall system management capability. As Perrow proclaimed "...it is no overstatement to suggest that humanity's future will be shaped by its capacity to anticipate, prepare for, respond to, and, when possible, even prevent extreme events [8]." For this reason, the National Critical Infrastructure Report devotes considerable attention to the use of modeling and simulation of crisis events to learn about system performance and response [12]. In a similar vein, we are using process simulation software to examine end-to-end performance of emergency response systems, including naturally occurring spikes in emergency service requests (such as during holiday periods) [4]. By examining these naturally occurring spikes, a more comprehensive understanding of the variable performance of emergency response systems could be achieved; including the role of new, innovative information systems during normal and peak conditions (see Figure 4.5.1, bottom row).

XML for Time-Critical Services

One example of a next-generation system can be found in the Blue Ridge Mountains of Virginia. Some 38 years after the call from Haleyville, a next-generation data call was made during a test of the Virginia E-Safety Network. The E-Safety Network is an emergency Web services information architecture that facilitates linkages between the various information systems used by law enforcement, dispatch (PSAPs), ambulance, fire, transportation, health care systems, and emergency management. It is based on open, non-proprietary standards in order to facilitate interoperability between disparate systems and agencies and to encourage the development of new applications by private companies competing for a national market [10].

Located in the greater Winchester, VA area, this system was designed based on a distributed computing (service-oriented) architecture utilizing XML-based standards and protocols. It was created to receive and distribute various emergency incident-related XML messages via standard interfaces. The intention from the outset was to support several time-critical functions as illustrated in Figure 4.5.2. For example, a series of notification services would selectively "push" new incident alerts to pre-identified and pre-authorized recipients on the basis of geospatial and other message content through an Intelligent Message Broker (IMB). The IMB integrates message alerts with geographic information system (GIS) data about an incident and sends them to registered users and associated applications.

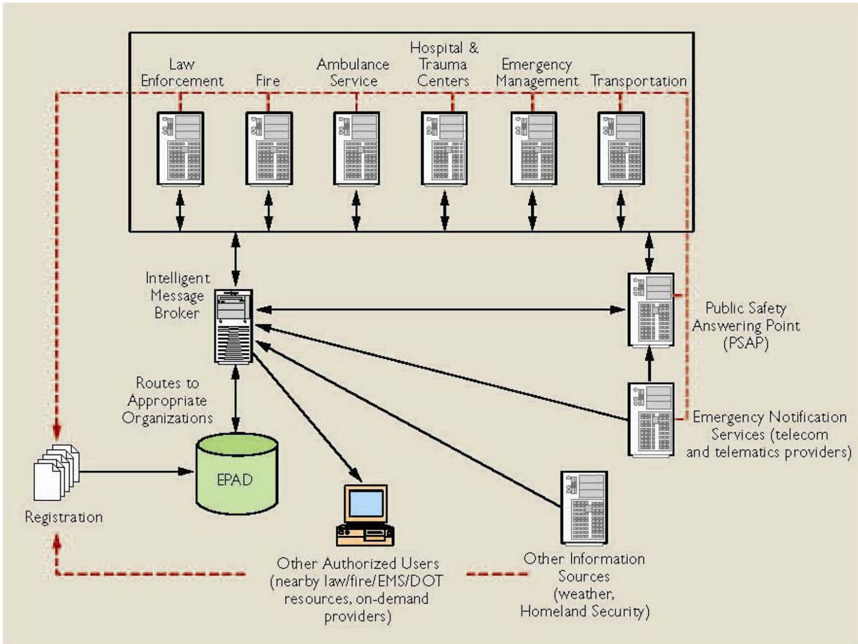


Figure 4.5.2 Virginia's E-safety network

Source: adapted from COMCARE

The IMB tracks the emergency response and facilitates real-time information exchange between response organizations.

Authorized users could also "pull" data to their systems or become data sources to provide critical information to the network. A system that has for many years been completely reliant upon serially transmitting emergency information using voice over two-way radio has now been designed with the capability to transfer data to and from multiple organizations dynamically.

The E-Safety design not only enables timely transmission of key information, it also supports interorganizational coordination and information sharing. This can be organized through a shared registry of authorized emergency response and associated agencies and organizations called the Emergency Provider Access Directory (EPAD). The purpose of EPAD is to facilitate registration with the network as well as data access levels for each registered and authorized organization.

Of course, the development and use of this EPAD directory hinges on interorganizational participation. In this regard, the E-Safety network has been expanded over the past few years to include the 13 hospitals and six counties that make up the Northern Virginia Hospital Alliance.

Looking Ahead

Brown and Hagel [1] have advanced the case for Web services and service-oriented architectures to produce new levels of supply chain performance. Similar to the literature on BPR and JIT, discussions of Web services tend to focus on private sector applications. But the Virginia E-Safety network suggests that such architectures can be of great use in developing time-critical information services for public agencies and services. There have been some recent efforts to develop XML-based standards for the collection, sharing, and reporting of EMS performance data, including a national coalition that has recently tested new data structures in various locales [6]. While these technical efforts are moving forward, there is a parallel need to derive methods and analytic devices that are geared toward developing interorganizational collaboration. In this sense, there is a need to devise “orgware” as well as software.

Fortunately, there is a burgeoning research effort that promises to shed much needed light on the value and functioning of time-critical information services. Beyond our own modest efforts, Turoff, Van de Walle and colleagues [11] have led a series of workshops and literature on the emergency and crisis response topic (see the special section in this issue). Funding agencies such as the National Science Foundation and the Department of Homeland Security have begun to sponsor integrated testing and assessments of emergency response systems. These recent technology-focused efforts benefit from a bevy of prior emergency response studies, a portion of which has been noted in our time-critical concept explication. Moreover, there are a variety of methods that have been applied to better understand dimensions of time-critical services, ranging from test beds [7] to usability testing [2] to field ethnographies [9]. These add to the qualitative and case study-based research that has formed a strong foundation for studying interorganizational cooperation in EMS, such as we have undertaken.

In conclusion, the post-September 11 and Hurricane Katrina environment is one in which practitioner and research communities have become more concerned and involved with emergency and crisis management. Due to advances in XML and related open systems, the technical basis now exists to develop time critical interorganizational information services for the public. The EMS application domain provides an emblematic case where such implementation has the possibility to achieve demonstrable public good in terms of lives saved. Despite the challenges ahead, this is an opportunity for computer and information professionals to innovate through new systems, applications, and research. The factors raised in our framework are but a starting point.

4.5.4 Acknowledgements

This work is supported in part by the National Science Foundation (Grant Nos. 508938 and 535273). The authors also gratefully acknowledge financial support from the U.S. Department of Transportation through the ITS Institute and State and Local Policy Program, University of Minnesota, as well as research support from Richard Burkhard, Ugur Kaplancali, Michael Marich, Denise McCabe, and David Aylward.

4.5.5 References

1. Brown, J.S. and Hagel, J. Flexible IT, better strategy. *The McKinseyQuarterly 4*, (2003).
2. Cai, G., Bolelli, L., MacEachren, A.M., Sharma, R., Fuhrmann, S., andMcNeese, M. GeoCollaborative crisis management: Using maps tomediate EOC-mobile team collaboration. In *Proceedings of the 5thAnnual NSF Digital Government Conference (*Los Angeles, CA, May23–26, 2004).
3. Hammer, M. Beyond Reengineering: How the Processed-Centered Organizationis Changing Our Work and Our Lives. Harper Business, New York,NY, 1997.
4. Horan, T., McCabe, D., Burkhard, R. and Schooley, B. PerformanceInformation Systems for Emergency Response: Field Examination andSimulation of End-To-End Rural Response Systems. *J. Homeland Securityand Emergency Management 2*, 1 (2005).
5. Horan, T. and Schooley, B. Interorganizational emergency medical services:Case study of rural wireless deployment and management. *InformationSystems Frontiers*, *7*, 2.
6. Mears, G., Ornato, J. Drew, E., and Dawson, D. Emergency medical servicesinformation systems and a future EMS national database. *PrehospitalEmergency Care 6*, 1 (Mar. 2002), 123–130.
7. Midkiff, S.F. and Bostian, C.W. Rapidly deployable broadband wirelessnetworks for disaster and emergency response. In *Proceedings of the FirstIEEE Workshop on Disaster Recover Networks*, (June 24, 2002), NewYork, NY; accessed June 27, 2005; www.cwt.vt.edu/research/detail/disaster_response/Midkiff_Bostian_DIREN02.pdf
8. Perrow, C. Extreme events: A framework for organizing, integrating andensuring the public value of research. Paper prepared for *Extreme Events:Developing a Research Agenda for the 21st Century (*Boulder, CO., July7–9, 2000).
9. Petterson, M., Randall, D., and Hegelson, B. Ambiguities, awareness andeconomy: A study of emergency services work. In *Proceedings of ACMCSCW (*New Orleans, LA, 2002).
10. Potter, J., Miller, L., Dubrueler, E., and Dubreuler, A. Interoperabilitynow: Integrating emergency communication and information—A Virginiacase study. *Topics in Emergency Medicine (*June 2004).
11. Turoff, M., Chumer, M., Van de Walle, B., Yao, X. The design of adynamic emergency response management information system. *Journalof Information Technology Theory and Application 5*, 4 (2004), 1–36.
12. White House. The National Strategy for Physical Protection of CriticalInfrastructures and Key Assets, Washington, D.C., Feb. 2003.

4.6 User Perspectives on the Minnesota Interorganizational Mayday Information System

Benjamin Schooley, MIS
School of Information Science
Claremont Graduate University
Claremont, CA 91711
F: +1 909/621 8564
E-mail: Ben.Schooley@cgu.edu

Thomas A. Horan, Ph.D.
School of Information Science
Claremont Graduate University
Claremont, CA 91711
F: +1 909/621 8564
E-mail: Tom.Horan@cgu.edu

Michael Marich, Ph.D.
School of Information Science
Claremont Graduate University
Claremont, CA 91711

4.6.1 Introduction

Information Systems for Emergency Medical Services (EMS) Systems

Nearly four decades ago, the National Academy of Sciences released a path-breaking report titled *Accidental Death and Disability*. The report was largely responsible for the creation of the first widespread 911, Trauma, and Emergency Medical Services (EMS) systems in the United States. One important recommendation in the report charged the medical field and public sector services and agencies to build modern communications systems to enable efficient and effective coordination of services across police, fire, dispatch, ambulance, and hospital organizations (NAC, 1966). As a result, the federal government responded by funding two-way radio infrastructure development projects across the United States. Since that time, emergency medical communications systems in the U.S. have largely remained a “voice-centric” infrastructure.

A little over a decade ago, the National Highway Transportation Safety Administration (NHTSA) (1996) identified the importance and need for a more “data-centric” communications infrastructure for EMS. It discussed a need to update communications systems to better utilize data to improve EMS processes and enable an infrastructure for performance analysis and reporting. This report, and other more recent studies and reports, have concluded that multi-organizational EMS continues to operate without a sufficient research basis to support many of its operational and information systems decisions and thus needs better data collection, analysis, and reporting systems (NHTSA, 2001; McLean et al., 2002; Sayre et al., 2003; IOM, 2006). Such an infrastructure would enable information sharing across the range of emergency response organizations and was thus identified as an important precursor to improving EMS research and system-wide services (IOM, 2006; NENA, 2001; NHTSA, 2001).

4.6.2 Background

Of course the need and motivation to improve EMS systems stems from a long-standing and noble cause: to reduce death and disability and the costs associated with doing so. Significant financial and human resources, for example, are dedicated to reducing traffic related deaths. In 2005 alone, there were over 43,000 traffic related deaths that resulted in an economic cost estimated at approximately $50 billion.[x] Common sense and the general public perception is that a faster emergency response will decrease the likelihood of death or disability consequences for an automobile crash victim. It’s not difficult to imagine a stranded, unconscious driver of a crashed vehicle unable to dial 911. Cell phones used by passing travelers and new automatic crash notification (ACN) systems on board vehicles provide innovative solutions. However, only a few empirical studies have demonstrated that faster emergency response times reduce the likelihood that a trauma patient will experience disability or death consequences. One of these studies found that accident victims are between 5 and 7 times more likely to die if arrival to a hospital exceeds 30 minutes from the time of an automobile crash (Grossman et al., 1997). Another study found that approximately 50% of trauma victims could benefit significantly from faster arrival to a trauma center, while the other 50% would not (Trunkey, 1983). Data from the Federal Highway Administration's Fatality Analysis Reporting System (FARS) shows that over 60% of fatal automobile crashes occur on rural roads largely due to the additional time required to respond and care for a patient located in a far away rural or remote location. The FARS data analysis shows that the length of time for an emergency response, including the time to answer the 911 call, arrive on scene, and transport a patient to a hospital, is approximately 52 minutes in rural areas compared to 34 minutes in urban areas (NCSA, 2006). The logical assumption that is often made is that these longer emergency response times in rural areas result in increased fatality rates.

A concerted focus on timely emergency medical service often translates into what is referred to as a "scoop and run" service, or the idea that the job of an ambulance is to pick up a patient and speed to a hospital as fast as possible with lights and sirens blaring. More recently, there has been a shift towards applying a higher level of quality health care and treatment at an incident scene and during the ambulance ride prior to a patient arriving at a trauma center. This practice is largely motivated by the belief and experience of trauma physicians and care givers that providing a higher level of care at the scene is better for the patient than providing a "scoop and run" service. There exists little empirical evidence in support of or against this perception. As a result, it remains largely unknown whether faster response times and higher quality pre-hospital care improve patient outcomes (Pons and Markovchick, 2002; Carr et al., 2006). There is a need for additional research to better understand the benefits and trade-offs of these two approaches to system improvement. There is a parallel need for more advanced interorganizational information systems to allow for such evaluation to occur.

More research and better data collection systems are needed to better understand the correlation between emergency response times and patient outcomes. And not just from a "fatality" perspective but also including disability consequences, length of hospital stay, severity of injury, and other outcomes associated with community, patient, provider, and payer costs of care delivery. This need for a more evidence based approach to EMS performance evaluation and improvement is one major motivation for developing new end-to-end information systems for EMS.

Table 4.6.1 DOT motivations for building IT for EMS

Minnesota Department of Transportation (DOT) Motivations for Building Information Technology Systems for Emergency Medical Services
• Reducing emergency response times. • Providing data to decision makers to make better resource allocation decisions. • Providing data to EMS and trauma center decision makers to enable higher level of quality care to patients.

The Minnesota Mayday system was originally construed as a working test project motivated by the needs described above, for more advanced IT systems in EMS. The goal was to build an advanced data-centric system in the State of Minnesota for the purpose of reducing emergency response times to automobile crashes. The case study reported herein refers to this innovative technical information system and interorganizational business process change that occurred, and how it influenced interorganizational information sharing. The analysis comes from the perspective of Mayday system users from a series of roundtable discussions and interviews with Mayday stakeholders. The remainder of this chapter proceeds as follows: first we provide a description about the methodology used in this study and the guiding analytical framework. The Emergency Medical Services (EMS) context is then discussed as background to summarize the current state of IT in EMS. We then explain the performance motives for building advanced IT systems and describe the Minnesota Mayday case study. Findings from end-user interviews and focus groups are presented followed by a discussion about the utility, needs, and design considerations for end-to-end EMS systems.

4.6.3 Methodology

The Mayday project implementation was conducted from October 15, 2004 to September 30, 2005. A test project technical evaluation was written and published by the Minnesota Department of Transportation (2006) as well as by Linnell et al. (2006) on behalf of the Minnesota Department of Transportation. It reports test results such as the system's ability to send, route, and receive data; data reliability, throughput and latency performance; and data storage and retrieval capabilities. The research team writing this report did not participate in the implementation and technical evaluation of the Mayday system. Rather, at the conclusion of the Mayday test project, the research team conducted an evaluation to understand benefits and challenges of the Mayday system from an end-user perspective. While the Mayday system project evaluation focused on the technical capability of the operational system to function as intended and designed, this case analysis focuses on end user perspectives and perceptions about the utility, effectiveness, challenges, and future opportunities of the system. Researchers accomplished this through a series of interviews and roundtable discussions with individuals from each participating organization (State Patrol, GM OnStar, Mayo Clinic, Minnesota Department of Transportation).

The importance of extracting end-user needs and perspectives has been articulated in the information systems literature for several years (Fahy and Murphy, 1996; Gunton, 1988; Taylor et al., 1998). Integrating end users into the design of information systems has been shown to increase the creativity of a solution, incorporate a greater degree of specialized organizational knowledge into a solution, and produce additional opportunities for new and innovative strategic information systems (Davenport, 1994; McBride et al., 1997). When the end-user needs are well understood, organizations can expect greater levels of acceptance and diffusion of technology, greater levels of satisfaction, and more effectively aligned systems with organizational needs (Katz and Kahn, 1978; Robson, 1997; Zinatelli et al., 1996). By understanding end-user perspectives, we gain a better understanding about how to integrate technology into the design and operation of socio-technical systems such as EMS. Thus, the need and focus of this chapter on understanding user perspectives about the Mayday system.

The end-user evaluation utilized on-site visits with each participating organization as well as individual interviews and roundtable discussions with participants. Participants included personnel from both management and non-management positions and included call center operators, medical dispatchers, State Patrol officers, paramedics, physicians, hospital administrators, and nurses. Appendix C includes a list of participant job titles and organizations involved. In particular, we were interested in understanding issues related to system usability and improvement, as well as performance implications that affect service timeliness and service quality from end-to-end. The types of questions asked are listed in Appendix C.

The evaluation was conducted in two overlapping phases. The first phase sought to understand the operational Mayday system as described by documentation and users. The analysis utilized business process documentation, Mayday performance data for the year, technical information system documentation, management reports, and performance reports, interorganizational agreements including formal and informal contracts, as well as field notes and supplemental interviews. This data was collected through field visits on location at each participating organization as well as through follow-up phone and email conversations. The organizations visited and positions of persons interviewed are outlined in Table 4.6.2.

The second research phase examined contextual issues about the Mayday operational processes and information exchanges. Interview and roundtable discussion participants were selected based simply on whether they interacted with the Mayday system and whether they were willing to participate. Semi-structured interview questions sought to understand dimensions to information sharing. In particular, we were interested in understanding operational, organizational, and governance issues related to system usability and improvement, as well as performance implications that affect service timeliness and service quality from end-to-end. The intention of the interviews was to understand what conditions inhibit or prohibit information sharing across organizations, the role information sharing (and technology) plays in the delivery of public services, and the role of information sharing to manage interorganizational service performance. Researchers took detailed field notes and summarized observations.

Guiding Framework

The overall study methodology and research process was guided by the time-critical information services model (Horan and Schooley, 2007). This framework was developed as a way to distinguish between different simultaneously ongoing streams of phenomena, some of which are organizational, some of which are performance-based, technological, time-dependent, etc., and frame them into an analytical lens for interorganizational systems (IOS) analysis. The conceptual model includes several levels of analysis for TCIS, both in regard to EMS specifically and other public services generally. These levels, shown in Figure 4.6.1, include (1) the time and information critical elements of a sequential public service process, (2) the interactions and information exchanges across multiple cooperating service organizations that include both qualitative organizational elements as well as "hard" information flow elements, (3) the end-to-

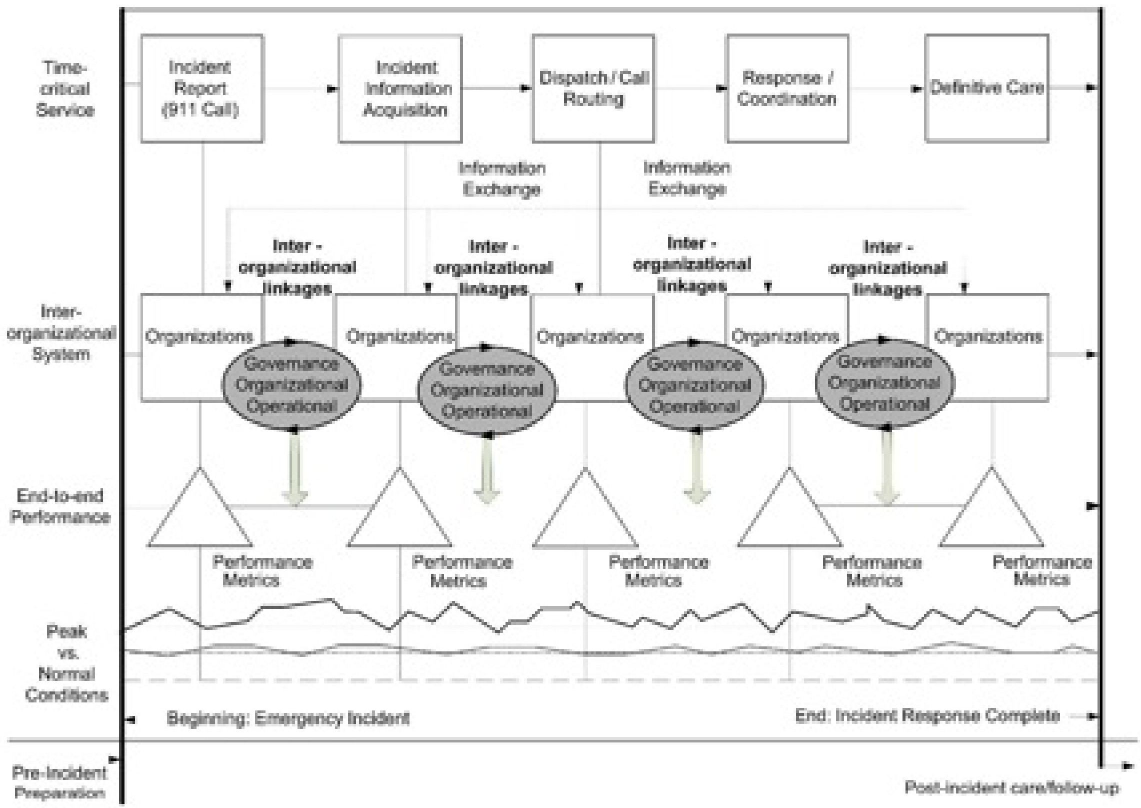


Figure 4.6.1 Time-critical information services framework (from Schooley and Horan, 2007)

end elements that consider performance metrics within and across the process flow, and (4) context variation elements such as normal versus peak conditions (in terms of service demand) (Schooley and Horan, 2007). The utility of TCIS has been demonstrated as a heuristic to analyze EMS systems from a patient-centered approach (see Schooley and Horan, 2007).

Defining the Three Dimensions of Information Sharing

This study focuses on the interorganizational information-sharing dimensions of TCIS (Figure 4.6.1, 2nd row from the top). This framework proposes a structure for understanding operational, organizational, and governance dimensions of interorganizational information sharing and integration to gain a deeper understanding about how information-sharing influences the design and improvement of time-critical public services, EMS service delivery, and information systems to support these services.

These three dimensions were defined in more depth in Schooley and Horan (2007) and are also summarized in Table 4.6.2below. The first of these dimensions, the *operational /technical* dimension of information sharing, includes business processes, procedures, and technological resources for sharing information and data across organizational boundaries. The *organizational* dimension includes issues related to trust, cultural differences, level of participation, power relations, and resistance to change. The *governance* dimension includes participant roles, decision-making processes, rules, and regulations, legal, political, and fiscal issues surrounding an interorganizational information sharing system. These three dimensions were explored within the end-user interviews conducted by researchers. The findings from interviews are reported below and are organized along these lines.

Table 4.6.2 Analytical dimensions of interorganizational information sharing

Operational Dimensions	
	Technical Systems (software &hardware)
	Business Processes (who, what, where, how)
	Communication flows (voice and data)
Organizational Dimensions	
	Power Relations
	Level of Participation
	Cultural, Sub-cultural Differences/Similarities
	Resistance to Change
	Trust
Governance Dimensions	
	Participant Roles
	Rules & Regulations
	Decision-Making Processes
	Political/Legal
	Fiscal

4.6.4 Discussion

A general complaint among users of EMS systems is that supporting information technology lags far behind private sector business capabilities. Nonetheless, significant technological changes have taken place in EMS over the past decade and continue to take place. Two-way radio systems continue to be an integral and important part of communicating during both day-to-day and large-scale incidents and certainly won't be completely replaced by data systems any time soon. Nor is it clear that they should. Rather, a significant amount of change has taken place to understand and implement data systems to replace *certain* voice communications where it "makes sense". And as discussed by users of EMS systems that we have worked with over the past several years, including the Minnesota Mayday system, it only "makes sense" to replace voice communications if the quality of emergency health care service given to a patient is not compromised by doing so. As one adamant paramedic explained, "If I have a choice between entering data into the laptop and stopping profuse bleeding, the choice is obvious. I'll communicate using my hands-free radio." But before getting into more depth with examples, it would be helpful to provide context about multi-organizational EMS systems. Using our time-critical information services (TCIS) model illustrated in Figure 4.6.1(Horan and Schooley, 2007), we look across the multi-organizational, end-to-end EMS process from a patient perspective.

The importance of looking at system operation from a patient perspective is this: a patient looks at an emergency incident as one single event that begins from the time of onset of a medical condition (e.g., heart attack) and continues through 1) incident notification (e.g., 911 phone call), to 2) the answering and reporting of the incident, to 3) the dispatch and arrival of service providers (e.g., police, fire, ambulance), to 4) response and coordination of medical services on-scene, through 4) definitive care at a health care facility (e.g., hospital emergency department), and ends when s/he is released from the hospital. In contrast, there are a multitude of public, private, and not-for-profit organizations that experience only a slice of the same incident. For example, a hospital typically views the incident only from the time that the patient arrives at the emergency department. The way in which we suggest the analysis of EMS systems should be is from the aforementioned end-to-end (i.e., start to finish) view (Schooley and Horan, 2007). As such, the primary information technologies used for EMS vary across the process as illustrated in the simplified description in Table 4.6.3..

Table 4.6.3 Common technologies used in EMS

TCIS Process Point	Typical Information Technologies Used
1. Incident Report	**Landline telephones** **Wireless telephones** **Internet Protocol (IP) telephones** – becoming more prevalent
2. Incident Information Acquisition & 3. Dispatch	**Landline telephones and PBX systems** – for forwarding, cueing, and answering 911 calls **Computer aided dispatch (CAD)** – for viewing and entering caller information, allowing for touch screen dispatching of emergency responders, tracking dispatched resources, and following incident progress. **Two-way radio** – for voice dispatching of emergency responders and providing ongoing support and coordination throughout the duration of an incident.
4. Response / Coordination	**Two-way radio** – for receiving instructions from dispatch, coordinating patient delivery to a hospital, and coordinating services with other response organizations (sometimes). **Patient Care Record systems** – for collecting incident, patient, and medical care information at an incident scene until delivery to a hospital. **Hospital Availability Systems**—for reporting the real-time capability for hospital's to accept new patients and to divert ambulances to other hospitals in the case that hospital wait times are too great.
5. Definitive Care	**Two-way radio**—for coordinating the hand-off of patients from ambulance providers to hospitals. Electronic Patient Registries—for recording the receipt of a patient into the emergency department, trauma center, critical care center, or other emergency care facility. Electronic Medical Records—for tracking patient status and medical care information throughout the length of stay at a hospital/clinic.

There are a number of other electronic systems used across the EMS process in various locations around the U.S. including electronic personal health records owned by individuals that can pre-populate the ambulance patient care record, geographic information systems (GIS) that geo-locate emergency response resources, voice activated systems for capturing data in the field, as well as performance management, business intelligence, and other data mining tools for conducting systems analysis and reporting functions, and a range of others. While there are many opportunities to implement new technologies and to integrate data systems across police, fire, ambulance, and hospitals and other organizations, the above list represents what we have seen as "typical" in many EMS settings in the U.S. We present this list to provide some meaningful context for the remainder of this chapter. Discussing the full range of advanced IT systems that have been implemented in the EMS setting would be far too exhaustive for this article.

4.6.5 Minnesota Case Study

The purpose of the following section is to provide case study context about the Mayday project including related statewide initiatives to reduce emergency response times, Mayday project organizational arrangements, business process changes that occurred, the information systems structure, and the operational performance of the Mayday system. The following information was collected and assembled through interviews with Mayday system designers, managers, and users; system documentation; and from technical reports written by the Minnesota DOT (2006) and Linnell and colleagues (2006). This section provides context for understanding the end-user perspectives extracted through interviews—which are reported later in this chapter.

Minnesota and the Goal to Reduce Traffic Fatalities

In 2005, the state of Minnesota had an estimated 5.1 million people living in 79,610 square miles. Approximately 2.6 million of those people lived in the Twin Cities metropolitan area, which covers less than 500 square miles (FedStats, 2006; Metropolitan Council, 2006). Approximately 50% of the residents and 95% of the land mass are located in rural and remote areas. In 2005, approximately 68% of all vehicle fatalities in Minnesota occurred on rural roadways. This is largely due to a number of geographical challenges associated with responding to rural and remote emergency incidents. There are often long delays before emergency responders are notified and long distances for emergency responders to travel to an emergency scene and then to an appropriate hospital. Though emergency response times vary a great deal across the United States, the National Emergency Medical Services Information System (NEMSIS) shows that many urban areas respond in less than 12 minutes. In contrast, some statistics show response times in rural areas to average over 50 minutes (NCSA, 2006).

While we are highlighting Minnesota, this is certainly not just a Minnesota problem. In the United States, over 50% of traffic fatalities occur on rural roadways. Furthermore, there were a total of approximately 43,000 traffic fatalities in 2005, placing traffic deaths as one of the leading causes of death in the U.S. (NCSA, 2007). This public health issue has become an important focus in the State of Minnesota as can be evidenced by the statewide Towards Zero Death (TZD) initiative. The TZD initiative includes the Department of Transportation, Department of Public Safety, Minnesota State Patrol, Federal Highway Administration, the Center for Transportation Studies at the University of Minnesota, and a wide range of other government and not-for-profit organizations dedicated to completely eliminating traffic fatalities in the state.[xi]

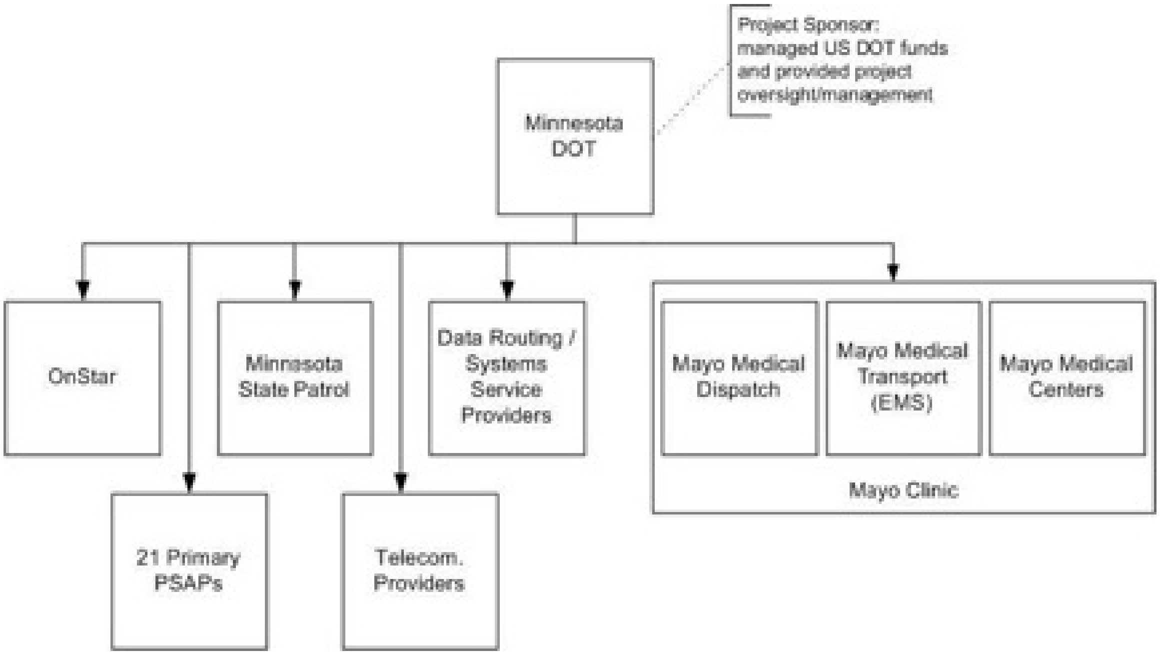


Figure 4.6.2 Mayday system interorganizational arrangement

The Minnesota Mayday 911 Project Overview

In 2004, the U.S. Department of Transportation (DOT) funded the Mayday project, a statewide test project for the Minnesota DOT to establish a multi-organizational collaborative relationship with emergency response organizations to design and build an innovative interorganizational information system. The goal of the project was to use information technology to reduce emergency incident notification times for automobile crashes and the overall Emergency Medical Services (EMS) response times for those crash incidents. The project included several counties in the state of Minnesota and proceeded from October 15, 2004 to September 30, 2005. The test project collaboration included the Minnesota Department of Transportation (Mn/DOT), the Minnesota State Patrol, GM OnStar, the Mayo Clinic including Mayo Medical Transport (air and ground ambulance provider), its dispatch centers, and trauma center, wireless and traditional wire line telecommunications carriers in the geographic region, data routing and information systems providers, and local and county 911 call centers commonly referred to as public safety answering points (PSAPs) (see Figure 4.6.2for organizational arrangement).

The geographic areas that participated in the project included 13 out of the 87 Minnesota counties (see Table 4), which included such populous counties as Hennepin, where parts of Minneapolis are located, and rural counties such as Renville and Meeker. The total estimated 2005 population in the test regions was approximately 2.5 million (FedStats, 2006; Metropolitan Council, 2006). Within these counties, there are 21 primary 911 call centers, or Public Safety Answering Points, and one medical 911 call center, or secondary PSAP, which participated in the project. These areas were representative of population density considerations; that is, 9 city, 13 county, 7 rural, and 15 urban/suburban call centers.

Table 4.6.4 Minnesota counties and county populations included in the mayday project

County	Population	County	Population	County	Population
Anoka	323,996	McLeod	36,636	Renville	16,764
Carver	84,864	Meeker	23,371	Scott	119,825
Dakota	383,592	Mower	38,799	Steele	35,755
Hennepin	1,119,364	Olmsted	135,189	Washington	220,426
Kandiyohi	41,119			**Total Population in Test Areas**	**2,538,581**

The purpose of the Mayday project was to develop and demonstrate a method for reducing the time required to notify emergency response providers of a stranded or disabled vehicle by relaying vehicle location and other critical information about the event to a wide range of EMS and transportation stakeholders (dispatch centers, State Patrol, ambulance providers, health care facilities, traffic management centers and the traveling public). An important project goal was to utilize a standards based, web services approach to achieve a national model for Mayday 911 event information delivery. Though the project was an operational test, the system was designed and developed for ongoing operations and remains in full operation as of this publication in 2007. The operational test included GM OnStar customers who were involved in real automobile crashes and whose crash data was automatically pushed to the GM OnStar call center, local Minnesota public safety 911 dispatch centers and Mayo Clinic emergency medical dispatch centers.

Mayday Information System Overview

The Minnesota Mayday system was designed to bring OnStar crash information (data) into the Minnesota Department of Transportation statewide traveler information and information exchange system as well as the Condition Acquisition and Reporting System (CARS) that is available to authorized DOT, State Patrol and other emergency response providers. The data generated by the emergency GM OnStar system is routed through the GM OnStar call center to a secure public Simple Object Access Protocol (SOAP) server which then distributes the data by automatically pushing it to the CARS system and data routing systems to dispatch call centers and traffic management centers (see Figure 4.6.3). Crash incident data is delivered using the Vehicular Emergency Data Set (VEDS), a standard Extensible Markup Language (XML) data set developed by the Automatic Crash Notification Working Group that contains standardized variables for crash time, severity, and location that can be exchanged with pertinent stakeholders (Mn/DOT, 2006; Linnell et al, 2006).[xii]

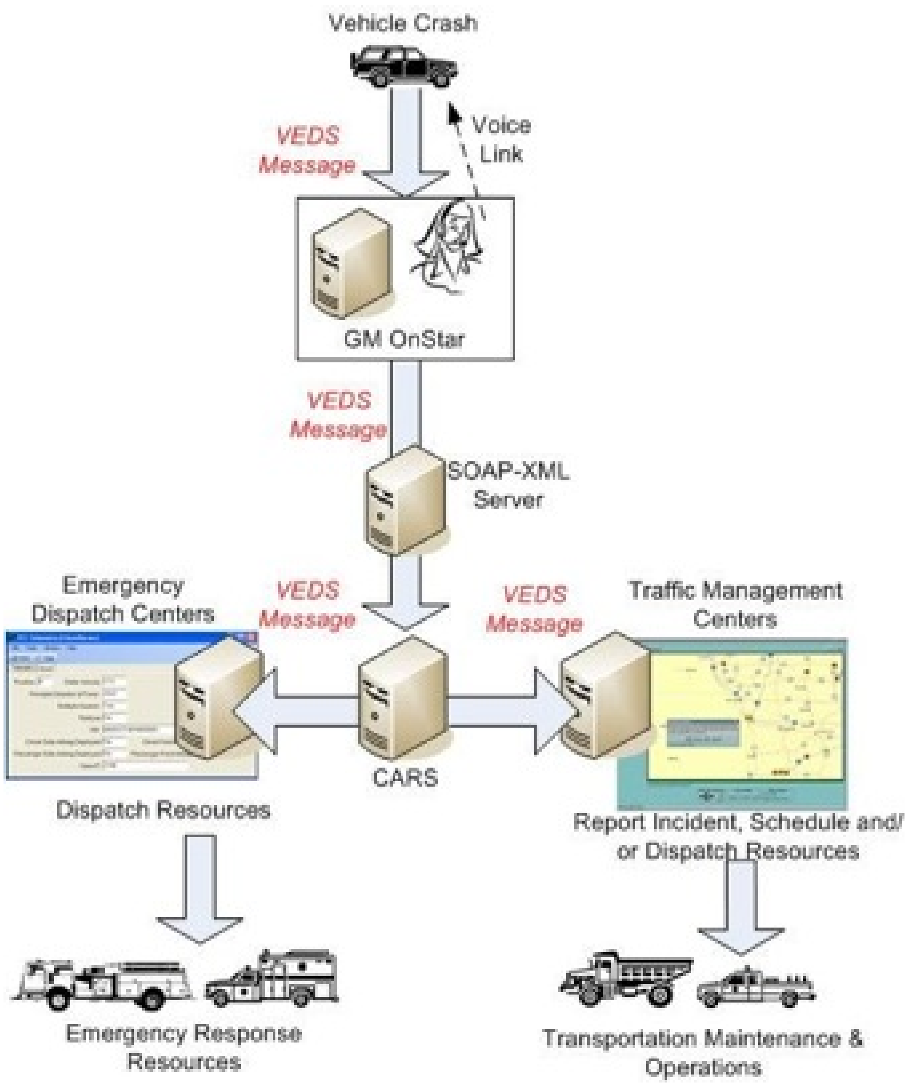


Figure 4.6.3 Mayday emergency data routing system overview (adapted from Linnell et al., 2006)

The system created a significant interorganizational business process change with the purpose of reducing redundancies inherent in cross-organizational EMS communications and creating communication efficiencies that did not exist prior to the change. For each individual incident all applicable organizations are sent incident notifications simultaneously allowing both law enforcement and medical dispatch centers to receive emergency notifications at the same time. It also provides traveler information to the public through a public web site that displays the locations of traffic incidents and resulting traffic congestion.

To better understand the significance of this business process change, Figure 4.6.4 displays the notification process in Minnesota *without* the Mayday system, and then the new process *with* the Mayday system. To help explain the significance of this change and the communications efficiencies that were created, we first describe the "without" scenario and then the "with" scenario below.

The 911process without the Mayday system starts with a vehicle or person notifying OnStar of an incident. In the case that a vehicle notifies OnStar, an OnStar emergency operator will try and contact the vehicle occupants to determine crash severity and their emergency needs. Whether they are contacted or not, the operator will then make or forward the call to the nearest public safety 911 call center to the incident. This can be very complicated since there exists an astounding 121 call centers distributed across the many small rural towns and counties in the State of Minnesota. Many of these call centers staff only 1-2 operators at a time and some are not functional 24-7. Though OnStar keeps a detailed database of these call centers and their geographic service areas, on occasion an incorrect public call center is contacted, which then requires the public call center operator to re-route the call to the correct call center. Despite rigorous efforts by the telecommunications industry to keep accurate 911 call center databases, 911 telephone calls are often mis-routed, which requires re-routing by live operators that can result in significant delays in emergency response efforts. In any case, when the call is answered by the public 911 call center, the OnStar operator will describe the incident to the public safety 911 operator providing the most important information first: incident location, phone number of caller, and some indication as to the severity of the crash. The occupants of the vehicle, if contacted, will then speak with the public safety 911 operator who is responsible for dispatching law enforcement resources (e.g., police, state patrol) and sometimes fire resources as well. The public safety 911operator will then forward the call to an EMS medical dispatch center where the incident will again be described to the operator of that center. The caller will again describe the incident giving more details as to the health condition of vehicle occupants. This third and final call center has the responsibility of dispatching medical emergency resources (e.g., air and ground ambulance, and sometimes fire). The typical 911 call process for most locations in the United States consists of verbal descriptions about an incident and the subsequent forwarding of a phone call to various call centers. It can be a time consuming and inefficient process, but there are of course many historical, political, organizational, and financial reasons—such as the historical need to separate law enforcement and medical emergency response crews due to the advanced skills set and knowledgebase needed to handle each separate and distinct type of call (IOM, 2006). However, the reasons have been well documented in other sources and are outside the scope of this chapter.

In order to utilize a more data driven approach to speed the notification and dispatch process, the Mayday system was born. The system notifies both public safety and medical dispatch centers simultaneously of an automobile incident thus eliminating redundant reporting of an incident. Furthermore, the initial incident notification is distributed via data communications and then supplemented with voice communications, enabling resources to launch quicker.

The 21 911 call centers that participated in the project are charged with dispatching local law enforcement, State Patrol, and fire resources while the Mayo Medical dispatch center dispatches ground and air ambulance resources. The Mayo Clinic, together with its dispatch center, ambulance transport service, and various medical and trauma centers participated in the response and delivery of patients to health care facilities.

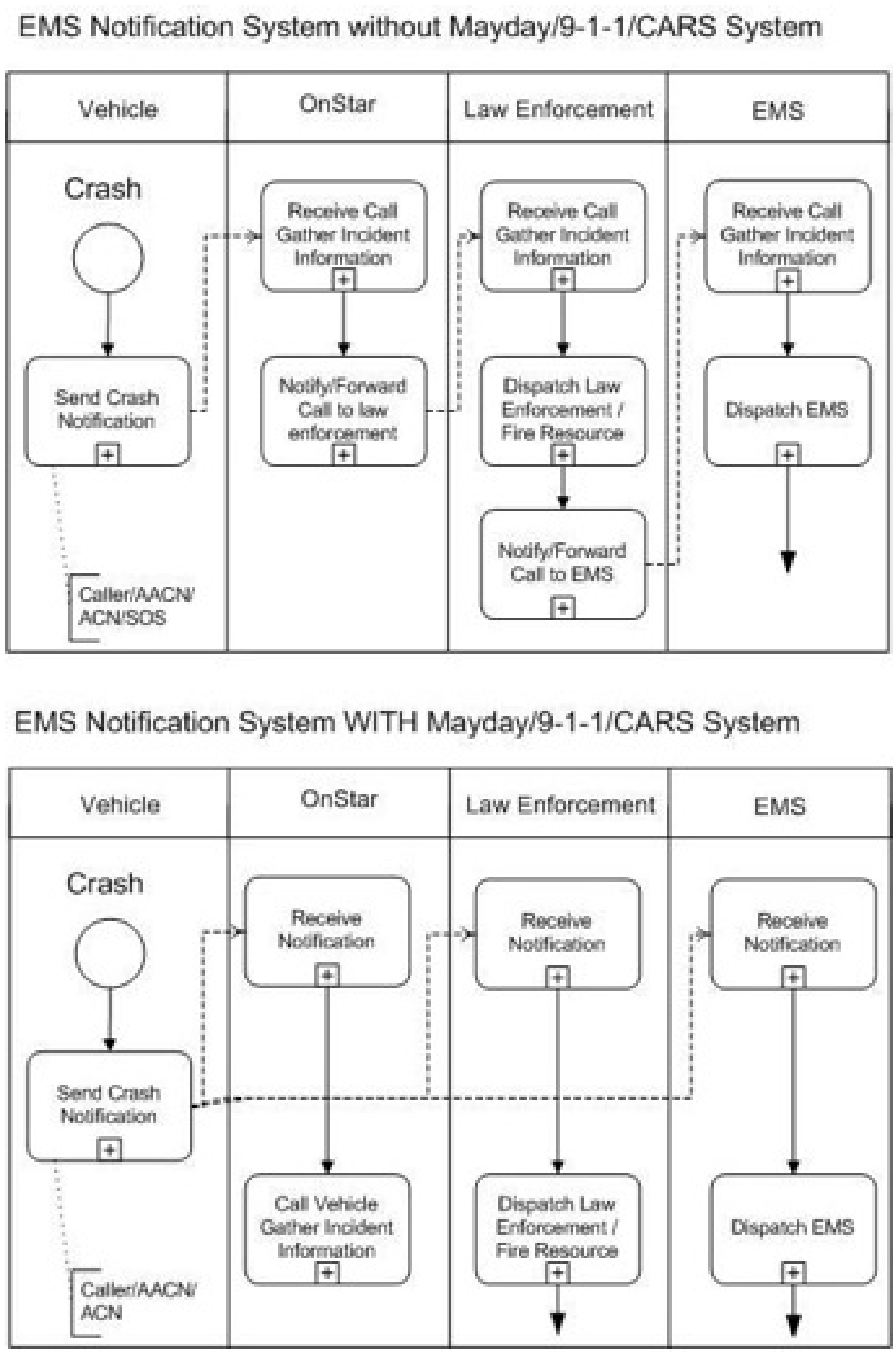


Figure 4.6.4 EMS notification business process change

Mayday System Performance Overview

Across the entire U.S., GM OnStar reports a monthly average of approximately 380 Advanced Automatic Crash Notification (AACN) calls, 1,000 Automatic Crash Notification calls (ACN), and 11,400 emergency push-button calls (SOS) (see Table 4.6.5 for definitions of AACN, ACN, and SOS). As reported by Linnell and colleagues (2006), during the Minnesota Mayday test period there were a total of 17 AACN calls, 137 ACN, and 1093 SOS calls (see Table 4.6.6).

Though not a large volume of calls, the test did provide an opportunity for end-users to experience the operational changes that the system offered. Table 4.6.6 shows total call volumes, weekly average call volumes, the maximum and minimum number of calls received each week, and the amount of time it took for Mayday messages to be sent to emergency call centers for each of the call types recorded during the test (AACN, ACN, and SOS). Table 4.6.5 provides explanations and definitions for AACN, ACN, and SOS calls as they pertain to the Mayday project. In terms of weekly incident volumes, there were very few AACN calls due to the very few vehicles that are equipped with such systems across the United States. However, the number of vehicles equipped with such systems is increasing dramatically each year increasing the impact that the Mayday system will have on drivers and emergency responders.

Table 4.6.5 Definitions of AACN, ACN, and SOS Calls

Term	Definition
AACN	Advanced Automated Collision Notification (AACN): Those events where an advanced collision notification system is on-board a vehicle. The system sends basic information to an OnStar advisor, such as the location of the vehicle, vehicle type, contact information, and whether an airbag was deployed or not; as well as additional crash information such as delta velocity (speed at impact), principal direction of forces, and whether a roll over occurred. A voice communication link is also established between OnStar and vehicle occupants.
ACN	Automated Collision Notification (ACN): ACN systems automatically send basic crash information such as the location of the vehicle, vehicle type, contact information, and whether an airbag was deployed or not. It creates a communication link between OnStar and the vehicle for voice communication to occur.
SOS	Emergency Key Press (also referred to as SOS): Those events where OnStar users (vehicle occupants) press the emergency key button in their vehicle. This establishes a voice communications channel only between the vehicle occupants and OnStar. No crash information is sent automatically. In the case of a crash, OnStar advisors receive the call and establish a three way call with the local 911 call center. • **Vehicle Location.** (latitude, longitude). Description of the route or roadway that the event occurs on, as well as the intersection (or mile marker). • **Incident Type/Description.** This key phrase contains the best description of the event. For example, automobile crash, heart attack, etc… • **Air Bag Deployed.** This specifies whether an air bag was deployed during the incident. • Advanced systems indicate which air bag/s. • **Time-stamp Data.** Incident start and stop times. • **Crash Information (Delta Velocity, Rollover, Multiple Severe Impacts).** Additional information concerning the crash facts may also be inserted. This is in accordance with what data OnStar is able to send. For the Mayday system, this data was limited to delta velocity (speed of vehicle at impact), whether the vehicle rolled over, and if there were multiple severe impacts. Data forwarded to systems that provide data to the general public are limited to basic information about the location and general nature of the event.

Table 4.6.6 Weekly Call Volumes (by type) and notification time intervals

	AACN Calls	ACN Calls	SOS Calls	Failures	Time Delivery (seconds)
Totals	17	137	1093	50	
Weekly Average	0.36	2.9	23.26	1.06	1.38
Range (Maximum)	3	15	46	36	23.25
Range (Minimum)	0	0	0	0	0
Weekly Median	0	2	26	0	0.835

Table 4.6.7 illustrates the types of data collected from each of the different types of OnStar calls. Different amounts and types of information are collected depending on whether the call is an AACN, ACN, or SOS call; with AACN providing the most in-depth description of an automobile crash incident. Further detail about the information provided in a Mayday message can be found below.

Table 4.6.7 Data types collected from each system type (ACN, AACN, and SOS)

Data Types	AACN	ACN	SOS
Vehicle Location	X	X	X
Incident Type	X	X	X
Air Bag Deployed	X	X	
Time-Stamp	X	X	X
Delta Velocity	X		
Rollover	X		
Multiple Severe Impacts	X		

In terms of Mayday system performance, notification times across the duration of the project ranged from less than 1 second to approximately 23 seconds, far shorter than the average voice notification. Interviews with State Patrol and Mayo Clinic experts indicated that the Mayday system reduced the emergency notification interval for ambulances by approximately 3 to 10 minutes and the overall ambulance response interval by approximately 2 to 9 minutes, depending on other crash factors such as the incident location, the capacity of the local 911 call center, and availability of an ambulance (Linnell et al., 2006). When comparing these time savings to the average time it takes to notify and respond to an emergency, it is clear that the Mayday system has the potential to significantly impact end-to-end efficiency. To put this in context, evidence indicates that average emergency response time intervals across the U.S. are as shown in Table 4.6.8 (NCSA, 2006).

Table 4.6.8 Average time intervals across an end-to-end emergency response

Time Interval	Rural Crash	Urban Crash	Mayday Time Benefits
Time of Crash to Notification of crash (911 call center receives call)	7 min.	3 min.	3-10 min.
Notification of crash to time of EMS Arrival at Scene (ambulance arrives)	12 min.	7 min.	2–9 min.
Arrival at scene to Time of Arrival at Hospital	36 min.	25 min.	NA
Total	52 min.	34 min.	7-19 min.

4.6.6 Mayday User Perspectives

Through interviews with system users, the overwhelming majority felt that the Mayday system proved a success. As is often the case with IT system assessments, there were plenty of complaints, suggestions for improvement, and new found requirements for the next version of the system. But comments were generally positive towards the end result largely because the project enabled and encouraged collaboration across traditionally disparate organizations. While there has been much written and spoken about collaboration, these groups were able to put such ideas into practice and gain benefits therein. One State Patrol Officer stated:

> We've been able to see and get information in ways that, well we really couldn't imagine like 10 years ago. Just the traffic and road maintenance reports from Mn/DOT, crash reports from OnStar, and we can see where the ambulance is.

This type of coordination, it's all very helpful for our officers in the field. A more in depth discussion about Mayday user perspectives follows. As mentioned in the methodology section previously, the following section discusses participant responses to a set of interviews and round-table discussions with users from each participating organization. The case study examples illustrate the interrelated, parallel, and overlapping importance of the three dimensions: operational, organizational, and governance.

User Perspectives: Operational Level

The following findings focus on operational dimensions of interorganizational information sharing across the Mayday system. As mentioned previously, these dimensions include technical software and hardware systems, business processes, and communication flows that influence cross organizational information sharing.

IT Quality Attributes: Usability vs. Security. The technical usability of the Mayday system came into question early on and throughout the duration of the project. The system went out of service several times as illustrated in the multiple message failures that occurred and were reported in Table 4.6.. These outages were due to the expiration of the security key required to ensure a secured connection between all routers. From one system administrator's standpoint, the system was at least very secured, but sacrificed usability and ease of management as a result. As such, the need to renew security keys on a regular basis was forgotten about on several occasions. Administrators stressed the importance of implementing rigorous management checks and processes to ensure a balance between ongoing functionality and security—both cited as top system quality attributes for such a highly privacy oriented and potentially life saving system.

Automation for Emergency Responders. Trauma care is highly variable in terms of the types of problems encountered and the care interventions administered. It can be dynamic, fast-paced, and involve a wide range of emotions on the parts of all involved. In these cases, emergency responders want to focus their attention on the needs of a patient. An important benefit of the Mayday system is that the automated and simultaneous data pushes from OnStar to both law enforcement and medical dispatchers allow emergency responders to focus on providing care to a patient rather than focus on gathering phone number, location, or other information. The automated data pushes were so appreciated in fact that participants began looking for additional opportunities for automation. One of the technical designers and builders of the system stated:

> The next step in automation is to have all OnStar calls automatically routed to the correct PSAP [local 911 call center]. During the project, some calls were forwarded to the wrong PSAP [911 call center] because we designed the Mayday system so that OnStar operators had to manually dial the 10-digit number to each local Minnesota call center. So what we can do now is take out human error by making it automated. Participants described the need for eliminating even the rare cases of human error out of the dispatch process and that automated dialing is one way that could help accomplish that.

Real-Time Decisions. The process of implementing and testing the Mayday system provided opportunities to re-assess real-time decision making processes across the service. In particular, it allowed for users to explore how crash data could provide advantages at varying decision points. For example, one Mayo trauma physician believed that certain data, such as vehicle type, the speed of the vehicle at impact, what part of the vehicle suffered damage, whether seatbelts were deployed or not, and other related information could be used to calculate the number of individuals involved and the probability of injury to each vehicle occupant. The Mayo trauma director stated:

> That kind of data could be extremely useful for preparing a trauma team prior to patient arrival. That alone could save a lot of time and lives. The Mayo dispatch manager thought that same data could be used by dispatchers to make better decisions on the number and type of ambulances (air or ground, advanced life support or basic life support) to deploy to accommodate one or multiple individuals with varying degrees of injury.

Incident Visualization. An additional valuable dimension of the Mayday system was how it enabled more effective communications within and between organizations. Participants discussed how emergency professionals were able to visualize other service providers from other organizations on the CARS user interface by looking at the GIS enabled map. A dispatch supervisor explained:

> The CARS interface gave us a visual map to see what DOT and the State Patrol were doing about an emergency incident. We could actually see which ambulances were closest to the incident and then dispatch them. A few times, we sent the dispatch to a competing ambulance provider because they were closer to the incident.

The system also allowed emergency personnel in the field to know who they could contact at the incident and what type of service the patient might be receiving. A Mayo dispatcher described an experience:

> I could see that a CPR trained State Patrol officer was the first responder at the accident and the report that we got was that either driver or passenger was having difficulty breathing. I told the paramedics who were en route that State Patrol was there and that he was CPR trained. I don't know if that information changed the overall patient outcome. That's hard to tell. But the paramedics like knowing what to expect when they arrive. They can at least know what equipment to bring with them, what procedures have been given, and what procedures they should prepare for.

In a like manner, State Patrol dispatchers could view the CARS interface to see the activity of Mayo EMS resources, for example, if an ambulance were alerted and en route to an incident or not. The CARS interface also allowed dispatchers to view traffic congestion and road maintenance information from Mn/DOT. This enabled ambulances to make better driving route decisions to avoid traffic. Participants noted how this information allowed collaborating organizations to make more informed decisions. For example, a State Patrol officer stated:

> If I know that an ambulance is on the way, I don't have to spend time wondering "when is the ambulance going to get here" or spending time telling EMS dispatch to send an ambulance that has already been sent. The new communication system was helpful in that sense.

In general, participants noted how their ability to visualize and see emergency and transportation resources enabled more effective communications, more informed decision making, and a higher degree of perceived service performance.

Streamlining Inter-organizational Processes. As stated previously, the Mayday system created a business process change. While it was agreed that change was good, participants discussed several operational challenges associated with the ongoing goal of establishing proven, tested, and standard business processes for delivering clinical care across service organizations. Standards of care often differ from one organization to another. For example, ambulance providers may adopt care practices based on their experience, or based on a different set of priorities (e.g., faster response times as a higher priority than care delivery), while the emergency department at a hospital may adopt a different practice based on other assumptions (e.g., new research). A trauma physician stated:

> Fortunately for us [Mayo Clinic], we purchased the ambulance provider just a few years ago, which has enabled us to have a more integrated approach to delivering care across the service. But it is still a very challenging task when you think of all the different types of cases we see. No case is exactly alike, which means we have to change our approach depending on the case.

Discussions alluded to the dynamic nature of emergency medicine and the need for information systems to support that environment. Some examples were provided. A trauma department administrator discussed how there are many complex interactions between human organs and biological systems, which makes diagnosis extremely difficult. Many injuries are discovered hours or even days after an automobile crash. In this regard, she explained that there exists a need to further develop quality of care standards and associated emergency health care processes that are adaptable based on scientific evidence. A Mayo trauma physician stated:

> Until there can be more agreement [on standards and processes], it will be difficult to create a comprehensive data set to share between emergency response and health care organizations.

In a related sense, participants discussed the need for a more comprehensive set of performance measures to adequately assess the end-to-end performance of EMS. The Mayo trauma department Clinical Director stated:

> One challenge is that there are numerous different descriptions of what good patient care means. Another is that there are so many variables involved in delivering health care that it makes performance measurement difficult. These challenges need to be addressed. A comprehensive look at quality care provision across the whole service has not been investigated in depth.

A data analyst felt that a discussion about medical care (business) processes, performance measurement, and quality of care needed to take into account the different types of data that are shared, or could be shared, across service organizations. There is "incident" data, "process" (timeliness) data, and then data that integrate the two. He explained:

> There is an important difference between "data about the incident" and "data about the process". Data about the incident includes information about the patient, his/her location, the incident type, care provided, resources dispatched and available and other related data. Data about the process includes time-stamp data, which is used to determine process and work activity costs that essentially determine process efficiency. What is less understood is how to use both "incident" and "process" data to measure quality of care.

In terms of performance implications, an emergency physician stated:

> Time from crash to definitive care remains the most important indicator of survival or patient outcome. But the right type of care will also improve response time, care delivery, and outcome.

In sum, participant comments highlighted challenges to implementing process changes across organizations—each stakeholder wanting to make sure his/her priorities are not sacrificed. While participants discussed challenges to establishing new business processes, they also pointed at the Mayday system as a "good fit" with both operational and clinical care needs. Participants felt that future process changes, in order to be successful, also need to have both operational and clinical care benefits in order to achieve long-term acceptance by users.

New Service Complexities. While the Mayday system provided new efficiencies, several users discussed some added complexities. Several State Patrol officers provided an example when only partial incident information was received from a call made by an OnStar customer. Soon after a customer pressed the OnStar emergency button while driving his/her black Escalade, the call was dropped. State Patrol officers did not know if the individual hung up the phone or the call was lost for some other reason. An officer explained:

> OnStar transferred the call to us but it was dropped before we could figure out the problem. The only information we received was the make, model, and location of the vehicle and the general coordinates where the call was made. But the car was moving and we didn't know which direction. So we were obligated to respond and of course, there was no black Escalade when we arrived. So, we pulled over every black Escalade we could find in the vicinity to search for the reported emergency. We never found it.

The State Patrol officers discussed their frustration with these types of incidents. They were obligated to respond to the call and look for the vehicle. Their obligation was based on the definition of their jobs as instituted by the State Patrol and Minnesota law, as well as a moral obligation. No exceptions had yet been made based on the circumstances of a dropped or lost call. They pointed out what they felt to be an important aspect of cross-organizational information systems—that quality and completeness of information relates to service efficiency. Though information systems (people and technology) may help improve service, it also creates complexities that can result in service degradation when it does not function as intended. In this case, having only partial information led to a long and extensive search. Concerning the above example, the State Patrol officer stated:

> Our time was wasted and an individual's health problem was possibly left unaddressed. It shows how important it is to have a complete set of information. We need to have policies in place that deal with incomplete information. We can't just ignore a dropped call. We have legal, not to mention ethical obligations to respond to every call. That's why we can sometimes seem pretty skeptical about taking on new technology.

Participant responses pointed to the need for new technology to be thoroughly tested and proven in order to ensure that the technology functions as intended, that the full range of intended information arrives at its destination so as to reduce user confusion and inefficiencies.

User Perspectives: Organizational Level

The discussion below relates organizational issues to information sharing. As outlined in Table 4.6.2, organizational and interorganizational dimensions of information sharing include trust, cultural and sub-cultural differences/similarities, effective communications, level of participation, power relations, and resistance to change.

Partnership Trust. Mayo Medical Dispatch valued their Mayday partnerships and discussed how trust developed during the project. A Mayo dispatch manager explained:

> OnStar does a good job at screening calls. It's nice that when we do get a call we know it has been screened well. And our dispatchers pretty much know that the call is an actual emergency that needs attention. And so that lets us dispatch an ambulance right off the bat before spending more time screening the call.

Dispatchers believed that the Mayday system and partnership allowed them to be more efficient. They acknowledged the difficulty in measuring the end-to-end benefits in quantitative terms because they did not know how long it took OnStar to carry out its job to screen and forward the call (except in the case of AACN and ACN) calls. But the perception from Mayo dispatch personnel was that the partnership provided a positive benefit to them because the information they received from OnStar was "trustworthy".

Easier to Trust Hard Data. A significant challenge to sharing data, according to paramedics and emergency department staff, rests between the "pre-hospital" and "hospital" environments. Participants discussed a lack of trust that can often exist between the two. A trauma center representative explained:

> Paramedics have to make assumptions about a patient's condition. But physicians question and discount those assumptions. So what I mean is a common perception that physicians have is that paramedics don't give accurate medical care information. And paramedics think physicians are arrogant. We do a pretty good job at Mayo to overcome that problem particularly when a patient comes into the ED [emergency department] from a Mayo ambulance. But we also get patients from the "scoop and run" [non-Mayo] ambulances and there can be definite trust issues in those cases.

According to Mayo physicians, a primary benefit of advanced data systems such as Mayday is the ability to distinguish, or at least separate, between qualitative human generated and machine generated data. More reliance on the latter helps to increase trust between service organizations. A trauma representative explained:

> Physicians and paramedics alike need to better understand that there is irrefutable hard data, like EKG readings, and then there is expert opinion that is subject to scrutiny.

The Mayday system provides an opportunity to advance hard quantitative data. Patient information, including demographics and health history, coupled with crash details offer the ability to create an objective description and predictive algorithms, which can be used to help determine proper care provision. For example, a predictive algorithm would include taking variables such as which airbags were deployed, the speed of the vehicle at impact, and which part of the vehicle was impacted, and would then calculate the probability and scale of injury to vehicle occupants. Based on that data, emergency responders could better assess how many and which emergency resources to send (e.g., helicopter, ground ambulance). A physician stated:

> That is one of the benefits of the Mayday system. We have hard data that says there has been a roll-over event and we have the delta velocity [speed of vehicle at impact] rather than a paramedic saying, "the vehicle was traveling fast and crashed really hard."

As noted by the above physician, the general notion from participants was that there are levels of distrust when humans relay health related opinions and/or impressions. And the hard data from instruments allows for performance improvement in terms of time, but also because human opinions and impressions are replaced by data from trusted instruments.

Aligning System Purpose with Organizational Culture. OnStar, the State Patrol, and the Minnesota Department of Transportation representatives all noted that their participation in the project was largely influenced by a shared belief that the Mayday system would both quicken response and allow for better health care decision making. This common and shared belief helped to mitigate typical interorganizational challenges to sharing information—particularly among the more integrated Mayo organizational units—Trauma center, Dispatch, and Ambulance service. Mayo participants felt that this was largely due to their organizational culture, which is both accepting of new innovations and focused on quality health care delivery. A Mayo representative stated:

> The Mayday system "fits" well with our philosophy and mission statement about putting the health of the patient first. We pride ourselves on providing quality care to every patient. It was a natural fit and received plenty of support. But we have to remember that we are spoiled here. We don't have many of the constraints that can be found in overcrowded, under funded urban hospitals.

While Mayo was able to participate and dedicate resources to the project, participants noted that most trauma centers in the U.S. don't have the resources to implement such a system. Hence the importance of developing standards based, duplicable models that can be implemented elsewhere.

Communicating Performance. Some participants discussed how an overemphasis on timeliness can act as an obstacle to communicating other valuable and pertinent information. Mayo Clinic personnel agreed that the time from dispatch to arrival to a hospital is paramount, but that there is too often an overemphasis on "timeliness" as a performance metric. A Mayo Clinic representative explained:

> How much and what type of care is given during that time period is equally (if not more) important. This is especially evident in rural and remote areas where response and transport time is lengthy, even for the most efficient responses.

Participants believed that the focus on response time had significantly impacted the regular exchange of response time data. And that a similar emphasis was needed to motivate the regular exchange of data that could impact health care decisions. A Mayo dispatch representative stated:

> There are many unknown, untested, but potential benefits to sharing "health care" information across organizational units. But we just haven't focused on that data enough. There is a lot of emphasis on the time-stamp data for good reason. We just need to extend the emphasis to other types of data. Mayday provides a good example of the potential benefits.

Cross-organizational Communication. Users discussed how one of the benefits of the Mayday project was that it stimulated managers to think more in terms of the future possibilities of using information technology to enhance EMS. A Mayo dispatch manager described an example:

> An automobile accident occurred on the outskirts of Rochester, an airbag was deployed, and an ACN message sent to OnStar and through the Mayday system. When responders arrived to the coordinates they couldn't find the crashed vehicle and almost gave up the search. It was finally found after an extensive search. It had driven off the road and into thick foliage underneath a freeway overpass. In our Mayday group meeting, we talked about that incident and it stimulated discussion from everyone [Mn/DOT, State Patrol, Mayo Clinic] about how technology could have been used to aid responders. Suggestions included sending repeated data messages until the vehicle was found and automated flashing lights or repeated horn honking to supplement data messages.

For participants, it was not the above solution that was most interesting, but the open conversation that took place between organizational representatives. The Mayday system created an atmosphere that facilitated cross-organizational communication.

Resistance to Change. Participants discussed how "immediate" and "observable" performance benefits helped overcome resistance to change. Dispatchers, State Patrol officers, and EMS professionals were able to see and experience immediate benefits from sharing information. A Mn/DOT project manager explained:

> Mayo dispatchers received incident notifications sooner than normal, and State Patrol officers were able to observe ambulances arrive on scene sooner than would normally be expected. Since they could see the benefits right up front, we really didn't experience much of the resistance to change that we sometimes experience when we deploy new technologies.

Mayo dispatch participants stated that dispatchers have been resistant to information technology changes in the past. Yet, they were enthusiastic about using the Mayday system soon after the first emergency incident messages arrived and observable improvements were discovered. Similar comments were received from State Patrol, trauma center, and Mayo transport participants. The ability to observe performance improvements significantly influenced perceptions about the value of interorganizational information sharing.

Aligning Technology with Human Needs. The Mayday system was designed for interface flexibility utilizing an XML data standard. The participating dispatch centers did not expend the resources to integrate the XML messages into their existing computer aided dispatch (CAD) systems and instead viewed automobile crash incidents through a separate web based graphical user interface (GUI). Participants explained that the desirable solution would integrate, or allow for Mayday data messages to be viewed within existing interfaces (e.g., computer aided dispatch (CAD), patient care record (PCR) systems, and hospital based decision support systems). A dispatch operator explained:

> We are often overwhelmed by the amount of information we must deal with on a regular basis and most dispatchers don't want yet another computer screen to look at.

Several State Patrol participants stated that most of the small public safety answering points would likely not participate in the Mayday project just for that one specific reason–that new information must be integrated in a useful and convenient manner or not at all. Mayo Dispatch and State Patrol operators are surrounded by 1-2 phones, 1-3 two-way radios, 2-4 active computer monitors as well as wall-mounted large screen monitors to display the status of emergency units in the field, view real-time status of multiple incidents simultaneously, view weather information, real-time video and graphically displayed traffic information, view the availability of emergency department and trauma centers, and to communicate via voice to provide a centralized support function. In short, they feel overloaded with information. One State Patrol representative stated:

> With so few OnStar users, it is difficult for many locales to justify the cost to implement technological systems that integrate with Mayday. They simply will not add another monitor or screen to the many already existing interfaces. It's a significant issue for dispatch centers.

A Mn/DOT participant disagreed with the challenges:

> The programming, coding, and technical infrastructure is already in place. It was accomplished through the Mayday test project. Data messages are pushed to all subscribers in a standard XML format, which means that local PSAPs just need a computer and an Internet connection to link in.

Even so, there are costs associated with getting the data into a form that end users are willing to utilize. Participants believed wide scale implementation would take place as AACN, ACN, and SOS technologies become more common in vehicles. Until then, agencies would not see the benefit of investing in systems to connect with the Mayday system.

In contrast to how the above professionals felt, information technology designers who developed the Mayday system felt that EMS professionals were reluctant to change. As stated previously, the Mayday information interface was a single large monitor shared by all communications center dispatchers. When an automated Mayday message came into the center, the interface map displayed the incident and an alarm signaled throughout the center. One system developer stated:

> The dispatchers don't want another screen to look at and they don't want the alarm. They were not enthusiastic during training. Several were skeptical and negative. But dispatchers stated that the resistance was not due to the data, but due to how they received the data.

A center manager explained:

> We have 3–4 monitors on the desk in front of us, a large screen on the wall with weather information, another with traffic, another with traffic video; we have a radio and a phone. We need the data to automatically enter our CAD system rather than have another interface to look at. It was OK during the test project and the OnStar data is very useful. But we need it to go to the next level and have it integrated with the CAD so we're not looking all over the room for information.

According to the dispatchers, the resistance was not due to the information being shared but due to the communication interface. Dispatchers felt that an appropriate interface would significantly impact their ability to use the data more effectively. However, issues still exist in terms of using new information effectively as emergency professionals are often overwhelmed with data and information. This tends to create a culture of resistance to new or additional information.

4.6.7 User Perspectives: Governance Level

The following discussion relates governance dimensions of information sharing across the interorganizational Mayday system including participant roles, legal definitions, policies, and rules and regulations.

Success Due to Well Defined Participant Roles and Responsibilities. Interviewees discussed how the overall success of the Mayday project was largely influenced by well defined participant roles and responsibilities. Many barriers to information sharing were overcome because the Minnesota Department of Transportation (Mn/DOT) had clear oversight and direct management over the project. A Mn/DOT project manager confirmed:

> The project was well designed, funded, it worked just as a field operational test works. We had an RFP process at the beginning and everyone reported to us. Mn/DOT funded the project through grant funding, which overcame budget issues for participating organizations. Mn/DOT had clear responsibility and accountability over system governance. This facilitated Mayday implementation and information sharing and enabled the performance efficiencies previously discussed.

However, participants from other organizations discussed their concern over who would own and manage the system and how it would move forward in the future. Ten months after the project ended, a Mayo Clinic representative stated:

> Mn/DOT has maintained the system since the project ended, but there is no indication as to how long it will do so. We know that the system that sent automated messages to ambulance pagers was shut off a few months ago. We don't know how long the system will keep working.

A Mn/DOT representative said funding was in place to keep the system running for a long time. Participants understood that the Mayday system started as a research project and demonstration and wondered about the sustainability of it. They were not aware of future plans, how long funding would continue, what the partnership arrangement would look like, or who would have financial/legal ownership. While clear governance structure was key to initial project success, uncertainty about its sustainability had created an atmosphere where some information sharing had terminated and there existed a standstill in terms of furthering information sharing initiatives. One State Patrol officer speculated:

> I think the system will continue to operate. Maybe they'll change the structure of it. OnStar could manage on a fee basis, or maybe DOT will continue to outsource management using public funds. But we haven't moved ahead with any new plans since the end of the project.

The Need for National Direction and Strategy. Participants discussed a much larger issue associated with information sharing at State and National levels, which has become increasingly important due to a focus on large-scale catastrophic events. A Mn/DOT representative explained:

> The Mayday system could have far reaching implications. It could be very valuable, for example, to find patients and deliver care during or after large-scale geographically concentrated events like terrorist attacks or earthquakes. Because messages are based on open XML standards, it could notify a wide range of state and national organizations when there is a sudden increase in service calls and if a local EMS system needs additional help to respond.

The Mayday system was built in a manner that enables OnStar to *send* data all over the United States. A large portion of the technical infrastructure is in place. But participants explained that there is a lack of national directive, or strategy on how to expand the system and include/integrate it with new systems such as the National Highway Transportation Safety Administration (NHTSA) "NextGeneration 911" initiative, an effort to create a more "data centric" 911 system. A Mayo Clinic Administrator stated:

> There is a lack of political leadership and collaborative efforts on what needs to be done to achieve such a National goal. We need "building blocks" and a plan that describes how to achieve next generation systems in an incremental fashion. Not just technically, but in terms of organizational and policy plans.

Participants discussed the need to define the parts and components of the National system, and what needs to be accomplished to build a foundation on which EMS systems can exchange data across local, state, and federal systems.

The Health Privacy Issue: HIPAA. Numerous discussions took place at the beginning of the Mayday project to determine what advanced automatic crash notification (AACN) data should or should not be sent from OnStar to other organizations. A Mayo representative described a topic of one of the meetings:

> We met with project representatives from each of the Mayday partner organizations. We discussed our "wish list" of data items we would like to receive, and then those items were discussed in terms of privacy and HIPAA regulations. There were some concerns that "travel speed at impact" and "seat belt engaged" data made available to State Patrol officers and insurance companies could cause major privacy concerns. But then we talked about how they already get that information. State Patrol experts conduct very accurate evaluations on highway accidents that produce that same information - travel speed at impact and seatbelt information. So why shouldn't we get that data from OnStar?

State Patrol and Mayo participants discussed two primary concerns with the "hard" data. First, OnStar would be providing quantitative computer generated electronic evidence that would be very hard to dispute, while the highway patrol analysis constitutes a human generated evaluation that could be more easily argued as inaccurate. Also, OnStar is a private company while the State Patrol is a public organization charged with conducting such analyses. They also noted that OnStar would want to take precautions to avoid litigation risks and a possible backlash from existing customers. Though sharing AACN data could provide valuable information to emergency and health providers and insurance companies, participants explained that the Mayday data was used with much more discretion to avoid liability and privacy concerns. An OnStar representative stated:

> OnStar's current policy is to default to data sharing. However, we maintain that customers have the ultimate right to decline data sharing with other organizations. Our legal counsel states that if a customer declines services, we are obligated to that decision. We understand there are many in the EMS community to disagree with that standpoint.

Legal concerns act as a deterrent to sharing information across organizations. There was speculation in regards to the performance implications. One trauma physician explained that data fields such as "seatbelt engaged" or "number of passengers in a vehicle" would have significant implications for EMS in determining the extent of injuries and the number of resources to dispatch, respectively.

A Legal Infrastructure for 911 Data Communications. The provision of 911 telecommunications services is regulated and monitored closely by the Federal Communications Commission (FCC). In order to establish the Mayday system data network arrangements and the ability to route 911 calls to PSAPs, a number of required forms and processes had to be filed. A network engineer stated:

> This was a significant challenge since the forms and applications are intended for telecommunications "phone" carriers rather than "data" carriers, and no simple guidance existed to help fill out the required documentation. Phone carriers have been around for a long time and know what they're supposed to do. At the same time, there's not a lot of new 911 carriers out there so finding someone to help was very difficult.

Participants stated that the regulatory atmosphere needs to be adjusted and adapted to more efficiently facilitate the establishment of "data" and Internet Protocol (IP) based 91-1 services similar to the Mayday system. The experience caused significant delays and acted as a barrier to information sharing.

4.6.8 Issues for Further Research

Responses from end-users of the Mayday system provided insight into several topics related to the design and development of interorganizational information sharing systems for EMS. These topics are summarized below.

Operational: New Interorganizational IT and Business Process Change

The operational Mayday system, including the underlying information technology, the business processes, and communication flows between organizations advanced information about an emergency incident in a new and different way throughout the end-to-end EMS service process. Participants observed performance benefits including reduced response times and the ability to deliver better and more quality health care by having and utilizing the data. The ability to dispatch an ambulance at the same time as the State Patrol provided one such obvious and observable advantage in terms of time savings. Yet participants noted several operational issues to improving information sharing. Issues included the need to integrate Mayday data into existing user interfaces, the need for additional automation for collecting and transmitting data, the need for data to better represent situational context, and the general need for the technology to more fully function as intended. Though system users noted challenges to understanding data, the potential benefits of using the Mayday data, and challenges to using the technology, the end-to-end delivery of EMS was improved through an innovative use of information systems. An operational theme for this case study is how information systems were used to create an inter-organizational business process change for a public sector service, which resulted in a clear improvement over the pre-existing system.

Organizational: Information Systems that "Fit" Organizational and Interorganizational Goals

Performance analysts in Minnesota have highlighted the time-value of the Mayday system. Participants agreed that reducing response time was an important accomplishment. But equally important, from the perspective of emergency health care providers, is the value of the data for improving the quality and appropriateness of care provision. The unique interorganizational aspect of this case study is the focus on cooperation. The State Patrol has cooperated and shared resources with the Mn/DOT for the past several years. In addition, the EMS dispatch center, the ambulance provider, and the "end" health care facility are all owned and operated by one organization—the Mayo Clinic. This allows for some commonality across these three functional units in terms of vision, goals, objectives, and the general culture that spans the larger organization. A major common goal that permeates the culture is the focus on delivering quality care. As such, sharing Mayday system data across organizations was viewed as a good "fit" with organizational and interorganizational goals. The result being that the new and additional information was accepted with few barriers. Participants noted several additional information sharing inhibitors related to interorganizational trust, effective communications, and overcoming resistance to change. They also noted some issues and challenges associated with sharing information—such as the resistance to change experienced by professionals who often feel overwhelmed with information overload. But the general theme across observations and interviews was the notion that information sharing was beneficial and fit well with organizational and interorganizational goals.

Governance: What Happens to Governance Oversight when the Test Project Ends?

For the Minnesota Mayday system, governance dimensions to information sharing include the looming legal and political concerns over data privacy and regulatory challenges associated with establishing a "data centric" 911information system. But the primary governance dimensions discussed center around the structure that facilitates information sharing. The Mayday system began as a test project that included clearly defined organizational roles, responsibilities, and funding sources. Now that the test has completed, the roles, responsibilities, and funding source are less clear, and the result is that continued system improvement has halted. Yet, participants noted the lack of guidance and uncertainty about a clear governance structure extends beyond the local or even the state-wide system. It includes a lack of well-defined directions and concentrated effort at the national level. And, on a related note, it is also the case that as a general matter the automobile companies have not traditionally been active partners in the emergency response arena. This is a notable omission, as experts agree that information about the crash (i.e., speed at impact) can be critical to understanding the nature and extent of crash injuries and consequent treatment course. In this sense, this case illustrates the opportunity to participate (for some organizations) and the opportunity to include others (for other organizations).

4.6.9 Challenges for Emergency Response System Development

Some unique challenges exist for those that are responsible for system development or the management of information systems associated with emergency response systems. Although the service aspect appears to be much like any other business process oriented system, there are distinctive operational, organizational, and governance structures that must be taken into account.

One challenge rests with deciding and deciphering which data sources will provide the most benefit to a wide range of users. In the Mayday case, there was a great deal of discussion across stakeholders about the operational and clinical benefits of automobile crash data. In the end, it was decided that the data would be useful and a system was designed to forward that data to multiple stakeholders. While automobile crashes are a significant public health issue, there are numerous other potential data sources that could be tapped to enhance emergency responses for other EMS related incidents—such as responses to non auto related trauma, cardiac, stroke, and a wide range of other health conditions. The use of cell phones to stream video, pre-existing data about patients that reside in hospital electronic medical records, electronic personal health records that are owned by citizens, and a range of other data sources could help to provide valuable information to quicken a response and allow for better clinical decision making. The initial challenge rests with understanding which data would be most valuable for performance improvement. In a related sense, defining "value" or "performance" from each stakeholder's point of view would be a parallel challenging activity. In this sense, a lack of standardization in the measures regarding system performance can lead to an inadequate ability to assess operational characteristics. Because performance measures are not comprehensive, justification for enhancements that may be thought to increase the impact of information technology associated with automation or process visualization may be difficult to quantify.

Beyond these initial challenges, decision makers must decide who should pay for the array of hardware, software, change management initiatives, and training programs needed to make use of the "valuable" data (once decided upon). To do so requires interorganizational agreement. Aligning organizational goals with system-wide goals is a challenge inherent in inter-organizational systems. The organizational structures must balance security due to privacy-oriented concerns with the overarching goal of system users, which is to save lives. Additionally, the organizations and the systems that they employ are varied due to factors associated with the diverse rural and urban settings in which they exist. While the Mayday project enabled and encouraged collaboration across traditionally disparate organizations, the historical evolution of the roles within typical organizations charged with emergency service provision tends to maintain strict separation of functions.

System developers must also be acutely aware of legal and political frameworks that impact information sharing systems. As discussed herein, emergency response systems and their users are governed by a number of federal, state, and local laws. For instance, 91-1 telecommunications services are regulated by the Federal Communications Commission, and data sharing between public and private entities is controlled through HIPAA regulations. Information systems that support emergency response must be designed to operate within the strict limits of a multitude of laws that must be clearly understood and embedded within the design of emergency medical response information systems.

The time-critical information services framework as applied to the Mayday system has shown how the key aspects of an emergency response system can be evaluated in terms of operational, organizational, and governance constructs.

4.6.10 Implications for Crises

A distinctive aspect of the Mayday project is that it introduced the automobile company as a partner in the emergency response system. By adding this partner, the system had better data on the crash and through this experience the test provided a taste of a new type of partnership that could be used to enhance emergency response systems. While the Mayday system, and this chapter, has primarily focused on the activities and perceptions of a "day-to-day" operational system, there are certainly implications for the disaster scenario.

The ongoing Mayday partnership provides a forum for emergency responders to address communications during large-scale emergencies. Automated data pushes allow dispatch centers a method for determining collisions involving multiple vehicles, such as in the case of the August 2007 bridge collapse on a major highway in Minneapolis. An early notification provides emergency responders the ability to assess volume and magnitude, rather than discovering the volume and degree of an incident through an onslaught of 9-11 phone calls—essentially taking more time and resources. A multi-organizational partnership also provides the opportunity to provide crisis information to citizens, or subscribers of wireless and in-vehicle systems. In this sense, it is important to note the recent partnership between GM OnStar and the American Red Cross to provide information to those in crisis situations.[xiii] OnStar can provide real time crisis information about centralized assistance for food, water, and shelter; share information from the American Red Cross database to connect subscribers to loved ones, family and friends; list GM OnStar subscribers on the American Red Cross Safe and Well website; and inform public safety, EMS, and other emergency responders when a GM OnStar subscriber is in need of disaster assistance. Such examples illustrate the opportunities and benefits made available through multi-organizational information sharing partnerships. Yet, despite the promise of new systems, the Mayday case study also reveals that all levels must be in play—operational, organizational, governance—if a promising system is to provide lasting value.

4.6.11 Acknowledgements

We gratefully acknowledge the transportation Center for Excellence in Rural Safety, Humphrey Institute, University of Minnesota for supporting this research. This work has also been supported in part by the National Science Foundation (Grant # 0535273). We are also grateful for the contributions of the many dedicated emergency professionals from the Mayo Clinic, Minnesota Department of Transportation, State Patrol, GM OnStar, and others from Minnesota.

4.6.12 References

1. Carr, B.G. BG, Caplan, J.M. JM, Pryor, J.P. (2006). A meta-analysis of prehospital care times for trauma. *Prehospital Emergency Care.* 10(2),pp. 198–206.

2. Davenport, T.H. (1994). Saving IT's soul: human centeredcentred information management. *Harvard Business Review* March-April 1994, pp. 119–131.

3. Fahy, M. & Murphy, C. (1996). From End-User Computing to Management Developed Systems. Proceedings of the 4th European Conference on Information System, Ed, Coelho, J.D., Jelassi, T, Konig, W, Krcmar, H, O'Callaghan, R., pp. 127–142.

4. Fedstats. Retrieved Aug. 3, 2006 fromhttp://www.fedstats.gov/qf/states/27000.html

5. Grossman, D.C., Kim, A., Macdonald, S.C., Klein, P., Copass, M.K., Maier, R.V. (1997). Urban-rural differences in prehospital care of major trauma. *The Journal of Trauma,* 42(4), pp. 723–729.

6. Gunton, T. (1988). *End User Focus.* Prentice Hall.

7. Horan, T. and Schooley, B. (2007). Time-Critical Information Services. *Communications of the ACM*, 50(3), pp. 73–78.

8. Institute of Medicine (IOM). (2006). *Emergency Medical Services: At the crossroads.* Washington, DC: National Academy Press.

9. Katz, D. & Kahn, R.L. (1978). *The Social Psychology of Organisations*. John Wiley, Chichester.

10. Linnell, M., Wainberg, S., Akundi, A. (2006, March). Mayday/911 Field Operational Test (FOT) Results and Path Forward Report (Document No. 232004). Prepared by Telcordia Technologies for Minnesota Department of Transportation.

11. McBride, N., Landau, R. & McRobb, S. (1997). Post-Modernist Information Management. *Proceedings of the 7th Annual Business Information Technology Conference,* Manchester Metropolitan University, November 5-6 1997.

12. McLean, S., Maio, R., Spaite, D., Garrison, H. (2002). Emergency medical services outcomes research: Evaluating the effectiveness of prehospital care. Prehospital *Emergency Care*, 6(2), pp. 172–179.

13. Metropolitan Council. Retrieved Aug. 3, 2006 fromhttp://www.metrocouncil.org/Census/census2000.htm

14. Minnesota Department of Transportation. (2006). Minnesota MAYDAY/911 Field Operational Test (FOT) Data Routing Project. Prepared for: USDOT ITS Joint Program Office—Public Safety Program and National Highway Traffic Safety Administration—Emergency Medical Services Division. Submitted by: Minnesota DOT. Last Revision: February 6, 2006.

15. National Academy of Science. (1966). Accidental Death and Disability: The Neglected Disease of Modern Society. The National Academy of Sciences Press: Washington D.C. Retrieved from www.nap.edu/openbook/POD716/html/1.html

16. National Center for Statistics and Analysis (NCSA), U.S. Department of Transportation. *Fatality Analysis Reporting System (FARS) Web-Based Encyclopedia*. Retrieved July 27, 2007 from http://www-fars.nhtsa.dot.gov

17. National Emergency Number Association (NENA). (2001). *Report card to the Nation: The effectiveness, accessibility and future of America's 911 service*. Columbus, Ohio: The RCN Commission and National Emergency Number Association. Retrieved on Sept. 15, 2005 from http://www.nena.org/Initiatives/RCN/RCN_FINAL_2-24.PDF

18. National Highway Transportation Safety Administration (NHTSA). (1996). *Emergency Medical Services Agenda for the Future*. Washington, DC: DOT.

19. National Highway Transportation Safety Administration (NHTSA). (2001). *National EMS Research Agenda*. Washington, DC: DOT.

20. Pons, P.T. & PT Markovchick, V.J. (2002). Eight minutes or less: Does the ambulance response time guideline impact trauma patient outcome? *Journal of Emergency Medicine.* 23(1), pp. 43–48.

21. Robson, W. (1997). Strategic Management and Information Systems. Pitman, London.

22. Sayre, M., White, L., Brown, L., McHenry, S. (2003). National EMS research agenda: Proceedings of the implementation symposium. Academic Emergency Medicine, 10(10), pp. 1100–1108.

23. Schooley, B. and Horan, T. (2007). End-to-End Enterprise Performance Management in the Public Sector through Inter-organizational Information Integration. *Government Information Quarterly,* 24(4), pp. 755-784.

24. Taylor, M.J., Moynihan, E.P & Wood-Harper, A.T. (1998). End-user computing and information system methodologies. *Information Systems Journal*, 8, 85–96.

25. Trunkey, DD. (1983). Trauma. *Scientific American.* 249,pp. 28–35.

26. Zinatelli, N., Cragg, P.B. & Cavaye, A.L.M. (1996). End-user Computing sophistication and success in small firms. *European Journal of Information Systems* 5, pp. 172–181.

Endnotes for 4.1

Towards End-To-End Government Performance Management

[i] 2005 estimated population retrieved from US Census Bureau on 8/17/06 at http://quickfacts.census.gov/qfd/states/06/06081.html.

Endnotes for 4.4

Performance Information Systems for Emergency Response: Field Examination and Simulation of End-To-End Rural Response Systems

[ii] A third related technical method involved creating a performance "ontology" for end-to-end systems. This ontology, as well are preliminary model construction, is presented in a more technical article (see Horan et al. 2004). The companion paper presented here focuses on the case study performance (vis a vis database) elements, including normal versus crises simulations and contextual findings.

[iii] The first three visits were: 1) Minnesota Statewide 911 Program, St. Paul, MN, June 19, 2003; 2) Baxter TOCC, Baxter, MN, June 20, 2003; and 3) Crow Wing County PSAP, Brainerd, MN, June 20, 2003. A fourth follow-up visit was made to the TOCC, on July 20, 2004.

[iv] The Rockwell Software Arena Simulation program was used in all EMS simulations.

[v] In a telling illustration of this data paucity, during one of the site visits, a large chart was posted in the Baxter-Brainerd TOCC conference room that portrayed increases in call volumes for the TOCC, though the data source was several years old. The TOCC manager acknowledged that indeed the chart represented the type of data (simple, not up to date) he had to work with in making plans and proposals for the TOCC.

Endnotes for 4.5

Time-Critical Services

vi Reported by the Alabama Chapter of the National Emergency Number Association (NENA). Accessed Oct. 30, 2006; www.al911.org/first_call.htm.

vii National Center for Statistics and Analysis (NCSA), U.S. Department of Transportation. Fatality Analysis Reporting System (FARS) Web-Based Encyclopedia. Accessed Oct. 24, 2006; www-fars.nhtsa.dot.gov.

viii Reported by the Cellular Telecommunications Industry Association (CTIA), retrieved Nov. 17, 2006; www.ctia.org/research_statistics/statistics/index.cfm/AID/10202.

ix Workshop information (accessed June 23, 2005); www.tcisresearch.org.

Endnotes for 4.6

User Perspectives on the Minnesota Interorganizational Mayday Information System

x National Safety Council. (2005-2006). Injury Facts. http://www.nsc.gov

xi http://www.tzd.state.mn.us/

xii See also http://www.comcare.org/VEDS.html

xiii See press release at: http://www.comcare.org/uploads/PR%20OnStar%20Press%206.7.07.pdf

CHAPTER 5

Lee Munnich: Industry Clusters and ITS

5.1 Applying Industry Clusters to Intelligent Transportation Systems

A New Framework for Analysis

Lee W. Munnich, Jr., Senior Fellow and Director
Principal Investigator
State and Local Policy Program
Hubert H. Humphrey Institute of Public Affairs
University of Minnesota
157 Humphrey Center
301 19th Avenue S.
Minneapolis, MN 55455
P: (612) 625-7357
F: (612) 626-9833
E-mail: lmunnich@hhh.umn.edu

James Lehnhoff
Research Assistant
State and Local Policy Program
Hubert H. Humphrey Institute of Public Affairs
University of Minnesota
Humphrey Center
301 19th Avenue S.
Minneapolis, MN 55455
P: (612) 626-0545
F: (612) 626-9833
E-mail: jlehnhoff@hhh.umn.edu

5.1.1 Executive Summary

Information and telecommunication technologies are transforming the way companies connect and interact with their workers, their upstream suppliers and downstream buyers, other businesses in the industry or local economy, as well as the end consumer. While these technologies have opened doors for new strategic models at the firm level, much less is known about their impact on industry clusters and regional economies overall. This is especially true in rural areas.

In order to understand how ITS is affecting rural economies, interviews with representatives from firms in the recreational transportation equipment industry and the wood products cluster in northwest Minnesota were conducted. The primary goal of this research is to analyze ITS use from the industry cluster perspective developed by Michael Porter rather than at the individual firm level. It is hoped that this analysis will lead to recognition of regional transportation, communication, and technology concerns and reveal avenues of further ITS development and implementation to keep rural economies competitive. This report has four primary objectives:

1. Apply the industry cluster analytical technique to better understand the role of transportation and technology in rural industries
2. Assess current ITS use in a rural industry cluster
3. Determine how ITS may be affecting a rural industry cluster
4. Explore future roles for ITS in rural industry clusters

ITS is enhancing the competitiveness of the cluster's rural location. Continued cooperation among large and small firms and continued support from ITS developers is needed to ensure the health of the cluster. ITS use ranges considerably between the two clusters but new technologies are being installed to make shipping, inventory tracking, and communication between companies inside and outside of the region more efficient, particularly in the recreational transportation equipment cluster. The larger firms are leading the way for the smaller firms in ITS adoption, but

smaller firms are reluctant to invest in new systems due to cost, lack of qualified personal, and perceived impracticality at their current level of business.

5.1.2 Introduction

Recent changes in information and telecommunication technologies have had a dramatic impact on the way we live and work. Information technologies—from desktop computers to remote sensors—have transformed how we collect, manage, understand, and communicate information. Meanwhile, telecommunications breakthroughs such as wireless technologies have granted us unprecedented flexibility in our ability to connect with others, all at decreasing real costs. Intelligent Transportation Systems (ITS) are designed to enhance information and telecommunications technologies for existing transportation systems, making them smarter, safer, more efficient, and laying the foundation for new modes of transportation.

For many businesses, information and telecommunication technologies are equally transformational. These technologies affect the way businesses connect and interact with employees, upstream suppliers and downstream buyers, other businesses in their industry or local economy, and the end consumer. While these technologies open doors for new strategic models at the firm level—supply-chain management and direct marketing strategies—much less is known about their impact on industry clusters and regional economies overall. This is especially true in rural areas. This study explores the impact of ITS technologies on rural industry clusters. There are four primary objectives of this study:

1. Apply the industry cluster analytical technique to better understand the role of transportation and technology in rural industries
2. Assess current ITS use in a rural industry cluster
3. Determine how ITS may be affecting a rural industry cluster
4. Explore future roles for ITS in rural industry clusters

To better understand the role that ITS technologies play in rural industry clusters, this paper presents a summary of ITS use and its affect on the recreational transportation equipment cluster and the wood products cluster in rural northwestern Minnesota. These clusters were identified in previous State and Local Policy Program (SLPP) research at the University of Minnesota's Hubert H. Humphrey Institute of Public Affairs.

5.1.3 Industry Cluster Framework

State and local economic development is fundamentally about increasing prospects for "place prosperity," or, in other words, for improving the economic outcomes for residents of a given city or region. Since a region's economic vitality has historically been linked to its industry mix, one of the most critical elements of state and local economic development policy is identifying, promoting, maintaining, and enhancing the competitiveness of industries that serve as the drivers of a regional economy.

With the gradual shift toward global and open economies, identifying the economic drivers of a region has become increasingly important and challenging. In the 1970s and 1980s, competition from both domestic and international, low-cost production locations overwhelmed established industrial regions and caused substantial de-industrialization of United States regions such as the Northeast and Midwest. In the face of this trend, a body of literature developed that endeavored to explain notable success stories, such as the shoemaking industry in northern Italy, industrial

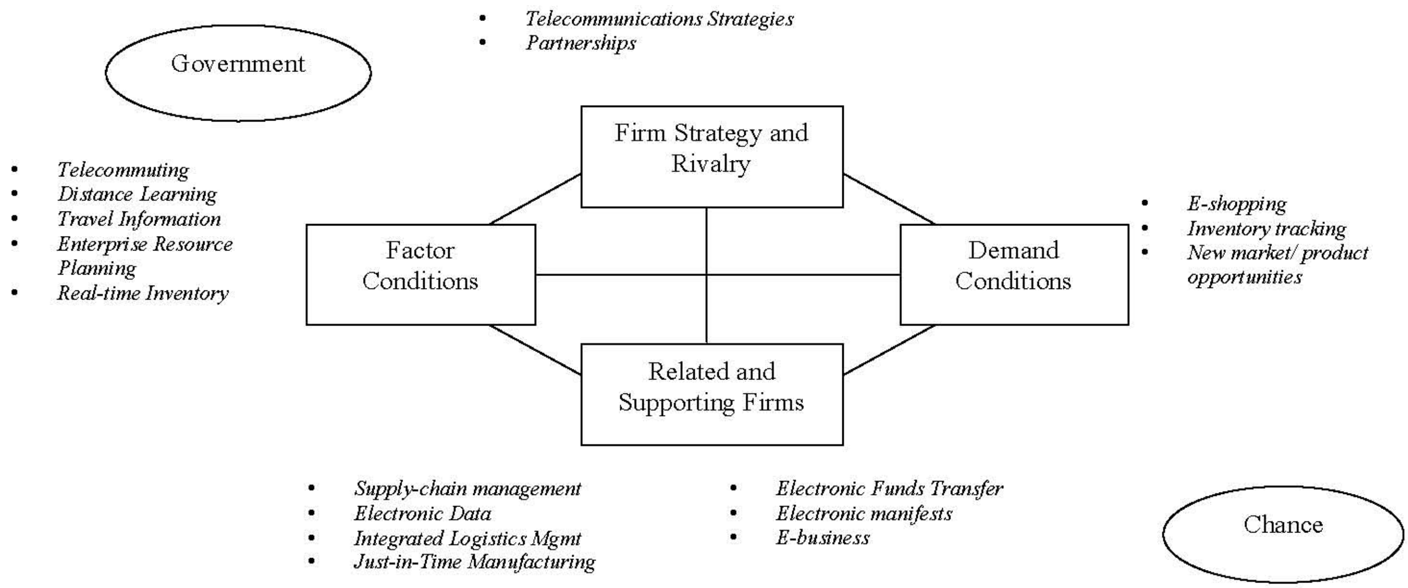


Figure 5.1.1 Diamond of advantage—potential applications of ITS to industry clusters

machinery in Germany and Japan, and high technology in Silicon Valley and Boston. In particular, these accounts noted the tendency of firms, both within a given industry and across related ones, to "cluster" spatially. This was most evident in industries where constant innovation in products and processes fostered self-sustaining regional competitive advantages on national and global levels.

The most compelling and lucid explanation of the cluster effect is from *The Competitive Advantage of Nations* by Michael Porter[1]. Industry clusters are, in sum, geographic concentrations of competitive firms in related industries that may or may not do business with each other but share similar needs for talent, technology, and infrastructure that, in turn, creates a source of jobs, income, and export growth for a region. In simpler terms, it is a localized form of agglomeration economies. Before Porter's theories on industry clusters, economists often discussed a region's comparative advantage as being based upon cheap inputs or low-cost labor. Porter, however, theorized that successful industry clusters could be explained and analyzed in terms of a "*diamond of advantage*" that drives innovation and results in a competitive advantage. This diamond consisted of four interrelated elements (Figure 5.1.1):

- *Factor conditions*—regional advantages such as human capital, physical resources, local specialized skills and knowledge, capital resources, and infrastructure can make a collection of firms more conducive to success; but disadvantages may also drive innovation.
- *Demand Conditions/Home Demand*—the nature of home demand for a product can dramatically affect development of a given product or service; strong home demand can lead to faster innovation among local firms vying for a local market.
- *Related and supporting industries*—when networks of buyers and suppliers are in close proximity, this can create faster and more active information exchange, collective learning, and supply-chain innovation.
- *Industry strategy, structure, and rivalry*—a climate that fosters both intense competition among localized producers, yet cooperation and collective action on shared needs, is most fertile for innovation and regional competitive advantage.

In addition to the four key elements, Porter also included a role for government and chance, which can play significant roles in the early development or location of industry clusters.Figure 5.1.1 illustrates several examples of ITS technologies within the diamond of advantage framework. These examples focus on potential uses of ITS technologies by rural industry clusters and are not all-inclusive.

Beginning in 1995, the State and Local Policy Program (SLPP) at the University of Minnesota Humphrey Institute of Public Affairs conducted industry cluster studies in five regions throughout Minnesota[ii]. In consultation with local officials, each study examined four industry clusters using Porter's diamond of advantage framework discussed above. The diversity of industries found in greater Minnesota is quite striking (Table5.1.1). While these industries are in various stages of maturity, each is important to their regional economy.

Given the difficulty of maintaining and enhancing rural economies and industry clusters, this research is but one step toward understanding what can be done to strengthen rural economies with the help of ITS applications. This research demonstrates that intelligent transportation systems are one part of improving the transportation and information infrastructure in rural areas; however, technology needs and progress very significantly from cluster to cluster and region to region.

5.1.4 The Industry Cluster Approach

Although Porter maintains that all four components of the diamond of advantage are needed for successful and innovative clusters, SLPP researchers have discovered that not all of the four components of the diamond of advantage must be within a small geographic proximity in order for the cluster to be functional[iii]. For example, with reliable telecommunication and transportation connections, firms can maintain relationships with customers and suppliers via email and fax, communicate complex information through a supply chain over the Internet or other private network, and utilize efficient just-in-time shipping. For an industry cluster to be functional and not geographically close, however, there must be an adequate infrastructure for both communication and transportation. This is especially true in rural areas where clusters are more likely to be spread over a larger distance.

Table 5.1.1 Rural Minnesota industry clusters studies by SLP

Twin Cities	Southeast Minnesota	Southwest Minnesota	Northwest Minnesota	Northeast Minnesota
Printing and Publishing	Composites	Computer and Electrical Components Manufacturing	Recreation and Transportation Equipment Manufacturing	Forest Products
Computers and Software	Food Processing	Value-Added Agricultural Cooperatives	Value-Added Agricultural Processing	Information Technology
Medical Devices	Printing, Publishing, and Software	Agricultural Equipment Manufacturing	Wood Products	Health Services
Machinery and Metalworking	Industrial Machinery and Computer Manufacturing	Dairy Processing	Tourism	Tourism
Financial Services				

Through interviews with firms in northwest Minnesota, this research analyzes ITS use within Porter's diamond of advantage framework. The survey questions focus on ITS's effect on demand and factor conditions, firm rivalry and strategy, and firm relationships in the cluster. By understanding how a cluster uses ITS technologies, it is hoped that the technological needs of the cluster as a whole, and hence the regional economy, can be made stronger, more efficient, and more competitive. Although the goal of this research is to assess a rural industry cluster's ITS use, the broader objective of this paper is to show how Porter's diamond of advantage framework can be used to analyze a specific aspect, such as technology and transportation, of an industry cluster.

5.1.5 Methodology

This research involves three tasks: selecting the industry clusters to study, consulting with rural economic development and industry cluster experts, and conducting a series of interviews with businesses in the clusters to evaluate current ITS use and its potential roles for further use.

Task 1: Identify rural industry clusters and select a cluster for a case study.

The recreational transportation equipment cluster, located in northwestern Minnesota, and the wood products cluster, located in north-central Minnesota, was selected for closer examination. The recreational transportation equipment (referred to at the RTE cluster from here on) cluster consists of two key manufacturers, Polaris and Arctic Cat, as well as a series of firms that supply one or more of these hub producers. The wood products cluster includes a small number of national manufacturers and many small, locally based loggers and manufacturers. In order to protect an individual firm's privacy, no firm has been directly identified in the summary of results.

These clusters were chosen for several reasons. First, the recreational transportation equipment cluster is a successful and growing cluster despite being located in the most sparsely populated region of the state. Additionally, this cluster is considered a "classic" industry cluster due to its strong manufacturing base, internal cooperation and competition between producers, local supply networks, significant economic importance to the region, strong local and national demand, and for not relying strictly on cheap labor or locally available raw materials. Finally, because the cluster is facing increasing competition from foreign and domestic markets, the region must find new ways to be innovative and more efficient.

The wood products cluster was chosen primarily because of its historic importance to the regional economy, direct ties to local raw materials, which is a contrast to the RTE cluster. The wood products cluster is also facing increasing competition, particularly from Canadian manufacturers, that make finding ways to increase efficiency and competitive advantages more important. These two clusters have significantly different needs and resources, but both are facing competition from outside producers.

Task 2: Convene national rural cluster experts for consultation.

National experts on industry clusters and rural development were assembled for a research roundtable in the Twin Cites in October 2001[iv]. The roundtable was convened to discuss the current state of academic and practitioner research, new theoretical models, and potential case studies for rural industry clusters. While the roundtable discussed many topics related to rural economic development, two findings were particularly important for understanding the inter-section between rural industrial clusters, information technologies, and transportation technologies:

- Scale and proximity of rural clusters: rural industry clusters tend to lack the agglomeration and scale more common in metropolitan clusters. They may be spread across wider geographic distances, resulting in greater reliance on transportation services to link buyers and suppliers and on information services for communicating with clients and collaborators.

- Rural disadvantages in producer services: producer services, such as financial, information technology, research and development, engineering, management consulting, and transportation services (particularly high-end or sophisticated services), remain highly concentrated in metropolitan areas. These services tend to be "catalytic" in nature, helping companies to innovate, collaborate, and research new markets. Thus, the relative disadvantage of rural areas due to small scale would tend to be self-perpetuating. This suggests the potential for local or region-wide collaborative solutions to overcome high costs associated with accessing new technologies.

The surveys, discussed in more detail below, strongly indicate that the problems identified by the roundtable participants are indeed barriers in northwest Minnesota and that ITS is playing a role in overcoming these disadvantages, though much more so in the RTE cluster.

Task 3: Conduct industry interviews regarding of supply chain relationships and examine the importance of the related transportation and information technologies.

Several firms of varying sizes were interviewed during January 2003 and March 2003. Questions were asked about current ITS use and how ITS use is affecting business relationships within their supply-chain. The types of questions asked include:

- How did transportation and communication costs affect your firm's location?
- How do you communicate with suppliers and consumers?
- How do you transport your products? What kinds of transportation and information networks are used to link firms to suppliers and customers?
- Are you involved in any type of computer or communication network that connects you to other firms in the region?
- Have your transportation needs changed in recent years?

One of the goals of this project is to determine if ITS technologies could have a greater role in rural economic development, particularly in relation to rural industry clusters. While this report will not recommend specific strategies for individual firms, the industry interviews did reveal some trends for potential ITS implementation that will need further exploration and discussion beyond this study.

5.1.6 Northwest Minnesota Geography

Northwest Minnesota is home to approximately 88,472[v] people in seven counties: Kittson, Marshall, Norman, Pennington, Polk, Red Lake, and Roseau (Figure 5.1.2). The western edge of the region is primarily farmland located in the Red River Valley and the eastern portion is mostly forests, lakes, and wetlands. No interstate highway crosses the region, though Interstate 29 runs north-south along the North Dakota border and Interstate 94 runs east-west immediately south of the region. US Highway 2 also runs east-west through the southern portion of the study area.

5.1.7 ITS Industry Cluster Evaluations

The first part of the industry cluster analysis focuses on a case study of the recreational transportation equipment cluster. The second part is a less in depth summary of ITS use in the wood products cluster due its smaller reliance on ITS technologies. These assessments are organized within the industry cluster evaluation framework: factor conditions, demand conditions, related and supporting industries, and industry strategy/rivalry. After the case study and analysis, there is an analyses of how ITS is affecting the clusters, which is followed by a brief examination at the potential for future ITS development in the cluster. While the number of firms surveyed was relatively small, certain trends did emerge. Throughout this analysis, the terms "larger firms" and "smaller firms" are widely used. Due to the small sample size, these general terms were adopted to broadly refer to businesses in northwestern Minnesota. Larger firms refer to those companies that ship a significant percentage of their products outside of the region, have more than fifty workers, and are drivers of the local economy. Smaller firms generally refer to those companies with fewer than fifty workers, sell a larger percentage of their products to other firms in the region, and supply the larger firms in the region.

ITS and the Recreational Vehicle Cluster

The recreational transportation equipment cluster in northwest Minnesota includes well-known final goods manufacturers such as Arctic Cat and Polaris as well as many smaller producers and suppliers. While the region is traditionally known for its snowmobile production, other products such as all-terrain vehicles (ATV's), jet skis, and track conversions for four-wheel drive vehicles are taking on an increasingly important role. The larger manufacturers have historically produced final consumer goods primarily for the upper mid-west; however, in recent years, their markets have expanded both nationally and internationally. The smaller manufacturers in the region produce supplies for other firms in the region and some limited final goods as well as products for other manufacturers in the rest of Minnesota and to a lesser extent the rest of North America and for international export.

For the most part, the interviewed firms reported expanding operations despite the recent economic slowdown. In fact, several of the firms expressed concern over keeping up with demand, locating additional qualified workers, and wanting to slow growth in order to re-evaluate their current position and customer base. Part of this re-evaluation for many firms has included updating, or at a minimum, assessing their current technology use.

Population (2000)	88,472*
Major Cities	
Crookston	8,192
East Grand Forks	7,501
Thief River Falls	8,410
Roseau	2,756
Population Density (pop/sq mi)	11
(Twin Cities	601
MN state:	62)
Population Growth (1990–2000)	–2%
MN non-metro	4%;
US non-metro	9%
Source	Census Bureau
Per Capita Income (2000)	$25,135
MN non-metro	$24,134
US non-metro	21,847
Per Capita Income Change (1990-2000)	56%
MN non-metro	54%;
US non-metro	48%
Job Growth (1990–2000)	16%
MN non-metro	25%
US non-metro	18%
Farm Employment (2000)	14%
MN non-metro	9%;
US non-metro	6%
Manufacturing Employment (2000)	16%
MN non-metro	15%
US non-metro	15%

Source: Bureau of Economic Analysis; income change data in nominal terms, not adjusted for inflation.

* Data in Table are for region that includes: Kittson, Marshall, Norman, Pennington, Polk, Red Lake, Roseau counties (Region 1).

Table 5.1.2 Northwest Minnesota key factors

Figure 5.1.2 Study region

Factor Conditions

- The distance from major markets and transportation connections has increased pressure for efficient, cost effective, and reliable transportation. This is a concern for the larger firms that ship considerably more final goods out of the region and for all firms that ship supplies into the region.
- All of the interviewed firms use some form of product, inventory, or supply tracking, though sophistication varies widely.
- All of the interviewed firms have access to high-speed Internet but not all firms have a web site or use the Internet to aid transportation or communication.
- There is an interest in distance learning by many of the firms and a few are considering long distance training accessible through the Internet.
- Given the distance between towns in the region and major markets outside of the region, manufacturers in the region are concerned with the current and future conditions of roads and highways.

- There is potential for increased use of travel information, specifically weather monitoring and road construction inside and outside of the region.

On average, most firms in northwest Minnesota are six or more hours from the nearest large metropolitan area, the Twin Cities. Although the cluster's location does help drive technology innovation and has increased the necessity of incorporating product and supply tracking technologies, competitive forces are equally important. According to the firms, competitive forces inside and outside of the region have required those technologies be adopted in order to remain viable. In some cases, the producers and suppliers require advanced telecommunication technologies be used in order to have any business-to-business relationship.

Since moving supplies and products in and out of the region is a major concern, businesses require fast, cost effective, and reliable shipping companies; however, on-time service was rated as the most important aspect of shipping. The larger firms ship products almost exclusively by truck with independent shipping businesses. The smaller firms do use independent shippers but also rely on nationally known package shipping services such as UPS and Federal Express. Through UPS and Federal Express, firms are able to use near real-time tracking services, however, this is not the case with all independent contractors. The larger firms in the region are moving toward, and some already require, shipping contractors to have real-time tracking or other similar services. This is being done for two reasons. First, it helps to streamline shipping costs and reduce backlogs. Secondly, it helps to ensure on time deliveries for distributors and end consumers. The larger firms have also begun to integrate real-time tracking into their just-in-time (JIT) shipping systems to increase the reliability of their JIT schedules.

While larger firms already use relatively sophisticated supply, inventory, and product tracking systems, the same is not true for many of the smaller firms, but the need for such software in the future is recognized. There were several reasons cited for not implementing full just-in-time methods or other electronic inventory management projects. Some firms simply did not see the need given their current business level; however, all firms agreed that the need is growing, especially for cost control and product tracking purposes. Quite often, the push for new technologies is due to business partners using such systems. However, all of the firms were particularly concerned about implementation costs, time, and support. One firm shared a story of the struggle over implementing a new software program that proved to be time-consuming to learn and vendor support was lacking. While most firms see the long-term necessity for using ITS-related tracking programs, many have found it difficult to justify the perceived short-term cost.

A recent report from Minnesota Technology[vi] found that Internet use and access is virtually ubiquitous throughout rural Minnesota. That has certainly proven true in northwest Minnesota for the RTE cluster. All of the firms surveyed in this cluster use high-speed Internet connections; however, not all of the firms have web pages for business-to-business or business-to-consumer sales and information. The larger firms tend to have both business-to-consumer and business-to-business web sites for sales and information distribution. Few of the smaller firms have web pages, but several of the businesses are considering adding a web page for information dispersion and possibly direct sales.

The physical infrastructure of the region is increasingly a prominent concern. Since northwest Minnesota is a considerable distance from most major markets and has relatively few major highways, road maintenance is extremely important. While ITS is helping firms maintain their business in northwest Minnesota, without well-maintained roads, heavy shipping would be hampered.

Though not addressed in detail and not widely used, a few companies also use tracking technologies to adjust schedules in case weather should interfere with shipping in production. Northwest Minnesota is known for its cold weather and blizzards, which have the potential to severely affect shipping for days at a time. While the last several winters have been relatively mild, firms using this technology are able to route drivers along the best roads in case of road closures due to storms as well as rail cargo to adjust schedules as needed.

Demand Conditions/Home Demand

- Company web sites are common among the larger firms and developing in the smaller firms.
- Demand conditions in the region, nationally, and internationally are forcing all of the firms to cut costs and make the entire manufacturing process more efficient.
- Foreign demand is increasing and ITS is helping to coordinate shipping.

With the Internet boom in the late 1990's, one would think that nearly every company would have a fully interactive web site with direct sales. This, however, is generally not the case and is perhaps another lesson of the tech boom. Larger firms such as Polaris and Arctic Cat have advanced and interactive web sites, but these sites are primarily for information distribution with some limited direct sales that are often for small accessories and clothing, which make up a very small percent of total sales. Several other firms that do not have web sites either do not see the need or are in the process of designing or deciding on a purpose for a new web site. Ironically, one of the smallest firms in the region is also the most dependent on the Internet with over 60% of sales linked to the firm's web site. The owner stated that without continued Internet access the business would not have the level of business it currently has.

Increased regional, domestic, and overseas competition is forcing all companies to cut costs wherever possible and to make manufacturing more efficient. Larger firms placed a particularly heavy emphasis on increasing transportation efficiency since a bulk of their final products are shipped out of the region and can add substantial costs to the final products. ITS tracking technologies are allowing larger firms to get products where they are needed, when they are needed more efficiently. ITS technologies are also helping firms make certain that supplies are delivered on time to ensure that manufacturing lines do not shut down and inventories sit idle—both of which add to the cost of the final product.

As much as 10 percent of total sales for larger firms are due to foreign sales and demand is increasing. While all of the firms primarily transport products by semi-trucks, companies also use rail and ship to get products overseas. This requires an increased level of coordination between trucks, rail, and ship, which is being aided by ITS.

Related and Supporting Firms

- Up to 30% of larger firms' supplies originate in northwest Minnesota.
- Relatively sophisticated supply-chain management technologies are common among larger and mid-size firms.
- Larger firms are moving toward systems that require suppliers, transporters, and distributors inside and outside of the region to connect to their computer systems and follow their production schedules.
- Almost all of the surveyed firms use network connections to transfer design files, though some are more advanced than others.

Larger firms tend to be the drivers of technological change in the region, which is likely due to their more advanced human resources and capital as well as their need to compete directly with international firms outside of the region. Some of the larger firms have only recently begun to require that most or all of their suppliers use compatible supply-chain management technologies; however, implementation of these technologies varies. When asked why smaller, local suppliers were not always required to participate in electronic supply-chain management systems, the most common response was practicality. Since the suppliers in the region are considerably closer than suppliers outside of the region, larger firms, at least until recently, simply telephoned, faxed, or emailed orders to their local suppliers. This is changing as larger firms move toward fully integrated supply-chain electronic management systems based on just-in-time shipping ideas. Although firms have different timelines for making this transition, it is occurring.

Since smaller firms have expressed concern over the cost and time of implementing such technologies, those that do participate often only connect to the firms that require it. In an attempt to move smaller firms into such systems, larger firms have offered limited training programs to educate suppliers in and outside of the region on the new systems. This, however, does not force smaller firms to add internal systems to track their suppliers, inventories, or final products. Smaller firms do recognize the need to install more advanced inventory control systems in the future but no definitive timelines were given.

In association with using ITS technologies to track materials through the supply-chain, all of the RTE cluster firms either were transferring data files via the Internet or expressed interest in learning how. All firms using this relatively simple technology lauded the convenience and efficiency of transferring CADD (computer aided drafting and design) over high-speed networks. One smaller firm stated that transferring CADD files over high-speed networks decreased the time needed to produce new parts because updated designs can be worked on in two locations and continuously revised without stopping production or waiting for redesigns from other firms.

Firm Strategy and Rivalry

- Although in its early stages, several firms have recently formed a cooperative association that could potentially be used to combine resources to make transportation more efficient and new technology and business training cheaper.
- Transportation is no longer considered a sunk cost by many of the firms but is instead a flexible cost that can be adjusted to make a company more competitive.
- History has rooted the firms in the region and none foresee any reason to change locations. ITS-related technologies are considered an important tool for remaining in the region.

The Manufacturers Association is a newly formed organization in northwest Minnesota. Although the Manufacturers Association is still in its early phase of creation, the interviewed firms that participated in the first meeting of this still-forming organization expressed interest in its potential as a better network for companies to share training opportunities, information, and strategies. One firm in particular was interested in learning new software for tracking inventory and CADD applications but was unable to do so individually because of the cost. It was this firm’s hope that the Manufacturer’s Association could organize firms to share costs on such training.

According to the larger firms, ITS technologies that track supplies, final products, and inventories have drastically changed the way their companies view shipping. Previously, shipping costs were considered a static part of doing business. Products needed to be shipped or there was simply no business. Although this basic premise has not changed, what has changed is that shipping costs are now considered more flexible. ITS technologies have made it easier to order and track supplies coming into the region. The larger firms reported losing several thousand dollars an hour if assembly lines shut down due to late arrival of supplies. One firm cited that new tracking technologies have decreased assembly line shutdowns by as much as 70 percent. By being able to quickly order and accurately track supplies, firms in the region are able to maintain JIT schedules without retaining large inventories. To make this change successful, however, requires ITS technologies to be installed at all levels of the supply-chain. Not all companies have taken this step. Smaller firms fear the cost and time commitments of installing these technologies, particularly if the technologies do not pay off. It is likely, however, that many smaller firms will be forced into these technologies due to their reliance on business from larger firms, which tend to use more sophisticated ITS technologies. Larger firms indicated a continued desire to decrease or, at a minimum, stabilize shipping costs into the future in order to remain competitive with international companies that often have lower manufacturing costs.

All of the interviewed firms plan to stay in northwest Minnesota and continuously update business methods to stay competitive. Tracking technologies have made traveling and shipping more efficient and firms are able to plan supply lines down to the hour. This would be virtually unattainable without real-time or near real-time supply and product tracking technologies. Firm rivalry both inside the region and outside the region is one of the primary drivers of these innovations.

ITS and the Wood Products Cluster

The wood products cluster is in a notably different position than the recreational transportation equipment cluster in terms of ITS usage, implementation, and planning. The RTE cluster has a larger market outside of northwestern Minnesota and a more complex supply-chain in terms of inputs and the number of companies in the chain. Although there are a few wood-based companies in northwestern Minnesota that sell products to the national and international market, most of the firms' sell their products in the upper Midwest or, more commonly, throughout Minnesota. While the RTE cluster is larger, the wood products cluster is no less important. Nevertheless, the wood products cluster certainly has different needs and is adopting technology at a different rate.

Despite the common need for wood, the products manufactured by the wood products cluster are quite diverse. The cluster manufactures products ranging from paper, hockey sticks, roof trusses, wood paneling, to entire homes. While the wood products cluster includes such well-known companies as Potlatch, Marvin Windows, and CB Hockey, these companies do not directly compete like the primary companies in the RTE cluster, and these companies have different, non-competing supply chains. Generally speaking, the wood products cluster does not have the same forces driving technological change as the RTE cluster.

Although northwestern Minnesota is home to abundant natural resources and industry growth during the 1980's and 1990's, the wood products cluster has been hurt by the current ailing economy and increasing competition from Canadian manufacturers[vii]. Large and small firms alike are feeling the pinch from cheaper Canadian and other foreign inputs in spite of the fact that northern Minnesota is rich in wood resources. Despite increasing competition and rising costs of business, the wood products cluster has not moved toward adopting ITS technologies at the same rate as the RTE cluster. The industry cluster analysis focuses both on ITS use and on factors that might be affecting ITS technology adoption.

Factor Conditions

- The majority of products created by the wood products cluster are sold in northern Minnesota.
- Few of the firms have adopted product, inventory, or supply tracking technologies.
- All of the firms use the Internet and email, and a majority of firms have a web site, but the Internet does not play a major role in business transactions.
- Physical transportation conditions are considered adequate by most firms.
- Most products are shipped via private or company owned trucking operations, but UPS and Federal Express are also major carriers for wood products.

Although the wood products cluster faces many of the same challenges as the RTE cluster, there are important differences. Both clusters are far from major metropolitan markets, face similar physical transportation limitations, are under increasing competition from foreign suppliers, and transportation prices are increasing. Since the major market for the wood products companies is, quite literally, their own back yard, few of the companies expressed the need for tracking technologies. Many of the wood products companies are relatively small, with fewer than ten workers, and cannot justify the cost of installing supply or product shipping technologies when they have only a handful of suppliers and consumers. Additionally, several of the firms ship their own products on company owned trucks due to special shipping needs for bulky products and unique product designs. Like the RTE cluster, UPS and Federal Express play a key role in shipping products in and outside of the region. UPS and Federal Express are also the primary providers of product and supply tracking via their web site. A few private shipping companies and rail shipping providers do offer tracking technologies, but none of the wood products companies expressed an intense desire or need for those technologies.

All of the interviewed firms use email and many of the firms have a web site; however, very few sales are completed via the Internet even for the largest companies. For those companies that have a web site, the most common purpose is for information distribution and various forms of advertising. None of the interviewed firms expected to increase direct sales via the web anytime soon.

Demand Conditions/Home Demand

- The major market for almost all wood products firms in northern Minnesota.
- There are a few international exports in the wood products cluster.
- Imports from Canada are a significant concern for nearly all interviewed firms.
- Many companies have a web site, but mostly for informational purposes and not for direct sales.
- Construction demand has increased the need for wood products.

As stated previously, the major market for the northern Minnesota wood products cluster is northern Minnesota itself. Additionally, the international export market for wood products out of this cluster appears to be shrinking while imports from Canada are increasing. According to the interviewed firms, imports were cheaper due to the high price of the American dollar compared to the Canadian dollar and not due to any technological advantage. Since many of the products are, for the most part, a basic commodity, there is little reason to ship products outside of the region, though a few companies do manufacture unique wood products that compete in regional and national markets.

Like the RTE cluster, the Internet is important for distributing information, but few of the companies see this as a source for direct sales. Several of the companies mentioned that their current buyers have been the same buyers for years and that the current buyers are unlikely to change anytime soon. A few of the companies do sell to distributors with larger markets, but, again, these sales are based on long-term relationships with little need for change, according to the surveyed firms.

Although the recent economic slowdown has dampened the need for construction materials, a strong consumer of wood products in northern Minnesota is the housing market. However, none of the companies expressed a need for a higher level of technology to take, build, or distribute wood products orders for the housing construction industry. The short distance for travel and the relatively small size of companies and sales does not necessitate a more integrated, electronic system.

Related and Supporting Firms

- The supply-chains and inputs for most companies are relatively small and consist of a few local or national firms.
- Only the largest firms use ITS tracking technologies such as just-in-time shipping between firms. When they are used, it is usually with parent or sister firms outside of the region.
- Supply-chain management technologies are not common among the wood products cluster firms.
- Unlike the recreational transportation equipment cluster, the wood products cluster does not have a collection of competing firms that drive technology adoption among other firms.

Since wood is the primary input for all of the wood products firms, most of the companies have standing orders for loggers both in and outside of Minnesota. Surprisingly, many of the interviewed firms import wood from other states and countries despite the abundant supply of wood in northern Minnesota. This is usually due to a need for a specific type of wood in the manufacturing process. Although the orders are often faxed and a growing number are emailed, only the largest companies use any ITS tracking technologies between firms. The small supply-chain length limits the necessity and complexity of supply tracking for most firms. While other supplies besides wood are needed in the manufacturing process, these supplies are also obtained with standing orders that are completed via phone, fax, and email. None of the companies indicated that they use an electronic ordering network or just-in-time shipping technologies to complete ordering or financial transactions.

Whereas the RTE cluster has large companies drive technology change down the supply-chain to improve efficiency and employ "lean" manufacturing methods, the wood products cluster has no single or collection of companies that drive technology adoption. Since the supply-chain is small and the number of inputs is limited, the firms did not see the transportation or communication process as an untapped source for increasing efficiency. The use of electronic file sharing, such as CADD files, is also uncommon among the wood products cluster.

Firm Strategy and Rivalry

- Transportation costs limit the market size for most wood products companies.
- None of the companies felt that changing locations would increase their competitive advantage.

The sheer weight of transporting wood products and the proprietary truck trailer design required to ship many of the products increases the cost of transportation. Because of this, each wood products company has its own market region that may or may not overlap with a similar company that also has its own market region a given distance away. At a certain distance, it becomes prohibitively expensive for a company to transport its product because a competing company will be able to charge a lower price due to lower shipping costs. In northern Minnesota, these competing markets are well established and most of the firms haul their own products, particularly wood products made for the construction industry. Although some of the companies use cell phone to check on the drivers, as already noted, none of the surveyed firms use a higher level of technology to track incoming supplies or outgoing products. The small size of the market and the well-established supply chains, whether local or national, does not warrant a higher level of technology use according to most of the firms.

Although many of the firms plan or desire to establish larger markets, none of the firms expressed a desire to relocate to gain a larger market. Nearly all of the locally owned companies were established in the region and have not considered nor do they plan to consider moving outside the region to gain a larger market. A few of the locally established firms have since been bought by companies outside the region, and the investments made by the outside companies keep those firms in northwestern Minnesota. Of all the companies surveyed, only one company expressed a concern over the slow technology adoption by most firms and strongly felt that this would negatively affect the industry in the future, though no specific information was given. However, several firms did comment on the aging nature of ownership and workers in the region and cited that this factor may be holding back technology implementation and adoption of new manufacturing, communication, and transportation techniques.

5.1.8 ITS Issues for Further Research

Clearly, ITS use, adoption, and implementation are not only different for each firm but for each cluster as well. Whereas ITS technologies are integral to the RTE cluster, the wood products cluster is much less dependent on such technologies. However, ITS technologies are but one of many parts of a successful rural industry cluster. While it is beyond the scope of this paper to prescribe specific ITS technologies that would benefit the two clusters, this section will draw upon themes from the interviews to show important ITS technologies being used now and to briefly describe how ITS needs are likely to grow.

The recreational transportation equipment cluster is growing, and so is the cluster's need for improved communication, information and inventory management, supply tracking, product tracking, and information distribution. However, while ITS has made crossing distances in the RTE cluster easier, there is still a certain level of convenience, historic connection, and cost-effectiveness to keeping suppliers and producers in relatively close proximity whenever realistically possible. The larger companies in the RTE cluster are driving technology adoption up and down the supply-chain, and that is likely to continue into the future. For smaller firms to survive, there must be continued technological support from ITS developers and from larger firms that are adopting the technologies first. Larger companies moving towards electronically integrated supply chains is intended to bring all suppliers closer electronically if not physically, however, suppliers in the region still have the potential to lose their competitive edge of being physically closer.

Since a great deal of the raw materials needed to make RTE-related product come from outside the region, in some cases as far as Europe and Japan, integrating just-in-time shipping with near real-time supply tracking makes it possible to rely on strict schedules that reduce inventories and ship final products to the proper place at the proper time. Firms in the region stated that moving to these technologies is essential to stay in business—not just because of the rural location. Without these technologies, smaller and larger firms would be significantly challenged by outside competition.

ITShas helped keep the rural RTE cluster competitive by increasing supply-chain transportation and communication efficiency. In order to continue the RTE cluster's success, companies will need to continue to innovate in the supply chain, adopt firm-size appropriate technologies, and communicate with other firms in the cluster to ensure efficient and compatible electronic systems. ITS has made it easier for the cluster to stay cohesive as well as competitive, but the various firms must continually monitor new ITS technologies to stay competitive with firms in and outside of the region.

The wood products cluster is in an entirely different situation. The small size of the market for many of the wood products firms limits the amount of capital investment and necessity for ITS technologies. While not every firm would benefit from implementing ITS technologies, there are firms in the region that, in the future, may benefit from ITS technologies. For this to happen, however, there must be a strong case for improved efficiency and low-risk to purchase ITS technologies. For example, basic inventory tracking and supply tracking technologies may be most beneficial. With many of the companies relying on raw materials from outside the state, a delay in shipment due to weather, road construction, or other problem could significantly hamper normal business activities.

Although there was not obvious distaste for new technologies in the wood products cluster, most of the firms simply did not perceive a need for sophisticated ITS technologies. In the immediate future, ITS adoption will likely come from larger firms that are supported by regional and national markets that require strict schedules in order to meet the needs of many consumers outside the region.

5.1.9 Conclusions

There is a future for ITS use in northwest Minnesota not simply because it is rural but because competition requires it. The challenge for ITS developers is to make user-friendly systems for smaller firms and to coordinate ITS technologies in cohesive, unified industry clusters. Arctic Cat, Polaris, Potlatch, and the myriad of other firms in northwestern Minnesota must continue to find rural northwestern Minnesota advantageous if they are to remain there. Simply providing them with ITS technologies will be insufficient—there must be support for ITS technologies, coordination between businesses, and expandability for growing businesses that may not need the most expensive technologies immediately. The smaller supplier firms must continue to increase efficiency to remain competitive with firms inside and outside of the region and that efficiency will likely run through the supply-chain to the larger companies.

Firms in northwestern Minnesota are not independent of each other. There are long standing and historical relationships between many of them. If the cluster is to remain viable, particularly for the sake of the smaller companies, there must be cooperation and communication in the supply-chain, particularly in the recreational transportation equipment cluster. Given increased worldwide competition, larger companies will only remain in the region and keep local suppliers as long as it is financially reasonable to do so. Larger companies and smaller companies in close proximity retain a synergy from working with each other, and it is this symbiotic relationship that ITS is helping to maintain.

The industry cluster approach analyzes a collection of industries rather than single firms to better understand how regional economy drivers function together. Using this approach to analyze technology use in a cluster is a useful technique for understanding how firms in the cluster are communicating as well as finding their strengths and weaknesses. By recognizing how technology use is changing within the cluster, these strengths can be built upon and weaknesses addressed. This differs from evaluating a single company because a cluster is attached to far more jobs than any single firm. By finding ways of keeping an entire cluster competitive, existing firms will strengthen and other firms will be drawn into the mix to support the cluster and connect into the infrastructure.

The industry cluster analysis evaluates the current situation, addresses how local and outside demand affects the cluster, assesses firm strategy and rivalry, and analyzes the individual companies in the cluster. The key part of industry cluster analysis is understanding how and why firms interact and, how ITS is affecting those interactions both inside and outside of the cluster. This paper has provided a brief investigation of the situation and illustrated interactions in northwestern Minnesota that can be used for future analysis. The primary implication of using cluster analysis is that once the existing situation is understood and how its likelihood of change is evaluated, ITS experts can use the information to make recommendations to technology policy-makers, ITS developers, and to firms within the context of sustaining a regional economy.

5.1.10 Acknowledgements

The Humphrey Institute of the University of Minnesota is hospitable to a diversity of opinions and aspirations. The Institute itself does not take positions on issues of public policy.

The University of Minnesota is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation.

The State and Local Policy Program thanks the companies that participated in the survey.

Thank you to Greg Schrock for his contributions in initiating this study and to Karen Cook and Marta Murray-Close for their advice and guidance. Additional thanks go out to the Minnesota Department of Trade and Economic Development, the Northwest Minnesota Foundation, Northwest Technical Colleges, and the Headwaters Regional Development Commission.

This study was coordinated under a multi-year research project started in 2000 by the ITS Institute of the University of Minnesota Twin Cities Center for Transportation Studies with funding from the Transportation Equity Act for the 21st Century (TEA-21) through the Federal Highway Administration.

5.2 The Industry Cluster as a Planning Construct for Freight ITS

Chandler Duncan (Corresponding Author)
Wilbur Smith Associates
325 Cedar St., Suite 700
St. Paul, MN 55101
P: (651) 291-8547
F: (651) 292-1482
E-mail: cduncan@wilbursmith.com

Mark Berndt
Wilbur Smith Associates
325 Cedar St., Suite 700
St. Paul, MN 55101
P: (651) 291-8547
F: (651) 292-1482
E-mail: mberndt@wilbursmith.com

Colbey Sullivan
Hubert H. Humphrey Institute of Public Affairs
State and Local Policy Program
University of Minnesota
157 Humphrey Center
301-19th Avenue S.
Minneapolis, MN 55455
P: (612) 625-7357
F: (612) 626-9833
E-mail: csullivan@hhh.umn.edu

Lee Munnich
Hubert H. Humphrey Institute of Public Affairs
State and Local Policy Program
University of Minnesota
157 Humphrey Center
301-19th Avenue S.
Minneapolis, MN 55455
P: (612) 625-7357
F: (612) 626-9833
E-mail: lmunnich@hhh.umn.edu

5.2.1 Introduction

As regional ITS architectures become more widespread and sophisticated, commercial ITS systems are becoming increasingly relevant in the concepts of operations for both public and private sector transportation and infrastructure providers. Identifying, defining and implementing the incentives and roles of industrial and business partners in ITS architectures is a critical component of effective freight ITS planning; however very few cases of truly integrated public and private systems have succeeded. The collaboration between private firms in different industries surrounding transportation technologies, and between private firms and public entities requires a conceptual framework rooted in both the fields of transportation planning and economic development, as well as supply chain management. The concept of the "industry cluster" has been widely researched and implemented in the economic development literature, but has yet to find widespread application in transportation planning. This paper (1) identifies and defines the value chain industry cluster as an appropriate construct for ITS planning (2) suggests a four-step industry cluster based approach to developing integrated public and private ITS concepts of operations and architectures within the USDOT guidance for Regional ITS Architecture and (3) offers examples of practical freight ITS issues which can effectively be addressed at the industry cluster level.

5.2.2 Literature Review

In the field of economic development, several researchers have examined the concept of the industry cluster as a valid construct and tool of analysis in economic development at all policy levels. Porter's [1] work continues to be the benchmark by which most current state and local cluster initiatives are measured. Others have taken the concept in different, arguably more robust and less geographically constrained directions. Roelandt and den Hertog [2] identified the 'value-chain' industry cluster as "a cluster consisting of an extended input-output or buyer-supplier chain comprised of multiple sectors or industries and including final market producers as well as suppliers at all tiers directly or indirectly involved in trade." This definition of an industry cluster is consistent with the industry cluster described two decades earlier by Czamanski and de Ablas [3]. In response to an abundance of literature detailing the geographic proximity and various agglomeration-related spillovers of industrial clusters, Bergman and Feser [4]show that increasingly industry clusters are not constrained by the spatial proximity of member actors and may in fact be less prone to distance-sensitive constraints on realized mutual benefit and synergistic competitiveness. The 'binding ties' between cluster members may in fact consist of linkages between geographically dispersed members of the same value-chain industry cluster. Further, Feser and Sweeney [5] determined that geographic agglomeration is really only crucial for knowledge-based or technology-sensitive clusters.

Significant work has also been completed on the role of information technology and sophisticated logistics management systems in industrial value chain relationships. McCann and Fingleton [6] discussed the spatial impact of Just-In-Time (JIT) buyer-supplier relationships on Japanese JIT manufacturing entities in the Scottish electronics industry. Tsao and Botha [7] developed system operating concepts and deployment sequences for the automation of inner-city trucking with the expressed intent of, among other things, increasing vehicle throughput and satisfying the numerous stakeholders involved. Mahajan and Vakharia [8] point out the importance of IT alliances between channel members and the positive spillovers that accrue to all value chain elements via the automation and/or reengineering of logistics processes downstream.

Research in the areas of transportation and intelligent transportation systems (ITS) has also focused on the use of technology and value-chain management systems to facilitate inter-firm transaction processing and streamline point-of-sale logistical protocol. Yoshimura, Kjeldgaard, Turnquist, and List [9] made the case for an expanded and more versatile transportation modeling system that would support both logistics and operations processes and integrate improved transportation decision tools in the areas of package allocation, fleet sizing, and routing/scheduling. Jensen, Williamson, Sanchez, and Mitchell [10] completed an evaluation of an ITS-enabled intermodal supply chain manifest system designed to increase air cargo security and streamline the logistics systems deployed at O'Hare and JFK international airports. The authors found that the technology did in fact create a secure electronic intermodal manifest system that both saved time and satisfied end users. Satisfied participants also recommended that the system would be even more useful if adopted by additional supply chain partners. Eisele and Rilett [11] exposed the lingering need for accurate estimations of travel time data to buttress both real-time and off-line transportation management applications—including those intended for use by freight haulers. Cambridge Systematics [12] detailed the four main categories of fleet management technology: automatic vehicle location (AVL) systems, mobile communications systems, on-board computers (OBCs) and routing/dispatching software. Opdam [13] completed a study of the efforts--and mixed results--of a U.S. cement company employing an advanced routing system to locate the closest truck to a delivery point and successfully navigate stagnant beltway traffic. Srour, Kennedy, Jensen, and Mitchell [14] completed a thorough evaluation of a major real-time freight information system designed by the New York Port Authority to merge the numerous sources of freight location and status information into one, easily navigable web portal utilized by end users to access cargo information and facilitate both transportation planning and logistics. Rao, Navoth, and Horwitch [15] examined the rise of third-party logistics and distribution providers, highlighting FedEx's efforts to capitalize on the urgent need to merge 'virtual-world' information technology and the ‘real-world’ physical delivery of products. Button, Doyle, and Stough [16] found that by implementing sophisticated dispatching software, a courier company increased driver productivity 24% and unexpectedly decreased the amount of stress born by the company’s dispatchers.

Relatively little research has been completed to date on the convergence between economic development’s industry cluster literature and transportation’s use of information technology to enhance value-chain transactions. Several (e.g. Von Thunen [17] Weber [18], Greenhut [1], Richards [20], Ziegler [21], Kilkenny [22]) have examined the role of transportation as a factor cost in industrial location decisions. Bergman [23] examines the technology adoption rate of members of the transportation equipment value chain and supports policies that work through the market structure of regional economies and a firm’s own value-chain to promote technology adoption. Munnich and Lehnoff [24] explored the effect of ITS on the spatially agglomerated recreational transportation cluster in rural northwestern Minnesota, finding that ITS investments occurred disproportionately in the cluster’s larger firms. However, none of these studies have explicitly defined the value chain industry cluster as an ITS planning construct, or set forth practical applications of cluster analyses in the development of ITS architectures and concepts of operations.

5.2.3 Defining the Industry Cluster in Supply Chain Terms

The term “industry cluster” has been used loosely by economic development and transportation practitioners to describe groups of businesses or establishments in different contexts. The concept of the industry cluster has often been misunderstood in the transportation planning community for two reasons. First, methods for integrating formal techniques of economic industry cluster analysis into transportation planning have not been widely implemented, and second, a straightforward process of using the cluster concept in transportation planning has never before been introduced.

Consequently, transportation-planning efforts attempting to utilize the “cluster” concept have often resorted to very simplistic studies that fail to grasp the transportation planning significance of true industry cluster relationships as defined in the literature above. For example, a recent freight study in Minnesota [25] used aerial photography to identify groups of non-residential buildings in close proximity and defined each group as a “cluster” without identifying the occupancy status, industry, type of business, or transportation and trading linkages between such businesses.

For this reason, effectively implementing the cluster concept in freight ITS planning requires (1) a working set of definitions for industry clusters of different types (2) an understanding of the theoretical basis of the cluster concept in economic terms and (3) an understanding of how collaborative, technology-based transportation and distribution relationships between firms in different industries can serve to catalyze industry clusters likely to have a stake in ITS planning and implementation.

Industry Cluster Definitions

In concept, industry clusters are defined in the literature cited above as groups of firms from multiple industries that are motivated by competitive pressures to form collaborative or competitive relationships. The dynamic of an industry cluster is determined by

1. Time. The cluster dynamic is affected by the period over which relationships among firms and industries evolves, as well as a function of the distance between their physical locations.

2. Space. The cluster dynamic is affected by the spatial proximity of firms to one another, and the degree to which space strengthens or hinders their relationships.
3. Linkage. The cluster is affected by the type of linkage between the firms, whether the firms are competitors, likely to be collaborators, or simply share the same resource pool.

Bergman and Feser [4] provide a comprehensive set of definitions for industry clusters and related concepts shown in Table 1.

Theoretical Basis of the Cluster Concept

Diamond of Advantage

Michael Porter's "Diamond of Advantage" [1] provides the theoretical basis for understanding intra-industry and inter-industry relationships in the industry cluster framework. While Porter's work viewed advantages as primarily national in nature, as opposed to regional within a domestic economy, certain key principles from Porter's "diamond" are useful for understanding industry clusters as relate to ITS and transportation related technologies.

The Porter "Diamond of Advantage" characterizes economic vitality of as a function of four conditions that drive innovation and productivity improvement:

1. Factor Conditions: Conditions that make available resources, technologies and opportunities unique to a select group of firms (such as culturally embedded tacit knowledge in the workforce, natural geographic resources such as proximity to an ocean or river).
2. Demand Conditions: The degree to which the environment provides viable markets for products and services.
3. Related and Supporting Firms: The degree to which the environment offers opportunities for mutualistic collaboration and cooperation among partners in different industries with complementary interests, needs or problems.
4. Firm Structure, Strategy and Rivalry: The degree and manner in which firms are able to enact strategies to compete with each other within and among industries.

Government and chance factors are understood as external to Porter's "Diamond" but affect economic vitality indirectly through influences on each of the four "corners" of the diamond. Munnich [24] (2003), interpreted Porter's "Diamond of Advantage" in terms of potential ITS applications as shown in Figure 5.2.1.

The broad categories of the activities in the boxes of Figure 5.2.1 represent opportunities for collaboration among industries that have the potential to support economic vitality by enabling firms to use the transportation system in a more efficient and innovative manner.

Table 5.2.1 Industry cluster definitions of Feser and Bergman [4]

Glossary Terms	
Concept	Definition
Sector (or Industry)	A sector or industry is a group of enterprises that manufacture similar products, as typically defined, for example, under the U.S. Standard Industrial Classification (SIC) system.
Industry cluster	A group of business enterprises and non-business organizations for which membership within the group is an important element of each member firm's individual competitiveness. Binding the cluster together are "buyer-supplier relationships, or common technologies, common buyers or distribution channels, or common labor pools (Enright 1997, p. 191)." See Porter (1990).
Regional industry cluster	A cluster whose elements share a common regional location, where region is defined as a metropolitan area, labor market, or other functional economic unit.
Potential industry cluster	A group of related and supporting businesses and institutions, that, given additional core elements, inter-firm relationships, or critical linking sectors, would obtain some pre-defined critical mass.
Value-chain industry cluster	A value chain cluster is an industry cluster identified as an extended input-output or buyer-supplier chain. It includes final market producers, and first, second and third tier suppliers that directly and indirectly engage in trade. It is comprised of multiple sectors or industries. (See Roelandt and den Hertog 1999). A "Value-chain cluster" is consistent with an "industry cluster" as defined by Czamanski and de Ablas (1979, p. 62): "a subset of industries of the economy connected by flows of goods and services stronger than those linking them to the other sectors of the national economy." May also be defined as potential, where enterprises may or may not presently trade with each other, although such trade could possibly occur in the future.
Business network	"A group of firms with restricted membership and specific, and often contractual, business objectives likely to result in mutual financial gains. The members of a network choose each other, for a variety of reasons; they agree explicitly to cooperate in some way and to depend on each other to some extent. Networks develop more readily within clusters, particularly where multiple business transactions have created familiarity and built trust (Rosenfeld 1995a, p. 13)." Ties between firms in networks are typically more formal than in clusters.
Italianate industrial district	A highly geographically concentrated group of companies that "either work directly or indirectly for the same end market, share values and knowledge so important that they define a cultural environment, and are specifically linked to one another in a complex mix of competition and cooperation (Rosenfeld 1995b, p. 13). Key source of competitiveness are elements of trust, solidarity, and cooperation between firms, a result of a close intertwining of economic, social, and community relations. See also Harrison (1992).
Industry complex	"A group of industries connected by important flows of goods and services, and showing in addition a significant similarity in their location patterns (Czamanski and de Ablas 1979, p. 62)."
Innovative milieu	Not a group of business or a region, but a "complex which is capable of initiating a synergetic process. . .an organization, a complex system made up of economic and technological interdependencies. . .a coherent whole in which a territorial production system, a technical culture, and protagonists are linked (Maillat 1991, p. 113)." See also Maillat (1988).
(Source: Regional Research Institute, with permission)	

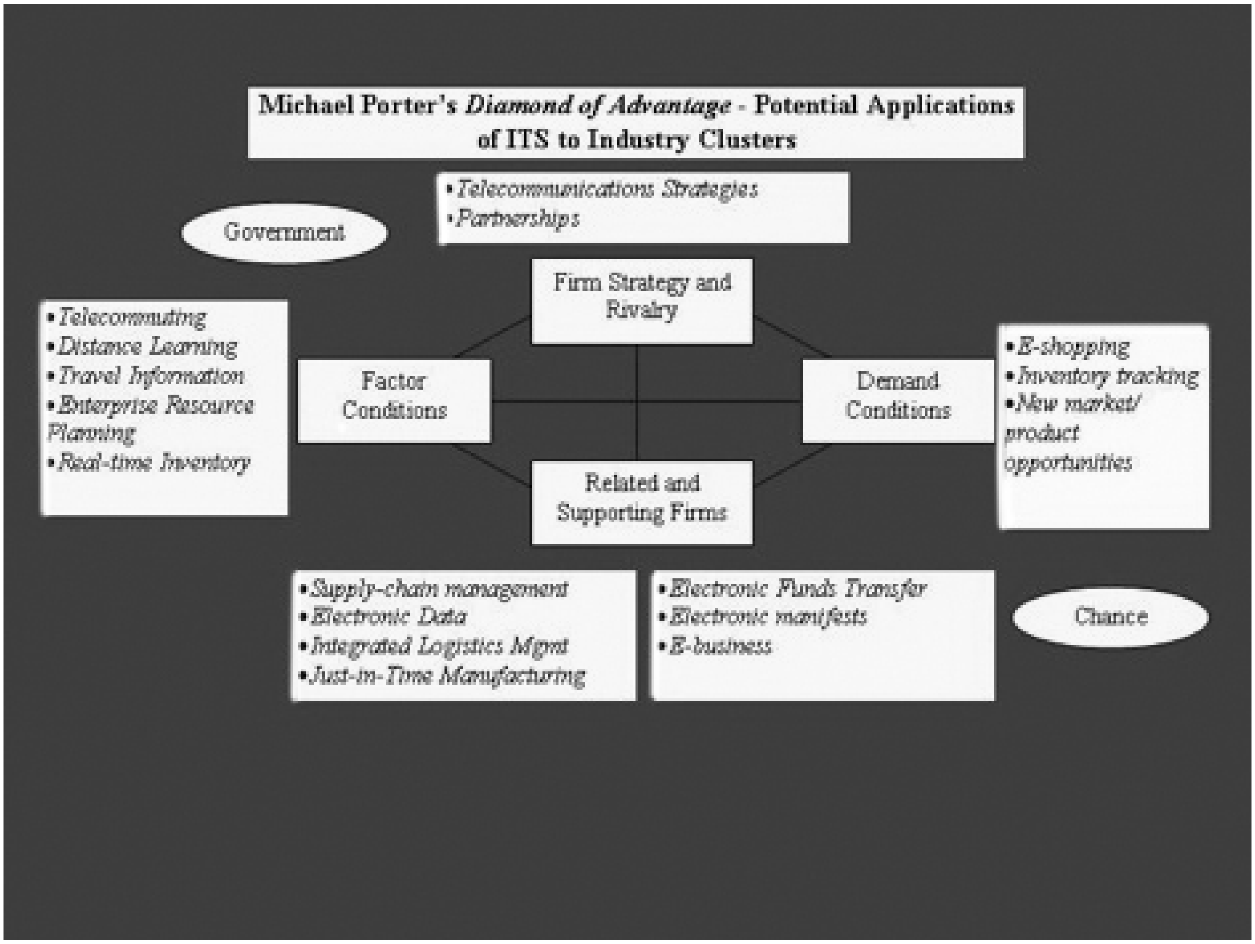


Figure 5.2.1 Porter's Diamond of Advantage as pertains to ITS opportunities [*23*]

Beyond the "Diamond"

As can be inferred by Figure 5.2.1, it is difficult to translate the general four points of Porter's "Diamond of Advantage" into specific types of relationships among firms using specific types of technologies and strategies. The "Diamond" framework addresses collaborative relationships among firms under the general category of "related and supporting firms" and classifies technology as a "factor condition," but does not specify the competitive pressures by which firms are motivated to enter into specific types of relationships. Furthermore, because government is exogenous to the "Diamond" the roles and relationships of public and private actors required for integrated ITS planning at the industry cluster level are ambiguous at best.

Ultimately, though Porter recognizes that clusters are not always spatially agglomerated, the "Diamond" has been widely interpreted by planners as a construct used to understand only spatially agglomerated industry clusters. This bias towards the industry cluster as a strictly local (or regional) concept poses limitations in its application to relationships among firms exchanging commodities across long distances through freight transportation activities. The strictly local (or regional) interpretation of the cluster concept has worked to the detriment of transportation planners seeking to develop ITS operational concepts and architectures. The strictly localized interpretation of the industry cluster is not appropriate for the global economy, in which decisions and partners far away, in other parts of the world can be more important to a business than partners in the local area. Consequently, ITS freight planning requires a more global view of industry clusters than has typically been applied. A cluster-based approach to ITS freight planning must recognize that the most critical economic relationships among firms in a transportation system are likely to be national and global in nature; with a common interest in transportation activities often serving as the catalyst for collaboration.

Carrying the argument further requires examining the relationships underlying the "Diamond of Advantage" at the more refined level.

To get beyond the limitations of the "Diamond" it is useful to consult Porter's "Five Forces" model.[26]The "Five Forces" model of industry competitiveness provides a more specific way of identifying and classifying the relationships implicit in the "Diamond of Advantage" in which transportation technologies can be found to play a role.

In the Five Forces model, Porter argues that the intensity of competition within an industry is driven by:

1. Strategic Rivalry Among Firms
2. Pressures from Buyers
3. Pressures from Suppliers
4. The Threat of New Entrants
5. The Threat of Substitute Products

The critical elements of competitiveness common to the "Five Forces" and the "Diamond of Advantage" are the pressures businesses face from buyers and suppliers; and the need for firms to innovate to competitively position themselves within the marketplace in response to these and other pressures.

In this context, the relationships between a firm in any given industry and firms in buying and supplying industries provide a key lever for competitiveness, and create a strong incentive for buying and supplying firms to collaborate. In manufacturing industries, this collaboration almost by definition must include collaboration on the physical movement of raw materials and products, as quantities, prices, schedules and other logistical factors are key determinants of costs and operational strategies on both sides of the buyer and supplier relationship. With reference to transportation, government finds its place in the "Five Forces" framework as a supplier of infrastructure.

Understanding exactly (1) how transportation and distribution related technologies engender such collaboration, (2) how the relationships implicit in such collaboration can be properly defined and understood in terms of industry clusters, and (3) how these relationships fit into an ITS architecture or concept of operations is important for ascertaining the significance of the industry cluster as a construct for ITS planning and policy.

Collaborative Supply Chains as Industry Cluster Catalysts

The evolution of advanced supply chain management technologies has resulted in collaborative groups of firms in different industries using information technology to more efficiently manage the distribution and transportation of inputs and commodities. Through these collaborative, technology-based relationships, firms in different industries and different locations function as industry clusters under the Bergman and Feser [4] definitions.

By the nature of their collaboration up and down the supply chain, clusters bound together by collaborative supply chain management strategies represent partnerships among firms, flows of

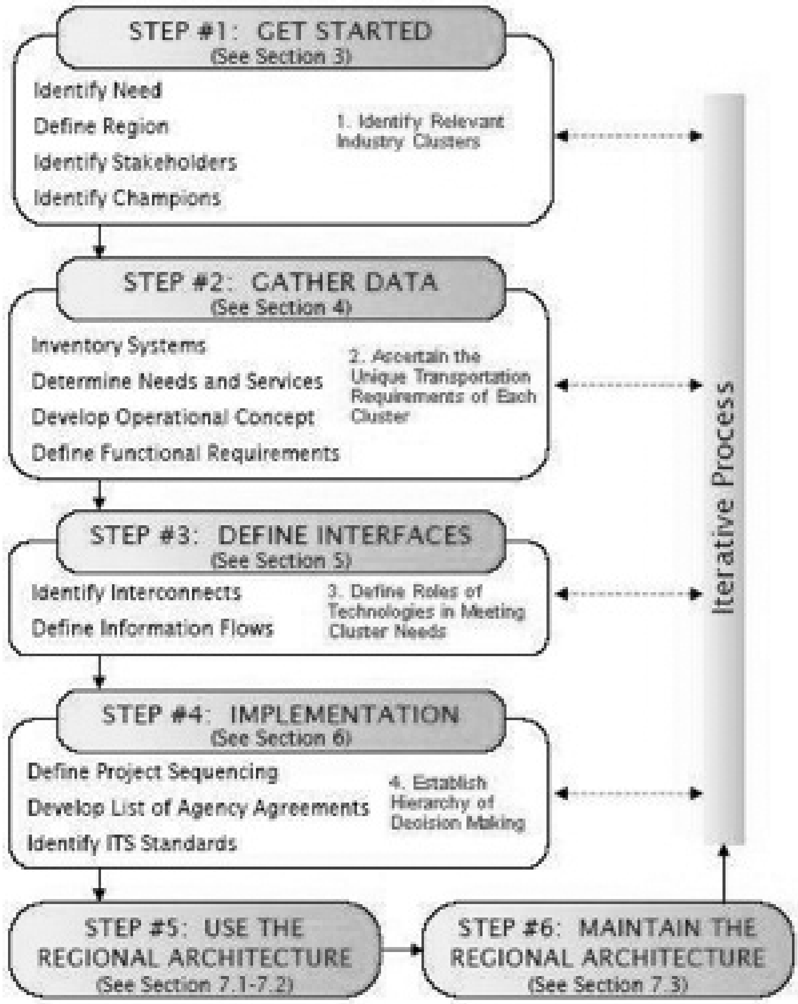


Figure 5.2.2 US DOT regional ITS architecture guidance [27]

information and a shared interest in the distribution and transportation activities that makes them important stakeholders in and potential contributors to freight ITS systems.

Five examples of collaborative, technology-driven strategies that form industries into clusters are presented in Figure 5.2.2 with a brief discussion of how the relationships function as clusters and the relevance of such clusters to ITS planning.

These are only five examples of the types of strategies that form supply chains into industry clusters. The examples illustrate ways in which it is not individual firms, or singular industries; but interactive clusters of industries that use the freight transportation system. The examples furthermore demonstrate the ways in which clusters represent distinctive transportation and technology requirements, decision-making processes and information flows with the potential to both contribute to and benefit from inclusion in regional ITS architectures and operational concepts.

Leveraging Freight User Technologies and Strategies in ITS Planning. Each of the five strategies described in Table 5.2.2 pertain to the management of the demand for freight services, and the use of technology to implement strategies in ways responsive to the demand for freight services.

"ITS" has often been implemented only from the perspective of infrastructure or service providers. For example, ITS systems are often implemented to enable public agencies operating and managing infrastructure to respond to problems assuming changes in demand are external to the system. Consequently, the key decision support capabilities, information flows and organization of ITS systems mirror the public agencies providing the infrastructure (agreements between federal, state and local enforcement, emergency response and highway organizations).

The information in these systems are not typically available to the private entities (shippers, carriers and logistics firms) using the system, and information flows utilized by system users (as shown in Figure 5.2.2) are rarely if ever available in the formal "ITS" architecture. Systems of inclement weather reporting, incident management and enforcement are common examples. Most of the parties exchanging and using information in such ITS systems are limited to infrastructure *suppliers*, collaborating to best meet a demand that is not comprehended by the ITS concept.

It is proposed that more comprehensive systems, built around the economic structure, trading relationships and information flows of system users could yield efficiencies and benefits to users and providers alike. The following section offers a European example (the Also Danube) of the type of system that is proposed.

The value-chain industry cluster is offered as a way of organizing users for inclusion in freight ITS planning process, identifying key relationships, technologies, information flows and opportunities on the demand-side of freight ITS. Systems built around value-chain industry clusters to bring freight demand information and technologies into the ITS framework would differ from current systems in the following ways:

1. They would be used by public agencies to manage infrastructure supply, *and* by private users to manage infrastructure demand.
2. Information from the supply and demand sides of freight transportation would be leveraged to support key decisions
3. Freight ITS would be developed cooperatively by government agencies, transportation carriers and affected industries to leverage information flows and technologies in support of key decisions.

Table 5.2.3 provides a comparison of existing practices in freight ITS planning and ITS planning processes developed around the value-chain industry cluster concept.

A European Example. A situation very much like the paradigm described above has been achieved in Europe for a logistics information system managing commodity flows on the Danube River. The Also Danube [27] system is a collaborative, multi-jurisdictional effort in the European community that involves key private industries and stakeholders in the development and application of collaborative public and private transportation decision support technologies.

1. The Also Danube project was developed by a consortium of:
2. Shipping Companies (sea and inland transport)
3. Inland waterway transport operators
4. Transport operators - rail and road
5. Ports, sea and inland (port authorities and transshipment operators)
6. Logistic service providers and forwarders
7. Industrial companies (consignor/consignee)
8. RTD-organisations

9. Telematic System Providers
10. Software and Consulting companies
11. National Authorities (e.g. customs, emigration authorities, public administration organisations for inland waterways, etc.)

The project commenced with the development of a strategic concept to provide a technology based logistics and transportation system for the Danube River. The concept had the objectives of:

1. Promoting the use of inland waterway transportation for value-chains
2. Implementing an advanced concept to manage inter-modal transportation chains
3. Establishing highly integrated logistic networks and operational platforms
4. Improving the efficiency of the waterway and
5. Demonstrating the functionality of logistic applications for the waterway

The approach was to develop WEB-based client applications, advanced EDI solutions, and innovative telematic technologies to be integrated, demonstrated, and evaluated in specific supply chains representing different transportation markets. This entailed the development of a common-source logistics database for all transportation and logistics information relating to the waterway, involving all members of the consortium. Applications and interfaces for specific logistic channels were developed to meet the needs of shippers, carriers, public entities and other channel members.

The result is a collaborative, technology-based intelligent transportation network built around a common-source logistics database with interactive links for information provided regarding: supply operations, logistics processing, planning and operation of infrastructure and ports, tracking and processing information from industries and carriers.

Table 5.2.2 Value chain industry cluster transportation and distribution strategies

Strategy	Description	Relationship to Value-Chain Industry Clusters	ITS Architecture Planning Implications
Electronic Data Interchange (EDI)	Electronic transmissions send real-time purchase orders, invoices, shipping notices, and payments between buyers and suppliers and firms and their customers.	Technology synchronization is a necessary prerequisite for EDI communication. Requires on-going cooperative relationships that improve information and merchandise flow and enhance competitiveness of participating cluster firms.	EDI-enabled clusters firms influence the design and utilization of ITS systems and transportation networks by impacting the timing and tracking of scheduled truck runs and the amount and type of good shipped in each load.
Cross Docking	Sophisticated just-in-time warehousing system where incoming pallets are divided into customer-specific segments while still on the loading dock and transferred directly to outgoing trucks. Requires close synchronization of all inbound and outbound shipments.	Cluster members utilize utilizing cross docking to coordinating closely with other cluster firms both up and down the value chain. Competitiveness is enhanced through reducing lead times and by freeing up working capital lost to overstocks and spoilage.	Efficient cross docking requires an ITS that supports real-time information transfer between cluster members (e.g. warehouse and carriers) to coordinate dispatch, routing, and scheduling. Transformed from a passive storage facility to a crucial node of information, warehouse information needs inform ITS architecture development and refinement.
Efficient Consumer Response (ECR)	In response to a more demanding customer base, ECR eliminates costs that do not add value and emphasizes continuous improvement in order to satisfy growing consumer demand for high quality, inexpensive goods and services.	Internal improvements only go so far. To be competitive and satisfy increasingly demanding consumers, cluster firms recognize they must work cooperatively with one another and with carriers to increase efficiency and effectiveness throughout the value chain.	A poorly planned ITS architecture that fails to support continuous improvements in supply chain efficiency can negate cluster efforts and erode both cluster relationships and competitiveness in the marketplace.
Quick Response	Collaborative arrangement between suppliers and retailers, distinguished from similar systems by a focus on sharing point-of-sale (POS) information with suppliers in order to more accurately forecast replenishment schedules.	Focused on the crucial link between supplier and POS retailers, Quick Response ensures that the fruit of cluster labor—the final consumer product—is available when and where consumers demand it and costly stock-outs are avoided.	The ITS architecture, designed with an understanding of inter-cluster transportation relationships would avoid transportation bottlenecks (e.g. tolling and commercial vehicle checks) that delay supplier initiated replenishment.

Continuous Replenish-ment Process (CRP)	While previous manual or electronic purchase orders introduced lag time into the product replenishment cycle, CRP uses algorithms based on past data to automate supplier-managed product replenishment and point of sale allocation from storage facilities.	Cluster buyer-supplier interactions are automated and efficient, requiring on-going but periodic collaboration to tweak the assumptions built into the shared CRP technology.	With access to the replenishment forecasting data shared between cluster firms, the public ITS architecture could inform traffic management and help public entities anticipate and manage the levels of truck traffic dictated by the automated system.

Table 5.2.3 Contrasts between value-chain cluster ITS planning and current US Practice

Planning Issue	Existing Practice	How Industry-Cluster Approach is Different
System Development	Firms and industries develop private commodity distribution systems (see Table 5.2.2) to manage their demand for freight transportation services within given infrastructure constraints. Public agencies develop public ITS systems to manage infrastructure supply separately.	Public ITS and transportation information systems would be developed for compatibility with private arrangements and systems supporting value-chain linkages among related private users. Private cluster members would have opportunities to collaborate in their private systems as well as with public systems within the context of an integrated architecture.
User Requirements	Private systems are designed to support decisions by shippers and carriers to manage the needs associated with the demand for freight transportation. Public systems are designed separately to provide information to support decisions and manage the supply of public infrastructure and services only.	Decisions and activities in both public and private sectors are supported by integrated ITS systems and information technologies. Public and private entities collaborate to develop strategic ITS, cohesively supporting critical decisions and activities on both the transportation demand and infrastructure supply sides of the system.
System Security	Private systems are accessible only to firms and their selected partners to prevent strategic breeches of confidence. Public systems ITS and transportation information systems may be available, but are not integrated with private (demand side) systems.	Information is conditionally shared between private and public systems as needed to support key inter-related decisions for managing transportation demand and infrastructure supply. Security needs are addressed as requirements and technologies in Step 2 and 3 (Table 5.2.4)

System Maintenance and Management	Public agencies manage and maintain ITS as part of the public infra-structure network. Private agencies operate and manage shared systems through joint ventures, strategic alliances and third party intermediaries (such as 3rd Party Logistics Contractors).	Integrated system is maintained by agreements, alliances, ventures and public-private partnerships established in Step 4 (Table 5.2.4). Freight ITS system functions with private support and participation as a quasi-public service.

The Also Danube system contains many features of both private sector logistics systems and public ITS systems in the US. The characteristic that makes the Also Danube example distinctive is the degree to which it incorporates transportation system information from carriers, industries and public entities in the development of a cohesive system to use transportation technology to support freight efficiency. Because the process began with the identification of key industries and value-chains, and was developed by members of these chains; it is fundamentally different from American-style systems that involve only information flows among public agencies supplying and managing infrastructure.

The involvement of value-chain industry groups in the initial concept of Also Danube was critical to its success. The common source logistics database at the heart of Also Danube would have never developed, and would not have users had key value-chains and their members not been involved early in the conceptual planning.

By identifying value-chain industry clusters, and incorporating industry groups into the conceptual planning process for freight ITS in the American setting; more integrated technology solutions can be developed for freight transportation. The key building blocks for developing such systems in the American setting are based on the application of the value-chain industry cluster concept within the US DOT guidance on regional ITS architecture development.

A Four-Step Process for Integrating Industry Clusters Into Its Planning

To incorporate industry clusters into the ITS planning process, four steps are recommended, in conjunction with the four steps of developing regional ITS architecture set forth in the USDOT Regional ITS Architecture Guidance Document [28] (Figure 5.2.2):

1. Identify Relevant Industry Clusters and Strategies
2. Ascertain the Unique Transportation Requirements of Each Cluster
3. Define the Roles of Transportation Technologies in Meeting Those Requirements
4. Establish formal and informal Hierarchies of Transportation Decision Making with appropriate information systems and standards.

The process can be applied for any entity for which a regional ITS architecture or concept of operations is required. Some examples may include trade or technology corridors, freight villages, inter-modal freight facilities and ports. These four steps are summarized below, with specific objectives, methods and potential technologies for each step given in Table 5.2.4.

1. **Identify Relevant Industry Clusters and Strategies**

The process of identifying relevant industry clusters and strategies for incorporation into an ITS architecture or concept of operations begins with identifying basic shipping industries using the infrastructure.

a. Identify Major Basic Industries in the Area Served by the Infrastructure using location quotients (ratio of regional industry share to national industry share).
b. Construct Supply Chains for Basic Industries using input-output multipliers to identify secondary and tertiary buyers and suppliers, including public infrastructure providers, emergency response organizations, enforcement agencies and other public entities as suppliers of transportation related services.
c. Identify Collaborative Supply Chain and Distribution strategies and technologies active within key supply chains by involving industry stakeholders.
d. Identify Information Flows Associated with these strategies and technologies

2. **Ascertain the Unique Transportation Requirements of Each Cluster**

By examining each cluster as a whole, critical decision points, transportation bottlenecks, chokepoints and critical paths can be identified (including not only capacity constraints, but time-lags for shippers, sources of delay for carriers and sources of inefficiency for both shippers and carriers).

3. **Define the Roles of Transportation Technologies in Meeting Those Requirements**

Based on the information flows between cluster members, and the potential public sector ITS systems, identify ways to supplement public and private information flows between cluster members to provide optimal information at key decision points, and minimize inefficiencies caused by lag times, infrastructure capacity constraints, and other sources of delay.

4. **Establish a Hierarchy of Decision Makers and Decision Making**

To identify specific decision support requirements and priorities, establish those decisions, entities making critical decisions with impact throughout the value chain. In this process, identify specific types of information available from information flows present in the cluster that would support those decisions and establish agreements, systems and standards to provide integrated decision support through the ITS effort.

5.2.4 Practical Applications and Advantages of the Value Chain Cluster Concept

Ultimately, the key to a successful integrated public-private cluster-based ITS architecture or concept of operations rests with the tangible payoffs in competitive business and policy terms for cluster members. The following are some examples of US transportation related business and policy situations in which private sector information flows within value-chain industry clusters could be supplemented and synergized with cluster level ITS planning efforts through cohesive architectures and operational concepts. The examples are offered to both support the merits of ITS planning at the industry cluster level, and to suggest how such an approach can offer tangible and attractive payoffs for all cluster members (including infrastructure providers).

Architectures Addressing Diurnal Operational Rhythms and Cycles

The time of day operating patterns of freight movement are determined by the diurnal cycles of business (manufacturing runs completed at different times of day). Freight carriers face the challenge of both responding to these different operating patterns among value chain members with routing and scheduling to avoid peak hour bottlenecks on the transportation infrastructure.

A value-chain based ITS resource with real-time scheduling information and infrastructure capacity and performance data available to shippers, carriers and infrastructure providers could provide significant leverage for managers seeking to minimize delay and transportation costs associated with time-of-day mismatch between less congested periods on the infrastructure, time-sensitive demand for commodities and operating rhythms of suppliers.

Functionally, this would entail a technology system whereby cluster members share information about schedules and operating cycles and chokepoints through a portal in the ITS system. This information would then also be accessible to carriers and infrastructure providers to better manage freight transportation infrastructure and services in conjunction with value-chain collaboration to manage freight travel demand.

Architectures Consolidating Asset Utilization Within the Cluster

Another practical situation in which ITS planning can offer benefits at the value chain industry cluster level involves optimizing the utilization of the assets of transportation related technologies and information flows. This goal has been suggested in recent discussion of an integrated approach to information systems international air cargo operations in the Twin Cities Metropolitan Area. An Air Cargo Task Force in the Mpls/St. Paul region has lead a number of studies that have identified more than 50 air cargo freight forwarders.

A significant issue for the region is that international air cargo drayed to Chicago is often interred for several days while a forwarder attempts to “fill” a truck. A study of the issue has proposed the creation of a consolidation center whereby all participating air cargo forwarders would operate on a “shared” information system and drayage fleet. This would increase the efficiency of not only the information technology, but of the trucks, warehouse facilities and other key transportation resources in the value chain.

By identifying the relevant value-chain clusters to be served by the facility with their relevant strategies and technologies, ascertaining their transportation requirements at the value-chain level, defining the roles and responsibilities of public and private entities and establishing key decision points and decision support requirements as described in the previous section, a cluster-based concept of operations for such a system could offer reduced costs and increased efficiencies for private sector cluster members, and an optimal use of the relevant ports and highways for public sector collaborators.

Architectures Combining Regional Proximity and Value-Chain Linkage

Increasing global trade is believed to have a positive impact on the United States’ economy and on gateway urban areas. However, the increase in trade together with air express, and just-in-time business expectations is often associated with concurrent increases in truck vehicle miles traveled (VMT), air pollution, urban congestion, and reductions in quality of life and regional competitiveness.

The concept of the Freight Village is a scenario in which inter-modal activities are concentrated by co-located shippers at inter-modal facilities just outside of major metropolitan areas. Naturally, when co-located shippers are members of the same value-chain clusters, they can significantly leverage their co-location to reduce the transportation times and costs in the value-chain.

Incorporating an integrated cluster-based ITS concept of operations and associated architecture in the transportation planning for a Freight Village may increase the viability and effectiveness of the Freight Village concept, by explicitly accounting for the industry cluster linkages when recruiting firms into the Freight Village and enabling the Freight Village to offer members something more than a location with proximity to both an inter-modal facility and a major trade center.

Table 5.2.4 Objectives, methods and outcomes associated with 4-steps of ITS planning for value chain clusters

STEP	OBJECTIVES	AVAILABLE METHODS	OUTCOME
(1) Identify Relevant Clusters	Determine the business relationships, information linkages and private sector resources aimed at optimizing use of the transportation system.	Economic Base Analysis Value-Chain Cluster Analysis (Input-Output Based) Commodity Flow Analysis (TRANSEARCH)	Definition of industry clusters, including business types and commodity flows where transportation efficiency is sensitive to technology. Inventory of existing information systems and other technologies used by businesses to optimize utilization of the transportation system.
(2) Determine Unique Transportation Requirements	Identify unmet needs for information and decision support in private sector transportation activities as well as public sector infrastructure system management. Identify inefficiencies resulting from inconsistencies in the application of technology among shippers, carriers, 3PL's and government agencies involved in the cluster.	Public Involvement Methods Such As: Interviews with Shippers and Carriers. Focus groups and round-table discussions with key cluster members. Formation of advisory committees and task forces addressing freight technology needs and opportunities.	Key users and user goals for ITS and supporting information systems established. Specific use-cases for integrated public and private ITS and transportation technology collaboration defined. Holistic, system-based perspective on technology requirements and opportunities is clarified.
(3) Identify the Role and Potential of Transportation Technologies	Provide a "menu" of available technologies to address needs identified in Step 2. Select appropriate technologies and supporting information systems to support user goals. Determine consistent requirements for managing data flows and data security among private and pubic sector cluster partners.	Data mining and review of secondary sources (available research). Development of data flow diagrams, summary use-cases (diagrams sowing relationships between individual use-cases)	Clear definition of existing and potential transportation technologies supporting the cluster. Clear understanding of the users and user goals met by each technology, and how these goals support the overall system. Clear definition of the systems and technologies to be implemented, with their requirements and objectives.
(4) Establish a Hierarchy of Decision Making	Prioritize the key decisions involved in optimal private sector utilization of the transportation systems. Prioritize decision support tools and related technologies from Step 3 to support critical decisions. Develop formal agreements defining authority and responsibilities of cluster members in the management and maintenance of ITS and freight related information for the cluster.	PERT Charts, and critical path analysis. Decisions trees, neural networks and other decision support tools. Memoranda of understanding, joint powers agreements, and strategic alliances among firms (which may also include public entities).	Roles and responsibilities of cluster members are well understood with regards to development of transportation technologies. Cluster is organized in such a way that key decisions and outcomes for all members are supported by appropriate technologies throughout the cluster. ITS and transportation technology needs and resources are defined and prioritized with respect to cluster members.

5.2.5 Issues for Future Research

By defining the value-chain industry cluster in terms relevant to ITS freight planning, and suggesting an outline for steps to include this construct in future ITS planning under US best practices, it is hoped this paper has provided a basis for additional research into the field. Methods and practices for implementing each of the 4 steps described in Table 5.2.4 is a potential area where applied research could significantly improve planning practice for freight ITS. The application of information management and decision support tools within the context collaborative freight ITS architectures has the potential to support the development of American systems similar to the Also Danube case cited in this paper. The potential of interactive simulation and travel demand models for freight is also a potential research area that could contribute significantly to the development of collaborative value-chain oriented decision support transportation technologies. Conceptual and applied research into business and stakeholder involvement in freight ITS planning and management could also greatly improve the prospects for collaborative ITS systems at the value-chain level.

5.2.6 Conclusion

The value-chain industry cluster is an established construct in the Economic Development literature which holds promise for transportation planners seeking to develop integrated public and private ITS systems for freight operations. Effective ITS planning and management can benefit from recognizing clusters in value-chain terms and incorporating the identification and analysis of value chain industry clusters in the scoping and implementation of ITS planning studies for technology corridors, freight villages, ports and other infrastructure systems. The Also Danube initiative is a European example of a cohesive freight technology system organized around industry groups and trading relationships. An industry cluster based approach to planning American ITS systems consistent with national best practices in regional ITS planning is likely to result in opportunities and payoffs for the affected transportation systems and their users. Consequently when scoping ITS planning studies and architectures, including steps to include and explicitly address industry clusters in each step of the ITS planning process can be an important element when the ITS system involves freight operations.

5.2.7 Acknowledgement

The authors acknowledge the University of Minnesota ITS Institute and the Federal Highway Administration for providing funding for this research under the STAR/TEA-21 Initiative. The authors also acknowledge Dr. Ed Feser of the University of Illinois, Urbana for his input and suggestions provided in the development of this paper and the Regional Research Institute for granting permission to use Table 5.2.1.

5.2.8 References

1. Porter, M. E. *The CompetitiveAdvantage of Nations*. Macmillan, London, 1990.
2. Roelandt, T. J. A., and P. den Hertog. *Cluster Analysis and Cluster-Based Policy Making: The State of the Art in Cluster Analysis and Cluster-Based Policy: New Perspectives and Rationale in Innovation Policy*. Organisation for Economic Cooperation and Development, Paris, 1999.
3. Czamanski, S., and L. A. de Ablas. Identification of Industrial Clusters and Complexes: A Comparison of Methods and Findings. *Urban Studies,* Vol. 16, 1979, pp. 61–80.
4. Bergman, E.M. and E. J. Feser. Industrial And Regional Clusters: Concepts And Comparative Applications. *The Web Book of Regional Science*, 1999, (http://www.rri.wvu.edu/WebBook/Bergman-Feser/contents.htm). Accessed June 2004.
5. Feser, E. J. and S. H. Sweeney. A Test for the Coincident Economic and Spatial Clustering of Business Enterprises. *Journal of Geographical Systems*, Vol. 2, Issue 4, 2000, pp. 349–364.
6. McCann, P. and B. Fingleton. The Regional Agglomeration Impact of Just-In-Time Input Linkages: Evidence from the Scottish Electronics Industry. *Scottish Journal of Political Economy*, Vol. 43, Issue 5, 1996, pp. 493–519.
7. Tsao, J. and J. L. Botha. Automated Highway System Dedicated to Inter-City Trucking: Design Options, Operating Concepts, and Deployment. *Intelligent Transportation Systems Journal*, Vol. 7, Issue 2, 2002, pp. 169–197.
8. Mahajan, J and A. J. Vakharia. Determining Firm-Level IT Investments to Facilitate Value Chain Activities: Should Spillovers Accruing to Value Chain Members Be Incorporated? *European Journal of Operational Research,* Vol. 156, Issue 3, 2004, pp. 665–683.
9. Yoshimura, R.H., E. A. Kjeldgaard, M.A. Turnquist, G.F, List. *Transportation System Modeling and Simulation in Support of Logistics and Operations.* Report Number: SAND-97-3098C, U.S. Department of Energy, 1997.
10. Jensen, M., M. Williamson, R. Sanchez, C. Mitchell. *Electronic Intermodal Supply Chain Manifest Field Operational Test Evaluation.* Report Number: FHWA-OP-02-XXX, U.S. Department of Transportation, ITS Joint Program Office, 2002.
11. Eisele, W.L., Rilett, L.R. *Examining Information Needs for Efficient Motor Carrier Transportation by Investigating Travel Time Characteristics and Logistics*. Report Number: SWUTC/01/473700-00005-1, Office of the Governor of the State of Texas, Energy Office and the U.S. Department of Transportation, University Transportation Centers Program, 2002.
12. Cambridge Systematics, ATA Foundation, U.S. Federal Highway Administration—Private Fleet Management Institute. *Commercial Vehicle Fleet Management And Information Systems. Technical Memorandum 3, ITS Fleet Management Technology Resource Guide.* Federal Highway Administration, Washington DC., 1997.

13. Opdam, P. Concrete Evidence: One Cement Delivery Firm Has Found a Way to Deliver on Time, but it's Not a Foundation For a Long-Term Distribution Policy. *Traffic Technology International*. October/November, 2002.

14. Srour, J., J. Kennedy, M. Jensen, C. Mitchell. Freight Information Real-Time System For Transport Evaluation Final Report. *UMTRI Research Review*, Vol. 10, 2003.

15. Rao, B., Z. Navoth, M. Horwitch. Building a World-class Logistics, Distribution and Electronic Commerce Infrastructure. *Electronic Markets*, Vol. 9, Issue 3, 1999, pp. 174–181.

16. Button, K., E. Doyle, R. Stough. Intelligent Transport Systems In Commercial Fleet Management: A Study Of Short Term Economic Benefits. *Transportation Planning And Technology*, Vol 24(2), 2001, pp. 155–70.

17. von Thünen, J. H. *Der isolierte Staat in Beziehung auf Landwirthschaft und Nationalökonomie.* Wiegandt, Berlin, 1875.

18. Weber, A. Ueber den Standort der Industrien. Tübingen, J.C.B. Mohr, 1909.

19. Greenhut, M. L. *Plant Location in Theory and Practice*. Harper and Brothers, New York, 1950.

20. Richards, H. A. Transportation Costs and Plant Location: A Review of Principal Theories. *Transportation Journal*, Vol. 2 Issue 2, 1962, pp. 19–25.

21. Ziegler, J. A. Location, Theory of Production, and Variable Transportation Rates. *Journal of Regional Science,* Vol. 26, Issue 4, 1986. pp. 785–792.

22. Kilkenny, M. Transport Costs, The New Economic Geography, and Rural Development. *Growth & Change*, Vol. 29 Issue 3, 1998, pp. 259–281.

23. Bergman, E. M. Innovation System Effects on Technological Adoption in a Regional Value Chain. *European Planning Studies,* Vol. 9, Issue 5, 2001, pp. 629–649.

24. Munnich, L. W., Lehnoff, J. Applying Industry Clusters to Intelligent Transportation Systems: A New Framework for Analysis. Presented at 13th Annual Meeting and Exposition of ITS America, Minneapolis, Minn., 2003, Conference Proceedings p. 23.

25. Adequacy of Freight Connectors to Interregional Corridors and Major Highways. Publication SRF0034739. Minnesota Department of Transportation, 2003.

26. Porter, M. E. *Competitive Strategy: Techniques for Analyzing Industries and Competitors*. Free Press, New York, 1980.

27. Also Danube Website (http://www.alsodanube.at/) accessed 11 November, 2004.

28. Regional ITS Architecture Guidance. USDOT EDL#13598. http://www.itsdocs.fhwa.dot.gov//JPODOCS/REPTS_TE//13598.pdf. Accessed on July 29, 2004.

Endnotes for 5.1

Applying Industry Clusters to Intelligent Transportation Systems: A New Framework for Analysis

---

[i] Porter, M. (1990). *The Competitive Advantage of Nations*. New York: The Free Press.

[ii] State and Local Policy Program and Metropolitan Council. (1995). *Twin Cities Industry Clusters Study*. St. Paul: Metropolitan Council. http://www.hhh.umn.edu/centers/slp/edweb/tcclust.htm

The Initiative Fund of Southeastern and South Central Minnesota and the State and Local Policy Program. (1996) *Southeast Minnesota Industry Clusters Study*. Owatonna: Southeastern and South Central Initiative Fund. http://www.hhh.umn.edu/centers/slp/edweb/seminn.htm

State and Local Policy Program and the University of Minnesota Extension Service. (1998). *Northwest Minnesota Industry Clusters Study*. Minneapolis: Humphrey Institute, University of Minnesota. http://www.hhh.umn.edu/centers/slp/edweb/nwclust.htm

State and Local Policy Program and the University of Minnesota Extension Service (1998). *Southwest Minnesota Industry Clusters Study*. Minneapolis: Humphrey Institute, University of Minnesota. http://www.hhh.umn.edu/centers/slp/edweb/swclust.htm

State and Local Policy Program and the Bureau of Business and Economic Research at the University of Minnesota Duluth (2001). *Northeast Minnesota Industry Clusters Study*. Minneapolis: Humphrey Institute, University of Minnesota. http://www.hhh.umn.edu/centers/slp/edweb/neclust.htm

[iii] State and Local Policy Program and the University of Minnesota Extension Service. (1998). *Northwest Minnesota Industry Clusters Study*. Minneapolis: Humphrey Institute, University of Minnesota. http://www.hhh.umn.edu/centers/slp/edweb/nwclust.htm

[iv] Policy Program at the University of Minnesota Twin Cities Humphrey Institute. Information available upon request.

[v] United States Census Bureau. (2000) Decennial Census. http://www.census.gov. October 2002.

[vi] Center for Rural Policy and Development & Minnesota Technology, Inc. (2002). 2002 Rural Minnesota Internet Study.

[vii] Minnesota Department of Trade and Economic Development. Minnesota's Forest Products Sector. http://www.dted.state.mn.us/PDFs/forest.pdf

CHAPTER 6

Lee Munnich: Privacy Issues of ITS

6.1 Thinking Privacy with Intelligent Transportation Systems

Policies, Tools, and Strategies for the Transportation Professional

Adam Kokotovich
Research Assistant
State and Local Policy Program
Science, Technology, and Environmental Policy Program M.S. Candidate
Hubert H. Humphrey Institute of Public Affairs
University of Minnesota
130 Humphrey Center
301- 19th Avenue S.
Minneapolis, MN 55455
P: (612) 879-8774
F: (612) 626-9833
E-mail: koko0013@umn.edu

Lee W. Munnich, Jr.
Senior Fellow and Director
State and Local Policy Program
Hubert H. Humphrey Institute of Public Affairs
University of Minnesota
157 Humphrey Center
301- 19th Avenue S.
Minneapolis, MN 55455
P: (612) 625-7357
F: (612) 626-9833
E-mail: lmunnich@umn.edu

6.1.1 Introduction

As intelligent transportation systems (ITS), or systems using electronic technologies to improve transportation, continue to grow and become more prevalent, there is an increasing need to address the many non-technical, social issues, such as privacy, that accompany many of these systems. In an attempt to help address the issue of privacy and ITS, the 2006 University of Minnesota Center for Transportation Studies conference's keynote address and following concurrent session engaged with this issue. Building on these sessions in addition to literature on the subject, this paper seeks help transportation professionals address privacy concerns relating to ITS through 1) providing a background to ITS relevant privacy theory, 2) describing tools and policies to help mitigate potential privacy concerns, and 3) analyzing privacy's relationship to ITS through the use of three case studies.

6.1.2 Background

Unlike many technical transportation planning issues which are predominantly discussed and dealt with by trained transportation professionals, the issue of privacy is one that is engaged with by a host of academic disciplines as well as by much of the general public. From law, political science, and other academic disciplines to privacy advocates and citizen groups to those developing the technology that may potentially affect privacy, these groups have different relationships to and views concerning privacy. The following sections on privacy provide an overview of the relevant privacy theory and the existing regulatory structure in the US, in addition to providing tools and policies that help achieve privacy protection.

Privacy Theory Background

In addition to the many anthropological and sociological studies that have found humans need a degree of privacy, such a conclusion is easily affirmed by realizing that nobody would want their every move monitored. The ultimate questions, then, are: what is the proper balance of privacy versus other considerations, how do you protect that level of privacy, and how are subsequent privacy risks distributed across society? Modern western claims to privacy have predominantly been based on the boundary between the public and the private, which separates individuals from other individuals, individuals from organizations, or individuals from the state. Such privacy considerations have relied upon construing civil society as a group of relatively autonomous, homogenous individuals which need a certain amount of privacy to fulfill the various roles of a citizen in a liberal democratic state [1]. Justifications for privacy have typically included considering it as: 1) a human right that everyone is entitled to, 2) a political value, allowing people to participate as citizens and providing a check against powerful state and private organizations, and 3) an instrumental value to insure that a person's data is used in the way they desire, so that trust is built between a person and an organization [2].

For comprehension and analytical purposes, five dimensions of privacy can be specified. These dimensions mark the different ways in which privacy can be considered. The first dimension, *spatial*, relies on the idea of public and private zones, where privacy can be expected in the private but not the public zone. The *behavioral* dimension understands certain behaviors as being private. The issue of surveillance, or of being monitored, is a privacy concern under this dimension. Under the *decisional* dimension, people have a right to make certain decisions without the intrusion of government or others. *Bodily* includes concerns about being intruded upon by institutions or technologies, such as issues pertaining to chip implants. And finally, *informational* includes information and data privacy and includes all other aspects of information or data collection [2]. The discourse around 'information privacy' emerged in the 1960s and 1970s about the same time that 'data protection' became a policy issue [1]. ITS privacy concerns fall predominantly under the informational and behavioral dimensions since most issues arise from concerns of surveillance and the large amounts of generated data.

As strategies were first developed for protecting information privacy, some notable assumptions were made. First, privacy is a highly subjective value. Concerns over the use of personal information vary over time, across jurisdictions, and across social status [1]. While members of a dominant social group may not oppose a mandatory national ID card program, those in marginalized social groups who would be more negatively impacted may find such a program a violation of their privacy. Secondly, the context of the use of the information is very important, making it very hard to prioritize data that is inherently worthy of greater protection [1]. Although the use of someone's name and address in a phone book may seem like an acceptable use of personal information, there are many people such as victims of domestic violence, police officers, doctors who perform abortions, and child protection staff, who may feel that this is a violation of their privacy. Because of these complexities, most information privacy policy is based on procedural rather than substantive tenets. As described by Bennett and Raab, such policy can,

> Put in place the mechanisms by which individuals can assert their own privacy interests and claims, *if they so wish*, and it can impose obligations on those who use personal data. But for the most part, the content of privacy rights and interests have to be defined by individuals themselves according to context." [1, p.9]

It is precisely the subjective and contextual nature of privacy that makes it such a complex issue.

Another important issue to clarify is the difference between data security and privacy. Many often equalize these two terms and believe that data privacy refers to the security of data against various risks such as being accessed by unauthorized people. Actually, data security needs to be a part, but not the entirety, of data privacy. An organization may keep data perfectly secure, but if the information should not be acquired in the first place, than an individual's privacy is still being violated [1]. On this point, privacy violations can occur as a result of technological failure, human failure, or both. A technically secure system could still be subject to a breach of privacy from human error. For example, at least 17.5 million people had their personal privacy compromised in May of 2006 when an employee from the Department of Veterans Affairs, under questionable circumstances, brought home a laptop and had it stolen. Although this laptop was later recovered without the data being violated, this event shows that even if the VA had the strongest technological data security systems in place, a simple lack of caution by an employee could put 17.5 million people's privacy at risk. [3]

Accounting for Social Effects

Two recent changes in dominant privacy theory are very pertinent to ITS privacy concerns: 1) social effects from privacy violations and 2) technology caused changes in privacy theory. As discussed in the previous section, dominant privacy theory considered those individuals affected by privacy to be autonomous and homogenous. Recent work, however, has helped realize that by considering all privacy subjects to be homogenous, many privacy effects are not recognized. Tracing privacy risks through social categories such as race, class, gender, and ethnicity expose important effects that may otherwise be overlooked. Privacy theorist Priscilla Regan argues that "recognition of the social importance of privacy will clear a path for more serious policy discourse about privacy and for the formulation of more effective public policy to protect privacy [1, p.40]." The mandatory national ID card example can be used to show the importance of considering social effects. A mandatory national ID card would not only affect different parts of the population to different degrees, but the degree would generally follow socio-economic boundaries. Immigrant communities and people of color would be most likely to face the negative impacts from such a policy due to the increased scrutiny of their "legal" status. Thus, they would be more likely than a dominant social group to feel that such a program was a violation of their privacy. Some may argue that the national ID card issue is not about privacy but national security, and that such a program would not discriminate as all people would be facing the same requirements. In reality, all questions of privacy are in opposition to competing interests such as security, safety, or efficiency, so whether an issue is classified as concerning privacy or another interest is a function of one's stance on the matter. In addition, just because people are faced with the same rule or program does not mean that they are all affected by it in the same way.

Arguably, all programs that threaten privacy affect a certain group of people more, whether it is immigrant communities and ID cards or predominantly white upper-middle class business people and ETC highways used to travel from their work in major cities to their homes in the suburbs. There are, however, two important points to emphasize when considering privacy's social effects. First, in order to fully understand a program's affect on privacy it is necessary to analyze the ways privacy risks are distributed across social categories such as race, class, gender, and ethnicity. Second, if such an analysis concludes that privacy risks disproportionately felt by a marginalized community, urgent attention needs to be given to mitigate this if negative impacts on social equity are to be avoided.

Bennett and Raab note that many systems for data protection provide avenues for redress, but many people may not be able to take advantage of them if they do not know their rights or have the means to exercise them. Access to such redress is very much a function of socio-economic status, so proactive steps should be taken to reach out to those who are disadvantaged in the process of seeking redress [1, p.47]. Overall, it is clear that a comprehensive understanding of privacy cannot be obtained without considering its social effects.

Effects from Changing Technology

Another privacy issue which is apt for discussions involving ITS is the effect which new technologies have on privacy theory. New technology is consistently challenging and rendering many dominant privacy assumptions irrelevant. For example, the distinction between the home (private) and the public sphere is challenged by technologies often found in the home such as the phone and internet which do not necessarily carry the guarantee of being private. In addition to individual technologies, the intersection of technologies leads to new privacy issues as well. An example of this is the way that information technologies have greatly increased data storage capabilities thus allowing for the accumulation of massive amounts of data. The ability to amass data presents new challenges for data privacy such as the security, accuracy, and longevity of databases. If data from credit card, phone, or similar source is collected and stored indefinitely without the owner's knowledge or ability to verify the data, any errors in the data which negatively affect the owner may go permanently unnoticed.

Lawrence Lessig also points out that privacy risks increase with the convergence of privacy affecting technologies. Although a single credit card bill, phone record, plane ticket, ETC record, hotel record may not be a significant risk to privacy, the combination of all these can provide quite the record of someone's existence [4]. Hence, it is important to consider not only the risk that an individual technology or program poses, but the larger context within which it is being released.

Regulation of Privacy

The United States does not currently have a comprehensive federal law covering privacy regulation. Instead, a patchwork of sectoral legislation is the main source of federal privacy regulation. Currently, federal sectoral regulation includes areas such as fair credit reporting, education records, video-rental records, polygraph tests, electronic funds transfer, financial privacy, and health information, among others [1]. This is in contrast to many European countries that have comprehensive privacy legislation or a combination of comprehensive and additional sectoral regulation. Since there is no sectoral or comprehensive coverage of ITS privacy considerations, many of those working with ITS data are left to state laws or their own policies to address privacy. This, as one can imagine, has led to quite the hodgepodge of ITS privacy policies across the US, and there have been many calls for national legislation to reduce such confusion. Given the post 9-11 political climate, however, national privacy legislation does not seem to be plausible anytime soon. The bottom line then, is that transportation professionals are responsible for and should be diligent in crafting privacy policies in accordance with state law that are appropriate for ITS applications. The following sections will examine ways in which this can be done.

Tools and Policies for Privacy Protection

Having provided a background on privacy theory and the current state of regulation, tools and policies which can be utilized to help provide data privacy are now discussed. One of the most well known sources of privacy problems is from data security breaches. Interestingly, most everyone agrees that such security breaches are problems. It is commonly realized that the hacking of databanks and the stealing of laptops are failures in the data privacy system that should be corrected. More complex situations arise when there is not agreement that privacy has been harmed. When trying to avoid these types of privacy concerns, organizations will find that much progress can be made by adhering to transparency and by explicitly addressing privacy questions. Trouble tends to occur when those involved with privacy do not ask the correct questions, not so much when they do not answer the questions "correctly." A good way of insuring that an organization addresses privacy concerns is to assign a single person or a group of people to be specifically in charge of privacy work.

The most well accepted set of standards for dealing with data privacy is what has come to be known as the "fair information principles"(FIPs). These FIPs are located explicitly or implicitly within all national data protection laws and appear in many voluntary codes and standards. These principles embody the procedural standards for proper data use and data privacy. While the specific wording varies, the principles state that a public or private organization:

- must be *accountable* for all the personal information in its possession;
- should *identify the purposes* for which the information is processed at or before the time of collection;
- should only collect personal information with the *knowledge and consent* of the individual (except under specified circumstances);
- should *limit the collection* of personal information to that which is necessary for pursuing the identified purposes;
- should not use or disclose personal information for purposes other than those identified, except with the consent of the individual (the *finality* principle);
- should *retain* information only as long as necessary;
- should ensure that personal information is kept *accurate, complete, and up-to-date*;
- should protect personal information with appropriate *security safeguard*;
- should be *open* about its policies and practices and maintain no secret information system;
- should allow data subjects *access* to their personal information, with an ability to amend it if it is inaccurate, incomplete, or obsolete. [1, p.12]

These FIPs should be incorporated into any organization's strategy for dealing with privacy. Not surprisingly, most of the issues with privacy occur because one or more of these principles are not being followed.

Since there is a lack of federal regulation for ITS, one of the tools used by organizations to protect privacy is self-regulation. Four general policy instruments used for self-regulation include: privacy commitments, privacy codes of practice, privacy standards, and privacy seals. *Privacy commitments* are the simplest form of self-regulation and are simply an organization's commitment to a set of privacy principles. *Privacy codes of practice* also include a set of rules for employees, members, or member organizations to follow. More than a simple commitment, privacy codes of practice provide guidance on proper procedure and behavior based on a version of the FIPs. They can take place on the organizational, sectoral, functional, or professional level. *Privacy standards* extend the code of practice by implying not only a common standard for performance measurement, but also an objective testing process to test an organization's claim of adherence. *Privacy seals*, the highest form of self-regulation, are awarded to organizations that are successfully registered or certified [1]. A commonly known seal is TRUSTe which was created to build public confidence on the internet. Since many privacy policies are integrated as inter-organizational policies, it is important that there is proper motivation or enforcement, so that members take them seriously.

In addition to self-regulation, there are many technological instruments which can be used to protect privacy. Public-key infrastructures, encryption tools, tools for anonymity and pseudonymity, and filtering tools all can be used to protect privacy [2]. It is important to note here that anonymous data may still carry privacy risks due to the fact that identity can often be determined through the combination of many different anonymous data elements. This risk has only increased with the growing ability and use of data mining, or utilizing powerful computers to search large amounts of data.

Organizational Policies to Protect Privacy

Bennett and Raab put forth that an organization's overall goals for data privacy protection, in addition to protecting privacy, should include balancing privacy with other objectives and promoting good computing practices. As described in the section above, the transparent use of the FIPs, self-regulation, and technical instruments is the best way for organizations to protect individual data privacy. Privacy considerations, in practice, will always be balanced with the other goals an organization is looking to fulfill. Bennett and Raab recognize that the idea of balancing privacy with issues efficiency, safety, or cost is similar to trying to compare apples and oranges due to their incompatibility. Given this reality, it is still important to analyze the competing interests so that the proper action for an organization can be determined. Proper action for an organization will be to neither ignore privacy concerns completely nor to privilege them over all other considerations, so some middle ground specific to each organization will need to be determined. Good computing practices are also important for organizations to achieve. These include practices involving the collection and maintenance of data such as adequate data security and data access procedures [1]. Actions similar to the potential breach of privacy at the Department of Veterans Affairs could have been avoided through proper attention to computing practices.

The criteria for analyzing the performance of data privacy protection goals listed above should include economy, efficiency, effectiveness, and equity. Economy refers to the cost of the input resources which are chiefly the money and staff deployed by the organization. Efficiency regards the relationship between the inputs and outputs, where outputs can be anything from the number of privacy complaints to the increased citizen access their personal data to the amount of educational outreach to the public. Effectiveness refers to the comparison of the output to the organization's overall objectives. Making sure that organizations have clear overall objectives with regards to privacy will help avoid situations where those working with privacy have to try to achieve a "balance between privacy and efficiency". Such vague objectives will only be a detriment, so thorough consideration should be given to determining meaningful, appropriate objectives. The final suggested criteria for determining the performance of data privacy protection is equity. Equity refers to embracing and utilizing the previous section's analysis of social effects. Looking at the ways that privacy risk is distributed across society can help realize and mitigate many of privacy's negative social equity effects. [1]

6.1.3 ITS and Privacy

The issue of privacy and ITS is unlike many of the issues dealt with by transportation professionals due to the competing technical and non-technical, subjective, social considerations involved. Due to the fact that privacy considerations have not historically been included and are difficult to compare to the typical transportation considerations of cost, efficiency, safety, and those involving the environment, it may be easier and more convenient to privilege those standard factors when analyzing ITS programs. The problem is that without adequately addressing the privacy issues involved with ITS technologies, the public will not support these novel technologies. Because public acceptance is essential for these systems to succeed, addressing privacy concerns is not something that transportation professionals can afford to neglect. In addition, addressing such an issue must not become merely trying to convince the public that they have nothing to worry about. A clear and transparent system of analysis and communication, using the tools and policies described in the previous sections, is needed to insure that the public is well informed and the concerns that they do have are well addressed. The following sections will analyze the privacy effects of three ITS applications: a safety program to reduce teen driving fatalities, a red light running law enforcement program, and the proposed Vehicle Integration Infrastructure project.

Safety ITS Applications

Many new ITS technologies have applications which can improve transportation safety. Max Donath, director of the Intelligent Transportation Systems Institute at the University of Minnesota, leads a project that looks at ways ITS can reduce traffic deaths in at risk age groups. Nationally, teenage drivers (16–19 years old) make up only 4.7% of all licensed drivers but are involved in 11.3% of all fatal crashes, giving them a higher fatality risk than any other driver age group on the road. In addition, although the annual teen fatality rate dropped from nearly 10,000 in the late 1970s to just over 5,000 in the early 1990s, the fatality rate has stopped decreasing and has actually slightly increased since then. Donath sees this as an indicator that new tools are needed to further decrease this number. Two of the major causal factors for teenage driving fatalities are speed and seatbelt use. Speeding or driving too fast for the conditions is a contributing factor to 23% of fatal teen driver crashes compared to just 15% for drivers age 30 to 50. [5]

Donath and his fellow researchers have identified three different ITS strategies for reducing teen fatalities and similar behavior modification applications: forcing behavior, driver feedback, and reporting behavior. Under the presumption that some unsafe conduct is habitual, forcing behavior requires specific actions to occur prior to or during vehicle operation. Driver feedback seeks to alert drivers to risks they otherwise would not be aware of, while reporting behavior looks to reduce the impulse people may have to take risks because they feel anonymous. Forcing behavior applications include seatbelt and alcohol interlocks that do not allow the vehicle to start if seatbelts are not in use or if alcohol is detected. Since over 50% of teen fatalities were not wearing seatbelts, and about 15% of 16–17 year old fatalities in addition to 30% of 18–20 year old fatalities had blood alcohol levels over .08, the legal limit for adults in many states, Donath believes these technologies could have very significant effects on teen fatalities [5]. Alcohol interlocks are already being used in New Mexico for first-time DWI offenders and there is legislation pending in at least 12 states that would require interlocks for some or all first-time offenders [6].

Another type of technology which could subsequently serve a forcing, feedback, or reporting function is intelligent speed adaptation (ISA). ISA employs a combination of GPS technologies and digital road maps to track driving behavior relative to a driver's road location. With ISA it is possible to decipher whether someone is driving 55 mph on a highway or on a residential road. ISA can be deployed in many ways including 1) a forcing fashion which overrides the driver to limit the vehicle speed, 2) a feedback function that would provide in-vehicle advisories and 3) a reporting fashion where the driving behavior of teenagers or at risk drivers could be monitored by parents or the appropriate authorities. The notification speed could be based on posted speed limits, account for construction zones and school zones, or even be based on hazards such as weather or traffic congestion. ISA has been evaluated in simulation and field studies in several countries, including Australia, Belgium, France, Germany, England, Netherlands, and Sweden. In general, these projects have shown a consistent reduction in speed levels, better awareness of speed limits, and improved compliance with speed limits. [5]

Donath is quick to point out that the best application of these technologies would be with young drivers while they are early in the learning process. The ISA or interlock technologies could be used to help teen drivers develop good driving habits, after which the technology would be removed. The specific, transparent application of these technologies to at risk drivers, such as teenagers and those convicted of DWIs, would help relieve some of the privacy concerns associated with this type of ITS. Although it would be expensive to install these technologies on cars which are already on the market, it would be much more affordable to install them during the vehicle manufacturing process. Another issue is how to set the performance criteria for the teens to pass their training period. Overall, however, Donath believes that these new technologies can help provide the needed boost to further reduce the teenage driver fatality rate and threats from other high-risk drivers. [5]

Privacy Analysis of ITS Safety Application

The project to reduce teen driving deaths is a great example of an ITS aimed at mitigating a public problem and subsequently reducing the privacy of a group of individuals. Many of the potential issues for this project are similar to those for comparable types of ITS safety projects. Although the two technologies proposed to reduce teen driving fatalities, interlocks and intelligent speed adaptation (ISA), would have different effects on privacy, there are many similarities in these technologies which can be drawn out for analytical purposes. Both technologies rely on mandatory participation of the group determined at risk, and the resulting privacy risks are distributed only among this group. For this type of program, where an at-risk group is forced to participate, questions often arise with talk about the program's expansion. In this case, many people may agree that teens should be forced to use interlocks or ISA systems for their and other's safety, but if the possibility of expanding this program to the elderly or to the public as a whole is brought up, many may have reservations. Thus, the issue of mission creep, or the changing of a programs scope after it is developed, needs to be thoroughly addressed. It is therefore important for transparent communication to take place about the benefits and risks of an ITS program before it is implemented so that the appropriate scope can be determined, communicated, and implemented. If the acceptance of an ITS program varies based upon how many groups of people fall under its jurisdiction, the decision about the cutoff level becomes more complicated. One would want to avoid potentially beneficial ITS programs from being opposed based on the unrealistic hyped-up threat of domain expansion, while making sure that programs that would normally be opposed are not snuck in under the disguise of smaller, less threatening programs. Also important is to make sure that described problem is likely to be mitigated by the proposed ITS program. There is no reason to threaten privacy if the problem is not well understood and not likely to be solved by the proposed ITS program. [2]

Law Enforcement ITS Technologies

An increasingly common use of ITS technologies is for law enforcement. Law enforcement officials can obtain ITS data directly from an application such as a red light camera, or in a secondary form such as accessing a collection of ITS data from a source such as a department of transportation. The upcoming section on the Vehicle Infrastructure Initiative will discuss the complexities surrounding law enforcement's use of secondary access of ITS data. This section will study the primary use of ITS technologies by law enforcement by examining the Minneapolis Police Department's Stop on Red program.

The Minneapolis Police Department implemented a red light running program in the summer of 2005 at 12 intersections in Minneapolis. Cameras were utilized to take short videos of the car and its license plate when sensors were set off by illegal entrance into the intersection. Fines of $142 were given to more than 26,000 people from June of 2005 until March of 2006. In March the program was deemed unconstitutional by a Minnesota court because it assumed that the owner of the car was the driver without having received enabling legislation from the Minnesota Legislature to make that assumption. This ruling is currently being appealed and the city hopes to eventually reinstate the program [7]. Although this program was at least temporarily struck down, the statistics from its short operation time highlight the benefits of such a program. As described by Lt. Gregory Reighnhardt, the head of Minneapolis's Stop on Red program, an average police officer in an urban setting issues 100 traffic violation tickets/ year, while with this program one officer sitting at a computer was able to issue 26,000 in nine months. In addition, Reighnhardt cited a 31% decrease in overall accidents at these intersections. For Reighnhardt, this program and ITS technologies represent the ability to free up officers to pursue more serious crime, while at the same time changing people's driving behavior to increase public safety. [Lt. Reighnhardt, "unpublished data"]

Law Enforcement Privacy Analysis

Although not currently in operation due to the court ruling, much can be learned from the introduction and operation of Minneapolis's Stop on Red program. Some interesting strategies were included in the implementation of the Stop on Red program to try and mitigate privacy concerns. The first was that all of the intersections used for the program were publicly announced, posted on the police department's website, and marked with street signs. This was completed for the sake of transparency, to encourage people to change their driving behavior, and so those people interested in learning about the program would not have to turn to sources critical of the program to learn of camera locations. The other step taken for the sake of privacy and transparency was that everyone who received a ticket from the program could go online and see the video of the infraction. In addition to the added transparency, this feature also reduced the motivation for challenging tickets [Lt. Reighnhardt, "unpublished data"].

The operation of this program generated large amounts of data, the resulting issues of which are interesting to examine. Data was collected via the license plate and registration information for all of the 26,000 tickets issued and for about another 30,000 triggered events. Triggered events are those in which the red light cameras observe a violation but there is no ticket issued either because the license plate cannot be read, the car was taking a legal left hand turn, or some sort of similar situation. For each of these tickets and triggered events, 35 to 40 pieces of information such as the vehicle owner's name, date of birth, and address are collected. Currently, all of this data is saved indefinitely in case the city has a lawsuit filed against it. Not surprisingly, those involved in the management of the data have considered the many possible secondary uses for the data. Since the address of the vehicle's owner is listed, the data could be valuable marketing information in determining traffic flows. Although the head of the program stated that he does not believe secondary uses should be allowed and that he was not pursuing any, he did say that there were no laws prohibiting it [Lt. Reighnhardt, "unpublished data"]. This is an example of the types of issues that someone using ITS technologies may encounter with the current regulatory structure. Looking at the FIPs for guidance in this situation shows that the retention and use identification principles could be employed here. There would be no question of secondary uses if the data was used only for the purposes originally identified. Secondly, not retaining the data indefinitely would result in fewer opportunities for data security lapses and data misuses.

And finally, this Stop on Red program is a good example of how one could look at privacy's effect on social equity resulting from an ITS program. Looking at the placement of this program's cameras is a starting point for analyzing the distribution of privacy risk. Although the red light cameras were said to have been placed at locations which were "high accident" zones, the resulting locations are in neighborhoods with significant populations of people of color. The 2000 census placed the non-white population of Minneapolis at 34.9% of the total population. Out of the twelve cameras put into operation, the four downtown vicinity cameras were located in neighborhoods with between 29.4 and 47.8% non-white population. Out of the remaining eight, five were located at intersections next to residential areas with over 78.1% non-white population. One had a non-white population between 47.8 and 78.1%, one between 29.4 and 47.8, and only one between 29.4 and 18.5. [8][9]

Five out of the twelve cameras were placed in neighborhoods with more than double the city's average percentage of non-whites, while only one camera was placed in a neighborhood with a smaller non-white population that the city average. This racially-correlated placement resulted in a disproportionate distribution of the risk of the loss of privacy, since these populations were exposed to increased surveillance from the cameras. Since the placement of the cameras—intentionally or unintentionally—was such that people of color were disproportionately exposed to reduced privacy, continuation of the program in this manner would potentially have a negative social equity effect.

It should be noted that the above analysis assumes that the people living near the cameras are more exposed to more risk. If, as may well be the case with the four downtown cameras, the population living around the intersections is a large user of public transit and there is a large volume of commuters from more racially representative outside areas, the resulting risk from these cameras may be equally distributed. This explanation is less likely to hold for the five intersections in non-downtown neighborhoods where the non-white population is twice that of the city average. Also, if the chosen intersections were so accident prone that there was an outcry for action by the adjacent neighborhoods, despite the knowledge of an increased privacy risk, then it could be argued that the increased privacy risk may be warranted. One challenge here would be to find the representative view of the neighborhood, since not everyone in a neighborhood is going to hold the same views as to the proper level of privacy to sacrifice for increased safety. If with this program, however, there were many other sites that could have been chosen and cameras were not seen as necessary to improve public safety by the adjacent neighborhoods, then the privacy risk inherent in these cameras should be more evenly spread throughout Minneapolis's population. This analysis has show the importance of studying privacy's social effects, in addition to providing an example of the types of questions that should be asked of ITS projects when studying their social effects.

Vehicle Integration Infrastructure

The Vehicle Integration Infrastructure (VII) is a United States Department of Transportation initiative which aims to deploy and enable an ITS infrastructure that supports vehicle-to-infrastructure and vehicle-to-vehicle communications for a variety of vehicle safety applications, transportation operations, and applications that support private interests [10]. The types of applications possible with the program are a result of the increased vehicle and infrastructure communication and include: accident reduction, increased efficiency through traffic signals, traffic re-routing capabilities, defect and warranty information transmission, real-time schedule transmission for transit vehicles, speed and lane departure warnings, roadway hazard detection, electronic toll collection, and traffic violation enforcement. Due to breadth of the VII plan, the VII Coalition's working group consists of US transportation related agencies, eight automotive manufacturers, state and local agencies including ten state DOTs, and several associations such as the Alliance of Automobile Manufacturers and ITS America. The group realized early on that the most difficult issues involved with VII would not be the technical issues, but the economic, institutional, and political issues. Because of the potential privacy issues accompanying a fully implemented VII program, the working group determined that rigorous privacy principles would be integral to the success of the program [10]. Although these privacy principles are still in the process of being created and approved, a draft form of them was outlined at the CTS conference. Parts of these principles will be analyzed in the next section.

Vehicle Integration Infrastructure Privacy Policy and Analysis

Although referred to by different titles, the VII privacy principles are very similar to the Fair Information Principles. There are a couple specificities in the VII privacy principles which deserve special attention. Their Notice Principle states that it is necessary to notify people about any opportunity to remain anonymous [Marthand Nookala, "unpublished data"]. The issue of anonymity is a significant one because much data can be collected and used for traffic management without being personally identifiable and thus not being such a threat to privacy. Many potential operations such as electronic toll collection or traffic violations would, however, require personal information. Deciding which applications people will have the option to remain anonymous for will be very significant, as many privacy concerns hinge on that decision.

Another noteworthy privacy policy is that VII will not be designed for non-transportation applications, such as law enforcement or national security issues, but such interests will be able to access VII data using normal mechanisms including warrants [Marthand Nookala, “unpublished data”]. Such a policy seems to be trying to address the concerns that a VII system may be used to increase government surveillance. If technically feasible to personally track every driver’s location, why wouldn’t law enforcement officials try to streamline this ability? This policy was created to insure that the infrastructure would be built only with transportation goals in mind, and that generally trusted legal mechanisms such as warrants will be the only way for data to be accessed. It is uncertain whether these policies will mitigate the privacy concerns involving such a massive program as VII. What is certain, however, is for the full potential of these draft privacy principles to be realized they will need to be passed with strong meaningful language and clear guidance for implementation.

Consumer and Commercial Opportunities and Risks

Although there is not adequate space for a full description and analysis, it is important to note possible consumer and commercial ITS uses. There are many ITS consumer products on the market today from well known products such as OnStar to the lesser known products such as those which can track teen drivers’ locations for their parents using GPS technology. Consumer product privacy concerns differ from those described previously in that, given proper notice by the product’s manufacturers, the consumer generally 1) is aware of the privacy risks at hand, 2) consciously makes the decision that the benefits of the product are worth the increased privacy risk (parent may be making this decision for teen’s privacy), and 3) is exposing themselves (or their teen) and not the general public to increased privacy risk. As long as there is transparency between manufacturer and user (which may not always be true), this type of privacy risk is understood, voluntary, and therefore not as problematic.

There are many commercial opportunities for ITS technologies which predominantly involve the ability to increase safety and efficiency. As the existing roadways continue to become more crowded for commercial vehicles, one way to adapt is through increased safety and efficiency via ITS technologies such as real-time GPS tracking or the proposed VII system. One privacy issue is that mobile workers who are using ITS tracking systems may be exposed to an increased level of surveillance during their work, and they should be adequately warned [2]. Other ITS issues for commercial interests include the lack of comprehensive national privacy legislation, the issue of ITS data access by competitors, and the lack of financial return on investment for new ITS technologies (Dan Murray, “unpublished data”).

6.1.4 Conclusion

The increasing importance of intelligent transportation systems and their resulting privacy risks will require that transportation professionals pay special attention to privacy concerns. This paper provided a background on ITS relevant privacy theory, proposed policies and tools to help protect privacy, and provided relevant examples and analyses to further the comprehension of privacy issues. Entities collecting ITS data should have a privacy policy on data use including the fair information principles such as accountability, purpose identification at the time of collection, informed consent for collection, use limit and disclosure, individual access and opportunity for correction, data quality and security standards, in addition to retention limitation. It is also important to analyze privacy’s social effects by considering the distribution of risk across categories such as race, class, gender and ethnicity, and to insure that any negative effects are mitigated. In addition, organizations should allocate a specific person or group of people and proper resources to dealing with privacy as well as utilizing technical and self-regulatory privacy mechanisms. Only through transparent thought and action regarding privacy concerns can transportation professionals involved with ITS insure that their programs are developed and implement effectively.

6.1.5 References

1. Bennett, Colin J. and Charles D. Raab. *The Governance of Privacy: Policy Instruments in Global Perspective.* The MIT Press, Cambridge, Massachusetts, 2006.
2. Bennett, Colin J. Privacy and Mobility: New Challenges for Technology-Enhanced Transportation. Presented at 2006 University of Minnesota Center for Transportation Studies Annual Conference, Minneapolis, MN, 2006.
3. Files, John. V.A. Laptop Is Recovered, Its Data Intact. *The New York Times.* June 29, 2006
4. Lessig, Lawrence. The Architecture of Privacy. Presented at the Taiwan Net conference in Taipei, March, 1998.
5. Donath, Max. et al. Using ITS to Reduce Teenage Road Mortality: Impact on Privacy. Presented at 2006 University of Minnesota Center for Transportation Studies Annual Conference, Minneapolis, MN, 2006.
6. O'Donnell Jayne. Will all autos some day have breathalyzers? *USA Today*. April 28, 2006.
7. Zack, Margaret. Minnesota Court of Appeals hears arguments for 'photo-cop' program. *Star Tribune.* June 27, 2006.
8. United States 2000 Census data. http://quickfacts.census.gov/qfd/states/27/2743000.html. Accessed July 5, 2006.
9. Percentage Non-Hispanic White, Hennepin County, 2000 United States Census. www.dataplace.org. Accessed July 5, 2006.
10. VII White Paper Series: Primer on Vehicle-Infrastructure Integration. The Intelligent Transportation Society of America, 2005.

CHAPTER 7

David Levinson: Travel Demand Modeling

7.1 A Model of the Rise and Fall of Roads

Lei Zhang and
Ph.D. Candidate
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-626-0024 F: 612-626-7750
E-mail: zhan0294@tc.umn.edu

David Levinson
Assistant Professor
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-625-6354 F: 612-626-7750
E-mail: levin031@tc.umn.edu

7.1.1 Introduction

In 1900 there were 240 km of paved road in the United States (Peat 2002), and this total had increased to 6,400,000 by 2000 (BTS 2002) with virtually 100% of the U.S. population having almost immediate access to paved roadways. The growth (and decline) of transportation networks obviously affects the social and economic activities that a region can support, yet the dynamics of how such growth occurs is one of the least understood areas in transportation, geography, and regional science. This lack of understanding is revealed time and again in the long-range planning efforts of metropolitan planning organizations (MPOs), where transportation network change is treated exclusively as the result of top-down decision-making. Non-immediate and non-local effects are generally ignored in planning practices because the complete network effects are incomprehensible with the current tools, which often results in myopic network expansion decisions. If one looks at the complexity and bureaucracy involved in transportation infrastructure investment, one might conclude that it is impossible to model the transportation network dynamics endogenously. However, changes to the transportation network are rather the result of numerous small decisions (and some large ones) by property owners, firms, developers, towns, cities, counties, state department of transportation districts, MPOs, and states in response to market conditions and policy initiatives. Though institutions make network growth (decline) happen on the surface, network dynamics are indeed driven by some underlying natural market forces and hence predictable. Understanding how markets and policies translate into facilities on the ground is essential for both scientific understanding and improving forecasting, planning, policy-making, and evaluation.

A transportation network is a complex system that exhibits the properties of self-organization and emergence. Previous research in dynamics related to transportation networks focuses on traffic assignment or traffic management. However, the dynamics of transportation network growth have not been adequately studied. If a transportation network is represented by a directed graph, there are several important questions yet to be answered: 1) How do the existing links (roads) develop and degenerate? 2) How are new links added to the existing network? 3) How are new nodes added to the existing network? This paper concentrates on the first question and focuses only on the rise and fall of existing roads, recognizing the inter-dependence of road supply and travel demand. The approach is microscopic in that network dynamics are modeled at the link level. The following key questions are examined:

1. Why do links expand and contract?
2. Do networks self-organize into hierarchies?
3. Are roads (routes) an emergent property of networks?
4. What are the parameters to be calibrated in a microscopic network dynamics model?
5. Is the model computationally feasible on a realistic transportation network?
6. Is the model capable of replicating real-world network dynamics?

One of the few previous studies (Yerra and Levinson 2002) in this area shows that even starting from a random or an uniform pattern, a transportation network tends to self-organize into a hierarchical pattern in which some roads attract more traffic, receive proper maintenance, and are gradually expanded, while some other roads are less popular, poorly maintained, and may eventually be abandoned. It is also demonstrated that although these hierarchies seem to be designed by planners and engineers, they are actually intrinsic emergent properties of networks themselves. However, the simulation model developed in that study is restricted in several ways. First, the links are assumed to have unlimited capacities. The impacts of network congestion on travel demand are ignored. The assumption also results in another unrealistic property of the model -the growth and decline of links are only reflected by changes of their free-flow speeds. Secondly, their model is tested only on simple hypothetical networks and it is not clear if the conclusions regarding road hierarchies hold on large-scale realistic networks. Those two restrictions are relaxed in this study. Travel demand is represented by a more realistic user equilibrium pattern. In the network evolution process, links exhibit dynamics in both free-flow speed and capacity. The improved model is then applied to the Twin Cities transportation network with nearly 8,000 nodes and more than 20,000 links, which allows us to examine computational properties and predictive value of the proposed microscopic network dynamics model.

The next section presents a brief review of related literature in regional science and economics. Though the reviewed studies have dissimilar objectives and methodologies, they all shed some lights on the nature of transportation network growth and its social-economical impacts. The following section develops a theoretical framework for studying the rise and fall of roads. The framework helps identify various influencing factors and inter-dependences among those factors, based on which a synthesis model of road expansion and contraction is developed in Section 7.1.4. The model is applied to the Twin Cities transportation network from year 1978 to 1998 with different model parameters and starting conditions. The results of those experiments are summarized, and the listed research questions answered in Section 7.1.4. Conclusions and future research directions are offered at the end of the paper.

7.1.2 Background

Few researchers have considered the process of transportation network growth at microscopic level, highlighting the importance of this research. Taaffe et al. (1963) study the economic, political and social forces behind infrastructure expansion in underdeveloped countries. Their study finds that initial roads are developed to connect regions of economic activity and lateral roads are built around these initial roads. A positive feedback between infrastructure supply and population was also observed. Barker and Robbins (1975) investigated the London Underground's growth, but did not develop a theoretical framework as we are considering here. Miyao (1981) developed macroscopic models to take transportation improvements as either an endogenous effect of urban economy or as an exogenous effect on the economy. Endogenous growth theory suggests that economic growth is a two-way interaction between the economy and technology; technological research transforms the economy that finances it (Aghion and Howitt

1998). The technology of transportation is unlikely to be an exception, suggesting transportation investment drives the growth that funds it. Macroscopically, the growth of infrastructure follows a logistic curve and that road infrastructure also has reached saturation levels in developed countries (Grübler 1990). Miyagi (1998) proposes a Spatial Computable General Equilibrium (SCGE) model interacting with a transportation model to study the interaction of transportation and the economy. Yamins et al. (2003) develop a road growth model to study co-evolution of urban settlements and road systems from an empty space with highly simplified travel demand and road supply mechanisms meaningful only for theoretical works.

Carruthers and Ulfarsson (2001) find that various public service expenditures like roadways are influenced by demographic and political characteristics. The New Jersey Office of State Planning (1996) also finds a similar pattern in roadways expenditure. A related line of research examines how transportation investment affects the economy at large, but tends to treat transportation (or highways) as a black box, and makes no distinctions between different kinds of transportation investment (Aschauer 1989, Boarnet 1997, Button 1998, Gramlich 1994, Nadiri and Mamuneas 1996). The input is investment in transportation (or infrastructure), and output is gross domestic product, measured at the state level. While this research provides no assistance in actually making management decisions, it suggests a way that a macroscopic network investment budget can be established.

Geography's central place theory seeks to explain how hierarchies of places develop (Christaller 1966). Models developed by Batty and Longley (1985), Krugman (1996), and Waddell (2001) consider land use dynamics, allowing central places to emerge. However, those models take the network as given. Clearly, there is a need for research that makes the network the object of study. In many respects, the hierarchy of roads is the network analogue of the central place theory.

7.1.3 Network Dynamics at the Microscopic Level

Regional economic growth is taken as exogenous for this study of transportation network dynamics because transportation infrastructure is not the only factor that drives economic growth and we do not yet have adequate other models to explain change in land use. It has long been known that transportation service and land use influence each other though iterative changes in accessibility and travel demand. However, land use dynamics are also treated as exogenous in the following network analysis, so that attention can be focused on transportation network growth, a process with enough complicated and unknown dynamics for one to start with. This limitation can be removed in future research. The dynamics of other factors involved such as travel behavior, link maintenance and expansion costs, network revenue, investment rules, link expansion and degeneration, are considered endogenously.

The foremost and probably also the most important constraint on future network growth is the existing network. In developed countries where transportation infrastructure has reached its saturated stage, it is rare to see new network growth from a *tabula rasa*. Even in an empty place without any previous developments, natural barriers such as rivers and mountains still impose constraints on future network growth. The current network connectivity determines whether two links complement (upstream or downstream) or compete (parallel) each other for demand. The existing network may or may not be at an equilibrium. It may still take years for road supply to meet existing travel demand even if no exogenous changes (e.g. population and economic growth) occur. The important question is how various forces drive the existing network to evolve, more than how long it takes.

Based on the current network, land use arrangements and individual socio-economic status, people make their travel decisions such as trip frequency, scheduling, destination, mode and route choices. These decisions transform into travel demand on the transportation network. This demand-generating process involves the existing network supply, congestion externalities, travel behavior, and link-level travel demand forecasting.

Transportation is a service and travelers pay to obtain that service in addition to spending their own travel times. In the US that payment is largely in the form of a fuel tax. However, if links were autonomous, they would set prices to maximize their profits in the form of a vehicle toll. In many real-world transportation networks, government agencies collect transportation revenue in terms of fuel taxes. We can set the price for using a link as a function of the length and the level of service (LOS) of the link. It is convenient to use a notion of link revenue. Revenues collected by individual links may or may not be pooled together for investment purposes depending on the underlying institutional structure of the network. Longer, faster, and high-demand (traffic flow) links should be able to generate more revenues. If not maintained appropriately, link LOS will decrease over time due to physical deterioration caused by the environment and traffic. Therefore, each link has a maintenance cost function. Link length, capacity, free-flow speed, and flow determine maintenance cost to a large extent. The amount of money required to expand an existing link can be calculated with a link expansion cost function. A previous empirical estimation of link expansion costs using network data in the Twin Cities metro area during the recent twenty years reveals that link expansion cost is positively correlated to lane-miles of expansion and road hierarchy (interstate, state highway, county highway, etc.), while negatively related to the distance from the nearest downtown (Karamalaputi and Levinson, 2003). Those results suggest that link length and capacity should be included in the link expansion cost function, and such function is also subject to local adjustments.

Specific revenue and cost structures in a transportation network provide inputs for investment decisions. Real-world observation suggests the hypothesis that decisions to expand transportation networks are largely myopic in both time and space, usually ignoring non-immediate and non-local effects. This myopic decision process, when applied sequentially, tends to improve the relative speeds and capacity of links that are already the most widely used, and thereby expand their use. The rate and extent of this process is constrained by the cost of those improvements and limited budgets (revenue). From a market economy point of view, transportation investment decisions induce supply (capacity) increases—as population grows and preferences shift, leading to higher demand, suppliers produce more of a good. While surface transportation decisions are often made in the political arena rather than the market, politicians and officials also respond to their customers—the voter and taxpayer. Although over the short-run transportation supply is relatively inelastic; in the long run it varies. However, it is not known to what extent changes in travel demand, population, income, and demography drive these long-run changes in supply. Answering this induced supply question in transportation is a critical step in understanding the long-term evolution of transportation networks. The output of the investment process would be an updated network where some links are expanded and some degenerated.

If a link is expanded, travel increases on that link both due to re-routing and rescheduling and due to what is often called induced or latent demand, a finding confirmed at both the macroscopic level (states and counties) (Noland, 1998, Strathman et al. 2000, Fulton et al. 2000) and at the microscopic level (individual links) (Parthasarathi et al., 2002). As travel costs for commuters are lowered, the number of trips and their lengths increase. The expanded link with increased travel demand can generate even more revenue, which may later result in further expansion on that link. Yet this loop, while positive, should have limits. The diminishing returns in the revenue structure and exponential increases of expansion costs will eventually stop this feedback loop. The opposite is true for degenerated links. All these suggest that reinforcement exists and transportation networks may self-organize into hierarchies. This hypothesis is subject to simulation tests in the following section.

Improving one link will also cause complementary (upstream and downstream) links to have greater demand, and competitors (parallel links) to have lesser demand (and be less likely to be improved). These network effects take time to propagate within transportation networks. They may get reinforced in complex transportation networks, create problems, leave little clue to planners as to the root of the problem, force planners to adopt myopic solutions which may crate even more problems. Such a condition has not been confirmed empirically but it is possible. This again highlights the importance of considering the full ramification of network expansion on future infrastructure decisions. Network effects both complicate the problem and suggest the analysis has to be iterative. Previous changes of the network, economy, demography, and even travel behavior cause a new travel demand pattern and hence new link costs and revenues. Accordingly, a new set of supply decisions will be made, generating new network changes. This loop is repeated until an equilibrium is achieved. When the constant exogenous changes in economy, technology, and population are considered, a transportation network may never reach equilibrium. The evolutionary microscopic network growth process should produce rich dynamics important to anyone who is interested in shaping the future transportation network in a better way.

7.1.4 A Network Dynamics Model

In this research, a network dynamics model is developed that brings together all the relevant agents and their interactions to simulate link expansion and contraction. Compared to the earlier network dynamics model due to Yerra and Levinson (2002), this improved model relaxes the assumption of unlimited link capacity, a necessary step that has to be taken to make the model be of any practical importance. The foundation for the model development is the microscopic network growth dynamics described in the previous section. The simulation model can be used to evaluate whether or not important system properties such as hierarchy, self-organization, and growth, actually emerge from decentralized processes. This purpose makes the principles of and modeling techniques for complex systems applicable. There is no universally accepted definition of a complex system. However, it is generally agreed that it consist of "a large number of components or 'agents', interacting in some way such that their collective behavior is not simple combination of their individual behavior" (Newman 2001), which is the case in transportation networks. Examples of complex systems include the economy—agents are competing firms; cities—places are agents; traffic—vehicles are agent; ecology—species are agents. In transportation networks, we model nodes, links, travelers and land use cells as agents. Cellular Automata (CA) and agent-based modeling techniques are commonly employed tools for modeling complex systems (von Neumann 1966; Schelling 1969; Wolfram 1994, 2002) that has been applied to model traffic (Schadschneider and Schreckenberg, 1993; Nagel and Schreckenberg, 1992). An agent-based structure is pursued wherever possible in the proposed network dynamics model.

An overview of model components and their interconnectivity is shown in Figure 7.1.1 A travel demand model predicts link-level flows based on the network, socioeconomical and demo–graphical information. Based on the demand forecasting results, links calculate revenues and costs. An investment module then operates and causes annual supply changes, producing an updated network. The modeling process does not have to iterate annually. Other updating intervals can also be used. But yearly supply changes correspond to budgets which are typically decided once every fiscal year. The transportation network is represented as a directed graph that connects nodes with directional arcs (links). The standard notation convention for directed graphs is adopted for the following presentation on the details of mathematical formulations of those sub-models. The directed graph is defined as: $G = \{N, K\}$ where N is a set of sequentially numbered nodes and K is a set of sequentially numbered directed arcs.

Travel Demand

Ideally, an agent-based travel demand model in which node, link, and travelers are modeled as interactive agents should be applied to estimate travel demand at the level of links, so as to keep the disaggregate model structure consistent. A previous study (Zhang and Levinson 2003) has proposed such a model with successful application to the Chicago sketch network. However, for two reasons it is not adopted here. First, in its current form, the agent-based travel demand model is not capable of incorporating congestion effects. The second and probably more important reason is that most urban planners currently do not use disaggregate approaches to predict future travel demand in their daily practices. Therefore, a traditional four-step forecasting model is used to predict travel demand at the link level, taking exogenous land use, social-economical variables, and the existing network as inputs. A zone-based regression structure is used for trip generation. The origin-destination (OD) cost table obtained from the previous year traffic assignment is used for trip distribution in the current year based on a doubly constrained gravity model (Haynes and Fotheringham 1984, Hutchinson 1974). The computation of the new OD demand table takes into account the historical impacts of past travel behavior. Travel demand in a given year depends on the demand in the previous year. Levinson and Kumar (1996) elaborate the idea of such a hybrid evolutionary model. In contrast to a traditional equilibrium model, the evolutionary demand updating procedure does not require supply and demand to be solved simultaneously. In this study, the new OD demand is updated by a process similar to the method of successive averages (MSA) (Sheffi 1985, Smock 1962) in traditional traffic assignment procedures. The weights in equation (1) are specified in such a way that OD demand tables in all preceding years are weighted equally toward the current year (t) OD demand.

$$d_{ij}^{t} = \left(1 - \frac{1}{t}\right) d_{ij}^{t-1} + \left(\frac{1}{t}\right) a_i O_i b_j D_j f(c_{ij}^{t}) \qquad (1)$$

where:

d_{ij}^t demand from origin zone i to destination zone j in year t

O_i number of trips produced from zone i

D_j number of trips destined for zone j

a_i , b_j coefficients in the gravity model

c_{ij}^t generalized travel cost of traveling from zone i to j

$f(.)$ travel cost impedance function in the gravity model; $f(c_{ij}^t)$

The resulting OD table is loaded onto the current year transportation network through the origin-based user equilibrium traffic assignment algorithm (OBA) developed by Bar-Gera and Boyce (2002). The generalized link cost function comprises two parts, a BPR travel time component and a vehicle toll.

$$c_k^t = \gamma \frac{l_k}{v_k^t}\left[1+\alpha_1\left(\frac{q_k^t}{C_k^t}\right)^{\alpha_2}\right]+\tau_k^t \tag{2}$$

where:

c_k^t generalized travel cost on link k in year t

γ value of travel time constant (dollar/hr)

v_k^t free-flow speed of link k (km/hr) in year t

C_k^t capacity of link k in year t (veh/hr)

l_k the length of link k (constant) (km)

q_k^t average hourly flow on link k in year t (veh/hr)

α_1 , α_2 coefficients of the BPR travel time function

τ_k^t link toll per vehicle (dollar, see equation 4 for details)

The OBA algorithm derives link flows at user equilibrium and generates a new OD cost table which will be used for trip distribution in the next year. In the traffic assignment step, if the relative excess travel cost is less than 0.001, the Wardrop user equilibrium (Wardrop 1952) is considered to be satisfied.

Revenue and Cost

Revenue is collected at the link level by vehicle toll. The annual revenue is simply the product of the toll and annual flow. The amount of the toll should be dependent on the length of the link and the level of service. Therefore, the following revenue equation is proposed:

$$\pi_k^t = \tau_k^t \cdot (\omega \cdot q_k^t) \tag{3}$$

$$\tau_k^t = \rho_1 \cdot (l_k^t)^{\rho_2} \cdot (v_k^t)^{\rho_3} \tag{4}$$

where:

π_k^t	revenue of link k in year t (dollar)
ω	coefficient to scale average hourly flow to annual flow
ρ_1	scale coefficient related to the toll level (dollar·hr$^{\rho 3}$/km$^{\rho 2+\rho 3}$)
ρ_2, ρ_3	coefficients indicating economies or diseconomies of scale

As the free-flow speed of a link increases, travelers are able to save travel time and hence willing to pay a higher toll. However, speed improvements have decreasing returns. For instance, if speed triples from 8 to 24 km/hr, time spent traveling 1 km drops 5 minutes from 7.5 min to 2.5 min. If speed increases 16 km/hr from 88 km/hr to 104 km/hr, the time drops from 41 seconds to 35 seconds—merely 6 seconds—which hardly seems worth considering. Therefore, coefficient ρ_3 should be between 0 and 1. Note that with appropriate values for those coefficients, the toll-based link-level revenue structure can also reasonably model centralized revenue collection mechanisms, such as fuel taxes ($\rho_2 = 1$ and $\rho_3 = 0$).

The link maintenance cost function has only two determining factors: link length and capacity:

$$\psi_k^t = \theta_1 \cdot (l_k^t)^{\theta_2} (C_k^t)^{\theta_3} \quad (5)$$

where

ψ_k^t	cost of maintaining link k at its present condition in year t (dollar)
θ_1	scale parameter (dollar·hr$^{\theta 3}$/km$^{\theta 2}$)
θ_2, θ_3	coefficients indicating economies or diseconomies of scale

It is also assumed that all links have the same link maintenance cost function. This assumption is obviously not realistic and should be relaxed when local link-specific data are available.

Link expansion cost function is not explicitly specified. If a link is autonomous and its annual revenue is higher than maintenance cost, the link will be expanded in the next year, assuming revenue is not spent elsewhere. If revenue falls below maintenance cost, the link shrinks in terms of capacity reduction and free-flow speed drop. As we will see later in the investment model, those ideas are actually incorporated into a link expansion/contraction function.

Investment Rules

The sub-model of network investment decisions can have two aims, describe reality or identify optimal policies. The emphasis in this paper is the prior one, which is in contrast to the long line of research on the Network Design Problem. The network dynamics model must be able to replicate what has happened in reality before it is applied for potential planning purposes. A prototype investment rule (link expansion and contraction function) is examined in which links manage themselves and do not share revenues.

$$C_k^{k+1} = C_k^t \left(\pi_k^t \middle/ \psi_k^t \right)^{\lambda} \qquad (6)$$

where

λ capacity change coefficient

Note that investment decisions in equation (6) are very myopic ones in that links only care about themselves, ignore network effects and spend all revenues immediately. The value of λ actually represents some properties of the link expansion process. If ⌊ is less than 1, it implies that there are diseconomies of scale in link expansion because doubled investment (π) would only produce less than doubled capacity. If ⌊ is larger than 1, economies of scale exists. Capacity changes of a link are usually associated with changes in free-flow speed. Vehicles are able to travel at faster speed on a wider road with less impedance. Free-flow speed and capacity data used by the Twin Cities Metropolitan Council in their regional transportation planning model on more than ten thousand roadway sections were used to study the co-evolution of speed and capacity. A regression model with a logarithm function is adopted (see Figure 7.1.2). R^2 of the model is 0.7 and both coefficients are statistically significant at level 0.01.

$$v_k^{t+1} = \mu_1 + \mu_2 \cdot \ln(C_k^{t+1}) \qquad (7)$$

The predicted free-flow speeds are plotted against data in Figure 7.1.2. Keeping component functions such as this one continuous and differentiable in the network dynamics model can save a lot of work for the calibration stage. This is also the reason why an explicit link expansion cost function is not specified and why it is assumed that links invest any extra revenue immediately. However, if these simple continuous functions cannot adequately replicate reality, more sophisticated modeling tools should be considered. For instance, one link expansion and contraction are in fact discrete events for which a choice model or catastrophe theory may be applied. With updated link capacity and free-flow speed, some factors influencing travel behavior such as link travel time and link toll change. Theses supply shifts, combined with preference, economical growth and demographical changes, give rise to the emergence of a new demand pattern.

So far, a complete cycle of the network evolution process has been modeled. This cycle repeats itself year after year. Simulation of these cycles can reveal various emergent properties of transportation network growth. The proposed network dynamics model can and should be calibrated and validated against observed time-series network and land use data. The calibration procedure may consist of two stages. The parameters in the sub-models (demand, revenue, cost, and investment) are estimated from empirical network data. These estimates then form a starting solution for an iterative optimization routine with an improving search algorithm. Finer adjustments to the model system and parameters should be undertaken based on an objective function, which can minimize the difference between the observed data and the model ability to predict which links were improved and by how much. In brief, the model parameters form a space which can be searched systematically to find a best fit between actual and predicted link expansions and contractions. The transportation network data in the Twin Cities metro area have been collected between 1978 and the present in digital format, while data collection work on corresponding land use and economical information is ongoing. In the most recent (2000) Twin Cities transportation network, there are 7976 nodes and 20914 links. A bit more than 600 link expansions have taken place since 1978, which implies the Twin Cites transportation network is mature.

Though a rigorous calibration work can not proceed unless all required data are collected, simulating the model with the available Twin Cities network data can still provide valuable information regarding the modeling concept, structure and feasibility on a large-scale realistic network. The values of model parameters in these preliminary runs are based on either empirical estimation or our best understanding of the economies and diseconomies of scale in the network growth process, which are summarized in Table 7.1.1. The simulation experiments also provide opportunities to examine some qualitative properties of network dynamics.

7.1.5 Simulation Experiments and Results

Four experiments are set up with different initial conditions and restrictions on link contraction. It is assumed in all experiments that there are no exogenous changes in land use, economy and population. The fixed land use, economy and population in the model are based on 1998 Twin Cities Metropolitan Council data. Let us imagine we were planners in 1978 who are interested in network growth twenty years from "now" (year 1998). The 1978 network thus becomes the "existing" network. So, in essence, these four experiments set up scenarios in which "estimated" land use twenty years from "now" is applied to the "existing" network. Using the real 1978 network as the initial condition for the simulation model (Experiments 1 and 2) allows us to observe whether and how this real-world network achieves equilibrium. The real 1978 network already exhibits hierarchies in that a few important roads carry the bulk of traffic while most roads have relatively low speed and volume. In order to see how network hierarchies emerge in the growth path, the other initial condition is the 1978 network with a uniform capacity of 400 vehicle/hour, which is the capacity of the narrowest link in the 1978 network. The adoption of two initial scenarios can also reveal if starting conditions significantly affect the future growth of a transportation network. In the investment model, link contraction occurs as long as the collected revenue is insufficient to maintain a link at its present condition. However, in reality links usually do not shrink—once you build it you can not easily abandon it. The presence of this practical constraint is considered and applied to two of the four experiments (Experiments 2 and 4). Comparison of simulation results with and without the link contraction restriction shed some light on future refinement of the investment rules. In the simulation, if the network does not change in two consecutive years, the simulated network evolution process stops and an equilibrium is achieved. It is also possible that the network does not converge and changes constantly among two or more distinct states.

The four simulation experiments are carried out on a personal computer with a Pentium 4 processor at 1.7 GHz, slower than standard personal computers currently sold in market. On average, it takes about twenty minutes for each simulation iteration. The traffic assignment algorithm consumes a major portion of the running time. There are a lot of link expansion and contraction activities at the beginning of the evolution process. As we can see in Figure 7.1.3, thousands of links are expanded and contracted in the first several years following 1978. However, the network settles itself very quickly, and after about 25 years fewer than a hundred links still experience (relatively small) changes in capacity and free-flow speed. In order to achieve the strict equilibrium defined as a network with no more capacity changes, it is necessary to continue the iterations for many more years at any network as large as the one in the Twin Cities metropolitan area. But all significant changes occur during the first 20 years. It is clear the network dynamics model is approaching an equilibrium smoothly. It is probably not practical (with this level of computer reality) to execute the model until a strict equilibrium is achieved. A goal function can be set up to determine the stopping point of the simulation. For instance, further iterations are not considered if the average percentage change of link capacity becomes less than 0.001. The remainder presentation of the simulation results only focus on the network dynamics between 1978 and 1998 since most important changes take place during this period.

Road hierarchies clearly emerge in all four experiments (see Figure 7.1.4 capacity and carry low flows, while only a few roads are expanded to very high capacities and carry the bulk of traffic. Experiment 1 and 2 start from the 1978 network with real capacity and hence the hierarchical structure is already present at the initial condition because the construction work of most freeways in the Twin Cites had been completed by 1978. It is, therefore, not very surprising to see the predicted 1998 network hierarchies conform very well with the observed 1998 data. With accurate network data in the starting year (1978) and good exogenous forecasts of land use and economic growth in a future year (we actually observed the 1998 land use), the proposed network dynamics model with very simple decentralized cost, revenue and invest functions provides satisfactory forecasts of link hierarchies in the future year. It is interesting to see that hierarchies also emerge in Experiments 3 and 4 where the starting condition is a uniform capacity network. The predicted hierarchies in these two scenarios are actually very close to the observed ones for lower-level roads. The results from Experiments 3 and 4 also suggest that if planners in the Twin Cities could design a brand-new network to serve the existing travel demand and replace the existing network, they would build many fewer roads with very high capacities, as seen on the right side of the two graphs. This finding may be somehow not very meaningful due to the big “if”. How the network arranges itself in a hierarchical pattern from a uniform status is a really interesting question. To answer that question, the growth path of the Twin Cities network in Experiment 4 is presented in consecutive maps where changes in road capacity are shown with lines with different weights (Experiment 3 gives almost the same results and is therefore not shown).

For those who are not familiar with the Twin Cities metropolitan area, a brief description of the features of the region may be helpful before the maps in Figure 7.1.5 are examined. Two traditional central business districts, downtown Minneapolis and downtown Saint Paul, are approximately ten kilometers from each other. The Minnesota River meets The Mississippi River right in the city. At the confluence point of the two rivers is the region’s international airport. A new suburban business area, downtown Bloomington, also emerges near the airport. The three downtowns, as well as the rivers are shown in the base year network, the first map in Figure 7.1.5. After four years, the model predicts that some roads are expanded. The location of these expansions tells us much about how road hierarchies emerge even from a uniform network. Natural barriers, such as rivers in this case, are sources of unbalanced road construction. It is clear that bridges are able to attract more flow than other roads in the network and hence get expanded first. Network effects then drive more flow to the roads emanating from bridges, for instance the roads along riverbanks. If one carefully examines the roads surrounding the airport, the circle just west of the river conflux, it is evident these roads are also able to generate more revenues than an average road and are expanded early in the evolution process. The rule of the airport here is much like some natural barriers such as mountains, because they all direct more flow to bypasses. The second source of hierarchy comes from activity centers. The three downtowns with high density of jobs and other activities are the areas with intense road expansions in the years following 1978. Finally, the fact that all major road expansions between 1978 and 1982 take place in the central area of the region suggest that boundary effects also contribute to the formation of road hierarchies. Though we live on a round globe, even the largest metropolitan area today is still better modeled as a planar surface. Travel demand on a limited plane is not uniform. Most trips originating from the edges of the city are inward trips and destined for activity centers located relatively closer to the geographical center of the region, while trips emanating from areas in the middle of the city are distributed along all possible directions. The asymmetry in demand patterns is the third source of road hierarchies identifiable from the second map. Again, network effects will help propagate the hierarchies created by those three sources throughout the whole network over the years. Twenty years later, road hierarchies can be found virtually everywhere in the network (see the third map).

Congestion is undesirable in a network and attracts a lot of attention in network analysis. In Figure 7.1.6, volume capacity ratios (VC ratios) of all roads in the network after twenty years of evolution are plotted in a histogram. The observed 1998 data suggest that most roads carry flows well below their capacity and a few roads operate at VC ratios near or slightly higher than one. Practically, over a long period of time, no road can carry flows more than its capacity. The presence of VC ratios larger than one in the model is the result of inadequate description of road travel delays and scheduling adjustments in the traditional four-step travel-forecasting model. Experiments 1 and 3 allow road degeneration, the results from which show a narrow range of VC ratios, suggesting a more uniform distribution of congestion in the network compared to the observed data. Note that the model does not say that at equilibrium a uniform distribution of VC ratios will be achieved. Roads have similar VC ratios in Experiments 1 and 3 but not the same. Experiments 2 and 4, with a constraint on road contraction, obviously predict congestion much better than their counterparts without the constraint. Once a road is expanded but demand later does not justify the capacity after the expansion, the road is still going to be maintained and capacity reduction is less likely to happen. Furthermore, in the real world capacity expansions are discrete (1 lane, 2 lanes), while here they are modeled as continuous. Therefore, a constraint on road degeneration in the model should make it more realistic. The spike near VC ratio of one is still present in Experiments 2 and 4. This is because the same revenue and cost functions are applied to all roads. In reality, it may be more expensive to expand some roads than others and hence different levels of congestion are observed. This suggests that cost and revenue functions in the model should be adjusted according to local conditions.

The need for differentiated cost functions is identified again from the comparison between the predicted road expansions from 1978 to 1998 and the ones that actually happened. In Figure 7.1.7 the prediction results from Experiments 1 and 2 are compared to the observed data. Although, the model successfully predicts construction on several freeway segments, it forecasts more expansions on roads already having high capacities (freeways) while fewer expansions on arterials than reality. Either the expansion costs of arterial roads are overestimated or the costs of freeways are underestimated in the model. At this point, we are not arguing that the model predicts what should have been done. It must be able to describe reality first before it can be used as a normative tool.

Finally, the impacts of starting conditions and constraints on the predicted network dynamics are examined in Figure 7.1.8. Clearly, they do matter. By comparing the four graphs vertically, i.e. Experiment 1 against 3, and 2 against 4, we find that different initial networks results in quite different networks at equilibrium. A horizontal comparison of the graphs reveals the influence of the presence of the restriction on road degeneration.

7.1.6 Conclusions

A transportation network is a very complex system that consists of a full spectrum of various sub-systems, the properties and behaviors of which are already hard to forecast. Efforts put into travel demand forecasting, network design problems, and revenue policies by numerous researchers are evidence of such difficulties. Predicting the growth of transportation networks is difficult because it requires us to consider almost all sub-processes involved in network dynamics. Understanding the true relationships between supply and demand in transportation networks is the crucial task in theoretical developments of network dynamics models. The difficulty also comes from practical issues, such as available data for model calibration and validation. Socio-economical, demographical, land use and transportation network data many years ago in an urban area must be collected and coded consistently over time. Several unresolved issues further complicate the problem and the foremost one—Is network growth simply designed by our planners or it can be indeed explained by underlying natural and market forces? In light of this debate, we would like to view this paper as a proof of concept that some important system properties, such as road hierarchies and self-organization in transportation networks, can be predicted through a microscopic evolutionary process, a demonstration that such a microscopic agent-based model of network dynamics can be feasibly applied to large-scale realistic transportation networks, and an enquiry into how this concept can be realized and produce useful modeling tools for planners. Growth of economy, population, and cities has been intensively studied and knowledge accumulated from such studies has greatly aided planners. Traditionally, transportation networks have been assumed to be static or predetermined in analysis of urban areas. A model of transportation network dynamics can reveal more completely the impacts of today's planning decisions in the future.

The present paper explores only the rise and fall of existing roads (maybe the rise and rise of existing roads given the preliminary simulation results), leaving the questions of how new roads are built and new nodes are created in transportation networks to be answered by future studies. The process of road development and degeneration at the microscopic level is analyzed and an agent-based simulation structure seems to be appropriate for modeling that process. In order to better describe reality, a systematic way to adjust cost and revenue functions based on area-specific factors such as type of roads, land value, and public acceptance should be considered. Calibration of the network dynamics model is still part of an on-going study.

7.1.7 Figures and Tables

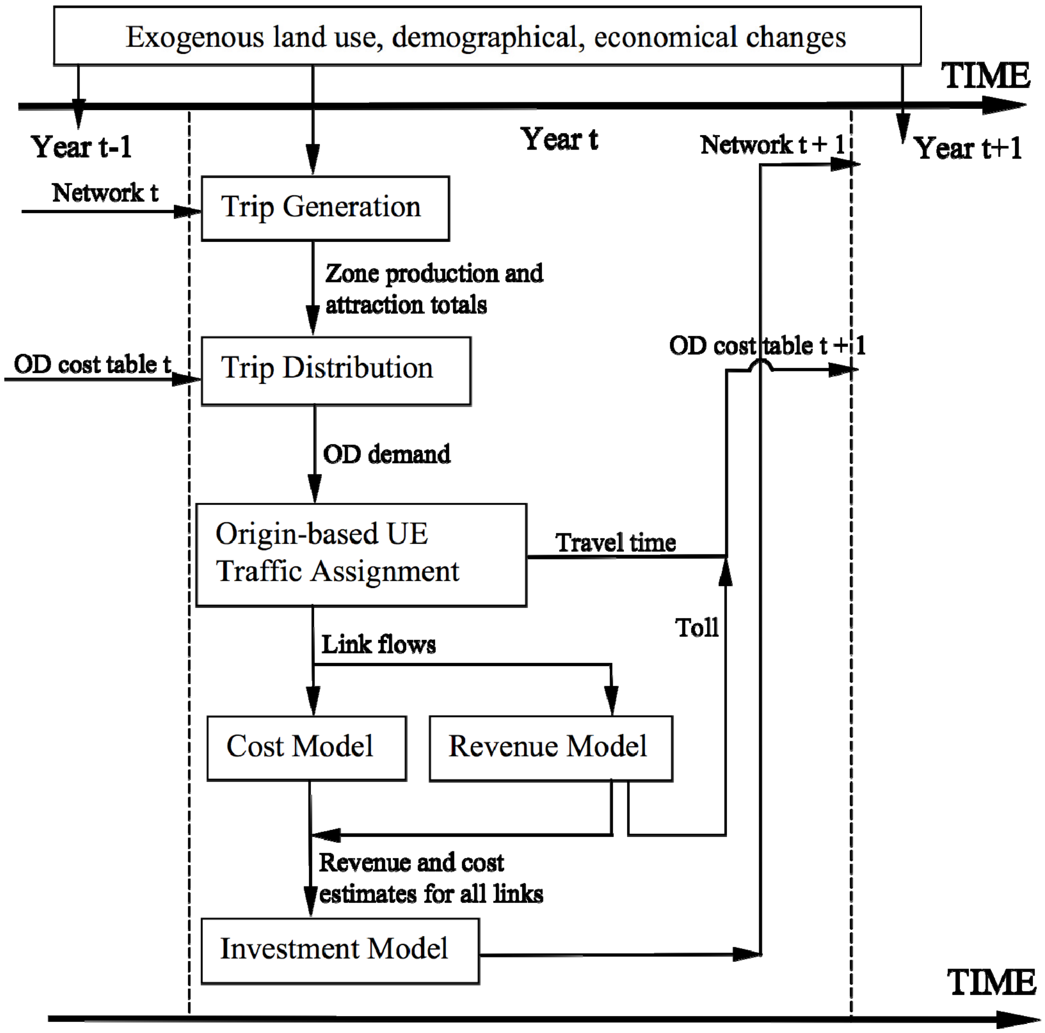


Figure 7.1.1 Flowchart of the transportation network dynamics mode

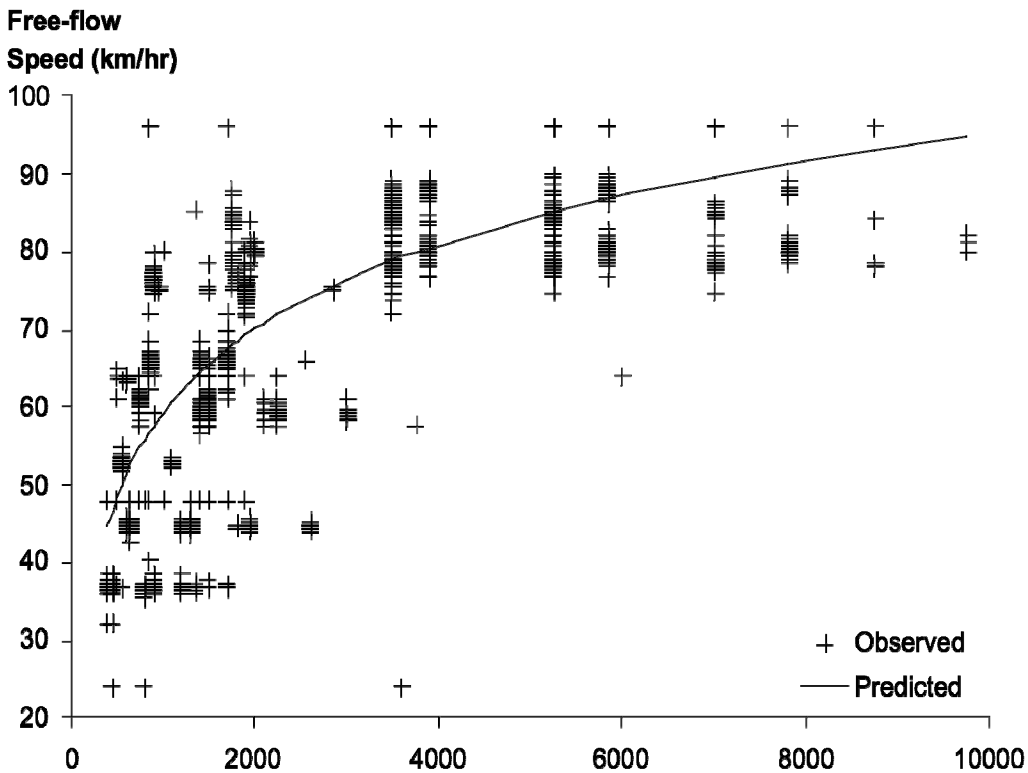


Figure 7.1.2 Link capacity and free-flow speed relationship: observed vs. predicted

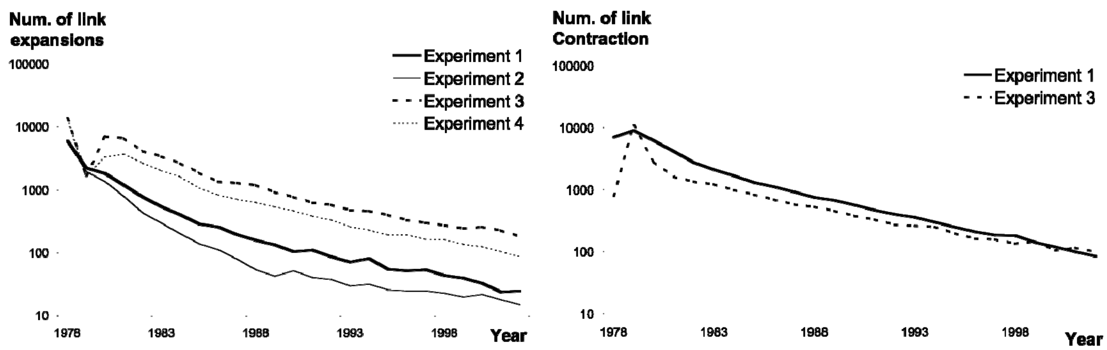


Figure 7.1.3 Convergence properties of the network dynamics model

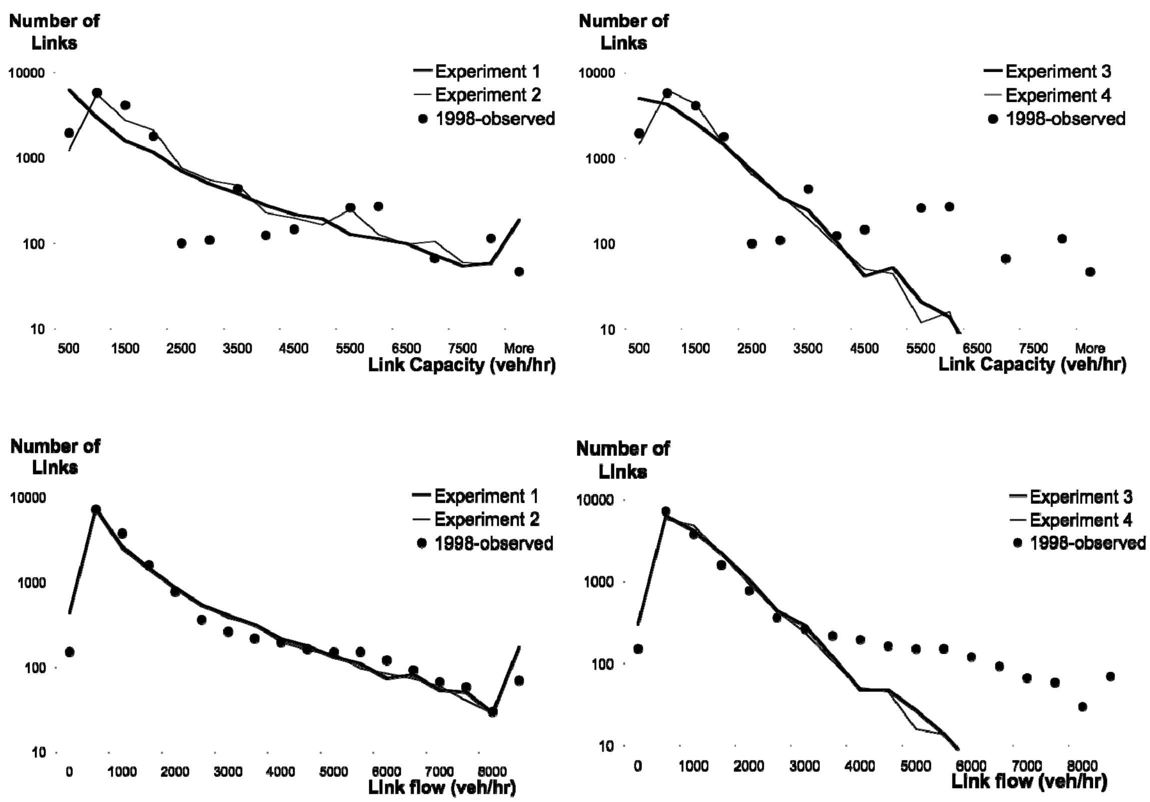


Figure 7.1.4 Link hierarchies after 20 years

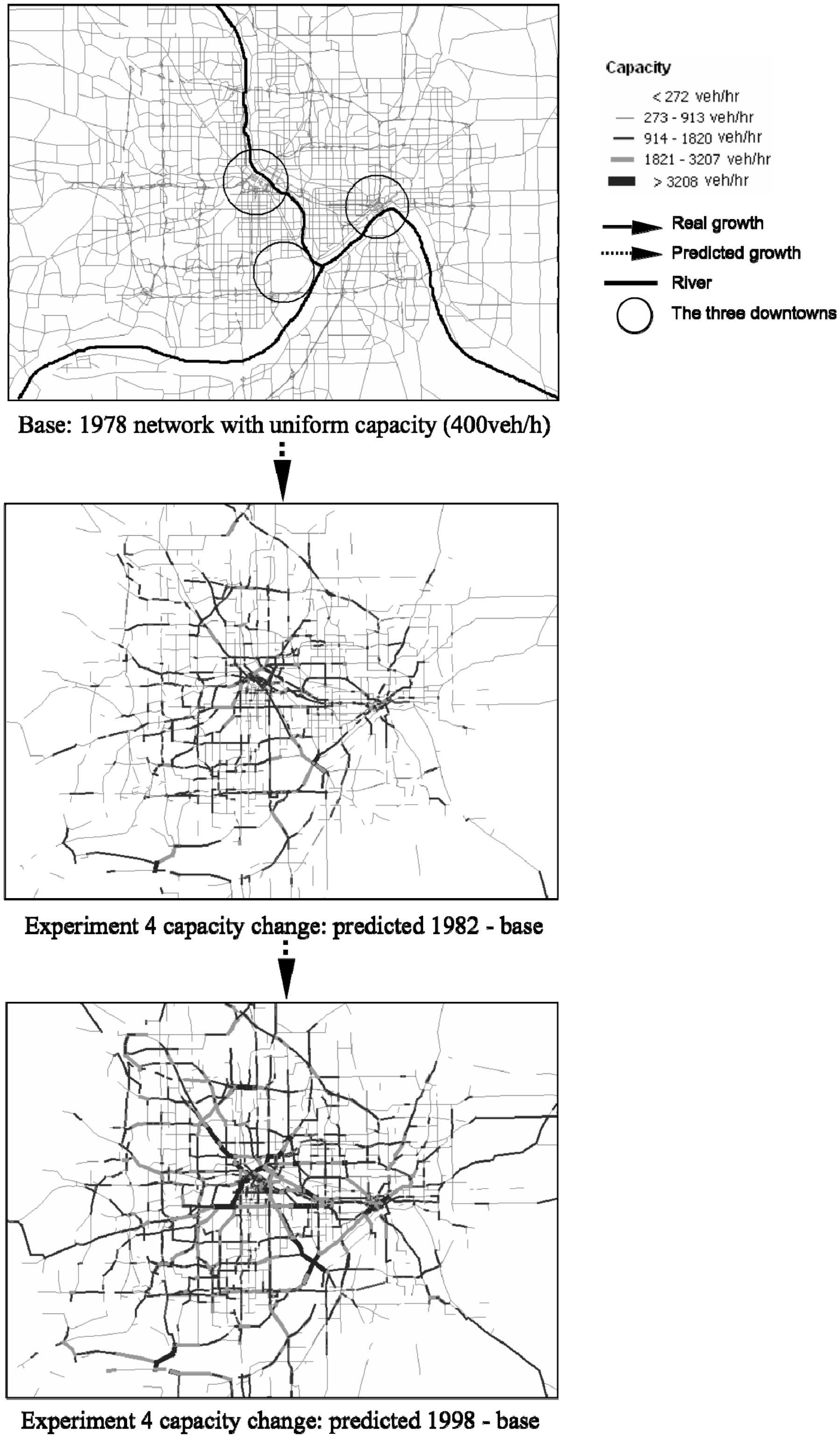


Figure 7.1.5 Emergence of hierarchies in experiment 4

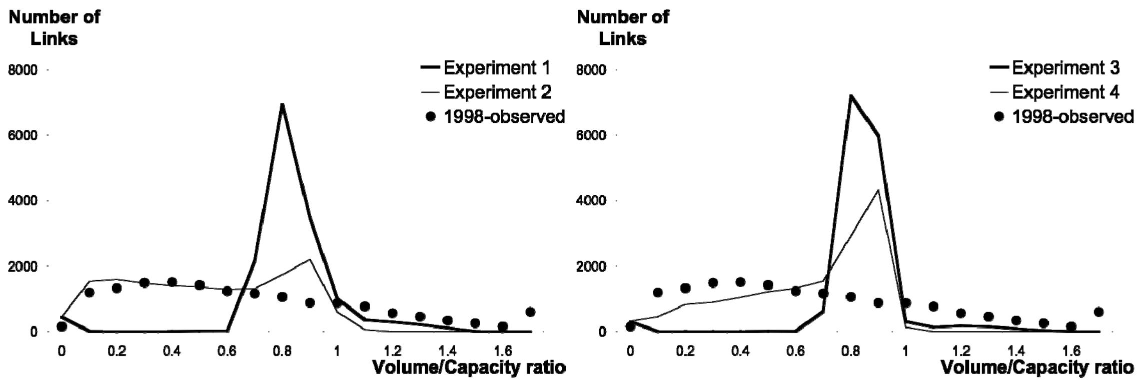


Figure 7.1.6 Network congestion after 20 years

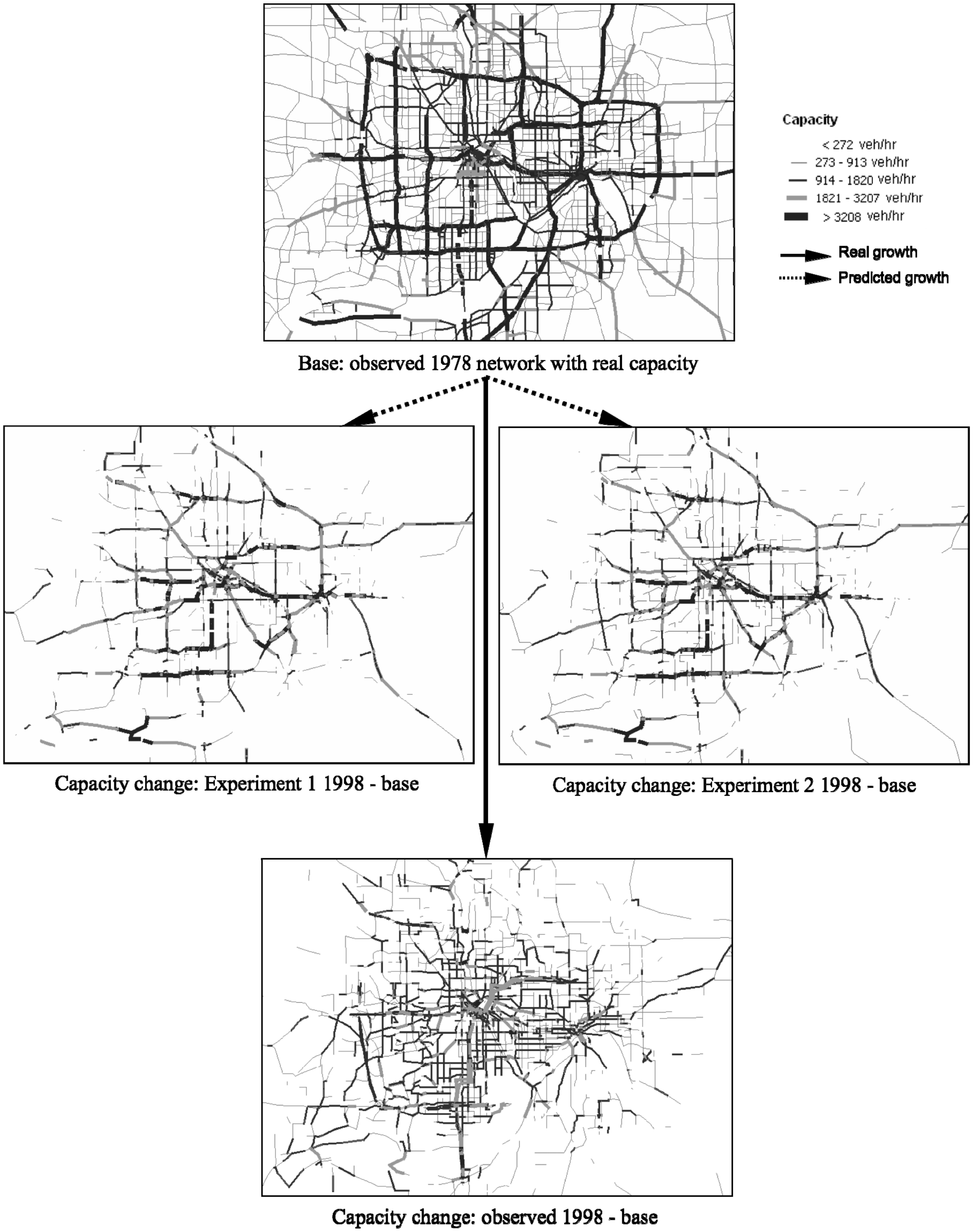
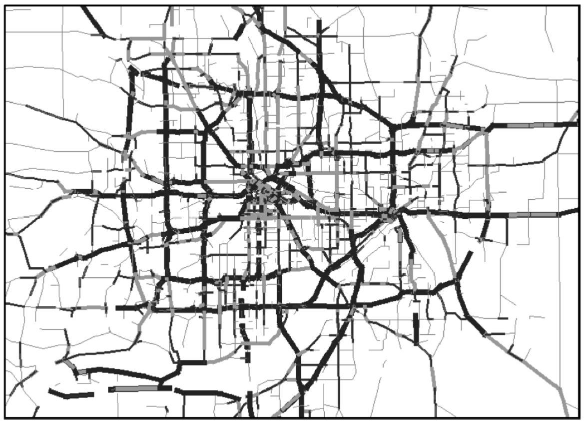
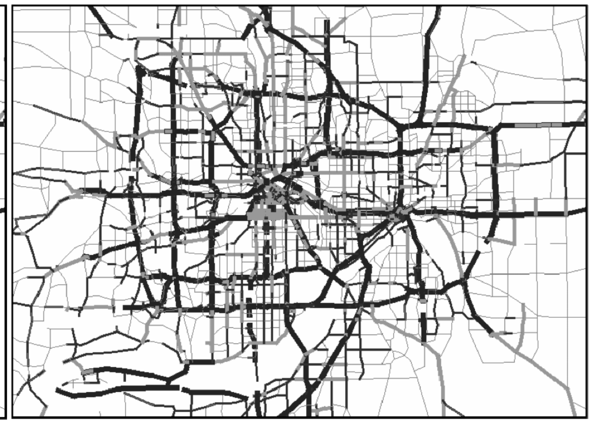


Figure 7.1.7 Experiments 1 and 2 vs. observed network growth after 20 years

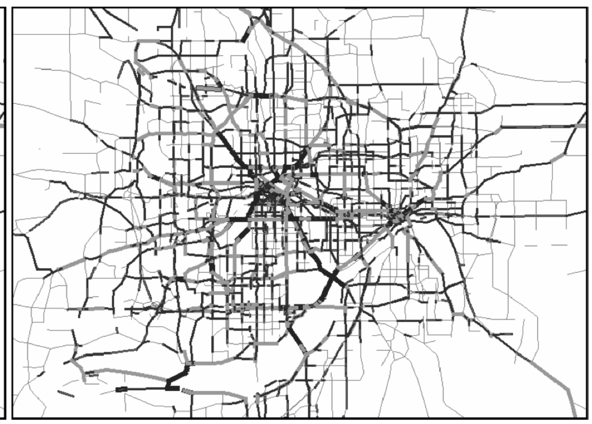


Experiment 1: predicted 1998 network



Experiment 2: predicted 1998 network

Experiment 3: predicted 1998 network



Experiment 4: predicted 1998 network

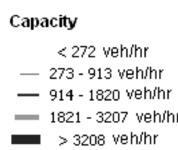


Figure 7.1.8 The impacts of starting conditions and constraints on network growth

Table 7.1.1 Coefficients used in the experimental runs of the network dynamics model

Parameter	Description	Value	Source
γ	value of travel time constant ($/hr)	10	Empirical findings
α_1, α_2	coefficients in the BPR function	0.15, 4	BPR
β	coefficient in the gravity model	0.1	Empirical findings
$\rho_1 \cdot \omega$	Combined scale coefficient in revenue model (dollar·hr$^{\rho 3}$/km$^{\rho 2+\rho 3}$)	1	Scale parameter
ρ_2	Power term of length in revenue model	1	CRS of link length
ρ_3	Power term of speed in revenue model	0.75	DRS of level of service
θ_1	Scale coefficient in cost model (dollar·hr$^{\theta 3}$/km$^{\theta 2}$)	20	Scale parameter
θ_2	Power term of length in cost model	1	CRS of link length
θ_3	Power term of capacity in cost model	1.25	IRS of capacity
μ_1, μ_2	coefficient in the speed-capacity regression model	-30.6, 9.8	Empirical estimate based on Twin Cities data
λ	capacity change coefficient	0.75	DRS in link expansion

CRS, DRS and IRS: constant, decreasing, and increasing returns to scale

Table 7.1.2 Four simulation experiments

Initial condition / Allow for link contraction?	Yes	No
1978 Twin Cities network with real 1978 capacity	Experiment 1	Experiment 2
1978 network with uniform capacity (400veh/h)	Experiment 3	Experiment 4

7.1.8 References

1. Aghion, P. and P. Howitt (1998) *Endogenous growth theory*. Cambridge MA, MIT Press.
2. Aschauer, D (1989) Is Public Expenditure Productive? *Journal of Monetary Economics* Mar 23a(2): 177-200.
3. Bar-Gera, H. and D. Boyce (2003) Origin-based algorithm for combined travel forecasting models. *Transportation Research* 37B(5): 405-22
4. Barker, T.C. and M. Robbins. (1975) *A History of London Transport, Volume 1 and 2*. Allen and Unwin, London
5. Batty, M. and Longley, P. (1985) *The Fractal Simulation Of Urban Structure* / by Cardiff, Wales: Department of Town Planning, University of Wales Institute of Science and Technology, Papers in planning research; 92.
6. Boarnet, M. (1997) Infrastructure Services and the Productivity of Public Capital: The Case of Streets and Highways. *National Tax Journal* 50(1): 39-57, March, 1997
7. BTS (Bureau of Transportation Statistics) (2002) *Transportation Indicators* http://www.bts.gov/transtu/indicators/Transportation_System_Extent_and_Use.html (Accessed June 12, 2002)
8. Button (1998) Infrastructure investment, endogenous growth and economic convergence. *Annals of Regional Science* 32(1): 145-62
9. Caruthers, J.I., and Ulfarsson, G.F., (2001): Public Service Expenditures: The Influence of Density and Other Characteristics of Urban Development. *The Pacific Regional Science Conference Organization.* Conference in Portland OR July 2001
10. Christaller, W. (1966) *Central places in southern Germany*. Englewood Cliffs, New Jersey Prentice Hall (Translated by Carlisle W. Baskin)
11. Fulton, L.M., Noland, R.B., Meszler, D.J. and J.V. Thomas (2000) A statistical analysis of the induced travel effects in the US Mid-Atlantic region. *Journal of Transportation and Statistics* 3(1): 1-14
12. Gramlich, E. (1994) Infrastructure investment: a review essay. *Journal of Economic Literature* 32: 1176-1196
13. Grübler, A. (1990) The rise and fall of infrastructures: dynamics of evolution and technological change in transport. Heidelberg: Physica-Verlag
14. Haynes, K.E. and A.S. Fotheringham (1984) *Gravity and Spatial Interaction Models*. Beverly Hills, CA: Sage Publications
15. Hutchinson, B. G. (1974) Principles of Urban Transportation Systems Planning. New York: McGraw-Hill
16. Karamalaputi, R. and D. Levinson (2003) Induced supply: a microscopic analysis of network growth. Forthcoming *Journal of Transportation Economics and Policy*
17. Krugman, P. (1996) *The Self-Organizing Economy*. New York: Blackwell

18. Levinson, D. and A. Kumar (1996) An evolutionary transportation planning model. *Transportation Research Record* 1493: 64-73

19. Miyagi, T. (1998) A spatial computable general equilibrium approach for measuring multiregional impacts of large scale transportation projects. *Network Infrastructure and the Urban Environment*, Springer, 224-44

20. Miyao, T. (1981) *Dynamic analysis of the urban economy*. Academic Press, New York

21. Nadiri, M. Ishaq and T. Mamuneas (1996) Contribution of highway capital to industry and national productivity growth. Federal Highway Adminstration

22. Nagel, K. and A. Schreckenberg (1992) A cellular automata model for freeway traffic. *Journal de Physique* 2: 2221–29

23. New Jersey Office of State Planning (1996) *Projecting municipal road costs under various growth scenarios*. Document Number 109

24. Newman, M.E.J. (2001) The structure and function of networks. *Computer Physics and Communications* 147: 40-45

25. Noland, R.B. (1998) Relationship between highway capacity and induced vehicle travel. *Transportation Research Board 78th Annual Meeting Preprint CD-ROM*, Transportation Research Board, National Research Council, Washington DC

26. Parthasarathi, P., Levinson, D., and R. Karamalaputi (2002) Induced demand: a microscopic perspetive. Forthcoming *Urban Studies*

27. Peat F.D. (2002) From certainty to uncertainty: the story of science and ideas in the twentieth century. National Academy Press: Washington DC p.xiii

28. Schadschneider, A. and M. Schreckenberg (1993) Cellular automata models and traffic flow *Journal of Physics A: Mathematical and General* 26: L679—83

29. Schelling, T. C. (1969) Models of segregation. *American Economic Review*59(2): 488–93
Sheffi, Y. (1985) *Urban transportation networks*. Prentice-Hall, Englewood Cliffs, NJ

30. Smock, R.J. (1962) An iterative assignment approach to capacity restraint on arterial networks. *Highway Research Board Bulletin* 156: 1–13

31. Strathman, J.G., Dueker, K.J., Sanchez, T., Zhang, J., and A.E. Riis (2000) Analysis of induced travel in the 1995 NPTS. Center for Urban Studies, Portland State University

32. Taffe, E.J., Morrill, R.L., and P.R. Gould (1963) Transport expansion in underdeveloped countries: a comparative analysis. *Geographical Review* 53 (4): 503–29

33. Von Neumann, J. (1966) *Theory of Self-Reproducing Automata*. Edited by A. W. Burks. University of Illinois Press

34. Waddell (2001) UrbanSim: modeling urban development for land use, transportation, and environment planing. Paper presented to the 17th Pacific Conference of the Regional Science Association, June 30-July4, 2001

35. Wardrop, J.G. (1952) Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, Part II, 1(36): 325-62

36. Wolfram, S. (1994) *Cellular Automata and Complexity*. MA: Addison-Wesley.

37. Wolfram, S. (2002) *A New Kind of Science*. Champaign, IL: Wolfram Media.

38. Yamins D., Rasmussen, S. and D. Fogel (2003) Growing urban roads. *Networks and Spatial Economics* 3: 69–85

39. Yerra, B. and D. Levinson (2002) The emergence of hierarchy in Transportation networks. Under review *Annals of Regional Science*.

40. Zhang, L. and D. Levinson (2003) An agent-based travel demand model. Submitted to the 83rd Transportation Research Board Annual Metting, Washington, DC

7.2 The Emergence of Hierarchy in Transportation Network

Bhanu M. Yerra
Thurston Regional Planning Council
2424 Heritage CT SW Suite A
Olympia, WA 98502, USA
E-mail: yerrab@trpc.org

David M. Levinson
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, Minnesota, 55455, USA
E-mail: levin031@umn.edu

7.2.1 Introduction

Transportation networks possess system properties like the hierarchy of roads, accessibility, topology, geometrical features (or network properties like the ratio of nodes to links), congestion, and so on. These properties can be assumed, or we can try to determine what underlying factors lead to them. This research aims to see if, without pre-specifying a hierarchy of roads, it can be seen as an emergent property, rather than a designed one.

For a given network topology and spatial land use distribution, trips from each network node to other nodes are calculated and these trips are placed on the network. Traffic on each link pays a toll to that link and this collected revenue is invested for that link's own development, thus changing the properties of links. Using these new link properties, which change the path of least cost travel between any given origin and destination, the whole process is repeated until it reaches equilibrium, or it is clear that it won't reach an equilibrium.

The results show that even when land use distribution and link length are uniformly maintained, hierarchies emerge. Moreover, these hierarchies approximately replicate the rank order rule of the hierarchies in real world networks. The results also show that roads are emergent properties of networks.

The next section presents analogous theories in regional science and economics that explain formation of hierarchies, and develops the hypothesis of this research. Section 7.3.3 describes the network dynamics model. Experiments are conducted and the results are presented in Section 7.3.4. The paper ends with conclusions and suggestions for future research.

7.2.2 Background

Roads (or road links) can be classified into hierarchies according to the flow of traffic they carry. Freeways, which are at the extreme end in the hierarchical classification, carry maximum traffic while their span is relatively small when compared to the total center line miles of roads in a given geographical area. The Interstate Highway System comprises one percent of all highway miles, but carries one fourth of the U.S. vehicle miles of travel (USDOT 2000). Conversely local roads, which comprise a larger share of road length, carry little traffic, mostly from nearby users.

There are mathematical models in fields like regional science, and economics that intend to explain the formation of hierarchies of their respective subjects. Regional science explores hierarchies of places or hierarchies in the system of cities (Beckman 1958; Beckman and McPherson 1961). Economics considers formation of income categories and firm sizes (Champernowne 1953; Roy 1950). But there are few in surface transportation, as it seems to be taken for granted that planners and engineers design the hierarchies of roads. Using the example of an underdeveloped country, Taaffe et al. (1963) explain the growth of the transportation network and emergence of hierarchies of roads, however, their study does not present a mathematical model. Garrison and Marble (1965) explained the order of rail network construction in Ireland by showing that nodes would connect to the nearest large neighbor. Yamins et al. (2003) model road growth co-evolving with urban settlements from an empty space with highly simplified travel demand and road supply mechanisms.

If hierarchies and roads really emerge, rather than being designed, then there is much to learn from the disciplines mentioned above. Any explanation of the formation of hierarchies of places cannot end without mentioning the central place theory developed from works of Christaller (1933) and Losch (1954). Christaller argues that central places emerge due to uneven distribution of facilities and transportation costs to reach these facilities. He then argues that these central places form a hierarchy. Lösch further suggested that these central places take a hexagonal shape. Krugman (1996) argues that these models don't qualify as economic models as they do not show emergence of these central places from ''any decentralized process.'' Recently, Fujita et al. (1999) formulated a mathematical model that explains the hierarchical formation of cities from a ''decentralized market process.''

Principles of complex systems have become popular among the fields that require modeling system properties from decentralized processes. There is no universally accepted definition of a complex system, however, it is generally agreed that it consists of ''a large number of components or 'agents', interacting in some way such that their collective behavior is not simple combination of their individual behavior'' (Newman 2001). Examples of complex systems include the economy—agents are competing firms; cities—places are agents; traffic—vehicles are agent; ecology—species are agents. Complex systems are known to exhibit properties like self-organization, emergence and chaos. Cellular Automata (CA) are commonly used to model complex systems (von Neumann 1966; Schelling 1969; Wolfram 1994, 2002). For the sake of this discussion complex systems are distinguished into two categories: One category has systems whose dynamics are modeled using networks or graphs and other has systems that are modeled using cellular automata. Literature in each of these categories is summarized below.

Watts et al. (1998) and Barabasi et al. (1999) proposed models that explain the properties and evolution of networks and graphs from localized link formations. Both works measure the formation of hierarchies of node-connectivity, defined for a node as number of adjacent nodes that are one link away, using power law distribution. In contrast, node-connectivity is not hierarchical in a typical road network; instead hierarchies are seen in link and node properties. It can be shown empirically that link growth and new link construction can be explained empirically by the position of the link (or potential link) in the network, its relationship with complementary (upstream and downstream) and competitive (parallel) links (Levinson and Karamalaputi 2003a,b). However, research into transportation networks can greatly benefit the Barabasi et al.'s (1999) concept of preferential attachment, which is akin to the concept of the "rich get richer", as the source of hierarchies.

Epstein and Axtell (1996) modeled social processes using agent based cellular automata. They modeled a "Sugarscape"—a landscape of resources—and placed interacting agents—rational decision makers—on it to simulate social processes. They showed the emergence of social wealth and age distributions (which can be considered as social hierarchies) from localized interactions of agents. They also modeled the dynamics of trade networks, credit networks and disease transmission networks using agent interactions (but did not model transportation networks). Nagel and Schreckenberg (1992) and Schadschneider and Schreckenberg (1993) have used cellular automata concepts to model traffic but did not model transportation network dynamics.

In the above-mentioned complex system studies, irrespective of how a system is modeled, simple local interaction rules are shown as sufficient to model the self-organizing properties of the system. In all those studies, the self-organizing behavior is measured using power law distributions. Measuring self-organization in emerging systems using power law distribution is well explained by Bak (1996).

7.2.3 Network Dynamics Model

In this research a network dynamics model is developed that brings together all the relevant transportation models to simulate network growth. An overview of transportation models and their interconnectivity is shown in Figure 7.2.1. Network structure, land use and demographic information, and user-defined events are exogenous inputs. In formulating the dynamics from a complex systems perspective, the network nodes, road links and travelers become agents. A schematic representation of the network structure and land use is show in Figure 7.2.2. The network structure and land use data are used by the travel demand model to calculate traffic flows on links. Traffic is assigned to links along the path with the least generalized cost of travel. A revenue model determines the toll the traffic must pay for using the road depending on speed, flow and length of the link. A cost model determines how much it costs to maintain the present level of service. Depending on revenue and cost, the investment model determines the link properties (speed) for the next time step. This process is repeated until an equilibrium of link speeds has been reached, or it is clear that no equilibrium can be reached. The output files produced are then exported to a visualization tool and the dynamics are viewed in a movie-like fashion. The details of mathematical formulations of these transportation models are presented in the following sub-sections.

Network Structure

The transportation network is represented as a directed graph that connects nodes with directional arcs (links). The directed graph is defined as: $G = \{N, A\}$ where N is a set of sequentially numbered nodes and A is a set of sequentially numbered directed arcs. An arc '*a*' connected from origin node m to destination node n is represented as $m \rightarrow n$. Let R denote a set of origin nodes and S denote a set of destination nodes. Note that in this network $R = S = N$, i.e. each node acts as both origin and destination. Let x_n and y_n represent the x and y coordinates of node $n \in N$ in Cartesian coordinate system. Let v^0_a be the initial speed on link $a \in A$. Let l_a be the length of the link a.

Travel Demand Model

Trip generation model. The geographical area under consideration is divided into cells (e.g., city blocks) and the land use is distributed among these cells. After reading the network, the trip generation model reads the size of the underlying land use layer in terms of number of cells and assigns the cells to the nearest network node. The land use layer is assumed to be static throughout the simulation. The land use layer is modeled as a square. Each cell is given two properties that represent the trips attracted (g_z) and trips generated (h_z) from that cell (z). Using the cell properties, trips produced at and trips attracted to a network node can be calculated by summing up the trips produced at and trips attracted to all cells that are nearest to that node. Let g_n and h_n be trips produced at and trips attracted to network node n.

Shortest path algorithm. Let t^i_a represent the generalized cost on link a for iteration i. This is calculated as the linear combination of link travel time and the toll (τ as shown in Eq. (1)), assuming the weights to dimensionally balance the equation are ones.

$$t^i_a = \frac{l_a}{v^i_a} + \tau(l_a)^{\rho_1} \quad \forall a \in A \tag{1}$$

where, ρ_1 is a coefficient.

From each node in the network, a least cost path to every other node is calculated using Dijkstra's Algorithm (Chachra et al. 1979). Let K_{rs} represents a set of arcs along the least cost path from origin *r* to destination s for iteration *i*. Let t^i represent the travel cost from origin *r* to destination *s* along the least cost path for iteration *i*. Then the relationship between K^i_{rs} and t^i_{rs} is:

$$t^i_{rs} = \sum_{a \in A} t^i_a \delta^i_{a,rs} \quad \forall r \in R, \quad \forall s \in S \tag{2}$$

where $\delta^i_{a,rs}$ is a dummy variable equal to 1 if arc *a* belongs to K^{r}_{si}, 0 otherwise.

Trip distribution. With the trip generation values and travel costs, a trip table (Origin Destination (OD) matrix) is computed using a gravity model (Hutchinson 1974; Haynes and Fotheringham 1984). Let q_{rs} be the number of trips from origin node *r* that are ending at destination node *s*. The gravity model indicates that q_{rs} is directly proportional to trips produced from origin node *r* (g_r) and trips attracted to destination node *s* (h_s) and is inversely proportional to the generalized cost of travel from origin node *r* to destination node *s* as shown below.

$$q_{rs} \propto \frac{g_r h_s}{d(r,s)} \quad \forall r \in R, \quad \forall s \in S \tag{3}$$

where, *d(r,s)* is called the friction factor. The friction factor function used in the gravity model is a negative exponential as shown below:

$$d(r,s) = e^{-\gamma \cdot t_{rs}} \tag{4}$$

where, γ is a coefficient that represents commuters disutility as costs of travel rise.

The resulting OD matrix is also incorporates the reverse trips from an origin to a destination to account for the evening traffic returning home. The resulting OD matrix q^*_{is} calculated as shown below.

$$q^*_{rs} = q_{rs} + q_{sr} \quad \forall r \in R \quad \forall s \in S \tag{5}$$

Traffic assignment. Using the OD matrix and shortest routes, traffic is assigned to each link. Flow (f_a) on each link is the sum of all the flows of paths between any origin and destination that passes through that link.

Revenue Model

This is a link-based model that calculates revenue for each link. Revenue is calculated by multiplying the toll and flow. Therefore the higher the flow on the link, the higher is the revenue. This model assumes that revenue is only collected by vehicle toll.

$$E_a = (\tau \cdot (l_a)^{\rho_1}) \cdot \psi \cdot f_a \quad \forall a \in A \qquad (7)$$

where ψ is a model parameter to balance the equation dimensionally and to convert the daily flow to annual.

Cost Model

This model calculates the cost to keep a link in its present usable condition depending on the flow, speed, and length.

$$C_a = \mu \cdot (l_a)^{\alpha_1} (f_a)^{\alpha_2} (v_a)^{\alpha_3} \quad \forall a \in A \qquad (8)$$

where,

C^i_a is the cost of maintaining the road at its present condition,

μ is the (annual) unit cost of maintenance for a link times the conversion factor to dimensionally balance the equation,

$\alpha_1; \alpha_2; \alpha_3$ are coefficients indicating economies or diseconomies of scale

Investment Model

Depending on the available revenue and maintenance costs this model changes the speed of every link at the end of each time step as shown in Eq. [9]. If the revenue generated by a link is insufficient to meet its maintenance requirements i.e. $E_a < C_a$, its speed drops. If the link has revenue remaining after maintenance, it invests that remaining amount in capital improvements, increasing its speed. This, along with a shortest path algorithm, embeds the ‘‘rich get richer’’ logic of link expansion. A major assumption in this model is that a link uses all the available revenue in a time step without saving for the next time step.

$$v_a^{i+1} = v_a^i (E_a^i / C_a^i)^{\beta} \quad \forall a \in A \qquad (9)$$

where, β is speed improvement coefficient.

With the new speed on the links the travel time changes and the whole process from the travel demand model is iterated to grow the transportation network until the network reaches equilibrium, or it is clear that it won’t.

7.2.4 Experiments and Results

The network dynamics model presented in the previous section provides a platform to conduct experiments on transportation networks to study their properties and dynamics. Several experiments are conducted and the results are presented in this section.

Base case

A base case is chosen and variations are made to this case to study how these variations affect the dynamics and the resulting hierarchies. The base case consists of an evenly spaced grid network in the form of a square with each link having the same initial speed. Each land use cell produces and attracts the same number of trips. The network structure and land use properties are chosen this way to eliminate network asymmetries as a confounding factor. Speeds on links running in the opposite direction between the same nodes are averaged in this case. Since Dijkstra’s algorithm does not list all possible shortest paths between any two nodes, symmetry conditions are externally applied. Table 6.2.1 shows the parameters used in base case and other experiments. The results for a 10 node by 10 node network are shown in Figure 7.2.3.a.

Figure 7.2.3.b shows the spatial distribution of speed for the network at equilibrium. The entire range of link speeds is divided into 4 equal intervals and these interval categories are used in drawing the figure with the line thickness and color representing the speed category. The spatial distribution of speeds depends on the parameters used in the model. Results clearly show that hierarchies and roads emerge from a localized link-based investment process. Despite controlling the land use and link length, we believe hierarchies are emerging because of the travel behavior induced by the presence of a boundary. Travel demand along the edges is inward while trips are evenly distributed along all possible directions in the middle of the area.

It can be argued that if edges are eliminated in the base case by carefully molding the geography into a torus while maintaining uniform land use and uniform spacing between links and adding new links to connect the edges, then any link from the resulting network becomes indistinguishable from any other link irrespective of its orientation along the meridian or longitudinal axis. This is the ideal case that produces no hierarchies or identifiable roads. Easing any conditions in this ‘‘ideal torus’’ case will result in hierarchies. In other words, the edges of the network are the force creating the hierarchy—the greater utility of central links for traffic increases their flow, and thus their speed. The investment model is also an important factor leading to the formation of hierarchies, which parallels Barabasi et al.’s (1999) concept of preferential attachment leading to formation of hierarchies.

Having examined the reasons for formation of hierarchies it is now time to reason for the converging solution. For a given static land use, the economies of scale in the cost function associated with traffic flow ($\alpha_2 < 1$) along with increasing cost for higher link speeds ($\alpha_3 > 0$) drives the system to an equilibrium. Diseconomies of scale in the cost function with respect to traffic flows result in an oscillating equilibrium.

Experiment A

Experiment A is similar to the base case except for randomly distributing the initial link speeds between 1 and 5. The network is evolved until an equilibrium is reached. Since random distribution of speeds makes the results stochastic, 20 such cases are performed and the average of the results are taken. A typical solution is shown in Figure 7.2.4.

Emergence of hierarchies and roads are clearly seen in this case. Random distribution of initial speeds produces non-symmetrically oriented roads with most of the faster links concentrated in the center of the geography. Belt or ring roads are common. Figure 4c shows the spatial distribution of traffic flow. Notice that there are fewer links that carry more traffic (thicker links in red) and many links that carry less traffic (thinner links in green), resembling a rank order rule. To investigate this rank order behavior of traffic flows of the network at equilibrium, results are compared with flow distribution of the Minneapolis and St. Paul (Twin Cities) network. Figure 7.2.5 compares the probability distribution of traffic flow for higher order networks for this case with probability distribution of 1998 Average Annual Daily Traffic (AADT) on links in the Twin Cities, where such counts are taken.

For this graph, the entire range of flow distribution is divided into eight intervals of equal size and the number of links falling in each interval is counted. These counts are divided by the total number of links in the network to get probabilities. These intervals are given a rank; the lower the flow range in the interval, the higher is the rank.

Notice in the graph as the size of the network increases, the probability distribution of the grid networks tend to get closer to the Twin Cities flow distribution. This graph concretely establishes that it is possible to replicate the global properties of real transportation networks by growing the network system using localized investment rules.

Experiment B

Experiment B is similar to the base case except for the treatment of initial speeds and land use characteristics. Land use characteristics of the cells are randomly distributed across the landscape between 10 and 15 trips. In this case, unlike the previous case, trips produced and trips attracted from a land use cell need not be the same. Link speeds in this model are dealt in two ways; firstly (experiment B1), initial speeds are assumed to be same for each link with magnitude 1, same as base case. Secondly (experiment B2), initial speeds are randomly distributed between 1 and 5 as it was done experiment A. Typical solutions for experiment B1 and B2 are shown in Figure 7.2.6.

Notice the similarity of this experiment with the previous case. We believe the differences in land use distribution (and initial speeds in case of experiment 2b) and the boundaries are responsible for the hierarchies in this case. Similar to the previous case, a rank order rule is observed in this experiment. The probability distribution of flows for this case are compared with 1998 AADT distribution of the Twin Cities in the Figure 7.2.7.

Notice that as the network size increases the behavior of traffic flow distribution is approaches the behavior of the Twin Cities, similar to the observation in experiment A.

7.2.5 Conclusions

This paper presents a transportation network dynamics model that includes localized revenue and investment models. This model can be considered as a ''bottom-up'' approach of modeling the emergence of hierarchies and roads, which are observed in several experiments. Therefore, it is possible to grow hierarchies and roads in a transportation network using decentralized investment rules. The network, no matter how random the initial speed distribution, when grown subject to localized investment rules produces order by self-organization.

If one looks at the complexity and bureaucracy involved in transportation infrastructure investment, one might conclude that it is impossible to model transportation network dynamics endogenously. But this research has shown that simple localized investment rules can be used to reflect the overall system properties. In fact, it is not the results that are most striking, but the simplicity of the investment rules in mimicking the system properties.

A new way of modeling and testing network dynamics is created by this research, which opens numerous opportunities of future research that can contribute immensely to our understanding of network dynamics. Research exploring different possibilities of investment rules that can not only reflect the global properties of networks, but also the network structure itself, can be considered. A realistic network can be used in these experiments instead of hypothetical grid networks and the coefficients can be estimated. More sophisticated travel demand models can be used. The cost model represented here can be made more realistic by introducing a construction cost function. A revenue sharing model—allowing links to share their revenue if they have excess—can be introduced and may produce much richer and realistic dynamics. Further examination of the rank-order rule is warranted. More research and a paper with application of this model to a real world network as a demonstration will encourage engineers and planners to adopt these kinds of models.

7.2.6 Figures and Tables

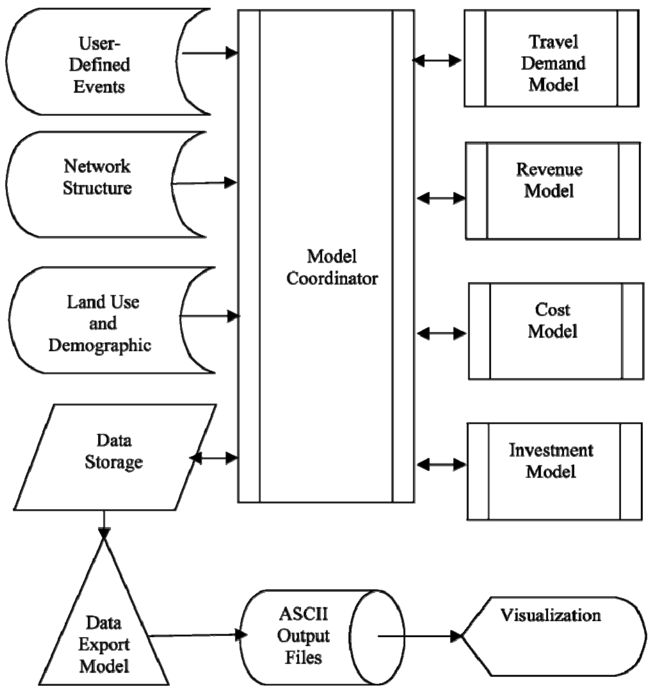


Figure 7.2.1 Overview of modeling process

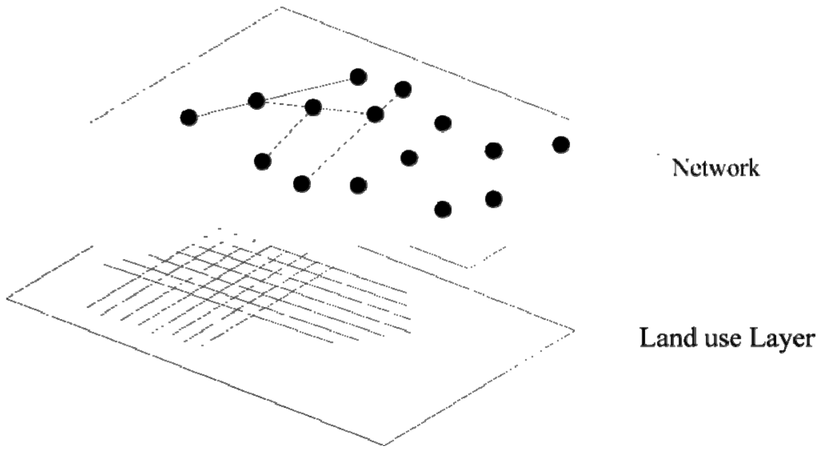


Figure 7.2.2 The network and land use layer

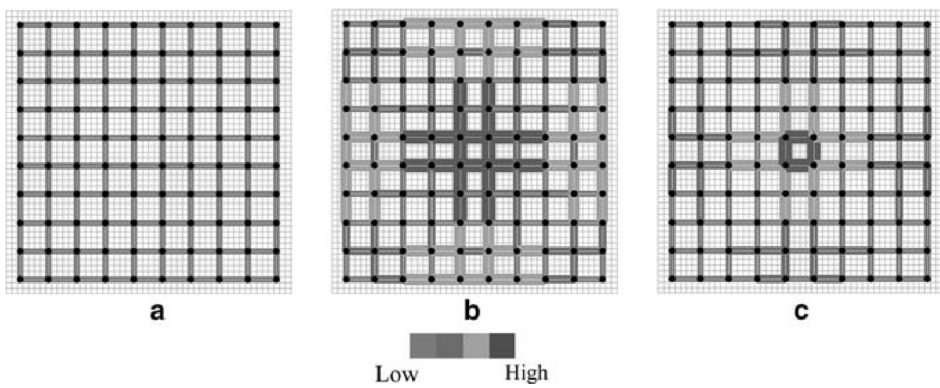


Figure 7.2.3 (a) Spatial distribution of speed for the initial network; (b) Spatial distribution of speed for the network at equilibrium reached after 8 iterations; (c) Spatial distribution of traffic flow for the network at equilibrium. The color and thickness of the link shows its relative speed or traffic flow.

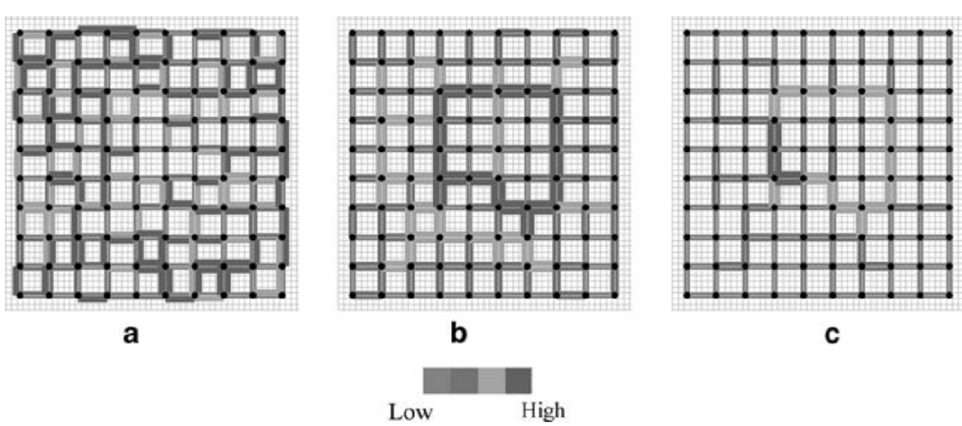


Figure 7.2.4 (a) Spatial distribution of initial speeds; (b) Spatial distribution of speeds for the network at equilibrium; (c) Spatial distribution of traffic flow for the network at equilibrium. The color and thickness of the link shows its relative speed or traffic flow.

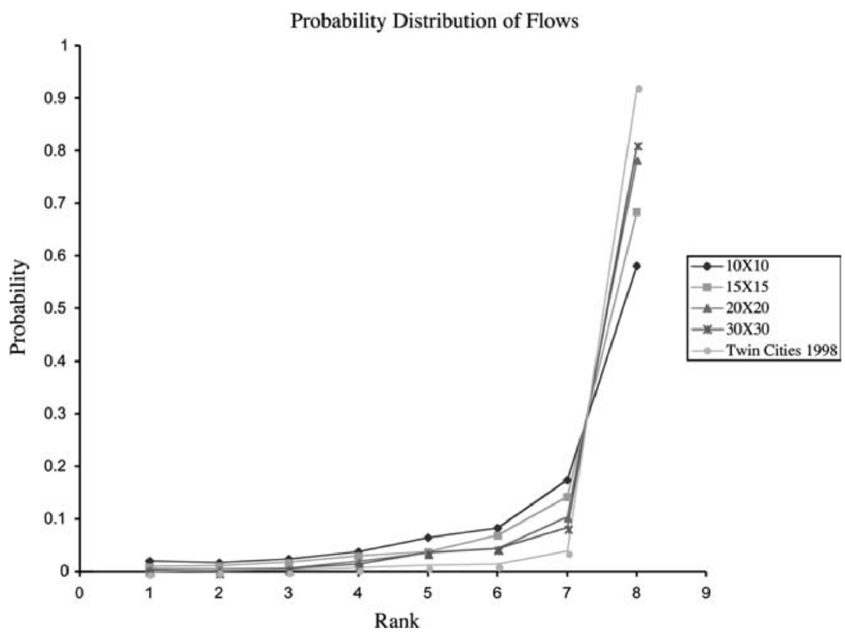


Figure 7.2.5 Comparison of traffic flow distribution of higher order networks with 1998 Twin Cities AADT distribution, Experiment A

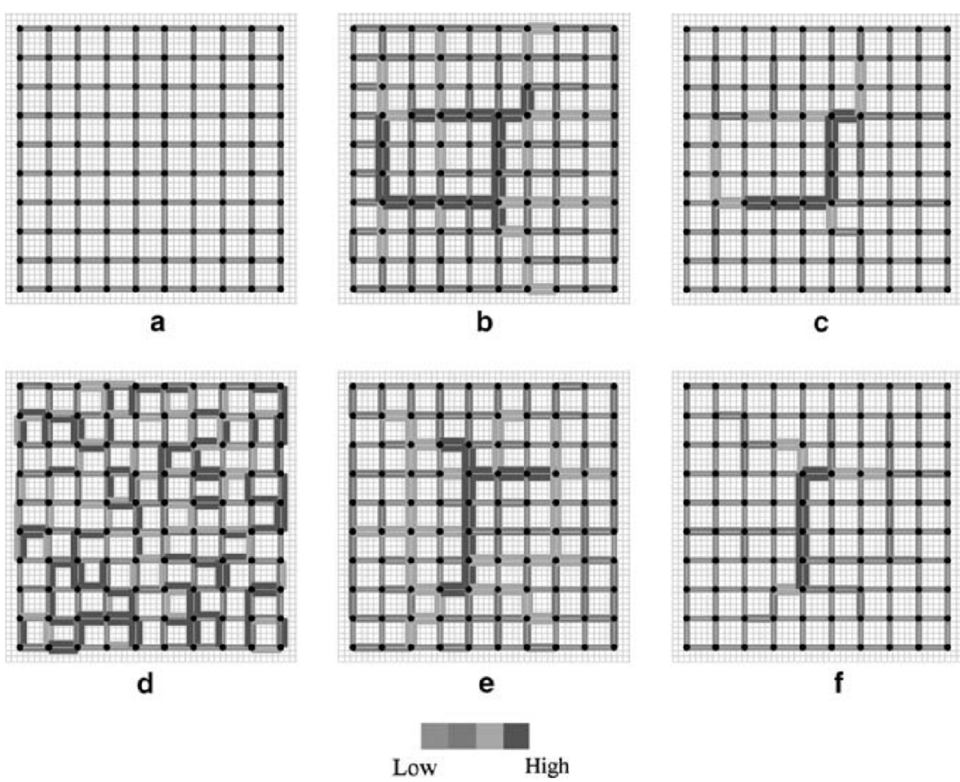


Figure 7.2.6 (a) and (d) Spatial distribution of initial speed for experiments B1 and B2 respectively; (b) and (e) Spatial distribution of speeds for the network after reaching equilibrium; (c) and (f) Spatial distribution of traffic flows for the network after reaching equilibrium. The color and thickness of the link shows its relative speed or flow.

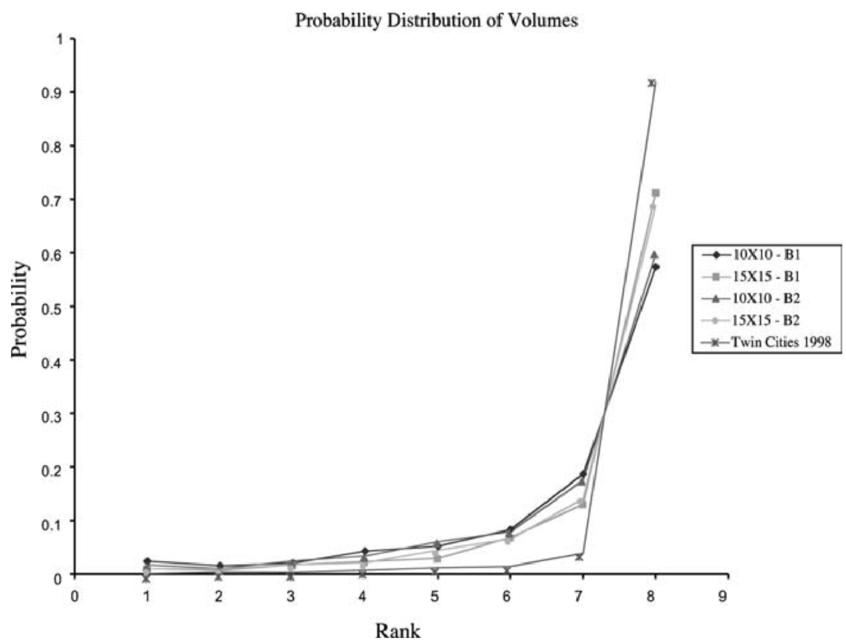


Figure 7.2.7 Comparison of traffic flow distribution of higher order networks with 1998 Twin Cities AADT distribution, Experiment B

Table 7.2.1 Model parameters and values used for experiments in paper

Variable	Description	Base case	Experiment A	Experiment B	
				Case-1	Case-2
v_a^0	Initial speed (integer)	1	1–5	1	1–5
g_z, h_z	Land use properties of cell z	10	10	10–15	10–15
γ	Coefficient in Eq. (4) trip distribution model	0.01	0.01	0.01	0.01
ρ_1	Length power in revenue model	1.0	1.0	1.0	1.0
τ	Tax rate in Eq. (7) revenue model	1.0	1.0	1.0	1.0
ψ	Revenue model parameter in Eq. (7)	365	365	365	365
μ	Unit cost in Eq. (8) cost model	365	365	365	365
α_1,	Length power in Eq (8) cost model	1.0	1.0	1.0	1.0
α_2	Flow power in Eq. (8) cost model	0.75	0.75	0.75	0.75
α_3	Speed power in Eq. (8) cost model	0.75	0.75	0.75	0.75
β	Coefficient in Eq. (9) investment model	1.0	1.0	1.0	1.0

7.2.7 References

1. Bak P (1996) How nature works: The science of self-organized criticality. Copernicus (Springer-Verlag), New York, NY
2. Barabasi A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286: 509–512
3. Beckmann MJ (1958) City hierarchies and the distribution of city sizes. *Economic Development and Cultural Change* 6: 243–248
4. Beckmann MJ, McPherson J (1970) City size distributions in a central place hierarchy: An alternative approach. *Journal of Regional Science* 10: 25–33
5. Chachra V, Ghare PM, Moore JM (1979) Applications of Graph Theory Algorithms. North Holland, New York
6. Champernowne DG (1953) A model of income distribution. *The Economic Journal* 63: 318–351
7. Christaller W (1933) Central places in Southern Germany. Fischer (English translation by C. W. Baskin, London: Prentice Hall, 1966) Jena, Germany
8. Epstein JM, Axtell R (1996) Growing artificial societies. MIT Press, Cambridge, MA
9. Fujita M, Krugman P, Venables AJ (1999) The spatial economy: cities, regions, and international trade. MIT Press, Cambridge, MA
10. Garrison WL, Marble DF (1965) A prolegomenon to the forecasting of transportation development. Office of Technical Services, United States Department of Commerce, United States Army Aviation Material Labs Technical Report
11. Haynes KE, Fotheringham AS (1984) Gravity and spatial interaction models. Sage Publications, Beverly Hills, CA
12. Hutchinson BG (1974) Principles of urban transportation systems planning. McGraw-Hill, New York
13. Krugman P (1996) The self-organizing economy. Blackwell, New York
14. Levinson D, Karamalaputi R (2003) Predicting the construction of new highway links. Journal of Transportation and Statistics 6(2/3): 81–89
15. Levinson D, Karamalaputi R (2003) Induced supply: a model of highway network expansion at the microscopic level. *Journal of Transport Economics and Policy* 37(3): 297–318
16. Lösch A (1954) The economics of locations. Yale University Press, New Haven
17. Nagel K, Schreckenberg A (1992) A cellular automata model for freeway traffic. *Journal de Physique* 2: 2221–2229
18. Newman MEJ (2002) The structure and function of networks. *Computer Physics Communications* 147: 40–45

19. Roy AD (1950) The distribution of earnings and individual output. *Economic Journal* 60: 489– 505

20. Schadschneider A, Schreckenberg M (1993) Cellular automata models and traffic flow. *Journal of Physics A: Mathematical and General* 26: 679–683

21. Schelling TC (1969) Models of segregation. American Economic Review, Papers and Proceedings 59(2): 488–493

22. Taaffe EJ, Morrill RL, Gould PR (1963) Transport expansion in underdeveloped countries: a comparative analysis. *Geographical Review* 53(4): 503–529

23. U.S. Department of Transportation, Bureau of Transportation Statistics (2000) The changing face of transportation. BTS00-007. Washington, DC Watts DJ, Strogatz SH (1998) Collective Dynamics of 'small-world' networks. Nature 393: 440– 442 Von Neumann J (1966) Theory of self-reproducing automata. In: Burks AW (ed). University of Illinois Press

24. Wolfram S (1994) Cellular automata and complexity. Addison-Wesley, Reading, MA

25. Wolfram S (2002) A new kind of science. Wolfram Media, Champaign, IL

26. Yamins D, Rasmussen S, Fogel D (2003) Growing urban roads. *Networks and Spatial Economics* 3: 69–85

7.3 Induced Supply

A Model of Highway Network Expansion at the Microscopic Level

David Levinson
Assistant Professor
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-625-6354 F: 612-626-7750
E-mail: levin031@tc.umn.edu

Ramachandra Karamalaputi
Graduate Student
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-625-6354 F: 612-626-7750
E-mail: levin031@tc.umn.edu

7.3.1 Introduction

Traffic demand is shaped by investments in new infrastructure and changes in public policy, while investment in highway network supply is itself induced by demand. Although highway agencies choose to expand small segments of the transport network, those investments are limited by decisions that have gone before; and perhaps more importantly, today's decisions constrain tomorrow's choices. This paper explicitly considers the growth of highway networks as an endogenous process, in contrast with current transport planning practice that strives to direct that growth exogenously. A 20-year database of network expansion is constructed to analyze the dynamics of network expansion decisions, which has been largely unstudied at the microscopic level.

Previous research on network expansion is limited. Grübler (1990) has looked at long-term technology diffusion issues, considering, for instance, the total length of roadway, vehicle kilometres travelled, or vehicles owned over the span of decades. Taaffe et al. (1963) study the economic, political and social forces behind infrastructure expansion in underdeveloped countries. Several studies have examined specific networks, for example the London Underground (Barker and Robbins, 1975), but no general theoretical framework has been given for network growth at the microscopic level.

Furthermore, increasing the capacity of a link in the network increases travel on that link due to re-routing and re-scheduling and also due to what is often called induced or latent demand (Noland 1999, Strathman et al., 2000). In contrast, the process of induced supply, how highway agencies respond to increasing travel demand, population, income, and demography has been largely ignored, but is crucial in our understanding of the decision process leading to infrastructure improvements. Transport infrastructure supply is inelastic in the short term but varies in the long term. Transport infrastructure and economic development are interrelated, particularly in under-developed and emerging regions.

The objective of this paper is to develop insight into the growth of transport networks. Specifically, to aid in planning and design, we want to know the investment rules governing agency decisions to expand transport networks. However, available (annual) budgets limit network growth. When an existing link is improved, we need to establish the conditions of single-lane versus double-lane expansion. We want to find whether improving one link will cause upstream and downstream links to have greater demand, and parallel links to have lower demand. We posit that the pressure to expand a link will decrease if we expand parallel links. The underlying question in this research is whether network changes can be predicted, and if so, to what extent. In a sense, this paper is about modeling the behavior of bureaucracies in response to the factors hypothesized to be the driving forces of network growth. The discrete nature of capacity expansion complicates the issue. This model can be used as a policy tool to predict how government transport agencies direct network growth as a function of projected future traffic, given demographic characteristics and budget.

The next section describes the underlying economic theory of network expansion. The following section consists of data used in this study. Some of the issues with regard to designating adjacent and parallel links in a transport network are dealt with. In the fourth section, a cost function is developed to estimate the cost of expanding a link given the year of construction, length of the section, number of lanes to be constructed, and hierarchical level of the road. A cost function is necessary to obtain a cost estimate for expanding links on which we do not have data. The fifth section describes the model used to predict network expansion, and poses the specific hypothesis. Results are presented in the sixth section while a final section summarizes and concludes.

7.3.2 Theory

The decision by transport agencies to increase the capacity of a link often comes in the wake of congested flow conditions on the link or as an attempt to divert the traffic from other competing routes. Alternatively, in undeveloped areas, capacity expansion anticipates the development of that area. However, expansion of links is constrained by the available annual budget for such purposes. Specifically, we want to test if capacity increases on the network depend on the capacity of the link in consideration, flow present and previous, and flows and capacities of connected and parallel links.

The traditional supply and demand curves for infrastructure supply are shown in Figure 7.3.1. The X-axis in the figure is new lane kms of capacity (C), the Y-axis shows the unit cost per lane km (e). Each of the above variables affects either the supply or the demand curve resulting in a new equilibrium. A higher annual budget (B) is agency income (allotted by the residents of the community the agency serves) that increases the willingness and ability to pay to expand or construct highways. A higher income results in a shift of the demand curve to the right. From the

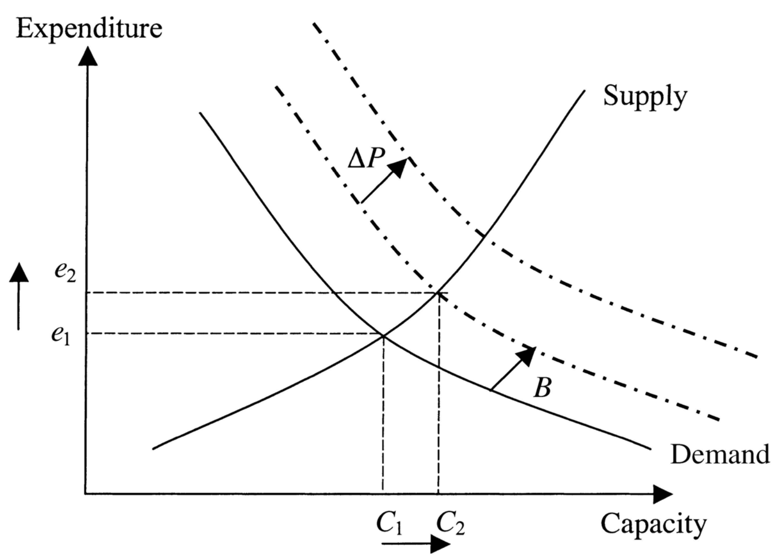
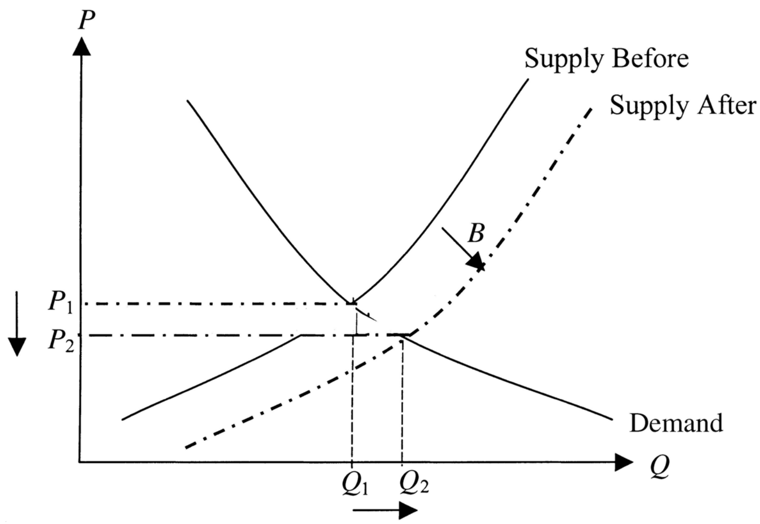


Figure 7.3.1 Infrastructure supply-demand curve

curve, it can be seen that higher cost of expansion per lane mile (*e*) decreases the willingness and ability of agencies to construct new infrastructure. Studies have shown that infrastructure growth rates in mature systems decline with time (Grübler, 1990), that is, the supply curve becomes more inelastic (more vertical) with time.

Due to diminishing marginal returns, the likelihood of widening a highway decreases with its capacity. There are diminishing returns because costs rise with the scarcity of land for expansion. Increasing road use over time due to population growth (ΔP) changes in travel demand is also

Figure 7.3.2 Induced demand and consumers' surplus



reflected by the outward shift of the demand curve. Due to the shift in the demand curve, a new equilibrium is reached with an increased supply of infrastructure (more capacity). Note that the new equilibrium has a higher equilibrium price (unit cost of infrastructure).

Changes in capacity of the network induce additional trips on that link due to re-routing and re-scheduling of the trips (Fulton, 2000; Noland, 1999). Although the presence of induced demand is now widely accepted, the exact relationship between a capacity increase and induced demand is not clear. Parthasarathi et al. (2003) has studied this relationship at the link level using the same dataset as we consider for this study.

The shaded area in Figure 7.3.2 is the consumer surplus resulting from the lower price and additional demand after an increase in infrastructure supply. Although consumers' surplus increases after construction, traffic is inconvenienced during construction. If the project takes a long time to be executed, the negative effects might overrun the consumers' surplus of future years. Duration of construction is then an important consideration in the benefit-cost analysis of the expansion. Since we do not have data on duration for unbuilt projects, length of the link is used as a surrogate. Regressions on available data showed length to be a good indicator of the duration of the project. In view of this, longer road segments are less likely to be expanded. Networks tend to grow more in the peripheries once they reach saturation levels near downtown areas. Land scarcity and heavy traffic in the downtown areas make it an inconvenient place to expand the network. Higher capacity is needed to cater to the traffic need, but again land acquisition problems in such areas act as a deterrent.

7.3.3 Data

To study network dynamics, data have been collected on construction of new links and on expansion of existing links spanning two decades (1978– 1998). Data were obtained from the following sources:

1. Network data from the Twin Cities Metropolitan Council.
2. Average Annual Daily Traffic (AADT) data on each link from the Minnesota Department of Transportation.
3. Link investment data was obtained from two sources: Twin Cities Transportation Improvement Program published by the Metropolitan Council. Hennepin County Capital Budget published by Hennepin County.
4. Population of Minor Civil Divisions (MCD) from the State Demography Center, Minnesota Planning.

Network data obtained from the Metropolitan Council gave a physical description of each link in the network in terms of number of lanes, length of the link, capacity of the link, type of highway, and its physical position. Each link is uniquely identified by its start node and end node. Each node is associated with a set of geometric coordinates that define the orientation of a link. The Twin Cities network has around 15,000 links of which 1,525 links are interstate highways, 2,362 links are trunk highways, and 4,394 are county highways. Of the county highway links, only those in Hennepin County are used for analysis, which reduces the number to 1,802 links, as investment data on other county highway links could not be obtained. Hennepin is the largest county of the seven in the Metro area and contains the city of Minneapolis. Remaining links are local roads and ramps to highways that are not considered for the analysis because investment data and AADT data could not be obtained for these links. Each type of road is analyzed separately because of the inherent differences in the utilization and financing of these roads. Figure 7.3.1 summarizes investment data, and Figure 7.3.2shows the number of links added in each hierarchy of the road during this period.

New construction projects follow different criteria and are not dealt with in this paper. Data in the four separate data sets were merged using ArcView GIS and through some custom computer programs. The database was then split by road type to form separate databases for interstate highways, trunk highways, and county highways.

Adjacent and Parallel Links in a Network

To input the surrounding conditions for each link in the network, we need to identify links connected to it and its most nearly parallel link. Each two-way link is divided into two one-way links. Adjacent links are divided into two categories: supplier links, and consumer links. Supplier links have traffic flow in the same direction as the link in consideration and are physically attached to the start node of the link under consideration. Consumer links are similar to supplier links but are attached to the end node of the link under consideration. A computer program was written to enumerate the adjacent and parallel links. Figure 7.3.3 shows supplier and consumer links of a link in a hypothetical network. For link 2–4: 1–2 and 3–2 are supplier links, and links 4–5 and 4–7 are consumer links.

Table 7.3.1 Summary of investment data

	No. of projects in TIPS*	Total cost of projects in TIPS ($1000's)	No. of projects in Hennepin County Budget*	Total cost of projects in County ($1000)
1979	11	229633	–	–
1980	7	80197	–	–
1981	5	132176	3	2780
1982	8	27120	1	6124
1983	2	38980	2	7760
1984	6	50711	3	13655
1985	3	214031	2	3677
1986	2	8538	4	13577
1987	0	0	1	2436
1988	0	0	3	10370
1989	1	55300	4	8338
1990	0	0	2	15312
1991	0	0	3	15443
1992	0	0	2	7158
1993	7	253700	3	38353
1994	1	8600	3	10380
1995	4	187500	5	16991
1996	0	0	7	35835
1997	0	0	8	35762
1998	0	0	4	22435
1999	4	268200	6	21647
2000	1	70000	4	18606

*Only projects costing more than $1 million or of more than 1 mile have been included.
Sources: *Local Transportation Improvement Program* and *Hennepin County Capital Budget.*
– Data unavailable.

Table 7.3.2 Number of links expanded or constructed by road type (1978–present)

	One lane expansion	Two lane expansion	New construction	Total
Interstate	104	43	35	182
TH	53	–	40	93
County Highway	86	–	3	89
Total	243	43	78	364

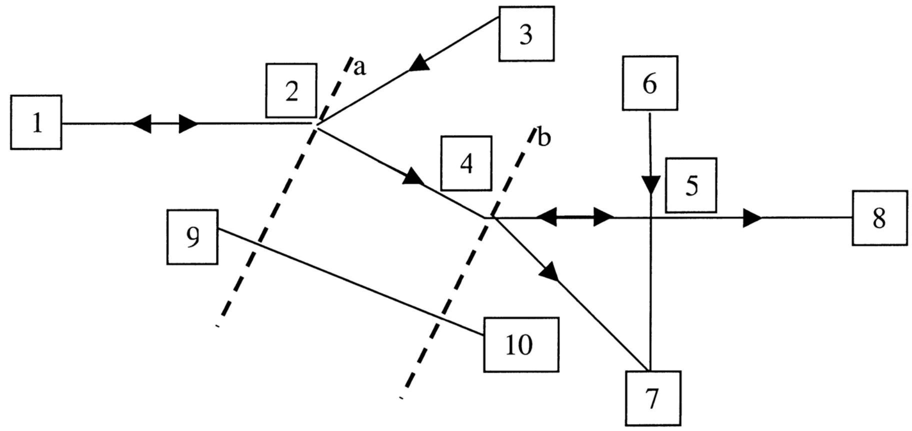


Figure 7.3.3 Adjacent and parallel links in a network

Parallel link description is a little more complicated. A parallel link can be thought of as the link that would bear the maximum brunt of diverted traffic if the link under consideration was closed. Note that a parallel link may or may not be in physical contact with the link in consideration. Merely looking at the mapped data and selecting a parallel link to a given link is a subjective choice. We need to define the attributes of parallel link to have a ''feel'' of the type of link to be identified as a parallel link. Crudely put, we are searching for a link that is in the proximity of the link L, approximately parallel to L in the literal sense, and of comparable length. Attributes have been defined based on this definition. The first attribute is how parallel the links are. The angular difference between the two links should be as small as possible.

The second attribute is the perpendicular distance from the mid-point of link L to the other link, divided by length of link L. The third attribute is the sum of distances between start nodes and end nodes of the two links being compared. The final attribute takes the ratio of lengths of the two links into consideration. Note that each of the attributes takes a range of values. Fuzzy logic has proven useful in this kind of case.

Fuzzy theory assumes a continuous truth-value rather than the deterministic Boolean values used conventionally. The sum composition method combined with appropriate weights has been found suitable for our purposes. In sum composition, the combined fuzzy output is obtained by computing the truth-value of each attribute and summing these values.

Here, we modified this method by weighing the truth-values of the attributes based on the importance of that attribute in relation to others.

The algorithm to find the parallel link needs some explanation. For a link L in consideration, we need to throw out links that are highly unlikely to be parallel. In no case should we identify a link in the continuous chain of links of which link L is a part as the parallel link. For instance, in Figure 7.3.3, link 5–8 should never be identified as parallel link to 2–4.

We need to remove all links of this type before evaluating attributes. One way to do this would be to drop perpendiculars from the nodes of the link L and check if the link we are comparing with has any point on it that falls between these perpendiculars. Further, all the links whose angular difference with link L is greater than 45 degrees are removed. The values of the attributes are given by:

1.	*Para*	= 1–(*angular difference*) = 45;
2.	*Perp*	= 1 – a * (*perpendicular distance*) = length of link L;
3.	*Dist*	= 1 – b * (*sum of node distances*) = length of link L
4.	Comp	= 1 – c * (lratio – 1);

where:
perpendicular distance is from the centre of link L to the other link; *node distances* are distances between the corresponding start nodes and end nodes; *lratio* is the ratio of length of probable parallel link to the length of link L or the inverse of it, whichever is greater.

Each of the attributes is also given a weight when determining the parallel link. It was found that giving a smaller weight to attribute Para led to favorable results. The values of parameters *a,b,c* and weights used are given in Table 7.3.3. The program was run for the Twin Cities network and the results have been manually inspected for accuracy.

Table 7.3.3 Values of Weights and Parameters of Parallel Attributes

Attribute	*Weight*	*Parameter*
Para	0.5	–
Perp	0.5	$a = 0.40$
Shift	1.0	$b = 0.25$
Comp	0.5	$c = 0.50$

7.3.4 Cost Function

To frame the network investment decision, a cost function is needed to estimate the cost of new construction or expansion of an existing facility. Data collected on previous investments are used to estimate a cost function for the Twin Cities network. Cost of construction can be modeled as:

$$E_{ij} = f(L_{ij} * \Delta C_{ij}, F, N, T, Y, D, X) \quad (1)$$

where:

E_{ij} = cost to construct or expand the link (in 1000's of dollars);
$L_{ij} * \Delta C_{ij}$ = lane kilometres of construction;
F = dummy variable for type of funding source;
N = dummy variable to check if it is a new construction or expansion;
T = dummy variables for interstate highways and state highways;
Y = year of completion — 1979;
D = duration of construction;
X = distance of the link from the nearest town centre (in km).

A Cobb–Douglas model is estimated to predict the cost of expanding a link. Initially, all the variables mentioned above were entered into the model. It was found that both funding source and road type were insignificant. However, examining the data reveals that most interstate highway construction is built under one funding source and non-interstate highways under a different program. To overcome this problem, funding source was dropped, as road type was sufficient for segregation. Also, subsequent models suggested that reconstruction, widening, and so on, could be grouped into one category. The model was re-estimated to give the implicit rate of inflation with respect to the highway construction sector, shown in Table 7.3.4 (Model 1).

Note that the modified year variable has been introduced into the model in a non-logarithmic fashion. If the model was written in exponential form, (exp.(0:078) – 1) * 100 = 8:14 gives the rate of inflation. This is quite high compared to the overall rate of inflation in the United States. One of the reasons may be due to an increase in the costs of transport materials at a rate higher than general goods and services. Increase in labor costs, additional attributes on roads such as bike and pedestrian features, and evolving safety measures, also contribute to it. Also, highways are being constructed with better materials due to improvements in technology resulting in an increase in cost. Cost per lane km was also considered as the dependent variable, but it did not improve results.

Finally, duration of construction (*D*) and distance from the nearest town centre (*X*) variables were also introduced into the regression.[b] In general, other than interstate highways, construction was only one year in duration. One might then be tempted to say that duration would also be accounted for by the road type variable. However, interstate construction projects took a variable amount of time. Distance from nearest town centre was negative and significant indicating that the project cost would decrease as we move away from town centre areas. The best model taking duration of construction and distance from the nearest town centre into consideration, with all the variables significant at the 5 per cent level, is shown in Table 7.3.4 (Model 2).

The coefficient of lane kilometres of construction is less than one, indicating economies of scale in construction. As can be expected, a new construction project is more costly than expanding an existing link. The cost of construction increases with the hierarchy of the road. This is so because of the greater thickness of pavements on higher-class roads, the use of concrete rather than asphalt, and their larger width including shoulders. Higher duration projects cost more and construction becomes costlier over time. Distance from the nearest town centre, which was entered as a linear variable, shows that the project cost would decrease as we move away from town centre areas. Town centre areas have higher traffic flows and land costs, and

Table 7.3.4 Coefficients of regression for cost models

$\text{Ln}(E_{ij})$		*Model 1*		*Model 2*	
Description of the variable	*Variable*	*Coef.*	*Std. Dev*	*Coef.*	*Std. Dev*
Lane kilometres of construction	$\text{Ln}(L_{ij} * \Delta C_{ij})$	0.48	0.114*	0.50	0.118*
Dummy for new constructions	*N*	0.38	0.184*	0.39	0.187*
Dummy for interstate roads	*Inter*	1.68	0.271*	1.97	0.300*
Dummy for state roads	*TH*	0.57	0.212*	0.56	0.226*
Year −1979	*Y*	0.08	0.009*	–	–
Log of year −1979	Ln(*Y*)	–	–	0.75	0.110*
Log of duration of construction	Ln(*D*)	–	–	0.16	0.142
Distance from nearest town centre	*X*	–	–	−0.03	0.016*
	_constant	6.17	0.248*	5.56	0.329*
	Number of Observations	110		76	
	Adj. *R*-squared	0.65		0.77	

*Significant at 90% confidence interval
– Variable not present in that model

hence restrict the construction flexibility justifying the extra cost. Note that in the final model, modified year has been entered as a logarithmic variable. This may be a better model due to the non-linear increase in costs with date of construction. According to Model 2, a lane-km of interstate highway taking three years to construct in the year 2000 would cost approximately 21.8 million dollars.

7.3.5 Model

Capacity in the next time period ($C_{ijt}+1$) can be modelled as a function:

$$C_{ijt+1} = f(C_{ij}, L_{ij}, Q_{ij}/C_{ij}, Q_p/C_p, \Delta(Q_{ij} * L_{ij}), B, \hat{E}_{ij}, Y, X, (Q_{hi}+Q_{jk}),$$
$$(C_{hi}+C_{jk}-C_{ij}), \Delta(Q_{hi}+Q_{jk}), \Delta(C_{hi}+C_{jk}), P, \Delta P) \quad (2)$$

where:

C_{ijt+1} = Capacity on arc ij (arc running from node i to node j) at time $t+1$;
C_{ij} = Capacity on arc ij;
L_{ij} = Length of arc ij;
Q_{ij} = Flow on arc ij;
Q_p = Flow on parallel link;
C_p = Capacity of parallel link;
$\hat{E}_{ij}$ = Unit expense of construction of improvements on arc ij (from cost function);
B = Budget for year t;
Y = Year of proposed construction −1979;
X = Distance from the nearest town centre;
Q_{hi} = Sum of flows on arcs hi;
Q_{jk} = Sum of flows on arcs jk;
C_{hi} = Sum of capacities on arcs hi (arcs supplying flow);
C_{jk} = Sum of capacities on arcs jk (arcs receiving flow);
P = Population of the surrounding Minor Civil Division (MCD).

All variables are vectors; Flows are bi-directional; indicates change between time period t and $t-n$:

Budget is pre-determined for a particular year but it varies over years. For each year, the budget is simply the total expenditure on the link expansions considered in the model. Only a few links are expanded in the network in a given year. There have been no instances of expanding the same link twice in the years considered for this study. Since previous expansion of a link predicts failure to expand in a given year perfectly, observations of such links are dropped from the data after its expansion. A strict capacity measure was unavailable and hence the number of lanes was used as a surrogate for capacity in modeling. Modeling was directed at predicting the increase in the number of lanes rather than predicting the absolute number of lanes in the next time period as a whole. This is because the number of lanes in the previous time period largely explains the number of lanes for a given period (since most links do not change each year). It should be noted that the increase in number of lanes is a discrete number and the increase is zero for most of the links in a given time period. To overcome this problem and to consider the discrete nature of the increase in number of lanes, discrete choice modeling was considered appropriate.

While logit models have been successfully used in a wide variety of multivariate discrete choice contexts, multinomial logit models exhibit the restrictive property of Independence of Irrelevant Alternatives (IIA). IIA states that the presence of a third alternative does not affect the relative probabilities of any two alternatives. Also, coefficients of variables are assumed fixed across individuals in this setting. In mixed logit models, the coefficients are random across individuals with a specified distribution (McFadden and Train, 2000; Train and Brownstone, 1999; Hensher, 2000). This relaxes the IIA property of multinomial logit models and allows for correlation across alternatives. Mixed logit models are estimated by integrating the likelihood function of the multinomial logit model over the distributions of the random coefficients. The mixed logit likelihood function is given by:

The parameters are estimated using simulated maximum likelihood due to the lack of a closed form solution to these integrals. Halton sequences, which are shown to be more efficient

$$\int \frac{\exp(\beta' x_{ij})}{\sum_j \exp(\beta' x_{ij})} f(\beta/\Omega)\, d\eta. \qquad (3)$$

and guarantee uniform coverage of the distribution, are used for this study (Bhat, 2001, 2002).

The presence of induced demand presents the problem of a circular relationship between capacity increase and change in demand, that is, the problem of endogeneity. The problem is overcome by substituting the endogenous variables in the model with instrumented variables. For the purpose of this study, change in demand has been instrumented using models estimated by Parthasarathi et al. (2003) on the same dataset. The change in demand predicted there was in terms of change in vehicle kilometres travelled and the same is used in our study. Based on the theory described above, the hypotheses are as follows:

- It is posited that the links with higher capacity (C_{ij}) and longer length (L_{ij})are less likely to be expanded.
- Congestion on a link (Q_{ij} / C_{ij}) increases the probability of its expansion or of the link parallel to it. The same is expected to be reflected in the congestion measure of the parallel link (Q_p / C_p).
- Increase in VKT on a link ($\Delta Q_{ij} * L_{ij}$) should increase the likelihood of its expansion.
- The higher the cost of a link expansion (E_{ij}), the lower is its probability of expansion.
- A higher budget for a year would result in more links being expanded and thus increase the probability of expansion for a particular link.
- Capacity expansion on a parallel link decreases the chances of the link under consideration to be expanded.
- Higher VKT of upstream and downstream links increases the chances of link expansion to facilitate the incoming traffic.
- Increase in capacity of a downstream link or an upstream link (ΔC_{hi} or ΔC_{jk}) would cause the link in consideration also to be expanded to take the burden of the resulting traffic.

- Chances of re-expansion of a link are assumed to be low since alternative routes will also be considered for expansion. . Distance from the nearest town centre favors the expansion of a link.
- Increase in population in an area (ΔP) would result in expansion of the links in that area to accommodate the excess traffic.
- Related and supporting industries – when networks of buyers and suppliers are in close proximity, this can create faster and more active information exchange, collective learning, and supply-chain innovation.

As can be observed, expansions are decreasing over time. This may be due to costs rising faster than budgets, or it may be due to some other factors (for example, network saturation or declining benefits of new expansions over time), and this should be reflected in the time variable (Y) of the regression.

Each of the road types was modeled separately due to their functional differences. The decision to construct or expand county highways differs from state highways, as they have separate funding sources. Also, in the data set considered, there were no two-lane expansions of trunk highways and county highways, warranting separate modeling for each hierarchy of the network.

7.3.6 Results

A multinomial logit model was used to model the increase in number of lanes (ΔC_{ij+1}) 1) over the previous year. Results of the regression for interstate highway are given in Table 7.3.5. Variables C_{ij}, L_{ij}, Q_{ij} / C_{ij}, E_{ij}, Chi + C_{jk}, C_{ij}, and Y are negative and significant while the variables $\Delta2(Q_{ij} * L_{ij})$, $\Delta24(Q_{ij} * L_{ij})$; $\Delta46(Q_{ij} * L_{ij})$, $Q_{hi} + Q_{jk}$, B, and ΔP are positive and significant. This shows that as the number of lanes increases, the probability of its expansion decreases and the probability of a two-lane increase is still lower in this case, supporting the hypotheses. Thus we find that links that already have higher capacities are less likely to be expanded due to decreasing marginal returns. Links with lower capacities would then be more likely to be expanded in order to achieve uniformity in the network. Long links that take more time to build tend to be overlooked for expansion in favor of other shorter links. It is difficult to divert traffic for the long period required for construction on longer links. It has been noted in previous studies (Miyagi, 1998) that the overall welfare can be negative in some cases if high volumes of traffic are inconvenienced for a longer period of time. Higher capacity on downstream and upstream links deters link expansion, again indicating decreasing marginal returns.

Table 7.3.5 Multinomial logit model for interstate highways

	$\Delta C_{ijt+1} = 1$			$\Delta C_{ijt+1} = 2$		
Variable	*Hypo.*	*Coef.*	$P > \|z\|$	*Hypo.*	*Coef.*	$P > \|z\|$
Cij	–S	−1.92E+00	5.25E−01*	–S	−2.22E+00	5.46E−01*
Lij	–S	−2.20E+00	9.98E−01*	–S	−3.54E+00	1.21E+00*
Qij/Cij	+S	−1.82E−05	8.28E−06*	+S	−3.84E−05	1.13E−05*
Qp/Cp	+S	1.82E−05	8.30E−06*	+S	1.79E−06	1.09E−05
$\Delta 02(Qij * Lij)$	+S	4.58E−04	1.22E−04*	+S	3.63E−04	1.01E−04*
$\Delta 24(Qij * Lij)$	+S	5.34E−04	6.97E−05*	+S	5.33E−04	7.27E−05*
$\Delta 46(Qij * Lij)$	+S	5.34E−04	6.97E−05*	+S	5.33E−04	7.27E−05*
$\Delta 68(Qij * Lij)$	+S	−8.04E−04	2.22E−04*	+S	−3.53E−04	2.06E−04
Eij	–S	−6.54E−09	1.29E−09*	–S	−3.19E−09	6.25E−10*
B	+S	4.16E−06	1.63E−06*	+S	7.83E−06	3.15E−06*
Y	–S	−1.12E+00	1.57E−01*	–S	−1.72E+00	2.59E−01*
X	+S	2.09E−01	6.04E−02*	+S	−1.54E−01	6.28E−02*
Qhi + Qjk	+S	1.17E−05	2.60E−06*	+S	1.13E−05	2.97E−06*
Chi + Cjk − Cij	–S	−7.18E−01	1.41E−01*	–S	−5.27E−01	1.52E−01*
P	–	7.78E−06	1.81E−06*	–	−3.09E−06	2.35E−06
ΔP	+S	1.63E−04	1.63E−05*	–	4.97E−04	2.37E−05*
_cons		3.02E+00	1.48E+00*		7.59E+00	1.64E+00*

Number of Observations: 10986
Initial LL = − 572.29   Final LL = − 277.65
LR $chi^2 = 575.55$ Psuedo R^2: 0.51

*** Significant at 90% confidence interval**

Increasing traffic demand for a particular link increases its probability for both one-lane expansion and two-lane expansion. Here we see the response of infrastructure supply to increases in travel demand. The probability of a link expansion increases if the flow on downstream and upstream links is high, showing again that links with greater inflow demand are expanded. Cost is negative and significant showing that links that involve higher expenditure are less likely to be expanded. The budget available is positive and significant as expected; a higher budget favors expansion of more links.

A very interesting trend comes into light with the linear variable year of construction (Y): The negative coefficient on the year of construction for two-lane expansion is considerably higher than that of single lane expansion. The expansion rate of the network has decreased and the relative probability of two-lane expansions (compared with one-lane expansions) declines with time.

Surprisingly, the congestion measure on the link is negative and significant for both one-lane and two-lane expansions. However, congestion on the parallel link is positive and significant for one-lane expansion. Interstate highways are less likely to be expanded and their parallel links, typically of a lower level of the hierarchy, tend to be expanded in the wake of congestion on an interstate highway.

Distance from a town centre is positive and significant for one-lane expansion and negative and significant for two-lane expansions. This implies that two-lane expansions are preferred near town canters where traffic demand is high and one-lane expansions are sufficient in the peripheries where traffic demand is comparatively low, although this may have to do with nonlinear effects of distance. Improvements may be favored in first or second ring suburbs over both the town centre and the exurban fringe. Population in the adjacent jurisdiction (MCD) has a positive effect on one-lane expansions. Since rights for land acquisition would be costly in such areas, a one-lane increase is feasible to cater to the traffic generated, but a two-lane expansion might overrun the budget. A higher population increase in a MCD favors expansion to meet the additional demand generated as expected.

Results with the trunk highway network are as expected. All the trunk highway expansions are one-lane expansions in each direction and only new construction had two lanes built in each direction. The results of this binomial logit model are given in Table 7.3.6. Capacity and length of the link are negative and significant as earlier. Flow and capacity variables of its adjacent links also behave in the same manner as for interstate highways. Again congestion on the link under consideration is negative and significant but it is insignificant on the parallel link. Some of the changes in VKT in the last eight years are positive and significant. Cost of expansion is insignificant and budget is positive and significant as usual. Both population and increase in population have the effect of favoring an expansion. Year is again negative and significant, indicating gradual decline in network growth with time, after controlling for budget and expenditure.

Results of the county highways network in Hennepin County are also given in Table 7.3.5. As before, capacity and length of the link are negative and significant. This is the first network in which the congestion measure is positive and significant. As in the case of trunk highways, the spillover effects on the parallel link are not significant. This is because trunk highways and county highways have lower capacity compared to interstate highways and their expansion does not interrupt traffic significantly. Change in VKT has a positive effect on its expansion. Higher flow on adjacent links decreases its chances of expansion. Cost of expansion is again insignificant and the reason might be the necessity to expand lower capacity links (compared to interstate highways) to facilitate smoother traffic flow within the metropolitan area. Surprisingly, both population and population increase have a negative effect on the chances of expanding.

Table 7.3.6 Logit model for trunk highways and county highways

		Trunk Highways		County Highways	
Variable	Hypo.	Coef.	Std. Dev	Coef.	Std. Dev
Cij	−S	−8.15E+00	2.07E+00*		
Lij	−S	−2.28E+01	3.91E+00*	−5.52E−01	5.46E−01
Qij/Cij	+S	−9.30E−05	3.74E−05*	1.97E−04	2.85E−05*
Qp/Cp	+S	1.05E−05	3.81E−05	−1.99E−05	1.53E−05
$\Delta 02(Qij * Lij)$	+S	1.67E−03	6.09E−04*	2.81E−03	2.99E−04*
$\Delta 24(Qij * Lij)$	+S	4.89E−04	3.40E−04	2.72E−03	2.41E−04*
$\Delta 46(Qij * Lij)$	+S	4.89E−04	3.40E−04	2.72E−03	2.41E−04*
$\Delta 68(Qij * Lij)$	+S	4.54E−03	1.10E−03*	3.80E−03	3.05E−04*
Eij	−S	−8.18E−10	6.15E−09	6.84E−10	2.49E−09
B	+S	7.32E−05	1.92E−05*	1.07E−05	4.44E−06*
Y	−S	−1.17E+00	5.02E−01*	8.37E−02	1.02E−01
X	+S	1.13E−01	6.47E−02	5.19E−02	3.69E−02
$Qhi + Qjk$	+S	2.37E−05	6.15E−06*	−1.71E−05	6.10E−06*
$Chi + Cjk - Cij$	−S	−5.49E−01	2.46E−01*	5.43E−02	9.88E−02
P	−	1.28E−05	5.35E−06*	−1.63E−05	2.00E−06*
ΔP	+S	1.17E−03	1.95E−05*	−1.40E−04	5.73E−06*
_cons	−	3.86E+00	4.73E+00	−1.40E+01	1.25E+01
		No. of Obs: 17926		No. of Obs: 6531	
		Initial LL = − 108.46		Initial LL = − 366.42	
		Final LL = − 41.34		Final LL = − 202.98	
		LR chi^2 = 114.75		LR chi^2 = 284.98	
		Psuedo R^2 = 0.61		Psuedo R^2 = 0.29	

*** Significant at 90% confidence interval**

Results of the mixed logit models are as given in Table 7.3.8. Standard deviations of only two variables, change in demand over the previous two years and length of the link, were found to be significant. Initially, models were estimated assuming the coefficients to be distributed independently and normally across the population. This resulted in non-significance of the standard deviations. The assumption of a normal distribution places an important constraint on the behavior of the links, although theory suggests the same sign on the coefficients for all the links. Another approach is to specify a distribution that may result in negative coefficients but does not necessarily impose such a situation. The triangular distribution satisfies this property and was chosen here. The two-year change in VKT ($\Delta 2(Q_{ij}*L_{ij})$ has a standard deviation comparable to the coefficient itself, indicating a wide range of response of individual links to this variable. In the time period considered, there were 73 lane expansions in the interstate highway network of 1,525 links. Of the 73 most probable link expansions predicted by the models, multinomial logit identified 52 links that were actually expanded, while mixed logit predicted the same 52 links and an additional 3 links correctly. Standard deviation of random variables for trunk highways is comparatively small although significant. County highways did not have any significant standard deviations in the coefficients. This shows that links on the hierarchy below interstates have a fixed response to the variables. Mixed logit models with unobserved taste as a random parameter failed to converge for all the highways. It should be noted that the final model significance depends on the significance of the input variables as well as the significance of the cost and induced demand models.

7.3.7 Conclusions

This paper developed for the first time a model to predict how transport agencies expand their networks as a function of traffic flow, flow on adjacent and competitor links, flow on parallel links, and estimated cost using data from the Twin Cities metropolitan area. In all the three models, capacity, length, change in VKT, total inflow, and budget have similar coefficients and significance. It is interesting to note the differences in the models. While expansion of interstate highways depends on both the budget and cost of expansion, the lower hierarchies of roads are seemingly unaffected by cost (over the range of values observed). Congestion positively and significantly affects agency decisions to expand county highways, while being negative and significant for other highways. Interstate and trunk highways show a decline in their growth with time that is not reflected in county highways. The elasticities of variables for trunk highways are generally higher, except for the population variables. These differences by type of highway give us a picture of the change in policy for each type of road.

The results are promising and suggest that a number of measurable properties drive network expansion. While it is obvious that politics factors into network expansion decisions, this model is based on empirically measurable attributes. The importance of this is in extension for modeling the implications of transport planning decisions. Any decision made today will lead to a chain of events, future network expansion decisions that are not considered in most static modeling frameworks. Endogenous network growth, and the pressures placed on future decision-makers because of today's decision are critical factors for planning and modeling.

Future research can be directed towards analysis of the allocation of resources in the Transportation Improvement Plans that US Metropolitan Planning Organizations are required to conduct (Crain and Oakley, 1995; Dueker, 2002). In particular, the model could be extended to treat the budget within the model endogenously.

Table 7.3.7 Comparison of elasticities of highway

Variable	Hypo.	Interstate Coef.	Interstate Elasticity	Trunk Highways Coef.	Trunk Highways Elasticity	County Highways Coef.	County Highways Elasticity
Cij	−S	−2.06E+00*	−4.04E+00	−8.15E+00*	−1.40E+01		
Lij	−S	−2.62E+00*	−9.39E−01	−2.28E+01*	−1.19E+01	−5.52E−01	−2.87E−01
Qij/Cij	+S	−2.50E−05*	−9.40E−01	−9.30E−05*	−1.48E+00	1.97E−04*	1.72E+00
Qp/Cp	+S	1.57E−05*	4.55E−01	1.05E−05	1.54E−01	−1.99E−05	−2.99E−01
Δ02(*Qij* * *Lij*)	+S	3.98E−04*	7.12E−01	1.67E−03*	2.03E+00	2.81E−03*	1.66E+00
Δ24(*Qij* * Lij)	+S	5.10E−04*	−9.63E+03	4.89E−04	−4.10E+03	2.72E−03*	2.96E+04
Δ46(*Qij* * Lij)	+S	5.10E−04*	9.63E+03	4.89E−04	4.10E+03	2.72E−03*	−2.96E+04
Δ68(*Qij* * Lij)	+S	−5.96E−04*	−1.55E+00	4.54E−03*	6.60E+00	3.80E−03*	3.74E+00
Eij	−S	−3.21E−09*	−1.22E+00	−8.18E−10	−8.25E−02	6.84E−10	4.32E−02
B	+S	4.38E−06*	6.68E−01	7.32E−05*	1.41E+00	1.07E−05*	4.05E−01
Y	−S	−1.09E+00*	−7.26E+00	−1.17E+00*	−7.77E+00	8.37E−02	5.50E−01
X	+S	−6.23E−03	−5.28E−02	1.13E−01	1.21E+00	5.19E−02	5.53E−01
Qhi + *Qjk*	+S	1.09E−05*	1.50E+00	2.37E−05*	1.52E+00	1.71E−05*	5.58E−01
Chi + *Cjk* − *Cij*	−S	−6.07E−01*	−2.64E+00	−5.49E−01*	−2.31E+00	5.43E−02	1.95E−01
P	–	2.56E−06*	3.24E−01	1.28E−05*	9.45E−01	−1.63E−05*	−1.47E+00
ΔP	+S	2.54E−04*	1.23E−01	1.17E−03*	7.26E−01	−1.40E−04*	−1.18E−01
_cons	–	5.22E+00	–	3.86E+00	–	−1.40E+01	–
			No. of Obs:10986 LL = −293.92 Psuedo R^2 = 0.46		No. of Obs:17926 LL = −41.34 Psuedo R^2 = 0.61		No. of Obs:6531 LL = −202.98 Psuedo R^2 = 0.29

* Significant at 90% confidence level

Table 7.3.8 Mixed logit model for interstate and state highway

		Interstate		*Trunk Highways*
Variable	*Hypo.*	$\Delta C_{ijt+1}=1$	$\Delta C_{ijt+1}=2$	ΔC_{ijt+1}
Cij	−S	−2.15E+00*	−2.27E+00*	−6.83E+00*
Lij	−S	−2.21E+00*	−3.69E+00*	−1.55E+01*
Qij/Cij	∓S	−1.12E−05*	−2.94E−05*	−6.17E−05*
Qp/Cp	∓S	2.12E−05*	8.62E−06	1.26E−05
Δ02(*Qij* ∗ Lij)	∓S	5.41E−04*	2.29E−04*	−1.20E−03*
Δ24(*Qij* ∗ Lij)	∓S	6.51E−04*	5.10E−04*	3.99E−04
Δ46(*Qij* ∗ Lij)	∓S	6.51E−04*	5.10E−04*	3.99E−04
Δ68(*Qij* ∗ Lij)	∓S	−1.24E−03*	−2.13E−04	3.51E−03*
Eij	−S	−7.39E−09*	−3.35E−09*	−1.22E−08
B	∓S	4.73E−06*	8.59E−06*	−1.34E−05*
Y	−S	−1.20E+00*	−1.82E+00*	−1.69E+00*
X	∓S	2.44E−01*	−1.46E−01*	8.62E−02
Qhi+*Qjk*	∓S	1.17E−05*	1.15E−05*	1.91E−05*
Chi+*Cjk*−*Cij*	−S	−7.35E−01*	−6.46E−01*	−4.80E−01*
P	−	7.65E−06*	−3.50E−06	1.12E−05*
Δ*P*	∓S	1.34E−04*	5.27E−04*	8.62E−04*
_cons		3.47E+00*	8.79E+00*	1.18E+01
Triangular Variances				
Lij		6.13E−02*	7.45E−01*	1.21E−07*
Δ02(Qij*Lij)		2.92E−04*	4.41E−05*	2.13E−10*
			No. of Obs:10986	No. of Obs:17926
			LL=−232.99	LL=−38.27

*** Significant at 90% confidence interval**

7.3.8 References

1. Barker, T. C. and M. Robbins (1975): *A History of London Transport,* vol. 1 and 2, Allen and Unwin, London. Bhat, C. R. (2001): ‘‘Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model.’’ Transportation Research, 35B, 677–93.
2. Bhat, C. R. (2002): ‘‘Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences.’’ *Transportation Research*.
3. Bhat, C. R. and S. Castelar (2002): ‘‘A Unified Mixed Logit Framework for Modeling Revealed and Stated Preferences: Formulation and Application to Congestion Pricing Analysis in the San Francisco Bay Area.’’ *Transportation Research,* forthcoming.
4. Caruthers, J. I. and G. F. Ulfarsson (2001): ‘‘Public Service Expenditures: The Influence of Density and Other Characteristics of Urban Development.’’ *The Pacific Regional Science Conference Organization.*
5. Crain, W. M. and L. K. Oakley (1995): ‘‘The Politics of Infrastructure.’’ *Journal of Law and Economics* 38, 1–17.
6. Dueker, K. J.: ‘‘A Critique of the Urban Transportation Planning Process: the Performance of Portland’s 2000 Regional Transportation Plan.’’ *Transportation Quarterly* 56, 15–21.
7. Friesz, T. L., S. Shah, and D. Bernstein (1998): ‘‘Disequilibrium Network Design: A New Paradigm for Transportation Planning and Control.’’ *Network Infrastructure and the Urban Environment, Springer,* 99–111.
8. Fulton, L. M., R. B. Noland, D. J. Meszler and J. V. Thomas (2000): ‘‘A Statistical Analysis of the Induced Travel Effects in the U.S. Mid-Atlantic Region.’’ *Journal of Transportation and Statistics,* 3, 1–14.
9. Grübler, A. (1990): The Rise and Fall of Infrastructures: *Dynamics of Evolution and Technological Change in Transport.* Heidelberg: Physica-Verlag.
10. *Hennepin County Capital Budget.* Hennepin County, Minnesota, 1980–2000. Hensher, D. (2000): ‘‘The Valuation of Commuter Travel Time Savings for Car Drivers:
11. Evaluating Alternative Model Specifications.’’ Working paper, University of Sydney, Australia.
12. Herbsman, J. (1986): ‘‘Model for Forecasting Highway Construction Cost.’’ *Transportation Research Record* 1056, 47–54.
13. Kosko, B., (1993): *Fuzzy Thinking: the New Science of Fuzzy Logic.* Hyperion, New York.
14. McFadden, D. and K. Train (2000): ‘‘Mixed MNL Models for Discrete Response.’’ *Journal of Applied Econometrics,* 15, 447–70.
15. MetroGIS DataFinder (http://www.datafinder.org/)
16. Metropolitan Council (http://www.metrocouncil.org/)

17. Miyagi, T. (1998): ‘‘A Spatial Computable General Equilibrium Approach for Measuring Multiregional Impacts of Large Scale Transportation Projects.’’ *Network Infrastructure and the Urban Environment,* Springer, 224–44.

18. New Jersey Office of State Planning (1996): *Projecting Municipal Road Costs under Various Growth Scenarios.* Document # 109.

19. Noland, R. B. (1999): ‘‘Relationships between Highway Capacity and Induced Vehicle Travel.’’ *Transportation Research Board 78th Annual Meeting,* Preprint CD-ROM, Washington DC.

20. Parsley, L., and R. Robinson (1983): ‘‘New TRRL Road Investment Model for Developing Countries.’’ *Transportation Research Record* 898, 7–10.

21. Parthasarathi, P. K., D. Levinson, and R. Karamalaputi, (2003): ‘‘Induced Demand: A Microscopic Perspective.’’ *Urban Studies* (in press).

22. State Demographic Center (http://www.mnplan.state.mn.us/demography/)

23. Strathman, J. G., K. J., Dueker, T., Sanchez, J., Zhang and A. E. Riis, (2000): ‘‘Analysis of Induced Travel in the 1995 NPTS.’’ Center for Urban Studies, Portland State University, Portland.

24. Taaffe, E. J., R. L. Morrill, and P. R., Gould, (1963): ‘‘Transport Expansion in Underdeveloped Countries.’’ *Geographical Review,* 53, 503–29.

25. Train, K. (1986): *Qualitative Choice Analysis.* MIT Press.

26. Train, K., and D., Brownstone, (1999): ‘‘Forecasting New Product Penetration with Flexible Substitution Patterns.’’ *Journal of Econometrics,* 89, 109–29.

27. *Transportation Improvement Program for the Twin Cities Metropolitan Area.* St. Paul, Minneapolis: Metropolitan Council, 1978–2000.

28. *Transportation Information Systems.* St. Paul, Minnesota: Metropolitan Council.

29. Whitman, J., and F. J. Wegmann, (1972): ‘‘Cost Estimating Model for Rural Interstate Highways.’’ *ASCE Journal of Transportation Engineering,* 531–46.

30. Zadeh A. L., (1992): *Fuzzy Logic for the Management of the Uncertainty.* Wiley Publications, New York.

7.4 Induced Demand

A Microscopic Perspective

Pavithra Parthasarathi
URS, 700 Thresher Square
Third Street South
Minneapolis, MN 55415
Fax: 612 373 6526
E-mail: pavithra_parthasarathi@URSCorp.com

David Levinson
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
Fax: 612 625 7750
E-mail: levin031@tc.umn.edu

Ramachandra Karamalaputi
Planner III, CTPS
10 Park Plaza, Suite
2150, Boston, MA 02116,
Fax: 617 973 8855
E-mail: knvrama@ctps.org

7.4.1 Introduction

Transport forecasts often assume limited or no response of demand to changes in supply. This supposition has been refuted by others—economists in particular. Anthony Downs (1992), for instance, famously postulated the idea of 'triple convergence' (also referred to as the 'Iron Law of Congestion') which says that when capacity is added, drivers switch routes, times and modes to take advantage of the more desirable capacity. Moreover, in the longer term, people move and change jobs, while activities relocate to absorb new capacity. Nevertheless, considerable controversy has existed over the existence and importance of the response of demand to supply, also referred to as induced or latent demand. According to the induced demand hypothesis as stated by Fulton et al. (2000), additions to highway capacity cause travel to increase. The presence of this relationship also raises the question of whether capacity expansions provide net costs or benefits to society, as additional travel generates negative externalities. Research on the induced demand hypothesis to date has mostly been carried out at the aggregate level, considering state, metropolitan or county-level traffic and capacity data. However, disaggregation of travel data is essential to understand where induced demand effects are the greatest (Noland, 1999). This study analyses the induced demand hypothesis at the disaggregate level.

The highway network for the metropolitan area of the 'Twin Cities' of Minneapolis and Saint Paul, Minnesota, is analyzed at the link level using a network from the regional planning agency and linking it with a traffic counts database and data from the capital improvement program. This microscopic focus allows the analysis of the induced demand hypothesis in greater detail. In addition to the flow and capacity variables, factors like the population of Minor Civil Division (MCD) and adjacent MCDs, and flow and capacity on parallel links are also taken as variables in the analysis.

The study focuses on network evolution while controlling for network interactions. This in essence means looking at how the network has grown over time. The Twin Cities network for the year 1995 from the metropolitan planning organization is taken as the base case and the network is built (or deconstructed) for the years from 1978 to 1998 using the investment data that were obtained from the capital improvement programs. The focus is on the effects of widening a road, in contrast with some earlier studies that emphasized new construction, or conflated the effects of new links with widened links. It is expected that the effects of the two are different in magnitude (though not direction), as link expansions serve existing populations more, while new construction tends to support future residents.

The rest of this paper is organized as follows. The next section provides a brief background and review of the literature on this topic. A description of the data-sets used for the analysis and the procedures required to format the data-set for use in this study are then described. This is followed by an explanation of the basic model used in this study and its functional forms. A brief introduction of selectivity bias in the models and its correction procedure is given in the next section. A description follows of the hypothesis formulated, the analyses done and the results obtained. Finally, the paper concludes with the key findings from the study.

7.4.2 Background

Transport planners and economists have acknowledged the presence of induced travel to changes in road capacity since the 1930s when benefit–cost analyses were first applied to road projects. Recently, however, there has been an increasing interest in the relationship and size of this travel response to changes in highway capacity. A great deal of attention has been placed on the potential of new construction to increase traffic (Kiefer and Mehndiratta, 1998). Strathman et al. (2000) feel that this attention has been caused by concerns about the relationship between highway construction, air quality and urban development patterns. Recently there has been increased emphasis on strategies that are demand-oriented or which better manage existing infrastructure. Urban road construction has suffered from diminished finances and concerns of adverse environmental effects. Studies have shown that there exists an indirect causal relationship between capacity increase and travel patterns, which induces growth in vehicle travel (Strathman et al., 2000).

Noland (1999) argued that increasing the capacity of a highway corridor attracts increased vehicle traffic. The basic theory of induced demand can be explained using the economic theory of supply and demand. From an economic perspective, the demand for travel is influenced by its cost. The cost of travel consists of the fixed capital costs of a vehicle inclusive of fuel and maintenance costs and the variable travel time costs. A highway capacity increase would reduce travel time and thus increase overall demand.

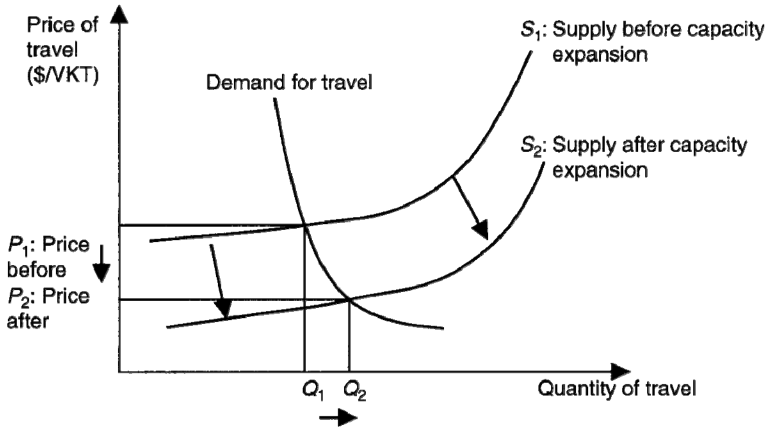
Figure 7.4.1 illustrates the induced travel hypothesis. S_1 represents the supply before the capacity expansion took place and S_2 represents the supply after the capacity expansion. S_2 is shifted downward with respect to S_1, indicating that the same demand is met at a lower cost than before the capacity expansion. The increase in travel or induced travel in terms of vehicle kilometres travelled (VKT) is indicated by Q_1–Q_2. The figure shows that the change in supply reduces the cost of travel and this in turn increases the amount of travel.

Noland (1999) also concluded that the induced demand hypothesis should not be rejected. This aggregate study used a cross-section of data from the 50 US states between 1984 and 1996. The results showed a significant relationship between capacity and distance traveled. Although there were other factors like population growth that contributed to increased distance traveled, increases in capacity accounted for nearly one quarter of this growth.

Fulton et al. (2000) looked at the induced demand hypothesis in the US Mid-Atlantic region using county-level data from Maryland, Virginia, North Carolina and Washington, DC. The Granger causality test that was conducted in this study showed that changes in travel were preceded by changes in lane kilometres.

Pfleiderer and Dieterich (1995) have argued about the presence of induced demand on the basis of the law of constant travel time budget. The question of travel time budgets is still unresolved, although evidence for a commute time budget is strong (Levinson and Kumar, 1994a; Barnes and Davis, 2001). A budget would suggest that, as capacity became available, link travel times would drop and much of the faster speeds would be absorbed in longer trips (and thus more VKT). Thus the speeds would still be better after the link expansion than before it, but not as much as a naïve model might suggest. Levinson and Kanchi (in press) developed a model to study the change in time use using the 1990 and 1995 NPTS data. The study identified how

Figure 7.4.1 The induced travel hypothesis. Source: Noland (1999).



travel time and activity durations change with capacity. The study concluded that increases in highway capacity cause small but significant changes in individual daily travel behavior and activity patterns.

Hansen and Huang (1997) estimated the relationship between state highway capacity and travel using panel data between 1973 and 1990 for the state of California. The authors conclude that increasing road capacity does not greatly reduce congestion in the long run. Aside from added capacity, other factors that contribute to increased travel include growth in population, decreasing household sizes, saturation of vehicle availability phenomenon and the effect of dispersion of residences and work (Kiefer and Mehndiratta, 1998). These factors play a role in determining demand and increasing growth in urban travel independent of the supply.

Mokhtarian et al. (2001) test the existence of induced demand using a matched-pair approach. Eighteen California Highway segments were considered and these segments were paired with control segments such that improved segments were paired with unimproved ones. The study concluded that there was no difference in the growth rates between the improved and unimproved segments. The authors believe, however, that the induced traffic effect due to capacity changes on existing facilities can be very small and hence its detection in such a study might not be possible without a large sample size.

Table 7.4.1 Tabulation of elasticity estimates obtained by different studies

Study	Elasticity
Dowling and Colman (1995)	0.3–0.5 per cent increase in trip generation due to congestion relieving projects
Fulton *et al.* (2000)	0.2–0.6 per cent increase in VKT due to 1 per cent increase in lane kilomctrcs
Hansen and Huang (1997)	0.6–0.7 per cent increase in VKT due to 1 per cent increase in lane kilometres (county level)
	0.9 per cent increase in VKT due to 1 per cent increase in lane kilometres (metropolitan level)
Noland (1999)	0.2–0.5 per cent increase in VKT due to 1 per cent increase in lane kilometres (short-run estimates)
	0.7–1.0 per cent increase in VKT due to 1 per cent increase in lane kilometres (long-run estimates)
Strathman *et al.* (2000)	0.29 per cent direct effect for 1 per cent increase in road capacity
	0.033 per cent indirect effect for 10 per cent increase in road capacity

From the studies, it seems difficult to reject the induced demand hypothesis. It is seen that increasing capacity reduces the cost of travel and results in greater travel and overall mobility. However, the Fulton et al. (2000) study indicates that the downside of increasing capacity is that the benefits from congestion reduction would be lost over time due to increased travel. A summary of the elasticity estimates obtained in the various studies is shown in Table 7.4.1.

7.4.3 Data

Four data-sets are used in the analysis. Network data for the year 1995 come from the Metropolitan Council for the Twin Cities, the region's metropolitan planning organization. The data contain information on the highway network given by links, which are defined by the start and end node respectively. In addition, the data contain information on the length of the link, the number of lanes on the link, capacity and modeled traffic flows on the link. Each link has a coded road type and location. A summary of the 1995 highway network data is given in Table 7.4.2.

Traffic counts data in the form of annual average daily traffic (AADT) are from the Minnesota Department of Transport (Mn/DOT). The AADT database contains the traffic volume for each roadway segment from 1978 to 1998. The AADT represents an estimate of the daily volume of all motorized vehicles for the segments, the segments being defined by Mn/DOT (Office of Transport Data and Analysis, Mn/DOT). Table 7.4.3 summarizes the change in vehicle kilometres traveled (AADT multiplied by link length) from 1980 to 1998. Figure 7.4.2 shows the change in lane km from 1980 to 1998.

Population estimates for the Minor Civil Divisions (MCDs) in the Twin Cities area from 1980 to 2000 are from the State Demographic Center at Minnesota Planning.

Investment data are obtained from two sources. The data from the Transportation Improvement Program for the Twin Cities Metropolitan Area provided information on the Interstates and Truck Highways in the seven counties of the metropolitan area. The data from Hennepin County Capital Budget provided information on the County State Aided Highways (CSAH) in the Hennepin County. The investment data obtained from these two sources provided

Table 7.4.2 Summary of the 1995 highway network of the Twin Cities Metro Area, classified by road type

	Interstates		Trunk Highways		Hennepin County Highways (CSAH)	
	Number of links	Total link length (km)	Number of links	Total link length (km)	Number of links	Total link length (km)
Total	1193	710	2138	1808	1658	1023

Note: Only links used in the analysis indicated.

information on the new links constructed and the expansion to the existing links between 1978 and 1998.

The data sources had to be merged to get a final network database. First, the network data were merged with the AADT data so that each link on the network had the corresponding AADT associated with it. This was done using a C program. The MCD data were merged with the network data using Arcview GIS. The network map and the MCD map in GIS format were obtained from MetroGIS. The MCD map was overlaid on the network map and from this it was possible to identify the MCD to which each link belongs. The results of this merging gave a final database which contained information about the links and their characteristics, the AADT on the links, the MCD to which it belongs and the population estimates for the MCD.

Each link in the network was represented as a one-way link. For each link in the network, the corresponding parallel link was identified and the AADT on the parallel link was obtained.[i] For analysis purposes, the AADT on the link was converted into vehicle kilometres traveled (VKT) by multiplying the AADT on the link by the link length.

For each MCD, the adjacent MCDs were identified from the GIS map. The corresponding population estimates were summed to get the adjacent MCD population. Hence each link in the

Table 7.4.3 Summary of ΔVKT

	Summary of ΔVKT ($VKT_t - VKT_{t-2}$)								
	Interstates			Trunk highways			CSAH		
Year	< 0	= 0	> 0	< 0	= 0	> 0	< 0	= 0	> 0
1998	61	25	1107	403	298	1437	677	78	903
1996	176	60	957	475	225	1438	554	127	977
1994	198	61	934	587	303	1248	647	173	838
1992	231	42	920	450	303	1385	888	160	610
1990	181	52	960	562	135	1441	483	142	1033
1988	186	0	1007	561	75	1502	655	85	918
1986	288	22	883	515	101	1522	379	82	1197
1984	79	8	1103	465	60	1613	464	30	1164
1982	330	4	859	736	46	1356	709	76	873
1980	696	17	480	894	90	1154	769	77	812
Total	2426	291	9210	5648	1636	14096	6225	1030	9325

Note: Only links used in the analysis are indicated.

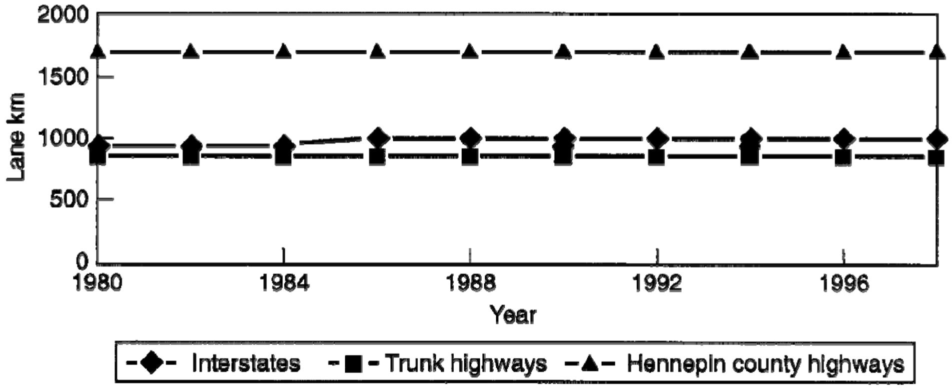


Figure 7.4.2 Summary of lane kilometers. Note: Only links used in the analysis are indicated.

network was associated with the population of both the MCD to which it belongs and the population of the adjacent MCDs.

The next step was to build (or deconstruct) the network for the various years. The 1995 network was used as the base year network. The investment data provided information on the lane additions in the links for the various years. The network was built forward from 1995 to 1998 and backward from 1995 to 1978. Using this information, lanes were added to or subtracted from the previous year's network to build the network for any particular year. The final built network contained 8281 links for a period of 20 years from 1978 to 19982[ii].

The analysis concentrated on the interstates and trunk highways in the seven counties and the Hennepin County state-aid highways (CSAH) because complete data could be obtained for these road types. The data-set was separated by road type. The traffic flow in terms of VKT was used for the years 1978 to 1998 for each road type. In addition to the network variables, the population estimates of the MCDs and adjacent MCDs were also considered. Since the population estimates were available from 1980 only, the final data-set used was a panel data-set of the network for the years 1980 to 1998.

7.4.4 Model Specification

The model predicts the change in VKT based on the flow and capacity conditions on the link and the neighboring links. The analysis focuses on how changes in capacity given by lane additions and expansion would have an effect on the traffic flow in the link. In addition to looking at the influences of the changes on the link on traffic flow, the analysis examines the changes on neighboring links.

The basic difference model used in this analysis is given below.

$$\Delta Q = f(\Delta Q_n, \Delta Q_{pn}, \Delta L_n, \Delta L_{pn}, \Delta P_n, \Delta P_{an}, Q_t, Q_{pt}, L_t, L_{pt}, P_t, P_{at}, D)$$

where:

t base year; ΔQ change in VKT on the link between time $t + n$ and t; Q_n change in VKT on the link between time t and $t - n$; ΔQ_{pn} change in summed VKT on the parallel links between time t and $t - n$; ΔL_n change in the number of lanes on the link between time t and $t - n$; L_{pn} change in the summed number of lanes on the parallel links between time t and $t - n$; t and $t - n$; ΔP_n

change in the population of the MCD to which the link belongs, between time t and $t - n$; P_{an} change in the summed population of the adjacent MCDs between time t and $t - n$; ΔQ_t VKT in the link at base year t, ΔQ_{pt} summed VKT in the parallel links at base year t; L_t number of lanes in the link at base year t; L_{pt} summed number of lanes in the parallel links at base year t; P_t population of the MCD to which the link belongs, at base year t; Pat summed population at the adjacent MCDs, at base year t; D dummy variable for the base years; n lag year 2, 4, 6, 8.

The model given above is estimated with a linear functional form using ordinary least squares regression. Other functional forms and transforms of the variables were tested with similar results and excluded for brevity (Parthasarathi, 2001).

The dependent variable is the change in VKT of a link (VKT_{t+n} VKT_t). The independent variables are the change variables of the previous years and the base year variables. The change variables refer to the difference in the value of the variable at time t and time t – n [$(t) - (t - n)$] while the base year variable refers to the value of the variable existing at time t and n refers to the lag variable.

The traffic counts obtained from Mn/DOT are updated only once every two years on many links. Due to this the value of the lag variable n was set at 2 and multiples of 2 namely 4, 6 and 8. Hence the model has 4 different sub-types: the 2-year lag model, the 4-year lag model, the 6-year lag model and the 8-year lag model. Also, only even years are considered in the analysis. The 4 lag models allow us to account for both the short run and long run effects that are likely to be seen due to changes in the network.

The analysis of each of the road types namely Interstates, Trunk Highways and CSAH roads in Hennepin County are done separately. Each road type is analyzed using the four different sub-types of the difference model given above. In total 12 different trials of the basic difference model are carried out. Each sub-type of the difference model is constrained by the data-set available. The independent variables in each version are calculated as [$(t) - (t - n)$]. The constraint for each sub-type of the difference model is that for the independent variables, $t - n$ should be $\geq$1980, which limited the years that could be considered and consequently the sample size.

Dummy variables have been introduced for each base year. This has been done to see the behavior of each base year dummy variable rather than using the variable year as a continuous variable in which the effect of the base year dummies cannot be clearly seen.

7.4.5 Selectivity Bias

When the dependent variable is observed only for a non-random set of observations, there is a possibility of selectivity bias. Levinson and Karamalaputi (2002) observe that expansion of a link depends on a number of factors including travel demand, adjacent and parallel link properties, project costs, budget constraints and demographic characteristics. Since expansion of a link is based on characteristics not captured entirely by the models in this study, we need to check for potential selectivity bias. This self-selection of links that are expanded induces a correlation between change in lanes and the error term of the regression. The estimated coefficients of variables are biased if selectivity bias is present. A chow test is performed to check if the coefficients of variables change significantly in the sub samples of lane expansion and no lane expansion. The test statistic is given by:

$$\frac{[E_c - (E_1 + E_2)]/k}{(E_1 + E_2)/(n_1 + n_2 - 2*k)}$$

where:

E_c is the error sum of squares in the model with complete observations, E_1 and E_2 are the error sum of squares of each of the sub sample models, k is the number of parameters estimated and n_1, n_2 are number of observations in the sub samples. Note that the change in number of lanes is not included in the above models. The test statistic follows a F(k, n_1 n_2 2*k) distribution. Using this test, it was found that only interstate highways have selectivity bias.

Selectivity bias is corrected by using a two-step estimation procedure. In the first step, a probit model is estimated to calculate the probability of link expansion. Using this probability, the Inverse Mills Ratio (IMR) is calculated for each of the variables (Greene, 1993). The second step is to include the IMR variable in the regression along with other explanatory variables to correct selectivity bias, so the interpretation of the coefficients of the variables is valid for the entire population. Expansion of a link can be modeled as:

$$\begin{aligned} y = \mathrm{f}[&C_{ij}, L_{ij}, \Delta(Q_{ij}*L_{ij}), Q_{ij}/C_{ij}, Q_p/C_p, \\ &E_{ij}, B, Y, X, (Q_{hi} + Q_{jk}), \\ &(C_{hi} + C_{jk} - C_{ij}), \Delta(Q_{hi} + Q_{jk}), \\ &\Delta(C_{hi} + C_{jk}), P, \Delta P] \end{aligned}$$

where:

y = dummy for link expansion; C_{ij} = capacity on link ij; L_{ij} = length of link ij; Q_{ij} = flow on link ij; Q_p = flow on parallel link; C_p = capacity of parallel link; E_{ij} = unit expense of construction of improvements on link ij; B= budget constraint; Y = year of proposed construction—1979; X = distance from the nearest downtown; Q_{hi} = sum of flows on arcs hi; Q_{jk} sum of flows on arcs jk; C_{hi} = sum of capacities on arcs hi (arcs supplying flow); C_{jk} = sum of capacities on arcs jk (arcs receiving flow); P = population of the surrounding Minor Civil Division (MCD). Note that Flows are bi-directional and Δ indicates change between time-period t and $(t – n)$.

Levinson and Karamalaputi (2002) estimated the cost of expanding a link using the same data as used for this study. The model is reproduced in Appendix D (Table D4). Probit models were estimated separately for interstate highways, trunk highways and Hennepin County Highways. The results are given in Tables 7.4.5, 7.4.6 and 7.4.7 respectively. It was observed that high capacity links are less likely to be expanded while links with high inflow and outflow demand are more likely to be expanded. Links with low cost of expansion are favored and a higher budget for a year results in the expansion of more links. The coefficient of year is negative and significant in two cases and positive and insignificant for Trunk Highways indicating that the expansion rate of the network decreased. The probabilities estimated by this model are used to calculate the probabilities of link expansion for 2-year, 4-year, 6-year, and 8-year intervals that are in turn used to calculate Inverse Mills Ratio.

7.4.6 Hypothesis

The specific hypotheses tested in the model are as follows:

- An increase in the number of lanes on the link will increase the VKT on the link, which is the basic induced demand hypothesis.
- An increase in the VKT in past years is expected to have a positive effect on the change in VKT, which reflects the continuing of trends in traffic flow over time.
- The higher the change in the number of lanes on the parallel links, the lower will be the VKT on the link. The parallel links can handle more flow due to the increase in capacity and they act as substitutes to the link.

Table 7.4.4 Probit model to predict lane expansion on interstate highways

Variable	Coefficient	z	P > |z|
Capacity of the link	− 7.89E-01	− 8.02	0.00
Length of the link	4.39E-01	4.53	0.00
Change in VKT over 2 years	− 2.48E-06	− 0.48	0.63
Congestion on the link	− 1.28E-05	− 6.28	0.00
Congestion on the parallel link	2.96E-06	2.32	0.02
Cost of expansion	− 2.64E-10	− 3.5	0.00
Budget	1.29E-06	5.84	0.00
Year—1979	− 5.45E-02	− 4.9	0.00
Distance from nearest downtown	− 4.20E-02	− 3.74	0.00
Flow on adjacent links	4.45E-06	8.92	0.00
Capacity of adjacent links	− 2.98E-01	− 9.55	0.00
Population	− 2.56E-07	− 0.85	0.40
Change in population over 2 years	− 1.62E-08	− 0.02	0.98
Constant	6.54E-02	0.28	0.78
Number of observations 23 609			
Log likelihood − 597.41			

Table 7.4.5 Probit model to predict lane expansion on trunk highways

Variable	Coefficient	z	P > |z|
Capacity of the link	− 5.34E-01	− 4.99	0.00
length of the link	5.97E-01	5.71	0.00
Change in VKT over 2 years	− 3.25E-06	− 0.57	0.57
Congestion on the link	− 1.38E-05	− 6.48	0.00
Congestion on the parallel link	2.62E-06	1.9	0.06
Cost of expansion	− 6.91E-10	− 5.45	0.00
Budget	6.83E-06	11.6	0.00
Year—1979	9.56E-03	0.71	0.48
Distance from nearest downtown	− 6.45E-02	− 4.93	0.00
Flow on adjacent links	4.36E-06	8.15	0.00
Capacity of adjacent links	− 3.06E-01	− 9.16	0.00
Population	− 5.89E-07	− 1.52	0.13
Change in population over 2 years	− 9.81E-05	− 2.05	0.04
Constant	− 6.14E-01	− 2.34	0.02
Number of observations 20 887			
Log likelihood − 511.55			

Table 7.4.6 Probit model to predict lane expansion on Hennepin county highways

Variable	Coefficient	z	P>|z|
Capacity of the link	−4.849E-01	−3.32	0.00
Length of the link	2.694E-01	3.04	0.00
Change in VKT over 2 years	−2.690E-05	−2.26	0.02
Congestion on the link	2.700E-05	3.31	0.00
Congestion on the parallel link	−3.860E-07	−0.19	0.85
Cost of expansion	−6.380E-10	−2.15	0.03
Budget	−1.430E-06	−2.65	0.01
Year—1979	−4.407E-02	−3.72	0.00
Distance from nearest downtown	−1.286E-02	−1.35	0.18
Flow on adjacent links	1.370E-06	1.45	0.15
Capacity of adjacent links	2.997E-02	1.16	0.25
Population	−9.450E-07	−1.92	0.06
Change in population over 2 years	−4.530E-05	−0.91	0.36
Constant	−1.942E+00	−6.61	0.00
Number of observations = 26 550			
Log likelihood = −449.02			

- An increase in the population of the Minor Civil Divisions (MCD) (and adjacent MCDs) to which the link belongs in the previous years causes the traffic flow on the link to increase due to the extra traffic that the rise in population generates. The population variables for the base year, both for the MCDs and adjacent MCDs are expected to have a negative influence on the dependent variable, as larger MCDs will have less growth, since they tend to be in mature areas.

7.4.7 Results

The linear model is estimated separately for each of the 3 road types and 4 lag periods resulting in 12 models in all. All results are attached in Appendix D. The results are summarized in Table 7.4.7. The table indicates the variables considered in the model, the hypothesis formulated for each variable and the summary of the results for each variable. The case column refers to the number of models (out of 12) in which the coefficient of the variables is consistent with that of the summary column.

The results indicate the behavior of the explanatory variables for all road types considered separately and for all lag models. Some of the variables have shown a variation in their behavior with respect to the road type indicating that the road type has an influence on the variable. The selectivity bias seen was corrected using the inverse mills ratio from a probit model predicting link expansion.

Considering the lane variable, it is seen at an overall level that the change in the number of lanes on a link in the previous years has a clear positive and significant effect on the change in VKT in 7 of the 12 models. This indicates that as the number of lanes in the link in the previous years increases, the VKT on the link can be expected to increase, which confirms the induced demand hypothesis.

The coefficients of the base year lanes are negative and significant in 5 of the 12 models.

Table 7.4.7 Results [dependent variable change in VKT between *t + n* and *t*(ΔQ)]

Variable $[(t) - (t - n)]$	Hypothesis	Summary	Cases
Change in lanes on the link (ΔL_n)	+ S	+ S	7/12
Number of lanes on the link in base year t (L_t)		− S	5/12
Change in VKT on the link (ΔQ_n)	+ S	− S	9/12
VKT on link in base year t (Q_t)		+ S	8/12
Change in lanes on the parallel link (ΔL_{pn})	− S	+ S	7/12
Number of lanes on the parallel link in base year t (L_{pt})		− S	8/12
Change in VKT on the parallel link (ΔQ_{pn})		− S	8/12
VKT on parallel links in base year t (Q_{pt})		+ S	8/12
Change in population of the MCD (ΔP_n)	+ S	+ S	7/12
Population of the MCD at base year t (P_t)	− S	− S	7/12
Change in the summed population of the adjacent MCDs (ΔP_{an})	+ S	NS	7/12
Summed population of the adjacent MCDs at base year t (P_{at})	− S	− S	11/12

Key: **− S = negative and significant, + S = positive and significant, NS = not significant.**

The change in VKT on the link in the previous years has a clear negative and significant effect in 9 of the 12 models, which in general contradicts our hypothesis. The results indicate that we may be seeing an equilibration process as travelers reroute over time to find the shortest time path in a changing environment.

The base year VKT of the link influences the VKT in a positive and significant manner in 8 of the 12 models. The results in general indicate a positive influence on the VKT and are in line with what was hypothesized.

The coefficient for the change in the number of lanes on the parallel links is positive and significant in 7 of the 12 models. An increase in the number of lanes on the parallel links seems to cause the VKT on the link to rise thus contradicting our hypothesis. A possible explanation is that the increase in lanes on parallel routes is a response to rising overall traffic levels, and thus is correlated with growing traffic levels.

The results for the base year lanes on parallel links indicate a clear negative and significant influence on the VKT on the link in 8 of the 12 models. The higher the number of lanes on the parallel links, the lower will be the VKT on the link in future years, which is in line with the hypothesis.

The coefficient for the change in VKT on the parallel links is negative and significant in 8 of the 12 models. This indicates that the change in VKT on the parallel links in the previous years has an overall negative effect on the VKT on the link in future years. The base year VKT on the parallel links has a clear positive and significant influence on the VKT on the link in 8 of the 12 models. This indicates that as the base year VKT on the parallel links increases, the VKT on the links in future can also be expected to increase.

The change in population of the surrounding Minor Civil Division (MCD) has a positive and significant effect on the VKT in 7 of the 12 models. Overall, an increase in the population of the MCD to which the link belongs causes the VKT on the link in future years to increase, which is what was hypothesized. The coefficients for the base year population of the MCD to which the link belongs are negative and significant in 7 of the 12 models. The results are consistent with the hypothesis and indicate that larger stable jurisdictions do not produce a change in VKT, while growing MCDs do.

The change in the population of the adjacent MCDs is insignificant in 7 of the 12 models. Thus, in general, the change in population of the adjacent MCDs doesn't seem to have an influence on the VKT on the link in future years. The coefficients of the base year adjacent MCDs shows a very clear negative and significant effect on the VKT on the link in 11 of the 12 models, which contradicts our hypothesis.

Another point is that the base year dummy variables considered are in most cases highly significant and negative indicating that the change in VKT on the link has decreased over the years. Thus as the road system becomes more mature, changes become less significant. It may also indicate a general slowing in traffic growth, as the major changes of the past several decades (female labor force participation rates, auto ownership) have worked themselves through the system.

The results indicate that the VKT on a link in future years is affected by the flow and capacity conditions existing on the link in the previous years. The influence of the conditions existing on the neighboring links doesn't seem to be very clear from the results.

7.4.8 Elasticities

The elasticity estimates for the change in lanes on the link in the previous years obtained from the model are summarized in Table 8. The elasticity estimates indicate the effect a 1 per cent per cent change in the number of lanes on the link in the previous years would have on the VKT on the link in future.

Overall these results are consistent with the summary results shown in Table 7.4.1, though somewhat lower. The difference between modeling link expansions and new construction may explain some of the difference. It does not indicate that all new capacity is used up. Thus there are likely to be real travel time-savings due to link expansion after accounting for induced demand. The results do indicate that the capacity is used up more in the long term than the short term, so a short-term measurement of induced demand will underestimate long term effects. Still, in general the elasticities are below 0.6, so that the capacity increases are not simply overwhelmed by new traffic, and the induced demand phenomenon is not insatiable.

Table 7.4.8 Elasticity estimates: percentage change VKT/percentage change lanes on the link

Model	Interstates	Trunk highways	Hennepin county highways (CSAH)
2-year lag model	− 0.042	− 0.003	0.014
4-year lag model	− 0.137	0.011	0.034
6-year lag model	0.208	0.005	Not significant
8-year lag model	0.552	Not significant	0.067

7.4.9 Conclusions

The focus of this study was to analyze the induced demand hypothesis at a disaggregate level and also identify the influence of network interactions and dynamics. Studies on the induced demand hypothesis to date have mostly been aggregate level studies.

The reason that the link level approach was chosen is that the effects on changes in the network were expected to be clearly seen at the link level. Any capacity changes on a link in the network are likely to have a very clear effect not only on the link alone but also on the other links. The output of a link, measured in terms of VKT here, is going to be affected not only by the capacity changes occurring on it but also by the capacity changes on the neighboring parallel links. The analysis aimed to understand and measure this neighboring influence in inducing traffic on a link.

The results indicate that the induced demand hypothesis is largely, but not universally corroborated. The results obtained indicate that the neighboring parallel links do affect the output of a link. The population estimates also tend to affect the output. Though key conclusions have been drawn for some of the variables like the changes in capacity and flow, the signs and the coefficients obtained for the many of the other variables don't seem to follow any particular pattern.

This analysis is limited to examining expansion on exiting links. That is only those links that already existed and underwent lane additions from the period from 1978–1998 are considered. The induced demand effect is more obvious for other links that were newly added to the network.

A key conclusion that has however emerged from this study is that the capacity expansions on a link in the previous years do have a positive effect on future Vehicle Kilometres Traveled (VKT), which is in line with the induced demand hypothesis. The elasticity calculations show that a 1 per cent increase in the number of lanes on the link in the previous years causes a 0.01–0.6 per cent increase in the VKT on the link in future years.

Many have argued that the conventional four-step Urban Transport Modeling System is not an accurate description of trip making (Levinson and Kumar, 1994b). In particular, the representation of network flow would not be accurate if demand elasticity is not considered in the theoretical framework.

Levinson and Kumar (1994a) also noted that over the long term, commute times to work tend to be stable. This has been corroborated over a period with significant population growth and lifestyle changes that increased travel demand far in excess of road capacity increases. The induced demand argument must be placed in the larger context of travel and activity patterns. As capacity is added, it may be used up. But that does not mean benefits were not gained. People using the new capacity are now able to pursue opportunities that were too expensive in terms of time budgets previously, and thus are better off, even if only examining the congestion record would not indicate so.

7.4.10 References

1. Baltagi, B. H (1999) *Econometrics*, 2nd edn. New York: Springer Verlag.
2. Barnes, G. and Davis, G. (2001) *Land use and travel choices in the twin cities, 1958–1990*. Report No. 6 of the Transportation and Regional Growth Study, Center for Transportation Studies, University of Minnesota.
3. Dowling, R. G. and Colman, S. B. (1995) Effects of increased highway capacity: results of household travel behavior survey, *Transportation Research Record*, 1493, pp. 43–149.
4. Downs, A. (1992) *Stuck in Traffic.* Washington, DC: Brookings Institution.
5. Fulton, L. M., Noland, R. B., Meszler, D. J. and Thomas, J. V. (2000) A statistical analysis of the induced travel effects in the U. S. Mid- Atlantic region, *Journal of Transportation and Statistics*, 3(1), pp. 1–14.
6. Greene, W. H. (1993) *Econometric Analysis*, 2nd edn. New York: Macmillan Publishing Company. HANSEN, M. (1995) Do new highways generate traffic, *Access*, 7(fall), pp. 16–22.
7. Hansen, M. and HuAng, Y. (1997) Road supply and traffic in California urban areas, *Transportation Research A*, 31(3), pp. 205–218.
8. Hennepin County, Minnesota (1980–2000) *Hennepin County capital budget*. Hennepin County.
9. Kiefer, M. and Mehndiratta, S. R. (1998) *If we build it, will they really keep coming? A critical analysis of the induced demand hypothesis*. Annual Meeting Preprint Paper No. 980937, Transportation Research Board.
10. Levinson, D. M and Kanchi, S. (2000) Road capacity and the allocation of time, *Journal of Transportation and Statistics* (in press).
11. Levinson, D. M. and Karamalaputi, R. (2002) *Predicting network expansion*. Preprint CD-ROM, *Transportation Research Board Conference*, Washington, DC, January.
12. Levinson, D. M. and Kumar, A. (1994a) The rational locator: why travel times have remained stable, *Journal of the American Planning Association*, 60(3), pp. 319–332.
13. Levinson, D. and Kumar, A. (1994b) Integrating feedback into the transportation planning model, *Transportation Research Record*, 1413, pp. 70–77.
14. Metrogis Datafinder (http://www.datafinder. org).
15. Metropolitan Council (http://www.metrocouncil. org).
16. Metropolitan Council (1978–2000) Transportation improvement program for the twin cities metropolitan area. St Pail, MN: metropolitan council.
17. Minnesota State Demographic Center (http:// www.mnplan.state.mn.us/demography)
18. Mokhtarian, P. L., Samaniengo, F. J., Shumway, R. H. and Willits, N. H. (2002) Revisiting the notion of induced traffic through a matched-pairs study, *Transportation*, 29, 193–220.

19. Noland, R. B. (1999) Relationships between highway capacity and induced vehicle travel. Preprint CD-ROM, Transportation Research Board 78th Annual Meeting, Washington, DC, January.

20. Office of Transportation Data And Analysis, Transportation Data Section, Minnesota Department of Transportation (http://www. dot. state. mn. us/tda/products2. html).

21. Parthasarathi, P. K. (2001) *A disaggregate analysis of induced demand*. University of Minnesota.

22. Pfleiderer, R. H. H. and Dieterich, M. (1995) New roads generate new traffic, *World Transport Policy and Practice*, 1(1), pp. 29–31.

23. Pindyck, R. S. and Rubinfeld, D. L. (1998) *Economic Models and Economic Forecasts*, 4th edn. New York: Irwin McGraw-Hill.

24. Strathman, J. G., Dueker, K. J., Sanchez, T. et al. (2000) *Analysis of induced travel in the 1995 NPTS*. Center for Urban Studies, Portland State University.

25. US Census Bureau (2002) *Decennial Management Division Glossary* (http://www.census. gov/dmd/www/glossary.html#M).

26. Whistler, D. (2000) *An online guide to SHAZAM* (http://shazam.econ.ubc.ca/intro/index. html).

7.5 Road Pricing with Autonomous Links

Lei Zhang
Ph.D. Candidate
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-626-0024 F: 612-626-7750
E-mail: zhan0294@tc.umn.edu

David Levinson
Assistant Professor
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-625-6354 F: 612-626-7750
E-mail: levin031@tc.umn.edu

Transportation Research Record: Journal of the Transportation Research Board, No. 1932, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 147–155.

7.5.1 Introduction

Roadway congestion, air pollution from cars, and the lack of resources to finance new surface transportation options challenge many nations. Road pricing, the practice of charging users a monetary toll in addition to the “cost” of time spent traveling, has been suggested as a solution to these problems. Although tolls are common on certain expensive facilities such as tunnels and bridges, they are less common on streets and highways. However, a new generation of private toll roads are being deployed, most recently SR-91 in southern California and the Dulles Greenway in northern Virginia. There have been a few trials of area-wide pricing schemes, such as in Singapore and London, and many other schemes have been proposed but not implemented. The combination of private and competing toll roads, ubiquitous over an area, would represent a comprehensive market-oriented approach to urban transportation problems, but its impacts are unclear. This research seeks to examine road pricing on a network of autonomous highway links with the goal of understanding the social welfare and equity implications of widespread adoption of road pricing and privatization under various circumstances. “Autonomous” refers to the links’ being competitive and independent and having the objective of maximizing their own profits without regard for either social welfare or the profits of other links, though possibly being subject to regulatory constraints. The basic approach taken is to begin with the link as the most elemental unit of analysis and aggregate to more complex interactions, including revenue and cost sharing.

A realistic network of highway links is not, in the economists’ terminology, perfectly competitive. Because a link uniquely occupies space, it attains some semblance of monopoly power. Although in most cases users can switch to alternative links and routes, those alternatives will be more costly to the user in terms of travel time. Theory suggests that excess profits will attract new entrants into a market, but the cost of building a new link is high, indicating barriers to entry not easily overcome.

Although roads are generally treated as public goods, they are both competitive when congested and excludable. These factors indicate that it is feasible to consider them for privatization. Several advantages are often associated with privatization: increasing the efficiency of the transportation system through road pricing, providing incentives for the facility operator to improve service through innovation and entrepreneurship, and reducing the time and cost of building and expanding infrastructure.

Most trials of road pricing suppose either tolls on a single facility, or area-wide control. Theoretical studies assume marginal cost pricing on links and do not discuss ownership structure. However, in other sectors of the economy, central control of pricing through either government ownership or regulation has proven to be less effective than decentralized control for serving customer demands in rapidly changing environments. Single prices system-wide do not provide as much information as link-specific prices. Links that are priced only at marginal cost, the optimal solution in a first-best, perfectly competitive environment, constrain profit. Although excess profit is not socially optimal in the short term, over the longer term, it attracts capital and entrepreneurs to that sector of the economy. New capitalists will both invest more in existing technology to refine its deployment and enter the sector as competitors, trying to gain from a spatial monopoly or oligopoly. Furthermore, new capitalists may also innovate and thereby change the supply (and demand) curves in the industry.

By examining road pricing and privatization from a decentralized point of view, the issues associated with a marketplace of roads can be more fully explored, including short- and long-term distributional consequences and overall social welfare. The main contribution of this research will be to approach the problem from a theoretical and conceptual level and through the conduct of simulation experiments. Specifically, an agent-based simulation model is developed in this study that incorporates travel demand estimation, road maintenance and construction cost functions, pricing and financing strategies of autonomous links, and network performance evaluation. The simulation model considers both short-term traffic equilibrium and long-term supply and demand equilibrium in a highway network; it therefore more completely assesses the consequences of alternative ownership structures and pricing strategies.

7.5.2 Literature Review

Gomez-Ibanez and Meyer [1] have reviewed transportation privatization at an empirical level, though the cases of roadway privatization are few and not entirely successful. Even if roadways were privatized, it is unlikely that their price structure would be left entirely to the private sector. In many ways, roadways are natural monopolies, in that their provision and use has a declining average cost (aside from congestion effects). The proposed model contains within it several different forms of networks. Most obvious is the transportation network as a physical system, which has been well developed in the transportation literature [2]. Boyce et al. [3] investigated the optimal network problem from a global perspective and developed algorithms based on optimal subset selection, which were later modified and applied to medium-sized networks by Rothengatter [4]. In the model proposed in this paper, the links are autonomous and can interact with each other. This provides a second level of network: an economic network that considers coordination between firms (links). Johansson et al. [5] describe various economic networks from an empirical viewpoint, while Nagurney [6] provides a computational framework that links analysis of economic networks (supplier–customer relationships) with algorithms developed for the analysis of physical networks. Economides [7] compares the economic structure of networks with vertically related industries. Because the model being described here is inherently dynamic, it may not contain a neat equilibrium solution. A modeling approach using cellular automata suggests specifying simple rules and allowing the system to evolve [8]. The extent to which links can set prices following those simple rules and still achieve a maximum profit level can be ascertained with the model. Recently, Verhoef and Rouwendal explored interrelations between pricing, capacity choice, and financing in transportation networks using a small network model [9].

Although the focus of this study is on the economic interactions between links and the consequences of price strategies, the travel demand and travel time components of the model need to be specified [10–12]. The monetary costs of infrastructure provision, user operating costs, and social costs on highways as a function of flow have recently been estimated by several studies [13, 14], and these costs will be considered by the links in profit maximizing and the prices they charge. They need to be integrated and solved in both traffic equilibrium [2] and long-term supply–demand equilibrium. Zhang and Levinson developed an exploratory agent-based travel demand model [15]. However, in its present form, the model cannot perform all functions of trip-based demand models.

7.5.3 Agent-Based Network Dynamics Model

Few researchers have considered the process of transportation network growth (or decline) at the microscopic level, although long-term transportation network dynamics are important for assessing alternative pricing policies and institutional structures. Analytical models of network growth are not practical except under simple, idealized conditions, represented by very small networks and analyzed using the principles of transportation engineering, microeconomics, game theory, and industrial organization. Zhang and Levinson proposed a model of transportation network growth with average cost pricing and myopic investment rules that demonstrates the

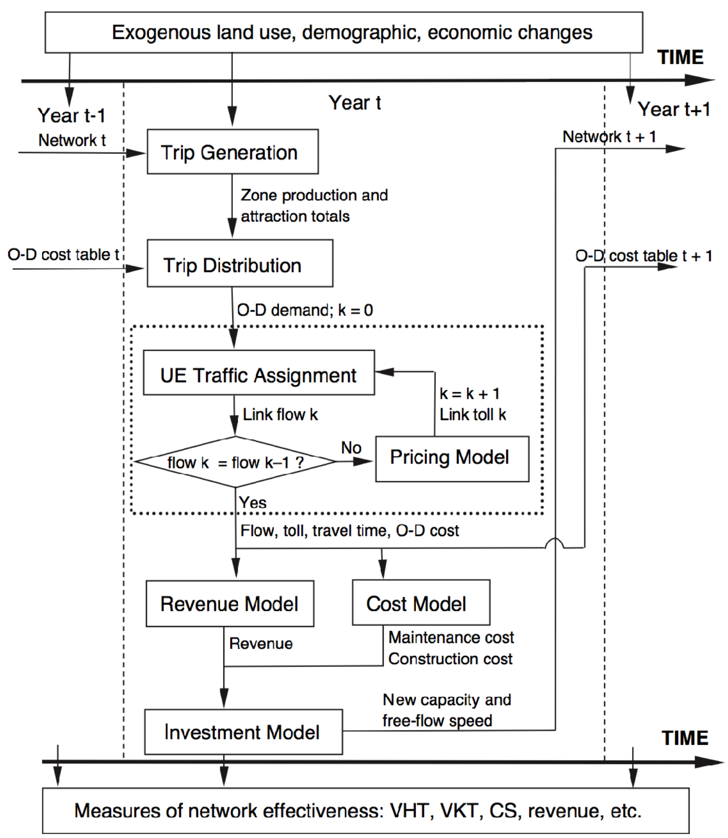


Figure 7.5.1 Flowchart of simulation model (UE = user equilibrium)

feasibility of an agent-based simulation approach for transportation-related policy analysis [16]. Their simulation model is extended in several ways in this study.

An overview of model components and their interconnectivity is shown in Figure 7.5.1. A travel demand model predicts link-level flows based on network, socioeconomic, and demographic information. Based on the demand forecasting results, links calculate revenues and costs. An investment module then operates and causes annual supply changes, producing an updated network. The transportation network is represented as a directed graph that connects nodes with directional arcs (links). The standard notation convention for directed graphs is adopted for the following presentation on the details of mathematical formulations of those submodels. The directed graph is defined as $G = \{N, A\}$, where N is a set of sequentially numbered nodes and A is a set of sequentially numbered directed arcs.

Notation

Notation used is as follows:

A_r	=	accessibility of zone r
CS^{0-i}	=	change in consumers' surplus from year 0 to i
$d(.)$	=	cost impedance function in the gravity model $d(t^i_{rs}) = e^{-\gamma t^i_{rs}}$
D_s	=	number of trips destined for zone s
DR^i_a	=	disposable revenue of link a in year i (dollars)
E^i_a	=	revenue (earnings) of link a in year i (dollars)
K^i_a	=	cost of expanding link a in year i (dollars)
f^i_a	=	average hourly flow on link a in year i (vehicles per hour)
F^i_a	=	capacity of link a in year i (vehicles per hour)
G	=	Gini coefficient of accessibility inequity
i	=	index of year
j, k	=	parameters in the decentralized pricing model
l_a	=	length of link a (constant) (km)
m_r, n_s	=	coefficients in the gravity model
M^i_a	=	cost of maintaining link a in year i (dollar)
O_r	=	number of trips produced from zone r
q^i_{rs}	=	demand from origin r to destination s in year i
t^i_a	=	generalized travel cost on link a in year i
t^i_{rs}	=	generalized travel cost from zone r to s
v^i_a	=	free-flow speed of link a (km/h) in year i
α_{1-3}	=	coefficients indicating (dis)economies of scale,
φ	=	scale parameter in expansion cost function,
γ	=	coefficient in the impedance function,
λ	=	value of travel time (dollar/h),
θ_{1-2}	=	coefficients of the BPR travel time function,
ρ_{1-3}	=	coefficients in the centralized pricing model,

σ_{1-3} = coefficients in the expansion cost model,
τ^i_a = link toll per vehicle (dollars; see Equation 4),
μ = scale parameter in maintenance cost function,
ω_{1-2} = coefficients in the capacity-speed model
ψ = coefficient to scale hourly flow to annual flow.

Travel Demand

A traditional four-step model is specified to estimate travel demand at the link level, taking exogenous land use, socioeconomic variables, and the existing network as inputs. Although the four-step model serves well for demonstration purposes in this paper, future studies should use more advanced travel demand models. For instance, combined travel demand models address inconsistencies in the sequential model by solving all steps in a coherent equilibrium [17]. Activity-based approaches [18] and agent-based microsimulation [15] improve behavioral representation in travel demand models. A zone-based regression structure is used for trip generation. The origin–destination (O-D) cost table obtained from the previous year's traffic assignment is used for trip distribution in the current year based on a doubly constrained gravity model [19, 20].

$$q^i_{rs} = m_r O_r n_s D_s \cdot d(t^i_{rs}) \tag{1}$$

The resulting O-D table is loaded onto the current year transportation network through the origin-based user equilibrium traffic assignment algorithm (OBA) [21]. The generalized link cost function comprises two parts: a BPR travel time component and a vehicle toll.

The OBA algorithm derives link flows at user equilibrium and generates a new O-D cost

$$t^i_a = \lambda \frac{l_a}{v^i_a}\left[1+\theta_1\left(\frac{f^i_a}{F^i_a}\right)^{\theta_2}\right]+\tau^i_a \tag{2}$$

table that will be used for trip distribution in the next year. In the traffic assignment step, if the relative excess travel cost is less than 0.001, the Wardrop user equilibrium [22] is considered to be satisfied.

Revenue and Cost Functions

Revenue is collected individually by autonomous links in a form of vehicle toll. The annual revenue is simply the product of the toll and annual flow. The amount of the toll depends on the pricing strategy adopted by an autonomous link agent. Therefore, the following revenue equation is proposed:

The link maintenance cost function has two determining factors in a Cobb–Douglas form:

$$E^i_a = \tau^i_a \cdot (\psi \cdot f^i_a) \tag{3}$$

link length and capacity. It costs more to maintain a link at its current level of service if the link is longer and carries heavier flow.

$$M_a^i = \mu \cdot (l_a)^{\alpha_1} (F_a^i)^{\alpha_2} \quad (4)$$

Link expansion cost is considered a function of link length, existing capacity, and additional capacity to be expanded. It is more expensive to expand a unit capacity on a link with higher existing capacity.

$$K_a^i = \phi \cdot (l_a)^{\sigma_1} \cdot (F_a^i)^{\sigma_2} \cdot (F_a^{i+1} - F_a^i)^{\sigma_3} \quad (5)$$

Investment Rules

Decentralized Autonomous Links. Two assumptions are made concerning the investment rules adopted by autonomous links. First, it is assumed that the system is closed and that all revenue will be spent to either maintain or expand links. Second, there is no incentive for links to save revenue (i.e., revenue accumulated in a year will be used in that year). These two assumptions could be relaxed if a bank agent is included in the simulation model and provides an endogenous interest rate. Disposable revenue is defined as the difference between total revenue and maintenance cost.

$$\mathrm{DR}_a^i = E_a^i - M_a^i \quad (6)$$

Disposable revenue of a link is used to expand that link. Therefore, one can substitute DR_a^i for K_a^i in Equation 5 and solve for the new capacity in year $i + 1$. It is possible that the disposable revenue of a link is negative because of previous overinvestment or competition. In that case, the link will shrink in the next year because total revenue falls short of maintenance cost. Note that this autonomous investment rule is myopic because links care only about themselves, ignore network effects, and spend all revenue immediately.

A capacity change is usually associated with a concurrent change of free-flow speed. Vehicles are able to travel at faster speeds on a wider road with less impedance. Free-flow speed and capacity data used by the Twin Cities Metropolitan Council in its regional transportation planning model on more than 10,000 roadway sections were used to study the correlation between speed and capacity. A log-linear model is specified and estimated. R^2 of the model is 0.7, and both coefficients are statistically significant at level 0.01.

$$v_a^{i+1} = \omega_1 + \omega_2 \cdot \ln(F_a^{i+1}) \quad (7)$$

With updated link capacity and free-flow speed, some factors influencing travel behavior, such as link travel time and link toll, will change. These supply shifts, combined with preference, economical growth, and demographical changes, give rise to the emergence of a new demand pattern.

Centralized Government Control. In contrast to decentralized investment decisions made by autonomous links, revenues collected on all links may be pooled together, and a central government agency may make all investment decisions. For comparison purposes, a centralized investment rule is examined. It is assumed that the central government can always adjust its pricing policy (see next section) so that total network revenue is higher than maintenance cost. The remaining network revenue is spent to expand existing links based on benefit–cost ratios.

The maximum possible benefit–cost ratio (BC_{max}) of expanding each link, as well as the corresponding optimal amount of expansion, is computed based on Equation 5 and the following assumptions: (*a*) Traffic increases by 4% every year; (*b*) interest rate is 3%; (*c*) value of time is \$10/h for all users; (*d*) the planning horizon is 30 years; (*e*) only local travel time benefits are considered. The network revenue is used to expand the link with the highest BC_{max}. Then, the link with the next-highest BC_{max} is expanded until the centralized revenue is exhausted. Similarly, a capacity change results in a new free-flow speed according to Equation 7.

7.5.4 Road Pricing

Decentralized Autonomous Links

Autonomous links seek to maximize their short-run profits in a competitive market. Because travel demand is elastic with respect to price, the profit-maximizing price is constrained by the market. Competing links also restrict the price that an autonomous link can charge and still maximize profits. It is anticipated that each link will have an objective function for profit maximization. However, depending on assumptions of whether the firm perfectly knows market demand and how the firm treats the actions of competitors, the Nash equilibrium solution to the problem may not be unique or even exist. Whether this system converges upon an equilibrium solution and whether that solution is unique are important questions that this research addresses.

The incompleteness of information is profound in the market comprised of non-cooperative competing autonomous links. The situation of incomplete information is further aggravated by the fact that the demand function on one link depends on its previous investment decisions and the pricing–investment decisions made by its competing and complementary links. How do autonomous links determine the profit-maximizing price in this dynamic situation? Underlying the decision of each autonomous link is an objective function, profit maximization given certain amounts of information, and a behavioral rule that dictates the amount and direction of price changes depending on certain factors. Once a link has found a toll that it can neither raise nor lower without losing profit, it will be tempted to stick with it.

Therefore, it is assumed that links try to achieve profit maximization in this interdependent and evolving system by adjusting their prices iteratively based on available information about link travel demand. In each iteration, a link determines its price based on prices and profits in the previous *k* iterations. Specifically, a link fits a quadratic curve in the profit–price domain. If the curve is concave, the new price is identified at the maximum point. If the curve is convex, the price corresponding to the maximum profits in the previous *k* iterations will be marked up or down by *j* percent to form the new price (see Figure 7.5.2). This pricing rule helps the link maximize profit and keep the price changes small. A myopic pricing rule is plausible when demand functions are unknown to autonomous links. The assumption of unknown demand will be checked in the later simulation experiment. If the demand functions turn out to be relatively stable from iteration to iteration (i.e., a reasonably accurate demand curve can be estimated after several trials), the proposed pricing rule needs to be revised because there are obviously better pricing strategies for profit-maximizing links.

However, a more intelligent link may realize that although it may have found a local maximum, because of the nonlinearities that make up a complex network, it may not be at a global maximum. Furthermore, other links may not be so firmly attached to their decision, and a periodic probing of the market landscape by testing alternative prices is in order. This too requires rules and should be explored in future studies.

It should be noted that a homogenous user group is assumed in this study. Several recent studies show that the ignorance of user heterogeneity and the possibility of product differentiation cause underestimation of the benefits of road pricing and decentralized control (*23–25*). The network growth model described above needs to use a multiclass travel demand model to account for variation in value of time, which should be pursued in future studies.

Centralized Government Control

Under centralized control, users pay a distance-based toll for using the roads. This method is similar to a fuel tax except that the variation of fuel efficiency among vehicles is ignored. Free-flow speed is also included in the centralized pricing model because it also affects fuel efficiency and hence the actual prices that users pay.

$$\tau_a^i = \rho_1 \cdot (l_a)^{\rho_2} \cdot (v_a^i)^{\rho_3} \tag{8}$$

Note that even a link-based congestion pricing rule would improve the efficiency of centralized control. Equation 8 is selected because it better describes the current road financing practice in most areas.

7.5.5 Simulation Experiments

So far, a complete cycle of the network evolution process has been modeled, and all elements in the flowchart have been specified. This demand-cost-pricing-investment cycle repeats itself year after year. It is possible to simulate the growth of the network with alternative pricing and investment policies. Measures of effectiveness collected from these agent-based simulation experiments are valuable for policy evaluation. The issue of whether a transportation network evolves better under centralized or decentralized control can be explored.

Estimation of Model Coefficients

Most coefficients in the specified network dynamics model are estimated based on empirical data

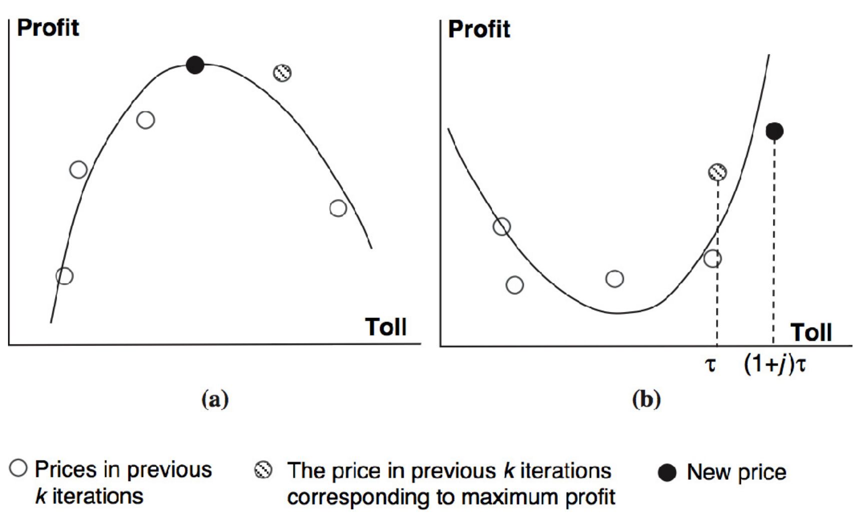


Figure 7.5.2 Pricing rule for autonomous links: (*a*) concave and (*b*) convex

in the Twin Cities (Minneapolis and St. Paul, Minnesota). Estimation of the cost function (Equations 4 and 5) is documented in detail in Levinson and Karamalaputi [14]. Several power coefficients are determined based on the authors' best knowledge of economies or diseconomies of scale in transportation network dynamics. A summary of estimation methods and results is shown in Table 7.5.1.

Measures of Effectiveness

The network dynamics model provides the following information for each year in the evolutionary process: population and activities at the zone level; demand, travel time, and generalized travel cost at the O-D level; and flow, capacity, speed, travel time, and toll at the link level. This information is used to develop several measures of effectiveness (MOE) for the evaluation of network performance over time. Total vehicle hours traveled (VHT) and total vehicle kilometers traveled (VKT) are fairly standard network MOEs. The change in consumers' surplus between year 0 and year *i* is approximated by the rule of half. Total net social benefit is the sum of changes in consumers' surplus and toll revenue.

$$\mathrm{CS}^{0-i} = \sum_r \sum_s \frac{1}{2}(q_{rs}^i + q_{rs}^0) \cdot (t_{rs}^0 - t_{rs}^0) \tag{9}$$

Accessibility to activities for residents in zone *r* is

$$A_r = \sum_s D_s \cdot d(t_{rs}) \tag{10}$$

That equity may be ignored is a concern in a privatized network in which toll revenue is the primary financing source. The Gini coefficient is used to measure the inequity of accessibility among different network zones, which falls between 0 (perfectly equitable) and 1 (perfectly inequitable). If the results show degraded equity with autonomous links, it is necessary to identify winners and losers.

$$G = \frac{\sum_r \sum_s |A_r - A_s|}{2N \sum_r A_r} \tag{11}$$

The average and distributional properties of link price and capacity will also be examined because they reveal the degree of network hierarchy that may differ under different institutional structures. Profitability for autonomous links is also of higher interest for its implications on viability of decentralized control.

Sample Network

The simulation system can be applied to any realistic roadway network. The execution time is mainly determined by the convergence speed of the traffic assignment model. A 10-by-10 grid network (100 nodes and 360 links) is used herein to explore the consequences of road pricing under alternative institutional structures. The same initial condition is specified for both simulation scenarios: centralized government control and decentralized autonomous links.

All links in the grid network are 4 km in length and have an initial capacity of 735 vehicles per hour (this value corresponds to a one-lane road according to a regression analysis using the

Table 7.5.1 Coefficients in Simulation Experiments

Parameter	Value	Source
λ	10	Empirical finding
θ_1, θ_2	0.15, 4	BPR function
γ	0.1	Empirical finding
ρ_1, ψ, ϕ	1	Scale parameters
ρ_2	1	CRS of link length
ρ_3	0.75	DRS of level of service
μ	20	Scale parameter
α_1	1	CRS of link length
α_2	1.25	DRS of capacity
σ_1	0.5	Empirical finding
σ_2	1.25	Based on empirical findings
σ_3	1	CRS of additional capacity
ω_1, ω_2	–30.6, 9.8	Empirical estimates
k, j	5, 0.2	Link behavior assumption

BPR = Bureau of Public Roads; CRS = constant returns to scale; DRS = decreasing returns to scale.

capacity and number-of-lane data in the Twin Cities). The initial network is heavily congested, with an average volume capacity ratio of 0.8 and an average speed of about 10 km/h (because road pricing and privatization are usually not considered for uncongested networks). The initial land use is uniform among all 100 network zones, with 10,000 trips originating from and destined for each zone, respectively. Convergence of the simulation model can be measured directly by the number of expansion activities in the network. Under centralized control, the network achieves the long-run supply–demand equilibrium if the total revenue is equal to the total required maintenance cost. With autonomous links, the equilibrium is achieved when the revenue is equal to the maintenance cost on each link.

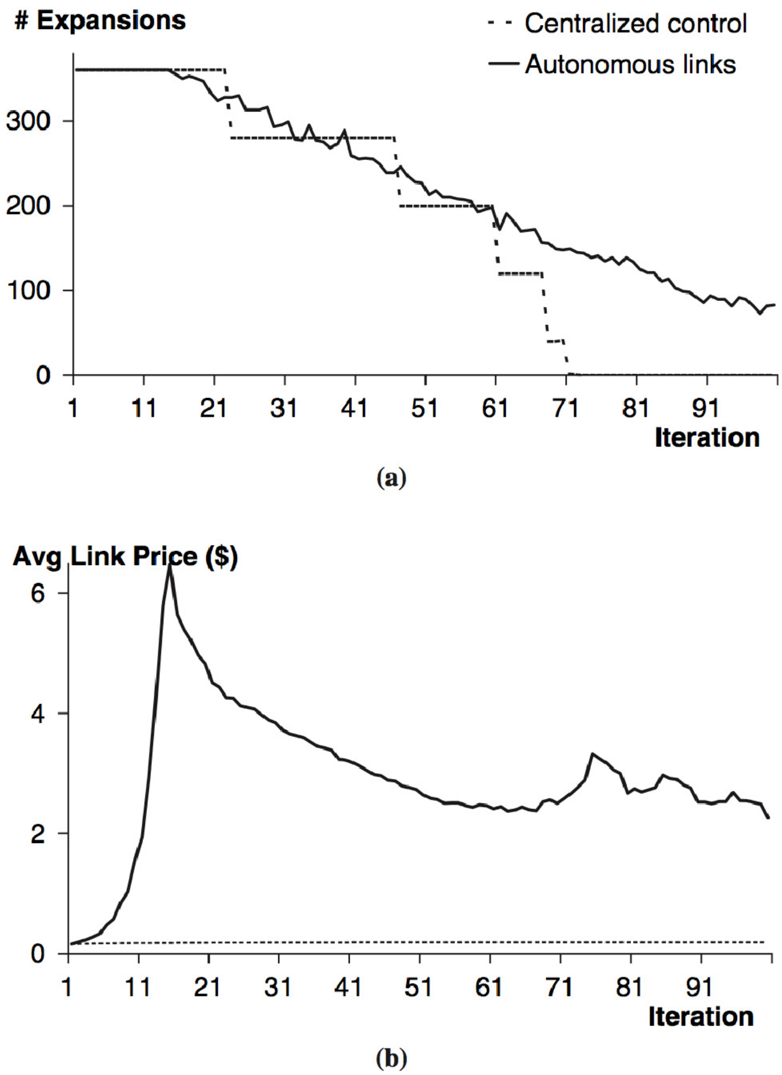


Figure 7.6.3 Convergence properties

7.5.6 Results

The long-term supply and demand in the grid network seem to equilibrate under both centralized and decentralized control. All links are expanded at the beginning of the evolutionary process because of initial congestion. After about 70 iterations (or years), a stable equilibrium is achieved under centralized control (see Figure 7.5.3). It takes longer for the scenario with autonomous links to arrive at an exact equilibrium, which is expected. Although there are still road expansions after 100 iterations, those expansions are characterized by extremely small changes in capacity. By examining the evolution of link prices, we can better observe the equilibrating process with autonomous links. When all links are privatized and start to make their own pricing decisions at iteration 0, there is a fast increase in link prices because of heavy congestion. After about 15 years of significant capacity expansion financed by abundant toll revenue, links must reduce their prices to maximize profits. The average link price continues to drop and eventually stabilizes itself around $2.5 after more than 80 iterations. There are still instabilities in the network in that any changes in individual pricing decisions may cause the system to fluctuate through ripple effects.

It is important to construct and understand link travel demand curves, which determine the pricing strategies autonomous links are likely to follow. The demand information at the link level is also valuable for the evaluation of alternative organizational structures for road financing and

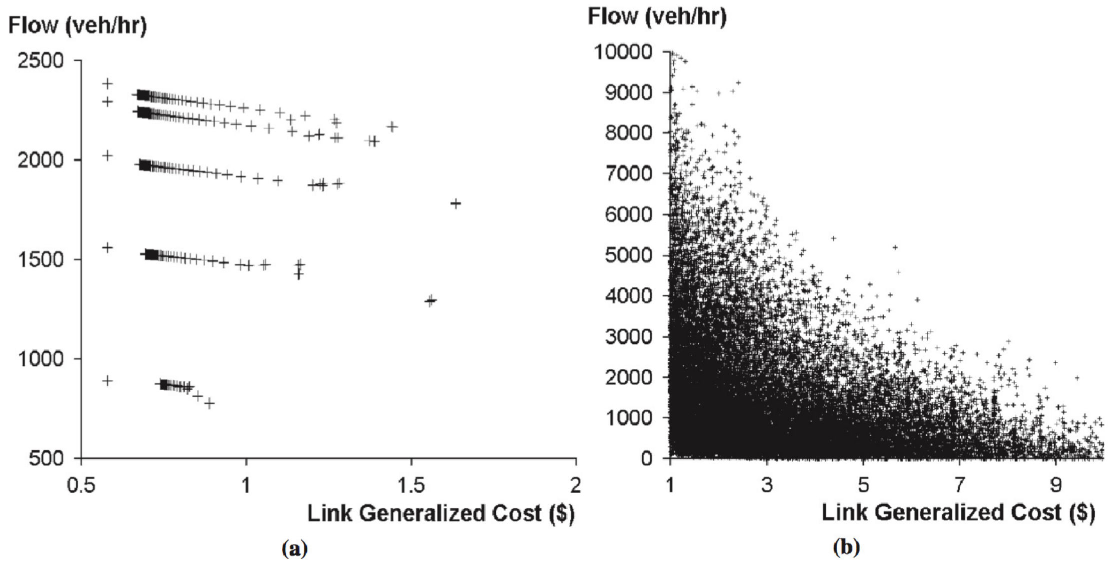
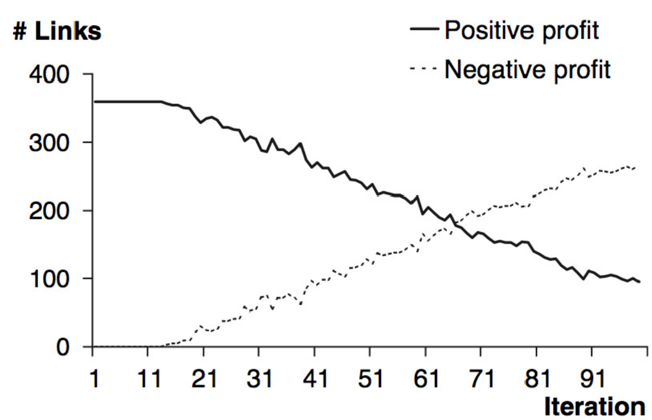


Figure 7.5.4 Link demand functions: (a) centralized control and (b) autonomous links

pricing. The two graphs in Figure 7.5.4 are created by aggregating flows and generalized travel costs of all links in all simulation iterations. Under centralized control in which prices are strictly determined by the length and the level of service of individual links, a linear relationship exists between demand and price. There are several parallel straight lines in Figure 7.5.4*a* because links naturally evolve into five categories based on their locations in the network. However, with autonomous links making pricing decisions non-cooperatively, the variation of demand at the same price level is so enormous that it is almost impossible for a link to identify the theoretically best price. The myopic pricing rule is reasonable given these results. With such a level of uncertainty and interdependency, links may be forced to adopt practical pricing strategies based on available information accumulated in their previous trials. Another implication is that in a situation in which several private profit-maximizing links compete with many public roads managed by a centralized government, it should be possible and rewarding for those autonomous links to estimate demand with a reasonable degree of accuracy.

Another interesting question is whether a set of autonomous competing links is a viable institutional structure for road financing and pricing. Can all links make and sustain a profit? The answer is negative if links do not properly manage their revenue. The number of links that can manage to generate a positive profit continues to drop over the years (see Figure 7.5.5). At the end of the simulation, only about 100 links still make money, whereas more than 200 others lose money. This is because autonomous links overinvest early in the evolutionary process when high prices bring in significant revenue, and they suffer high maintenance costs later on. This undesirable situation can be avoided by using a price ceiling regulation or heavy regulation on road expansions. Autonomous links should also be advised to use toll revenue to invest other sectors with a potentially higher rate of return. The overinvestment phenomenon under decentralized control could be partly attributed to the assumption that autonomous links expand capacity until short-run profit is zero. Future studies may develop more intelligent investment rules under decentralized control by modeling learning behavior. An interesting note is that, historically, overinvestment in capacity or a specific transportation technology is often related to decentralized control and competition (e.g., the U.S. railroad industry in the 19th century). Figure 7.5.6 presents the equilibrium tolls and capacity under both ownership structures. The existence of spatial monopoly is evident: the autonomous links at the corners and on the edges face less competition and lower demand elasticity, whereas links in the center must charge low

Figure 7.5.5 Profitability of autonomous links



tolls because of the existence of many parallel competitors.

Hierarchy has been long observed in road networks. Most roads have low capacity and carry low flows, whereas only a few roads are expanded to very high capacities and carry the bulk of traffic. Many believe road hierarchy is designed intentionally by planners. The results presented in Figure 7.5.6 clearly suggest that such structure also emerges in a free market driven purely by profit-maximizing behavior. A hierarchical structure is usually more efficient but has serious reliability and vulnerability problems. This infrastructure design issue is addressed in Zhang and Levinson [26].

As stated previously, one of the most important tasks of this study was to examine the welfare implications of road pricing with autonomous links. Various measures of network effectiveness developed in the previous section are computed for both centralized control (CC) and autonomous links and are summarized in Figure 7.5.7. Traditional network efficiency measures favor privatization. Average network travel speed is consistently higher with autonomous links over time (total VKT is comparable between the two scenarios, whereas total VHT is lower with autonomous links). Accessibility is, however, lower with autonomous links because faster speed is achieved mainly by pricing some users off the roads or forcing them to travel shorter distances. According to the computed changes in consumers' surplus, consumers actually gain more under CC than with autonomous links. The changes in CS are negative because the base case CS is calculated with the assumption of free-flow speed. On the supply side, autonomous links collectively charge much more to the users than does a central government agency, as shown by the revenue comparison. Net social benefit, defined as the sum of toll revenue and monetized changes in CS, is also smaller with autonomous links than under CC. The reason that the pricing strategy of autonomous links is inferior to completely regulated pricing under CC is manifold. First, the road market is not perfectly competitive. Spatial monopoly exists and some links (e.g., those near the corners of the grid network) are more dominant than others. Second, autonomous links do not have reliable demand information because of intractable spatial dependencies on competing and complementary links. In the simulation experiment, this lack of information leads to myopic non-optimal pricing behavior.

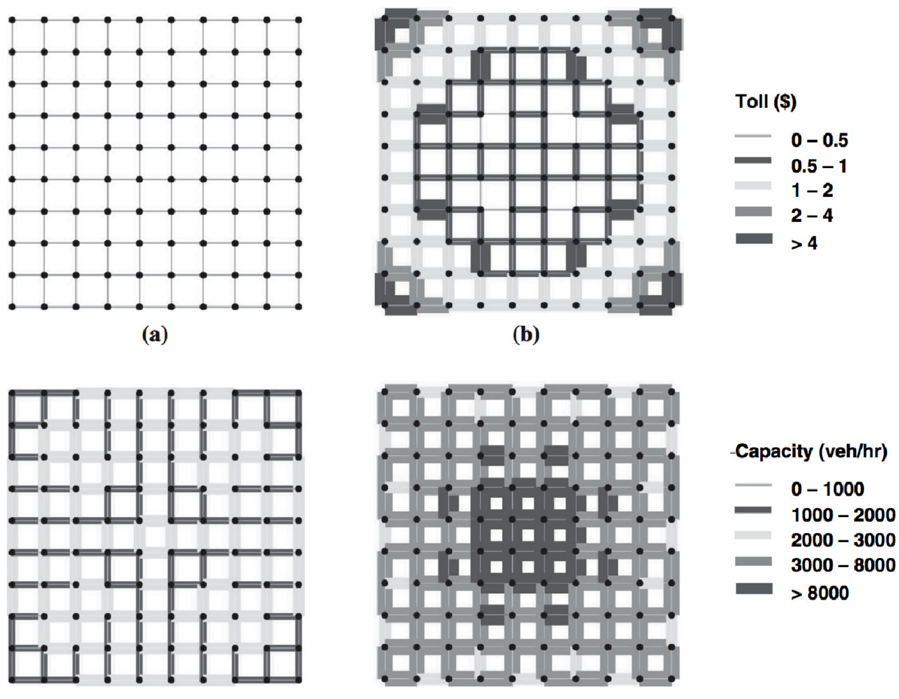


Figure 7.5.6 Equilibrium tolls and capacity: (*a*) centralized toll, (*b*) decentralized toll, (*c*) centralized capacity, and (*d*) decentralized capacity

Third, autonomous links adopt myopic investment behavior because of a lack of foresight that leads to overinvestment. Finally, in the proposed model, links are not allowed to cooperate or consolidate into more efficient structures. Revenue or cost sharing may be beneficial for individual links and the system as a whole.

Another observation is that the network is less equitable with autonomous links. The Gini coefficient is significantly higher in a privatized market. Therefore, the equity issue should be addressed when road pricing with autonomous links is considered. Some kinds of "basic access" criteria may improve equity but require government intervention.

7.5.7 Conclusions

This research develops an agent-based simulation model to study the problem of road pricing on a highway network composed of independent, profit-maximizing links. It addresses some issues around road pricing and privatization that have not previously been seriously considered, and it compares welfare and profit consequences of alternative organizational structures. The proposed modeling system integrates an equilibrated travel demand, route choice, and travel time model with a repeated road pricing game between autonomous links. Although a game theoretical approach seems to be more appealing, it is extremely difficult to model the payoff structure in a general network. The agent-based system can serve as a test bed for assessing long-term consequences of various transportation network investment and pricing policies and institutional structures. Another possible application of the system is to explore "free" roads that compete with toll roads and the consequences of regulatory constraints. Although this paper focuses on highway networks, the agent-based simulation approach could be used to analyze other types of networks with appropriate demand and cost functions.

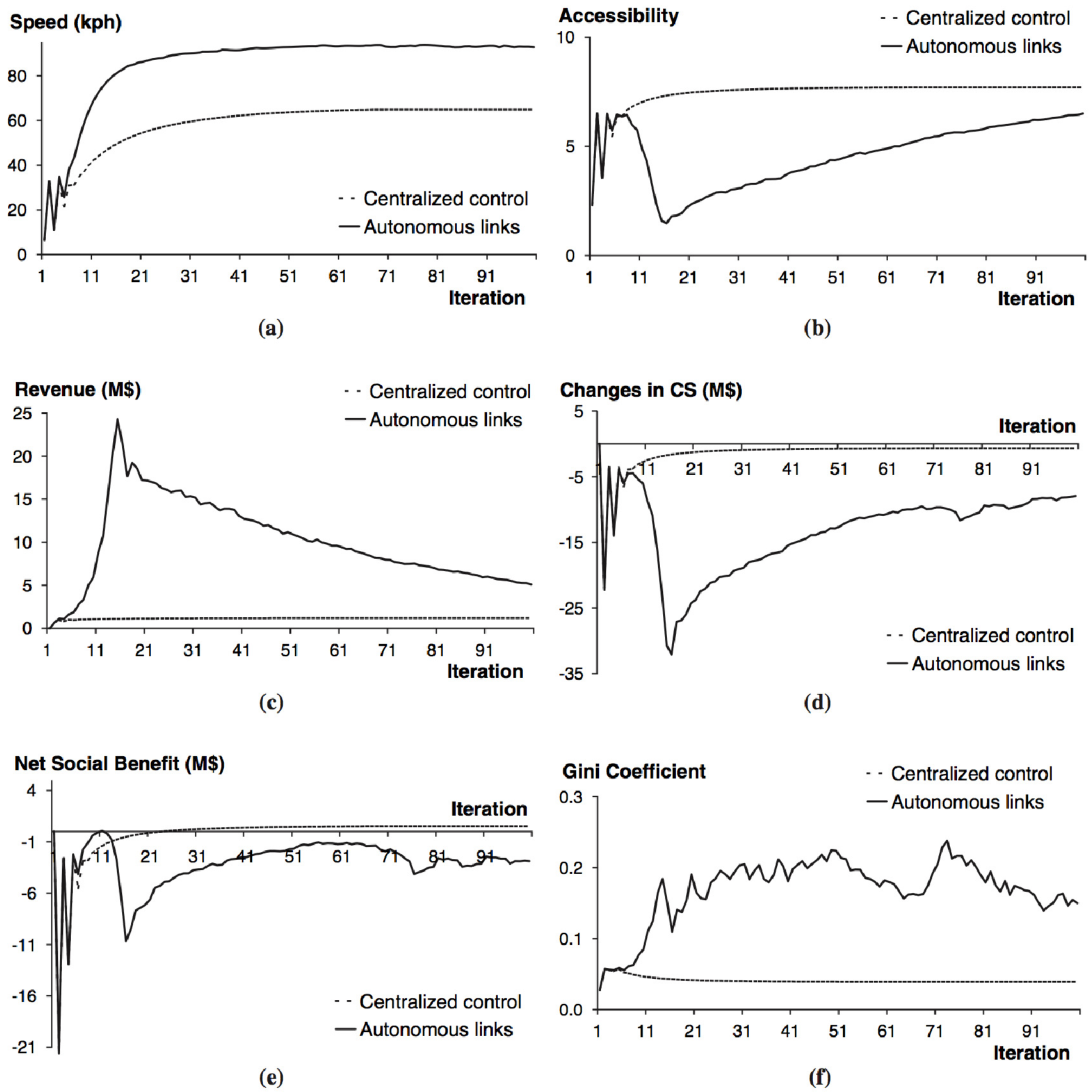


Figure 7.5.7 Results: measures of effectiveness

The existence of spatial monopoly, spatial dependence, and demand uncertainty may force independent links to adopt myopic non-optimal pricing and investment strategies, which in turn results in inferior social welfare compared to centralized control. For the same reason, many autonomous links eventually lose money in the pricing game. Some degree of government intervention in the form of price ceilings or restrictions on road expansion may prevent overinvestment. However, this is not to say that a market solution to highway financing and pricing is worthwhile. Besides non-cooperative independent links, there are alternative organizational structures that may improve both private and social welfare.

One limitation of the research is that cooperation among autonomous links is assumed away. Just as airline networks seem to have evolved a hub-and-spoke hierarchy, a specific geometry may be optimal in a private highway network. There may be advantages to both the private and social welfare if vertical integration of highly complementary links is allowed in the system. However, the degree of complementarity for which integration serves both public and private interests remains to be determined. As in other multi-agent systems, the critical issue here is the behavior of the decision makers: the autonomous links. How do coalitions between links form? In what circumstances will links pursue revenue and cost sharing? An interlink negotiation process must be developed to answer these important questions. The link pricing rule itself may be adjusted in the evolutionary process through learning and adaptation. None of these efforts is easy work. However, it would be interesting to see what kind of organizational structure emerges to take advantage of economies of scale in the network.

7.5.8 References

1. Gomez-Ibanez, J., and J. Meyer. *Going Private: The International Experience with Transport Privatization.* Brookings Institution, Washington, D.C., 1993.
2. Sheffi, Y. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice Hall, Englewood Cliffs, N.J., 1985.
3. Boyce, D. E., R. Farhi, and R. Weischedel. Optimal Subset Selection: Multiple Regression, Interdependence and Optimal Network Algorithms.
4. Springer-Verlag, New York, 1974.
5. Rothengatter, W. Application of Optimal Subset Selection to Problems of Design and Scheduling in Urban Transportation Networks. *Transportation Research,* Vol. 13B, No. 1, 1979, pp. 49–63.
6. Johansson, B., C. Karlsson, and L. Westin (eds.). *Patterns of a Network Economy.* Springer-Verlag, Berlin, 1993.
7. Nagurney, A. *Network Economics: A Variational Inequality Approach.* Kluwer Academic Publishers, Norwell, Mass., 1993.
8. Economides, N. The Economics of Networks. *International Journal of Industrial Organization,* Vol. 16, No. 4, 1996, pp. 673–699.
9. Langton, C. (ed.). *Artificial Life.* Addison-Wesley, Redwood City, Calif., 1989.
10. Verhoef, E. T., and J. Rouwendal. Pricing, Capacity Choice, and Financing in Transportation Networks. *Journal of Regional Science,* Vol. 44, No. 3, 2004, pp. 405–435.
11. Ben-Akiva, M., and S. Lerman. Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge, Mass., 1985.
12. Train, K. Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand. MIT Press, Cambridge, Mass., 1986.
13. Newell, G. *Applications of Queuing Theory,* 2nd ed. Chapman and Hall, New York, 1982.
14. Levinson, D., D. Gillen, J. Mathieu, T. Reinhold, and A. Kanafani. *The Full Cost of Intercity Transportation: A Comparison of Air, Highway, and High Speed Rail in California.* Research Report UCB-ITS-RR-96-3. Institute of Transportation Studies, University of California–Berkeley, 1996.
15. Levinson, D., and R. Karamalaputi. Induced Supply: A Model of Highway Network Expansion at the Microscopic Level. *Journal of Transport Economics and Policy,* Vol. 37, No. 3, 2003, pp. 297–318.
16. Zhang, L., and D. Levinson. Agent-Based Approach to Travel Demand Forecasting: Exploratory Analysis. In *Transportation Research Record: Journal of the Transportation Research Board, No. 1898,* Transportation Research Board of the National Academies, Washington, D.C., 2004, pp. 28–36.

17. Zhang, L., and D. Levinson. A Model of the Rise and Fall of Roads. Presented at 50th North American Regional Science Council Annual Meeting, Philadelphia, Pa., Nov. 20–22, 2003.

18. Boyce, D. E. Is the Sequential Travel Forecasting Paradigm Counterproductive? *Journal of Urban Planning and Development,* Vol. 128, No. 4, December 2002, pp. 169–183.

19. Kitamura, R. An Evaluation of Activity-Based Travel Analysis. *Transportation,* Vol. 15, 1988, pp. 9–34.

20. Haynes, K. E., and A. S. Fotheringham. *Gravity and Spatial Interaction Models.* Sage Publications, Beverly Hills, Calif., 1984.

21. Hutchinson, B. G. Principles of Urban Transportation Systems Planning. McGraw-Hill, New York, 1974.

22. Bar-Gera, H., and D. Boyce. Origin-Based Algorithm for Combined Travel Forecasting Models. *Transportation Research,* Vol. 37B, No. 5, 2003, pp. 405–422.

23. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. *Proc., Institution of Civil Engineers, Part II,* Vol. 1, No. 36, 1952, pp. 325–362.

24. Arnott, R., A. de Palma, and R. Lindsey. Route Choice with Heterogeneous Drivers and Group-Specific Congestion Cost. *Regional Science and Urban Economics,* Vol. 22, No. 1, 1992, pp. 71–102.

25. Shmanske, A. A Simulation of Price Discriminating Tolls. *Journal of Transport Economics and Policy,* Vol. 27, No. 3, 1993, pp. 225–235.

26. Small, K. A., and J. Yan. The Value of "Value Pricing" of Roads: Second-Best Pricing and Product Differentiation. *Journal of Urban Economics,* Vol. 49, No. 2, 2004, pp. 310–336.

27. Zhang, L., and D. Levinson. Investing for Reliability and Security in Transportation Networks. Presented at 84th Annual Meeting of the Transportation Research Board, Washington, D.C., 2004.

28. The Road Pricing Subcommittee of the Taxation and Finance Committee sponsored publication of this paper.

7.6 Agent-Based Approach to Travel Demand Modeling

Exploratory Analysis

Lei Zhang
Ph.D. Candidate
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-626-0024 F: 612-626-7750
E-mail: zhan0294@tc.umn.edu

David Levinson
Assistant Professor
Department of Civil Engineering
University of Minnesota
500 Pillsbury Drive SE
Minneapolis, MN 55455
V: 612-625-6354 F: 612-626-7750
E-mail: levin031@tc.umn.edu

7.6.1 Introduction

Contemporary models of urban passenger travel demand date from the 1950s [1, 2]. Aggregate demand models that relate the consumption of goods to the attributes of the goods, the competing goods, and consumer characteristics were found inappropriate for travel demand modeling because of both their inability to test some important transportation-related policies and the complexity of the transportation system itself. Therefore, a disaggregate or behavioral approach has attracted most of the research interest in the past several decades. Disaggregate travel demand models directly assume the behaviors of real-world decision-making units such as an individual or household. Discrete choice analysis based on random utility theory has been widely adopted, and individuals are assumed to always select the alternative that maximizes their utilities [3].

Urban travel demand results from a multidimensional hierarchical choice process. A list of such choices includes residential and business location, automobile ownership, and when to make a trip, with whom, from where to where, by which mode, and by which route. Some studies suggest that travelers, by developing heuristics, may only be able to find a feasible, not necessarily global, optimal solution to the choice problem subject to a set of constraints [4–6]. However, it is difficult to consider all these choices in one single model, although an integrated model is the final goal. Also, even with today's computing power, an integrated model will inevitably require some strict assumptions that will reduce its application value to local specific problems. The classical way to forecast the results of such a complex choice process is to divide it into simpler subprocesses in a logical and tractable way. Models for these subprocesses are then developed individually, and the hope is that they can eventually be assembled to provide useful predictions for decision makers. The past half-century has witnessed several different methods of disentangling the complex travel decision-making process. Two major approaches have emerged over time: trip- and activity-based approaches.

The traditional four-step travel forecasting models are often referred to as trip-based approaches in that they treat individual trips as the elementary subjects. In so doing, the four-step model tends to ignore the diversity among different individuals and considers aggregate travel choices in four steps—trip generation, trip distribution, mode split, and route assignment. Other choices are either treated as exogenous (e.g., land use and automobile ownership) or extremely simplified (e.g., trip scheduling). An up-to-date summary of the achievements in this field can be found in the book by Ortuzar and Willumsen [7]. There is some disagreement about how to assemble these four sub-processes in travel forecasting. Some researchers are of the opinion that the four steps should be solved in a coherent network equilibrium instead of sequentially. Boyce [8] provides a thorough review of the origin and the recent development of that issue.

An important nature of travel demand ignored by trip-based approaches is that travel is a derived demand—travel is desired to participate in other activities, not for its own consumption value. In view of this and other inadequacies of the four-step model, activity analysis has been applied to travel demand analysis since the 1970s. Activity-based approaches describe which activities people pursue, where, when, and for how long given fixed land use, transportation supply, and individual characteristics. A trip is generated to connect two spatially separated sequential activities. In activity-based approaches, every individual is a decision maker who confronts a huge choice set of various activity patterns in the time-space domain. Each combination of activities and their locations, starting points, and durations forms a unique activity pattern. Individuals select (or at least intend to select) the patterns that maximize their utilities by somehow solving a large-scale combinatorial optimization problem conditional on others' decisions. Different from the trip-based models, activity-based approaches deem individuals' decision making as sub-processes of the emergence of travel demand. These sub-processes are typically assembled by microscopic travel simulation to form aggregate travel forecasting. At the current stage of activity-based approaches, route choice and sometimes mode choice are still modeled by external modules such as dynamic traffic assignment algorithms. Several publications mark the milestones in the advance of activity-based approaches [9–12]. More-recent research progress is reported by Ettema and Timmermans (*13*) and McNally and Recker [14], among others.

After a half-century of continuous development, travel demand models now play an important role in urban planning and transportation–land use policy evaluation. However, there is still much room for improvement, notably that understanding the nature and dynamics of individual travel behaviors and their interactions is not adequate; the trend of disaggregate modeling requires faster solution algorithms as more and more complicated travel behaviors are modeled. Of course, these problems cannot be solved in a single study. Improving the existing travel demand models is not the purpose here. Rather, a new agent-based travel forecasting paradigm and a pilot agent-based travel demand model are proposed that may open a new door to solution of the problem.

Agent-based modeling methodology has a long lineage, beginning with von Neumann's [15] work on self-reproducing automata. Modern agent-based models employ methods from many fields, including artificial intelligence, cellular automata, genetics, cybernetics, cognitive science, and social science. The agent-based structure, flexibility, and computational advantages have made them powerful tools in modeling complex systems. In general an agent-based model consists of three elements: agents, an environment, and rules. Agents are like people, who have characteristics, goals, and rules of behavior. They are the basic unit of activity in the model. The environment provides a space in which agents live. Behavioral rules define how agents act in the environment and interact with each other. The characteristics of the environment itself also change in response to agent activities. Agent-based modeling techniques have found many applications in transportation. A recent special issue of *Transportation Research* [16] is dedicated to this topic. Microscopic traffic simulation can be viewed as an example of agent-based models. Vehicles are agents in the simulator, and a static road network is the environment. Vehicles are "born" at the entrances of the network and "die" at the exits. Rules, such as free-flow driving, car-following, and lane-changing, define how a vehicle behaves and interacts with other vehicles and the road network.

To apply the agent-based modeling method to a transportation demand system, one needs to define first the agents involved in the system and then the characteristics of each type of agent. Rules of agent behaviors need to be properly constructed in order to make the resulting model useful in travel forecasting. Given an initial condition, all the agents will behave on the basis of their "personal" characteristics, learning, and interacting rules. The transportation system will then evolve to a pattern, perhaps an equilibrium, from which useful macro-level information can be extracted. In this sense, travel demand would be the result of an evolutionary process.

An agent-based travel demand model is developed in the next section, followed by the application of the proposed model on both a hypothetical grid network and a realistic metropolitan area. With these two examples, computational properties and possible calibration procedures of the model are explored. Potential extensions of the model and future research directions are discussed.

7.6.2 Model

An agent-based travel demand model is formulated for a monomodal transportation network. Several agents in the transportation system are identified, as well as their characteristics and interacting rules, which enable the model to perform trip distribution and route assignment.

Agents and Their Characteristics

A transportation network in the model is fully represented by nodes and arcs as in a directed graph. The model considers three types of agents: traveler, node, and arc.

Traveler Agents. There are a certain number of traveler agents in the system. The goal of each traveler agent is to find an activity and to reach the activity with the lowest travel costs. Hence the first property a traveler agent has is status, which is a binary variable: an activity found (1) or not (0). In the process of searching for an activity, each traveler visits a set of nodes at which opportunities (potential activities) are located. At each step, each traveler moves from its current node to another through the connecting arc and decides to either accept or reject the opportunities at the new node on the basis of some rules, explained in the next section. Travelers learn arc costs along their search path when traveling on the network. Therefore, by adding arc costs, travelers know the total cost of a path from any node in their search path to each of the subsequent nodes, which is then added to the exchangeable knowledge base.

Node Agents. Nodes contain "demographic" and "social-economic" information of the system in terms of a_i number of travelers and b_i number of opportunities at node *i*. If a directed arc originates from Node 1 and is destined for Node 2, then Node 1 is called a supply node of Node 2 and, alternatively, Node 2 is a demand node of Node 1. Each node has a vector of S supply nodes $\boldsymbol{S}(s_1, ..., s_S)$ and a vector of D demand nodes $\boldsymbol{D}(d_1, ..., d_D)$ based on the transportation network structure. A node is also a supply node and a demand node by itself.

Node agents have two primary goals. First, whenever information exchange is possible, each node wants to either learn from travelers the shortest paths from other nodes to itself or distribute that information back to travelers, depending on whose knowledge is superior. For that purpose, nodes must store shortest-path knowledge and be able to exchange information with other agents. The second objective of node agent i is to provide turning guidance to travelers through an ($S \times D$) matrix P_i:

$$P = \begin{bmatrix} P_{s_1,d_1} & \cdots & P_{s_1,d_D} \\ \cdots & \cdots & \cdots \\ P_{s_S,d_1} & \cdots & P_{s_S,d_D} \end{bmatrix}$$

Node subscript i is omitted from the P matrix for simplicity, which should not create any confusion. Each element in P, ps, d, is the probability that a traveler coming from supply node s will move to demand node d, which can be affected by many factors including the traveler's personal characteristics (Ω_t), the number of opportunities at the current node i (b_i), the number of opportunities at each demand node of i (b_d), the quality of the opportunities (Q), and the ease of reaching the opportunities (A):

$$P = f(\Omega_t, b_i, b_d, Q, A) \tag{1}$$

In the current model, a simple functional form of f (*) is specified, and $p_{s,d}$ is computed on the basis of the following equations:

$$p_{s,d} = 0 \qquad \text{if } s = d \tag{2}$$

$$p_{s,d} = \frac{b_d}{\beta b_i + \sum_{d \in D \& d \neq i} b_d} \qquad \text{if } s \neq d \text{ and } d \neq i \tag{3}$$

$$p_{s,d} = \frac{\beta b_i}{\beta b_i + \sum_{d \in D \& d \neq i} b_d} \qquad \text{if } s \neq d \text{ and } d = i \tag{4}$$

Equation 2 ensures that if a traveler comes to node i from a supply node s, it will not go back to s in the next movement, which prevents a direct cyclic movement. Equation 3 states that the possibility that a traveler coming from supply node s will move to demand node d at the next step is proportional to the number of opportunities at node d (b_d). Equation 4 gives the probability that a traveler will accept an opportunity at node i; that is, the traveler agent stops its search process and no longer moves in the network. β is a weighting coefficient to be calibrated using trip length distribution data. A smaller β implies that on average a traveler agent needs to travel longer in order to find an activity because it is less likely to accept an opportunity at the current node. Theoretically, β can be any positive value. If $\beta bi + \sum bd = 0$ (i.e., there are no opportunities at any demand nodes), travelers will randomly select a demand node for the next movement. According to this specification of the turning guidance matrix, a traveler's search behavior is completely myopic in that the next movement is only based on the opportunities at the current node and its adjacent demand nodes.

Equations 2 to 4 with iterative execution actually provide a disaggregate algorithm for trip distribution that is in principle similar to the intervening-opportunities models [17–19] since trip making is not explicitly related to distance but to the relative accessibility of opportunities that satisfy the objective of the trip. Travelers consider available opportunities at increased distances from their origins. The agent-based trip distribution algorithm is more flexible than the intervening-opportunities model in two ways:

- Travelers consider only opportunities they have been exposed to along their search paths, whereas in the intervening-opportunities model, it is assumed that travelers have information on all opportunities in the region and are able to rank all destinations in order of increasing distance from their origins.
- In the intervening-opportunities model, the probability that a traveler will be satisfied by any opportunity is constant regardless of circumstances. The agent-based structure allows the probability to be dependent on the dynamic distribution of opportunities around the traveler.

Arc Agents. Arc i_1–i_2 connects origin node i_1 to destination node i_2 without any intermediate nodes. Its characteristics include capacity (C), length (l), free-flow speed (v_f), flow (q), and other costs (O), for example, tolls. Arc cost (c) is a function of those five factors:

$$c = g\left(C, l, v_f, q, O\right) \tag{5}$$

g (*) can take the form of an appropriate arc performance function. The current model assumes infinite arc capacities. Therefore, the arc cost becomes a constant, and congestion effects are not considered.

Route Cognition

The three types of agents, as just defined, enable one to examine some travel decision-making processes under an agent-based framework. Traveler agents' goal and behavior in many aspects are associated with real-world behavior of individuals. However, one limitation is that each traveler agent, as just defined, only pursues one particular activity. In reality, travelers may have multidestination tours with several activity types. The goal of the traveler agent must be expanded to accommodate activity chains. The arc agents are almost identical to physical road segments connecting intersections in the real world.

The node agent presented in the model needs to be elaborated a bit more. On the one hand, a node agent corresponds to a real network node at which arcs intersect and activity opportunities are located. On the other hand, a real-world intersection obviously does not know anything about the shortest paths within the network. The node knowledge should be interpreted as pooled, collective knowledge from some travelers who are familiar with the local area surrounding the node. For instance, an individual residing near a node knows the shortest paths from other nodes in the network to that node better than other individuals do who are unfamiliar with the area. Several studies on route cognition have shown that real-world travelers are only familiar with routes in the direct environment of their homes and activity centers that are frequently visited (a very limited part of the whole network), but, in general, they have limited knowledge about the routes in the remaining part of the network [20, 21]. Therefore, when knowledge exchange occurs between traveler agents and node agents in the model, it actually represents information exchange between different real-world travelers. How travelers learn about alternative routes in the network is a very important question.

Interaction Rules

Some interaction rules were pointed out when the agent characteristics were introduced. For instance, travelers acquire arc costs from arc agents and obtain turning guidance from node agents. Nodes communicate with each other so that each node knows the availability of opportunities at its demand nodes. Arcs update their flows based on travelers' search paths. These rules are simple since they only involve communication and no learning activities.

It is also necessary to define some learning rules for the model to be useful. Specifically, for the model to be able to realistically approximate trip distribution in the real world, a traveler agent should examine opportunities farther and farther away from its origin instead of making circular movements around the origin as the search process proceeds. Also, real-world individuals tend to choose the shortest paths for their trips, which require the traveler agents to have the ability to learn shortest paths between origin–destination pairs. An interaction rule defining learning activities between traveler and node agents in the model can meet both requirements. This learning rule applies whenever a traveler moves to a node *i.* It is assumed that a traveler has already visited a set of *n* nodes. Both the traveler and the node want to learn the shortest paths to travel from these *n* nodes to the current node *i* according to agent characteristics. The mechanism of this learning rule is not complex: for each node i' of the *n* nodes visited by the traveler, both the traveler and node *i* know a path to travel from i' to *i,* respectively. So they compare the lengths (or the generalized costs) of the two paths, and the agent who knows the longer path will learn the shorter one from the other.

The following example illustrates the traveler-node learning rule graphically. In the example, all arc costs are assumed to be 1 for simplicity. A traveler originating from Node 1 just moved to node *i* (Figure 7.6.1*a*). Therefore, the traveler-node learning rule is applied between these two agents. In this case the traveler has already visited three nodes ($n = 3$, $i' \in \{1, 3, 4\}$) before arriving at *i.* The traveler's knowledge is represented in the diagram by the solid line, and the node's knowledge is depicted by three types of discontinuous lines. The traveler and the node first compare the paths from Node 4 to *i* because Node 4 is the one most recently visited by the traveler. Since they know two equally short paths ($1_{node} = 1_{traveler}$), there is no learning activity between them (Figure 7.6.1*b* shows their respective knowledge after this comparison). Then they compare the two paths from Node 3 to *i* and find that the node knows a better path ($1_{node} < 2_{traveler}$). Thereby the traveler learns from the node (see Figure 7.6.1*c* after this round of learning). Finally, they compare paths between the traveler's origin Node 1 and node *i.* This time, the node learns from the traveler because the traveler knows a shorter path ($3_{node} > 2_{traveler}$; see bottom diagram in Figure 7.6.1*d*).

The result of this traveler-node learning rule and Equation 2 is that, once a traveler agent finds an activity, the path it used will also be the shortest path from its origin to the destination, based, on the traveler's best knowledge. Node agents also possess the knowledge of shortest paths identified by the model. If there are enough travelers in the transportation system, the shortest path found by the model approximates the real shortest paths, as will be seen in two examples presented later. In this sense, the learning rule in this model could be viewed as an asymptotic shortest-path algorithm based on distributed learning.

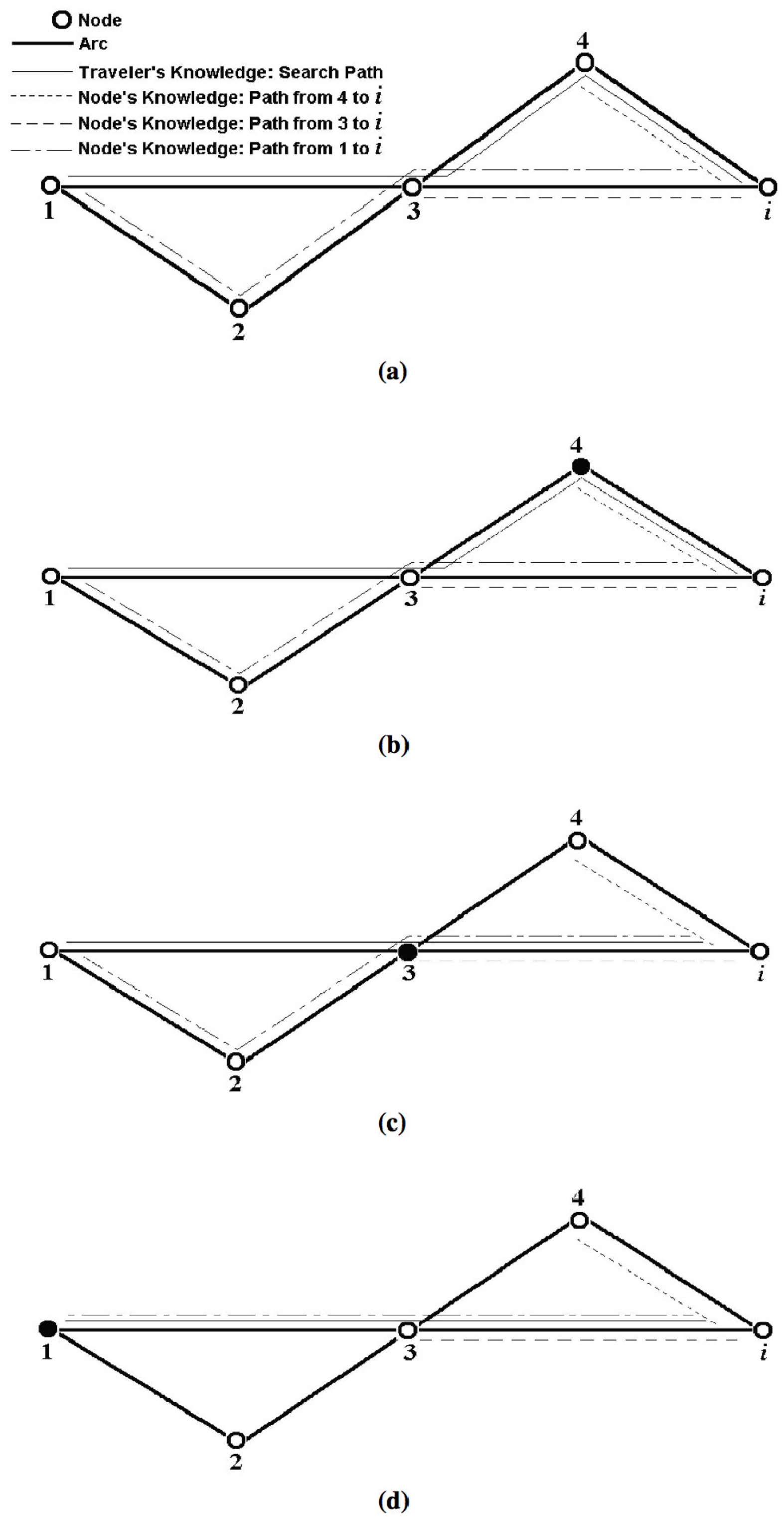


Figure 7.6.1. Traveler-node learning rule between traveler and node *i*

Transportation planners are familiar with the application of discrete choice analysis in travel forecasting [3]. Individuals' route selection behavior is modeled as the outcome of a cross-sectional choice process that contains two steps: choice-set generation, in which several alternative routes are identified, and choice-making, in which a "best" route is selected on the basis of utility trade-offs. The learning rule just described is an example of another paradigm for modeling routing decisions and can be interpreted as follows. A traveler agent is able to identify at least one route toward the activity destination (e.g., the traveler's own search path); however, without any learning activities with the nodes (holders of localized network information) along the traveler's own search path, the selected route will be by no means satisfactory. The traveler agent also recognizes that fact and wants improvements. However, in contrast to discrete choice analysis, it is not assumed that travelers are capable of identifying several alternative routes. Rather, it is assumed that travelers adjust their current route on the basis of localized network information.

Such localized network information can come from other travelers or from travelers' own experience. When a traveler agent arrives at a node in the model and learns a better shortcut from the node agent, in the real world this situation can be interpreted as one in which someone tells someone else a better route. If alternatively the node agent learns from the traveler agent, the traveler agent actually shares its own experience to improve the collective understanding of the network. This phenomenon can also frequently be observed in the real world. For instance, a good route from a traveler's home to the shopping center that is frequently visited by the traveler is very likely to become a part of the path selected by the same traveler for a trip from home to another destination near the shopping center.

The improvement or adaptation paradigm, in which travelers are assumed to adjust their decision until a certain aspiration level is achieved, was adopted in previous models of travel decision making, such as AMOS [22, 23] and SMASH [24]. Bowman and Ben-Akiva [25] generalize the decision-making processes in those models as repetitive execution of choice-set generation and choice making. However, the intensive learning and adaptive behavior may be better modeled under an agent-based framework.

Emergence of Travel Demand Through Evolutionary Process

On the basis of the specified agent characteristics and interaction rules, a transportation system is ready to evolve given a transportation network, an initial distribution of travelers, and activity opportunities in the network, which can be the outputs of any trip generation process. The evolutionary process is illustrated by the flowchart in Figure 7.6.2. The probabilities specified in Equations 2 to 4 can be realized through Monte Carlo simulation. The convergence of the evolution process can be directly measured by the number of residual travelers, that is, travelers who have not yet found an activity or the number of residual opportunities, whichever reaches zero first (the model does not require an equal number of travelers and opportunities). When all travelers are settled with activities, the transportation system reaches a stable pattern since there will be no more movements or interactions. All agents and their knowledge will remain constant thereafter. Therefore, this stable pattern is considered the end point of the evolution or, for simplicity, an equilibrium.

If each traveler corresponds to one or more trips, the trip distribution and assignment problems are solved simultaneously in the system equilibrium. The route each traveler takes is the shortest path from the origin to the destination based on the traveler's best knowledge of the network travel costs. That knowledge is accumulated through interactive, iterative learning with multiple node agents in the network. The result of route assignment in the agent-based travel demand model is in a sense similar to an all-or-nothing assignment since arc capacity constraints are not considered in the current model, although the two algorithms are based on completely different assumptions about travel behavior. The only coefficient that needs to be calibrated in the model is β in Equation 1, which can be interpreted as a traveler's willingness to travel further.

7.6.3 Application Examples and Calibration Procedures

Computational Properties

Before the discussion proceeds to numerical examples, the computational properties of the model are summarized analytically. For a transportation network with I nodes and T travelers, there are at most $I*(I-1)+T$ paths in the model since each node can at most keep information on $I-1$ paths from all other nodes to itself, and each traveler has one search path. All knowledge must be stored in the model, and hence the theoretical maximum memory consumption is proportional to the number of travelers and the square of network size. In practice, the actual memory requirement is much less because if no traveler travels between a node pair, the shortest path between the two nodes is not necessary and will not be stored by any node. In a large network, many node pairs will not be visited by travelers. As the system starts evolving, the number of paths further decreases since a traveler's search path is no longer useful and can be deleted once an activity is found.

An examination of the evolutionary process (Figure 7.6.2) would reveal a good property of the model—the computational time is only proportional to the number of travelers and is not sensitive to the size of the transportation network. The running time of the model will still increase as the network size increases since on average travelers will search more nodes to find activities. The travel-node learning process will take more time, but it will not increase

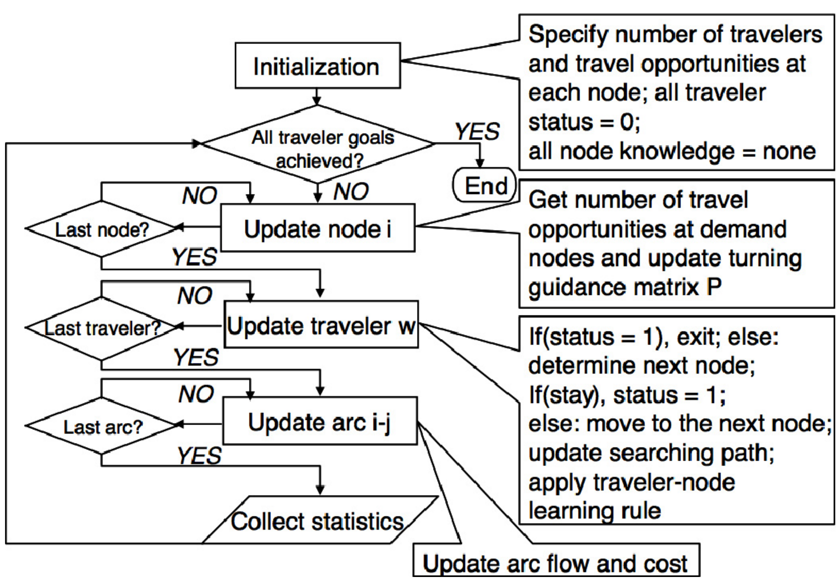


Figure 7.6.2 Flowchart of evolutionary algorithm

exponentially because in the agent-based model, information exchange and agent learning activities substitute for standard shortest-path algorithms, and thus path enumeration is not required. Another aspect related to running time is the ease of the calibration procedure, which will be discussed next along with two examples.

Numerical Examples

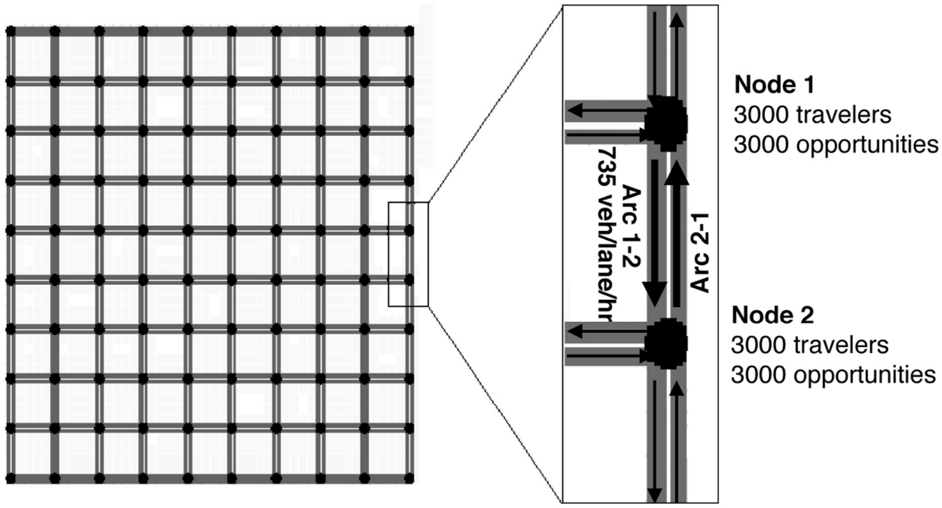
Example 1: 10 × 10 Grid Network

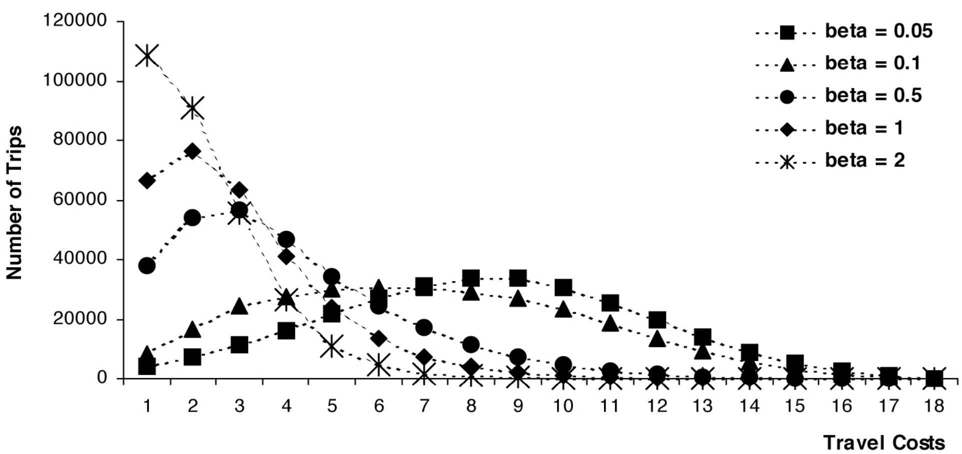
The first example uses a simple 10 × 10 grid network with 300,000 travelers and an equal number of opportunities to demonstrate the model. The travelers and opportunities are uniformly distributed among all nodes. The arc cost is 1 unit for all arcs (see Figure 7.6.3). The structure of the agent-based travel demand model can be implemented with any object-oriented programming language (Java was used in this study). Five different β's are tested ranging from 0.05 to 2. For each β, the resulting travel length distributions and the convergence properties at the five equilibria are summarized in Figures 4 and 5, respectively.

The model can approximate a variety of trip length distributions with negative exponential (large β) and normal distribution (small β) at the two extremes. In this small network with a moderate number of travelers, the evolutionary process quickly reaches the equilibrium. As travelers travel farther away from origins to find activity opportunities, it takes longer for the system to achieve the equilibrium. In all five scenarios, at equilibrium the shortest paths identified by the models are the real shortest paths between node pairs, which is not surprising in a small network. As the ratio of number of travelers to the size of the network decreases, some shortest paths learned by the travelers in the model may be longer than the real shortest paths, as will be seen in the next example. The selection of the initial random seed for Monte Carlo simulation has almost no impact on the resulting trip length distribution and shortest paths at the equilibrium, probably because the large number of random decisions and learning activities in the model tends to average out the initial variability due to different random seeds.

Example 2: Chicago, Illinois, Sketch Network and Model Calibration

In the second example, the agent-based travel demand model is applied to the Chicago sketch network, consisting of 933 nodes and 2,950 links, a fairly realistic yet aggregated representation of the Chicago region developed by the Chicago Area Transportation Study (CATS) (*1*). There are more than 1.26 million travelers in this test network according to the trip generation data, with each traveler representing one trip. The only coefficient β in the model is calibrated against CATS 1990 Household Travel Survey (HTS) data. The estimated travel time distribution with various β's and the observed distribution are plotted in Figure 7.6.6. The spikes on the observed travel time distribution reveal survey participants' tendencies to round their actual travel times to 30, 45, and 60 min. The mean square error (MSE) between the estimated and the observed distribution is plotted against βin Figure 7.6.7. It is clear in this graph that the MSE distribution is a unimodal one, and therefore simple one-dimensional search methods can be adopted to calibrate the model coefficient. The following is an applicable calibration procedure based on golden section search:





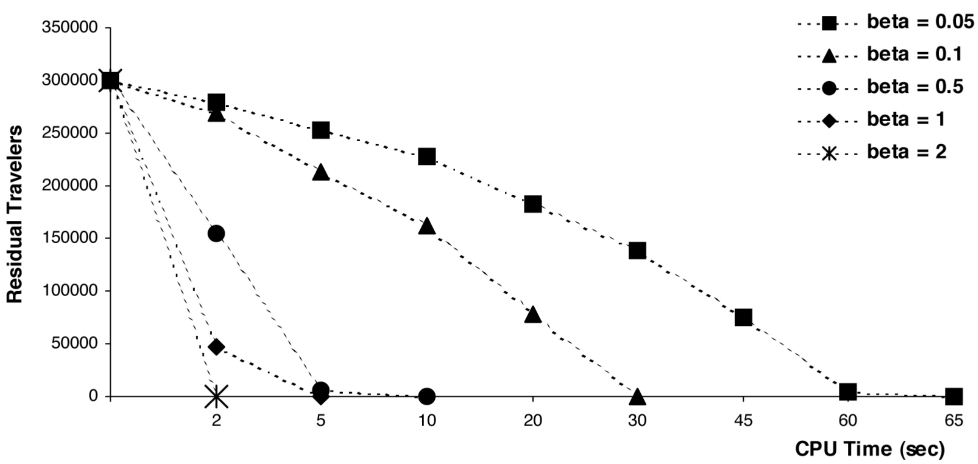


Figure 7.6.3 Uniform distribution of travelers and opportunities for 10 X 10 grid network

Figure 7.6.4 Trip length distribution with various β's for 10 X 10 grid network

Figure 7.6.5 Convergence speeds with various β's for 10 X 10 grid network

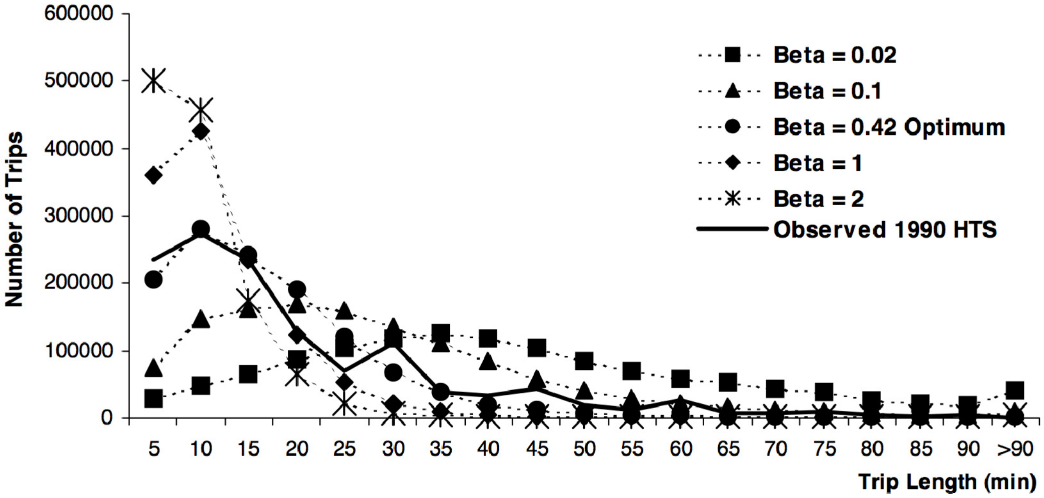


Figure 7.6.6 Chicago sketch network: trip length distribution with various β's

Step 0—Initialization: Lower-bound β^- =0.1 and upper-bound β^+ =1. (Theoretically, β can be any positive value, but for all practical purposes, [0.1, 1] should be a safe starting interval for the golden section search.) Determine stopping tolerance $e > 0$. Iteration counter t =0. Compute $\beta^1 = \beta^+ - 0.618\,(\beta^+ - \beta^-)$ and $\beta^2 = \beta^- + 0.618\,(\beta^+ - \beta^-)$. Evaluate the MSEs at all four points.

Step 1—Stopping: If $(\beta^+ - \beta^-) < e$, stop, and the optimal $\beta^* = 0.5\,(\beta^+ + \beta^-)$. Otherwise, proceed to Step 2.

Step 2—Iteration: If MSE (β^1) < MSE (β^2), narrow the search to the left part of the interval by updating $\beta^+ = \beta^2$, $\beta^2 = \beta^1$, $\beta^1 = \beta^+ - 0.618\,(\beta^+ - \beta^-)$, and evaluate the new MSE (β^1). If MSE (β^1) > MSE (β^2), narrow the search to the right part of the interval by updating $\beta^- = \beta^1$, $\beta^1 = \beta^2$, $\beta^2 = \beta^- + 0.618\,(\beta^+ - \beta^-)$, and evaluate the new MSE (β^2).

$t = t + 1$. Return to Step 1. □

Other one-dimensional search methods can be used as well, but the golden section search in general provides an efficient procedure. A more detailed discussion of unimodal function optimization may be found elsewhere [26]. The foregoing calibration procedure was applied to the Chicago sketch network with e =0.05 and the optimal β* was found to be 0.42 after five golden section search iterations (i.e., six executions of the model with different β's since the first iteration requires the evaluation of model MSEs twice), which took about 70 CPU minutes on a Pentium IV, 1.7-GHz personal computer. At the equilibrium with β*, travelers discovered 99.1% of all origin– destination paths, of which more than 98% are real shortest paths.

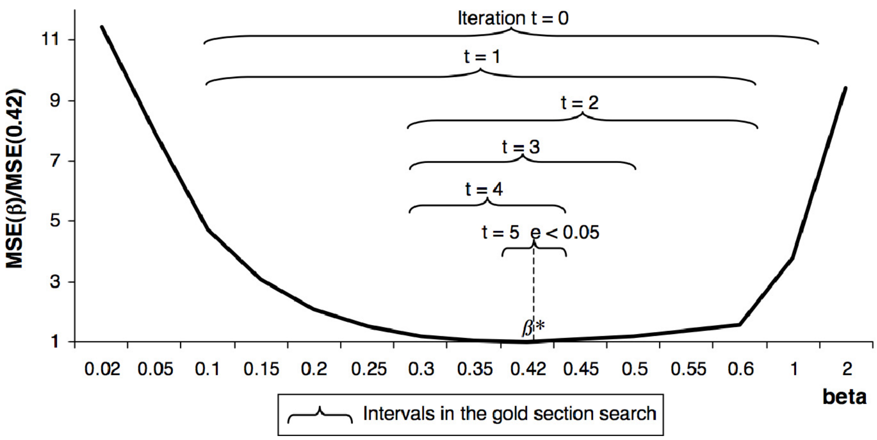


Figure 7.6.7 Unimodal MSE function and model calibration using golden section search

However, the β* estimated on one network is not directly transferable to other networks. For instance, the coefficient estimated for a sketch network that only includes major highways should not be used without further calibration for a full network with all types of roads. One needs to be consistent in coding the network when applying the proposed model. Of course, it is suspected that the coefficient also varies from city to city. From a computational point of view, the model transferability is not a big issue because it does not take much time to calibrate β for a specific network, but the estimated β's for different urban areas are not comparable. The β-coefficient is sensitive to the detail of the network used for calibration because in the proposed model, travelers base their next movements only on the relative distribution of activity opportunities at surrounding nodes. To improve the transferability of the model, one needs to relate the probability that a traveler agent will accept an activity opportunity with the actual distance (or duration) of travel.

The agent-based model after the calibration procedure distributes trips from origins to destinations in a disaggregate manner with a trip length distribution reasonably close to the observed one and assigns most traffic to the shortest routes. The model provides output statistics including arc flows, the origin and destination of each individual trip, the path of each individual trip, and turning proportions at all intersections.

7.6.4 Possible Extensions of Model and Future Research Directions

Though the proposed agent-based travel demand model is a novel and interesting way of forecasting travel demand, it has not achieved the scope of existing travel forecasting methods. Several extensions can be incorporated to improve the current model.

More Agent Characteristics and Knowledge

With only three types of agents and minimum agent characteristics, the proposed agent-based travel demand model is able to accomplish two critical steps in the travel forecasting process—trip distribution and traffic assignment—within a short amount of time. It would be worthwhile to extend the basic model so that mode split can also be incorporated and more-realistic traffic assignment algorithms can be approximated. One way to enable modal split in the agent-based model is to expand the node knowledge to path costs of all modes and embed a mode choice rule into travelers' characteristics. Congestion effects should be taken into account in future versions of the model, which requires an expansion of arc characteristics. However, with limited arc capacities, the shortest paths become dependent on travelers' choices. How traveler agents learn shortest paths in this new dynamic situation must be carefully modeled, probably through repetitive information exchange and learning from day to day.

More Types of Agents

Besides travelers, nodes, and arcs, other agents in the transportation system have significant impacts on travel demand. For instance, it is necessary to define agents that represent transit links and railways if these modes are to be incorporated in the model. Another extension of the current agent-based model would be the introduction of land use agents. The interaction between transportation and land use has been long recognized and studied. A metropolitan area can be divided into many land use cells, and each cell can be modeled as a special type of agent that has its own characteristics and behavioral rules. In an agent-based model, interaction rules between land use cells and transportation agents such as nodes and arcs, if appropriately defined, may be able to reasonably replicate the feedback between transportation and land use. The problem then becomes the calibration of these rules. Also, under the agent-based modeling framework, simple rules may well explain complicated real-world phenomena such as the transportation–land use feedback loop. Alternatively, urban land use can be modeled as the environment in the agent-based model. These possibilities should be examined in future studies. Transportation management policies, such as pricing schemes and financing strategies, have already been modeled by proper agents and their characteristics in several previous studies [27, 28].

More-Realistic Rules of Agent Behaviors

In constructing an agent-based model there are two major steps: (*a*) identify agents and their characteristics and (*b*) specify their behavioral rules. Different modelers may come up with different sets of agents for the same system. Some may be more useful in terms of facilitating the second step, rule specification, which is usually the challenging part. The model developed in this study employs only local rules according to which agents interact only with other adjacent agents. Local rules have been successfully used in many cellular automata applications, such as the cell transmission model for freeway traffic [29, 30]. In general, drivers make car-flowing and lane-changing decisions on the basis of the traffic conditions around themselves, and therefore local rules may be a realistic specification of their interactions. However, in the case of travel decision making, it is known that travelers sometimes rely on maps, media, and even route guidance systems when making decisions. This aspect implies that information sharing is beyond the local level.

Although occasionally global knowledge sharing, information flow, and learning activities can be reasonably approximated with local rules, that is not always the case. Do travelers find their activities and choose routes using the same methodology in the proposed agent-based travel demand model? Will a small deviation from real behavior significantly affect the resulting equilibrium of the evolutionary process? These questions are yet to be answered. In the two examples given in the previous section, travelers have no difficulty in finding the shortest path for their trips because there are so many travelers in the system and the intensive local learning activities solve the shortest paths for travelers. Had there been only one traveler agent in the model, it would definitely fail to find the shortest paths since no learning activities would happen. But because a single traveler in the real world can identify the shortest route for a trip (or at least a route not much longer than the shortest one) without interacting with other individual travelers, global knowledge sharing may need to be incorporated somehow into the agent-based travel demand model.

The progress made in travel behavior studies can be readily incorporated into the agent-based model with an update of agent behavioral rules. The only problem with more-realistic behavioral rules is their possible requirements for more computational resources. Finding and applying realistic behavioral rules of agents while at the same time keeping the model computationally feasible is the real challenge. This challenge should be kept in mind in future development of similar models. Because the human brain has a limit on complex computation, this problem may not be as serious as it seems.

7.6.5 Conclusions

An agent-based travel demand model is developed. Travel demand emerges from the interactions of three types of agents in the transportation system: node, arc, and traveler. Simple local rules of agent behaviors are shown to be capable of efficiently solving complex transportation problems such as trip distribution and route assignment. The model also provides an asymptotic shortest-path algorithm based on distributed agent learning activities. Possible extensions to the basic model are also discussed. The generic and flexible structure of the agent-based modeling method makes it easier to develop new models and to expand existing models. By giving agents intelligence and allowing them to learn, modelers can accomplish more with less modeling effort. The method also takes full advantage of the fast-growing computational power now available.

Compared with trip-based approaches, activity-based approaches represent a new paradigm for travel demand analysis. The proposed agent-based technique, however, does not imply another paradigm shift. Rather, it is a powerful modeling tool to disentangle complex systems. In general, agent-based models emphasize, at the microscopic level, searching and learning behavior, agents' perception of the environment, information flow, interagent interactions, and heuristics and, at the macroscopic level, self-organization, hierarchy, and other evolutionary properties. It is difficult and unnecessary to draw a line between agent-based travel demand models and activity-based approaches. The modeling needs for interpersonal linkages, person–environment interactions, and longitudinal aspects of travel behavior discovered in recent practice of activity-based travel analysis actually provide a stage for agent-based modeling techniques. Some recent activity-based microsimulation studies in which learning behavior [31] and activity interactions [32] are explicitly modeled have demonstrated the increasing popularity of agent-based methods.

This study pushes the application of agent-based methods for travel analysis beyond the scope of origin–destination demand estimation and into the realm of traffic assignment. It is possible that even the traditional equilibrium assignment process could be replaced with an agent-based model. A completely agent-based travel forecasting system is worth pursuing in the future. Though the proposed model is rudimentary in its current form, the authors hope that it can attract more research interest in applying agent-based modeling techniques to travel forecasting.

7.6.6 Acknowledgments

The authors thank David Boyce and Hillel Bar-Gera for their provision of the Chicago sketch network and related transportation planning data. Several smaller transportation networks on Bar-Gera's Traffic Assignment Problem website (www.bgu.ac.il/□bargera/tntp) were used in an earlier stage of the study.

7.6.7 References

1. State of Illinois, Cook County, and City of Chicago. *Chicago Area Transportation Study: Final Report.* Bureau of Public Roads, U.S. Department of Commerce, 1959–1962.
2. Beckman, M., C. B. McGuire, and C. B. Winsten. *Studies in the Economics of Transportation.* Yale University Press, New Haven, Conn., 1956.
3. Ben-Akiva, M., and S. R. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* MIT Press, Cambridge, Mass., 1985.
4. Recker, W. W., M. G. McNally, and G. S. Root. A Model of Complex Travel Behavior, Part I: Theoretical Development. *Transportation Research,* Vol. 20A, 1986, pp. 307–318.
5. Recker, W. W., M. G. McNally, and G. S. Root. A Model of Complex Travel Behavior, Part II: Operational Model. *Transportation Research,* Vol. 20A, 1986, pp. 319–330.
6. Kitamura, R., E. I. Pas, C. V. Lula, K. Lawton, and P. E. Benson. The Sequenced Activity Mobility Simulator (SAMS): An Integrated Approach to Modeling Transportation, Land Use and Air Quality. *Transportation,* Vol. 23, 1996, pp. 267–291.
7. Ortuzar, J. D., and L. G. Willumsen. *Modelling Transport,* 3rd ed. Wiley, New York, 2001.
8. Boyce, D. Is the Sequential Travel Forecasting Paradigm Counterproductive? *Journal of Urban Planning and Development,* ASCE, Vol. 128, No. 4, 2002, pp. 169–183.
9. Pas, E. I. State of the Art and Research Opportunities in Travel Demand: Another Perspective. *Transportation Research,* Vol. 21A, 1985, pp. 431–438.
10. Carpenter, S., and P. Jones, eds. *Recent Advances in Travel Demand Analysis.* Gower, Aldershot, England, 1983.
11. Jones, P., ed. Developments in Dynamic and Activity-Based Approaches to Travel Analysis. Avebury, Brookfield, Vt., 1990.
12. Kitamura, R. An Evaluation of Activity-Based Travel Analysis. *Transportation,* Vol. 15, 1988, pp. 9–34.
13. Ettema, D., and H. Timmermans, eds. *Activity-Based Approaches to Travel Analysis.* Pergamon Press, Oxford, United Kingdom, 1997.
14. McNally, M. G., and W. W. Recker. *Activity-Based Models of Travel Behavior: Evolution and Implementation.* Elsevier Science, Amsterdam, Netherlands, 2000.
15. Von Neumann, J. *Theory of Self-Reproducing Automata* (A. W. Burk, ed.), University of Illinois Press, Urbana, 1966.
16. Intelligent Agents in Traffic and Transportation (special issue). Transportation Research, Vol. 10C, 2002, pp. 325–527.
17. Stouffer, S. A. Intervening Opportunities: A Theory Relating Mobility and Distance. *American Sociological Review,* Vol. 5, 1940, pp. 845–867.

18. Schneider, M. Gravity Models and Trip Distribution Theory. *Papers of the Regional Science Association,* Vol. 5, 1959, pp. 51–56.

19. Haynes, K. E., and A. S. Fotheringham. *The Gravity Model and Spatial Interaction.* Sage Publications, Beverly Hills, Calif., 1984.

20. Benshoof, J. A. Characteristics of Drivers' Route Selection Behavior. *Traffic Engineering and Control,* Vol. 11, No. 12, 1970, pp. 604–610.

21. Bovy, P. H. L., and E. Stern. *Route Choice: Wayfinding in Transport Networks.* Kluwer Academic Publishers, Dordrecht, Netherlands, 1990, 56 pp.

22. Kitamura, R., C. V. Lula, and E. I. Pas. AMOS: An Activity-Based, Flexible and Truly Behavioral Tool for Evaluation of TDM Measures. *Proceedings of Seminar D, 21st PTRC Summer Annual Meeting,* PTRC Education and Research Service, Ltd., London, 1993, pp. 283–294.

23. RDC, Inc. *Activity-Based Modeling System for Travel Demand Forecasting.* U.S. Department of Transportation; Environmental Protection Agency, 1995.

24. Ettema, D., A. Borgers, and H. Timmermans. Simulation Model of Activity Scheduling Behavior. In *Transportation Research Record 1413,* TRB, National Research Council, Washington, D.C., 1993, pp. 1–11.

25. Bowman, J. L., and M. E. Ben-Akiva. Activity-Based Travel Forecasting. In Activity-Based Travel Forecasting Conference, June 2–5, 1996: Summary, Recommendations and Compendium of Papers,

26. Report DOT-T-97-17, U.S. Department of Transportation, 1997.

27. Rardin, R. L. *Optimization in Operations Research.* Prentice Hall, Upper Saddle River, N.J., 1998.

28. Yerra, B., and D. M. Levinson. The Emergence of Hierarchy in Transportation Networks. Presented at Western Regional Science Association Meeting, Rio Rico, Ariz., 2003.

29. Zhang, L., and D. Levinson. A Model of the Rise and Fall of Roads. Presented at 50th North American Regional Science Council Annual Meeting, Philadelphia, Pa., Nov. 2003.

30. Daganzo, C. F. The Cell Transmission Model: A Dynamic Representation of Highway Traffic Consistent with the Hydrodynamic Theory. *Transportation Research,* Vol. 28B, No. 4, 1994, pp. 269–287.

31. Daganzo, C. F. The Cell Transmission Model, Part II: Network Traffic. *Transportation Research,* Vol. 29B, No. 2, 1995, pp. 79–93.

32. Arentze, T., and H. Timmermans. *ALBATROSS: A Learning-Based Transportation Oriented Simulation System.* Eindhoven University of Technology, Netherlands, 2000.

33. Rindt, C. R., J. E. Marca, and M. G. McNally. *An Agent-Based Activity Microsimulation Kernel Using a Negotiation Metaphor.* Paper UCIITS-AS-WP-02-7. Center for Activity Systems Analysis, University of California, Irvine, 2002.

Publication of this paper sponsored by Passenger Travel Demand Forecasting Committee.

Endnotes for 7.4

Induced Demand: A Microscopic Perspective

[i] For each link in the network, the parallel links were obtained and the results were checked with the network map. Parallel links are those that are not connected to the link in question. These links bear the maximum traffic when the link in question was to be eliminated (for details, see Levinson and Karamalaputi, 2002).

[ii] There were a number of erroneous links. Many links in the network had no traffic count although the link existed. Also, many links had traffic flows on them even though the link did not exist in the regional planning network. Such links were removed from the data-set and were not considered for analysis. A summary of the erroneous links in the network is given in Table A5. The final network used for analysis consisted of 4989 links belonging to the following road types: interstates, trunk highways and county state aid highways (CSAH) in Hennepin County.

Appendix A

Spatial Patterns of Information Workers in Six United States Metropolitan Areas

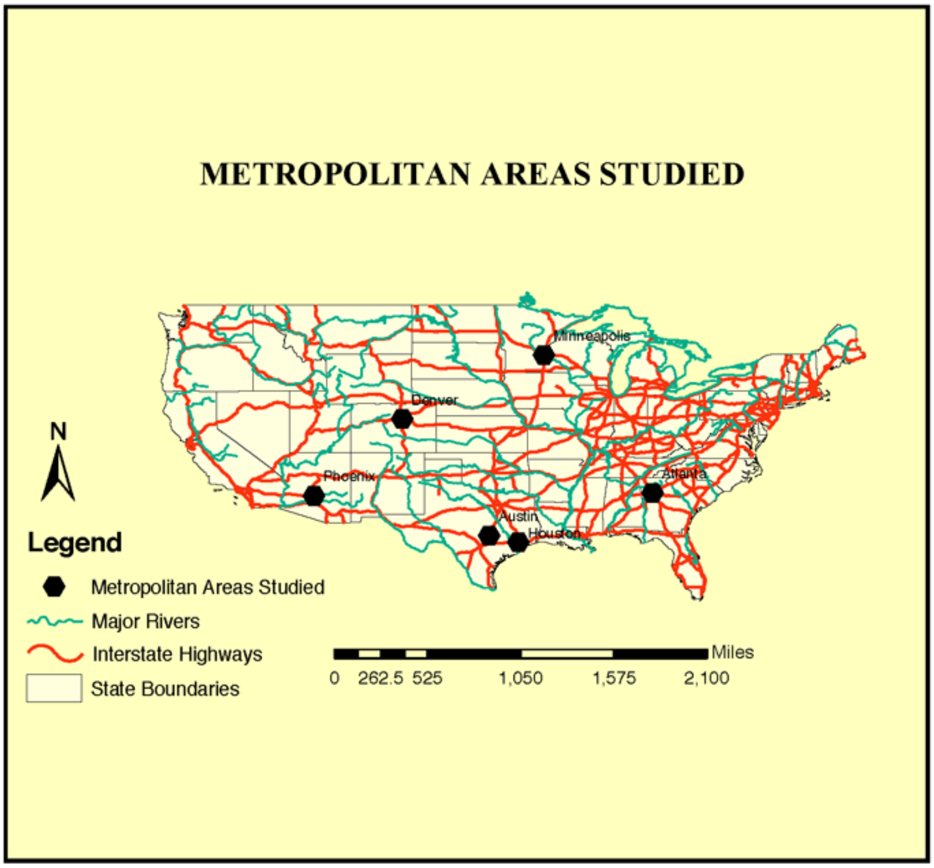
1990 Census Transportation Planning Package Occupation Categories
Managerial and professional specialty occupations (000-202)
Executive, administrative, and managerial occupations (000-042)
Professional specialty occupations (043-202)
Technical, sales, and administrative support occupations (203-402)
Technicians and related support occupations (203-242)
Sales occupations (243-302)
Administrative support occupations, including clerical (303-402)
Service occupations (403-472)
Private household occupations (403-412)
Protective service occupations (413-432)
Service occupations, except protective and household (433-472)
Farming, forestry, and fishing occupations (473-502)
Precision production, craft, and repair occupations (503-702)
Operators, fabricators, and laborers (703-902)
Machine operators, assemblers, and inspectors (703-802)
Transportation and material moving occupations (803-863)
Handlers, equipment cleaners, helpers, and laborers (864-902)

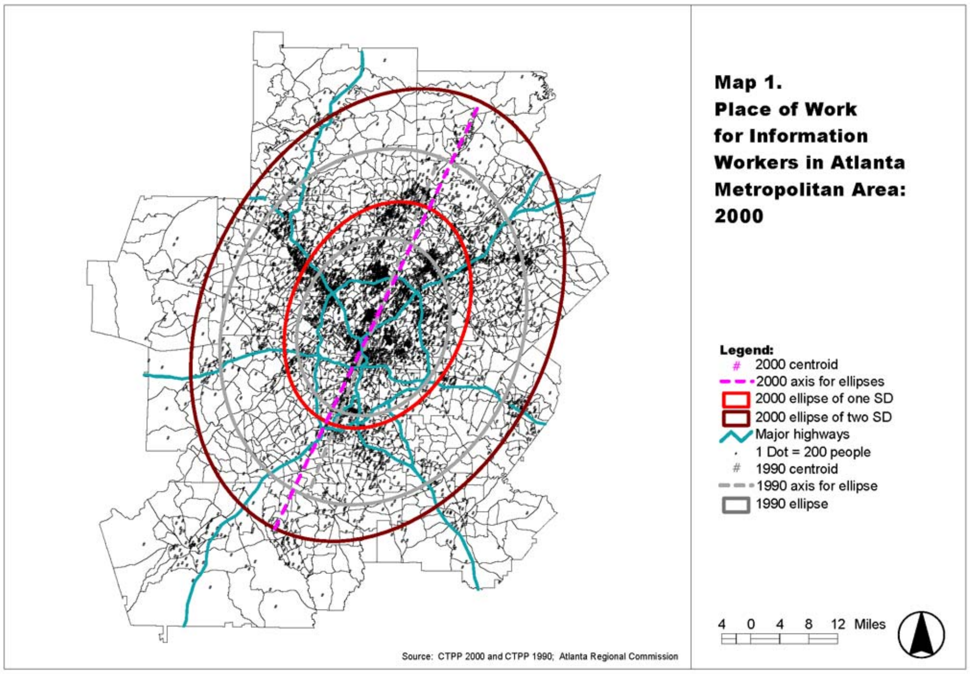
2000 Census Transportation Planning Package Occupation Categories
Management, professional, and related occupations
Management, business, and financial operations occupations
Management occupations, except farmers and farm managers
Farmers and farm managers
Business and financial operations occupations
Business operations specialists
Financial specialists
Professional and related occupations
Computer and mathematical occupations
Architecture and engineering occupations
Architects, surveyors, cartographers, and engineers
Drafters, engineering, and mapping technicians
Life, physical, and social science occupations
Community and social services occupations
Legal occupations
Education, training, and library occupations
Arts, design, entertainment, sports, and media occupations
Healthcare practitioners and technical occupations
Health diagnosing and treating practitioners and technical occupations
Health technologists and technicians
Service occupations
Healthcare support occupations
Protective service occupations
Fire fighting, prevention, and law enforcement workers, including supervisors

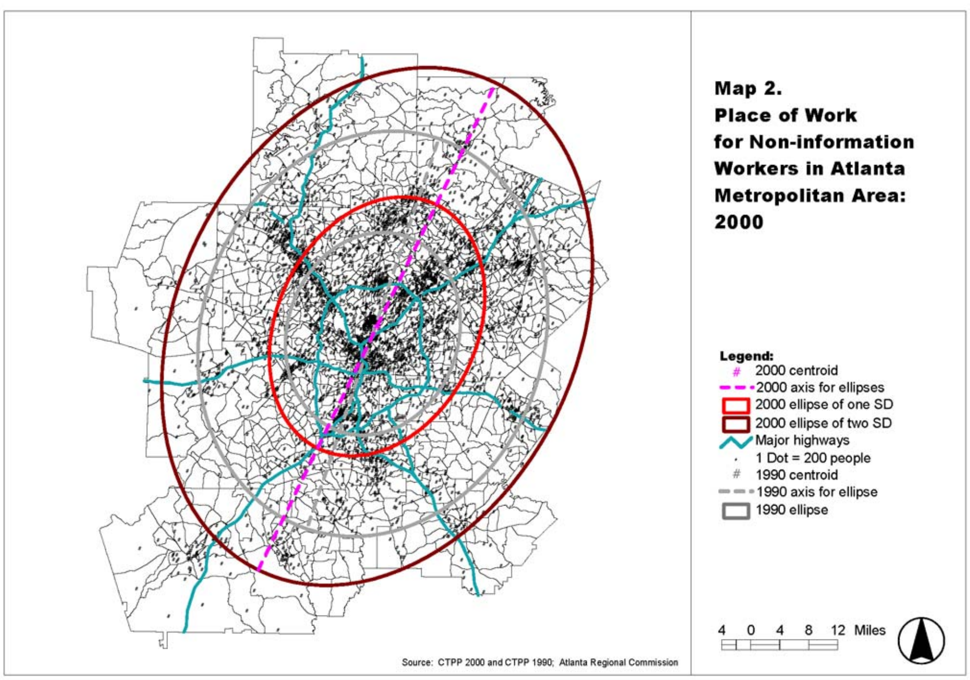
2000 Census Transportation Planning Package Occupation Categories
Other protective service workers, including supervisors
Food preparation and serving related occupations
Building and grounds cleaning and maintenance occupations
Personal care and service occupations
Sales and office occupations
Sales and related occupations
Office and administrative support occupations
Farming, fishing, and forestry occupations
Construction, extraction, and maintenance occupations
Construction and extraction occupations
Supervisors, construction and extraction workers
Construction trades workers
Extraction workers
Installation, maintenance, and repair occupations
Production, transportation, and material moving occupations
Production occupations
Transportation and material moving occupations
Supervisors, transportation and material moving workers
Aircraft and traffic control occupations
Motor vehicle operators
Rail, water and other transportation occupations
Material moving workers

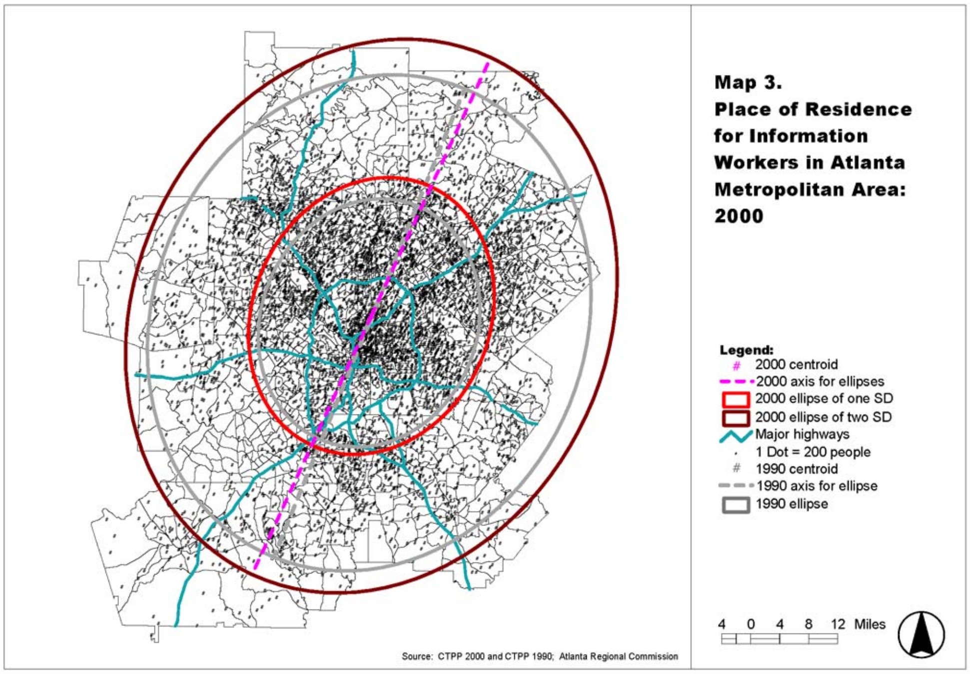
Appendix B

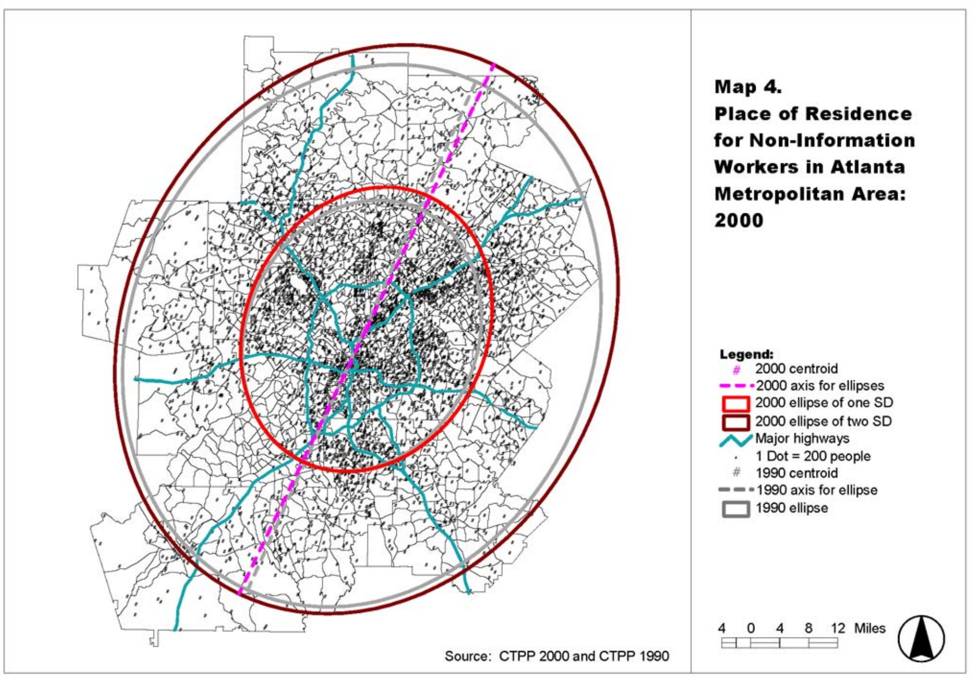
Spatial Patterns of Information Workers in Six United States Metropolitan Areas

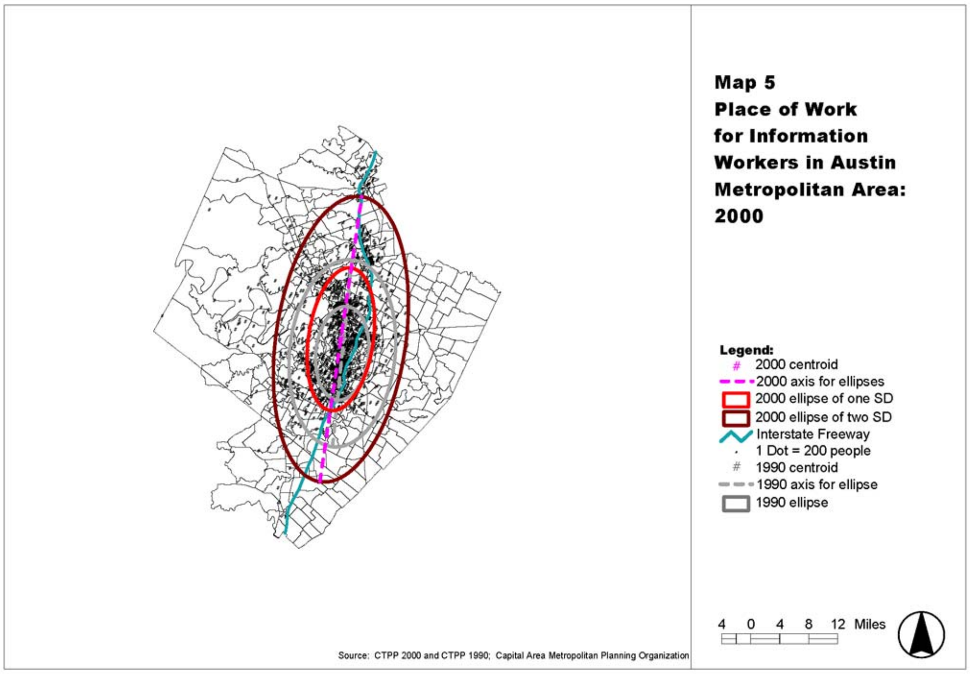

METROPOLITAN AREAS STUDIED
Minneapolis
Denver
Phoenix
Atlanta
Austin
Houston
N
Legend
Metropolitan Areas Studied
Major Rivers
Interstate Highways
State Boundaries
Miles
0 262.5 525 1,050 1,575 2,100

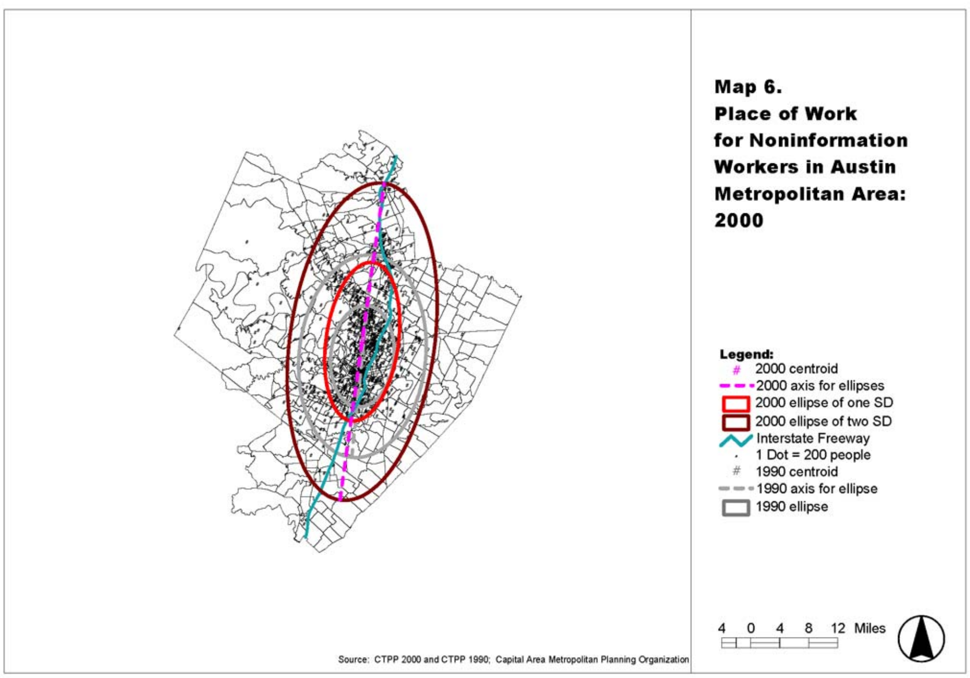

Map 1.
Place of Work
for Information
Workers in Atlanta
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Major highways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Atlanta Regional Commission

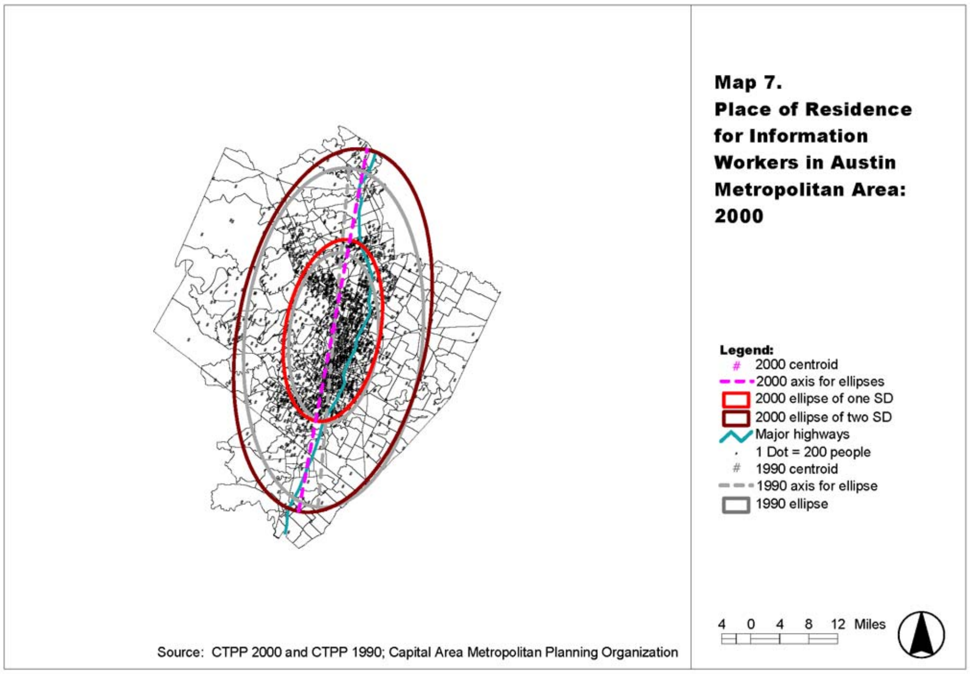

Map 2.
Place of Work
for Non-information
Workers in Atlanta
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Major highways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Atlanta Regional Commission


Map 3.
Place of Residence for Information Workers in Atlanta Metropolitan Area: 2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Major highways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Atlanta Regional Commission

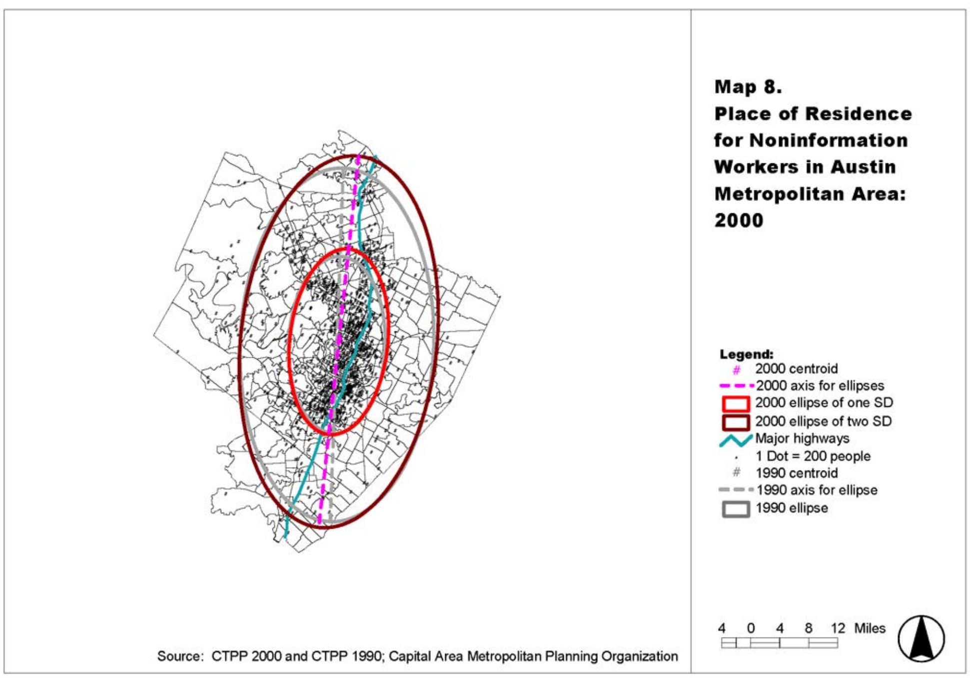

Map 4.
Place of Residence for Non-Information Workers in Atlanta Metropolitan Area: 2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Major highways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990


Map 5
Place of Work
for Information
Workers in Austin
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate Freeway
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Capital Area Metropolitan Planning Organization

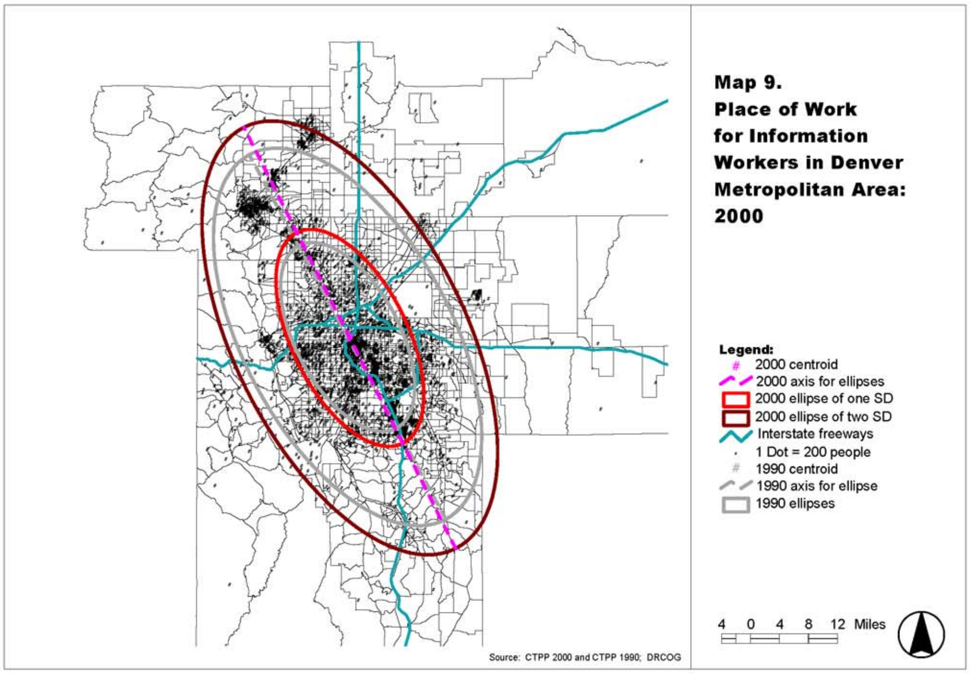

Map 6.
Place of Work
for Noninformation
Workers in Austin
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate Freeway
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Capital Area Metropolitan Planning Organization

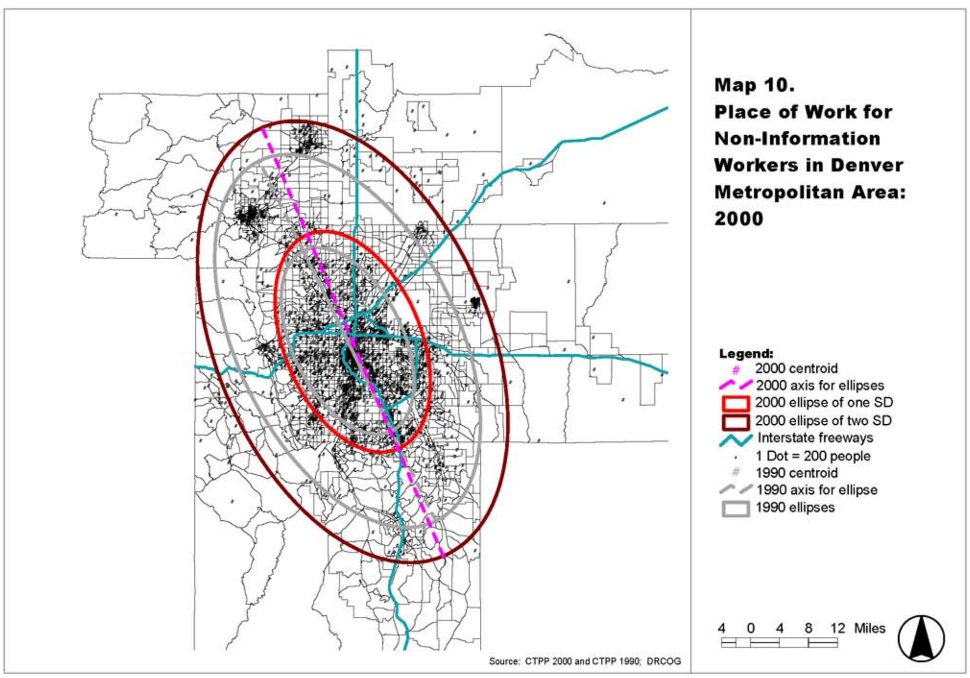


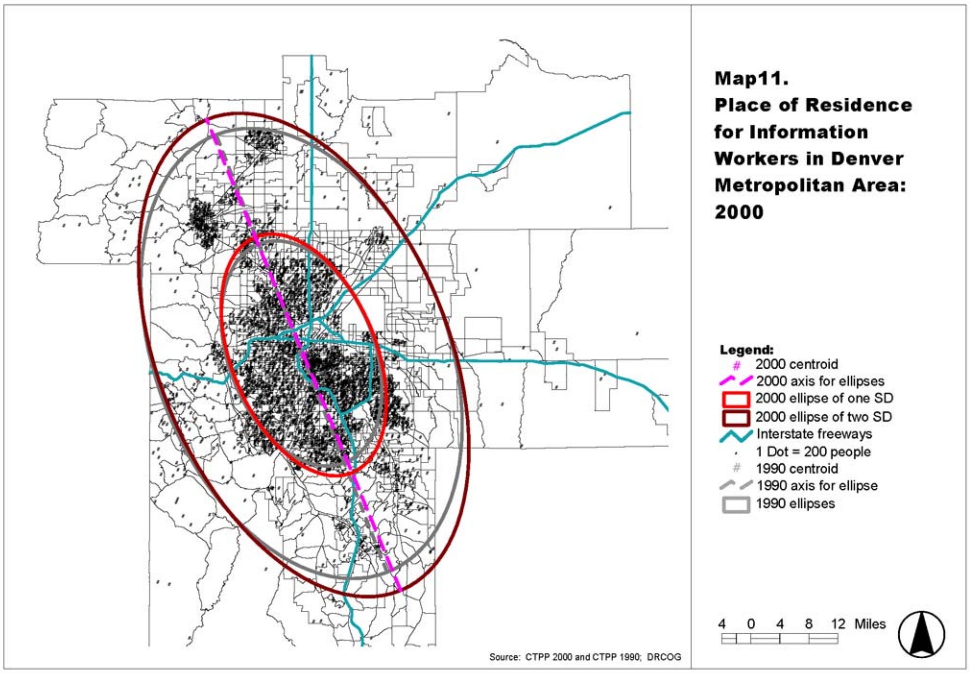
Map 7. Place of Residence for Information Workers in Austin Metropolitan Area: 2000

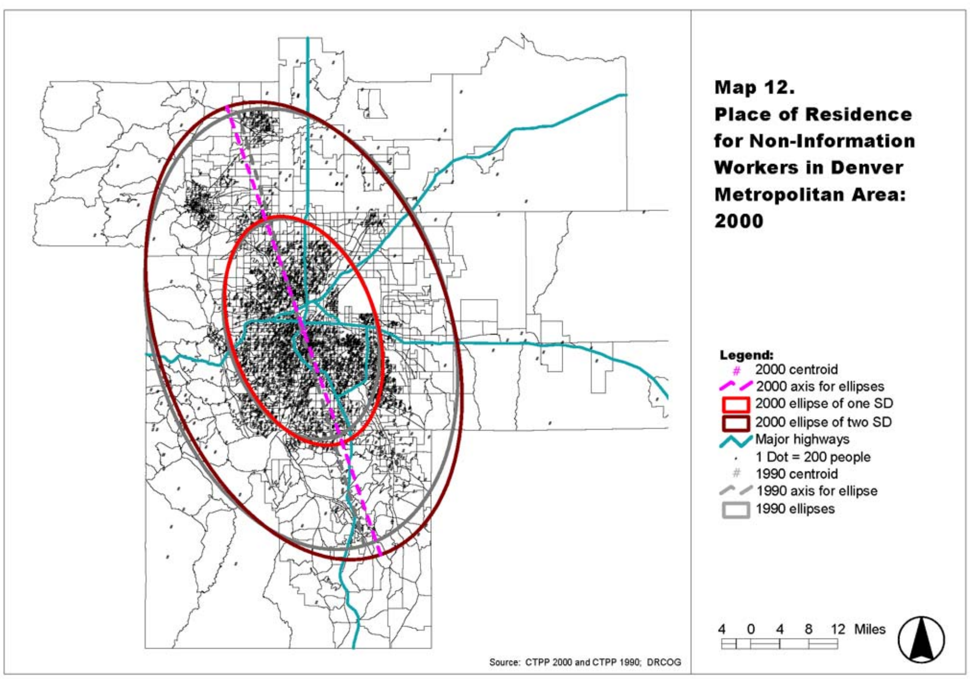


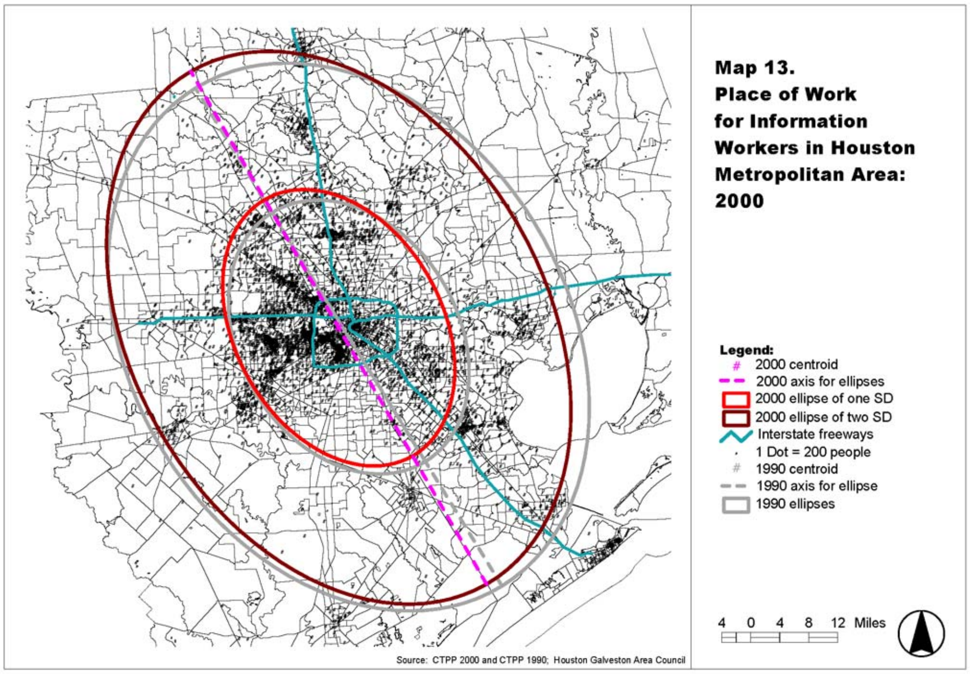
Map 8. Place of Residence for Noninformation Workers in Austin Metropolitan Area: 2000

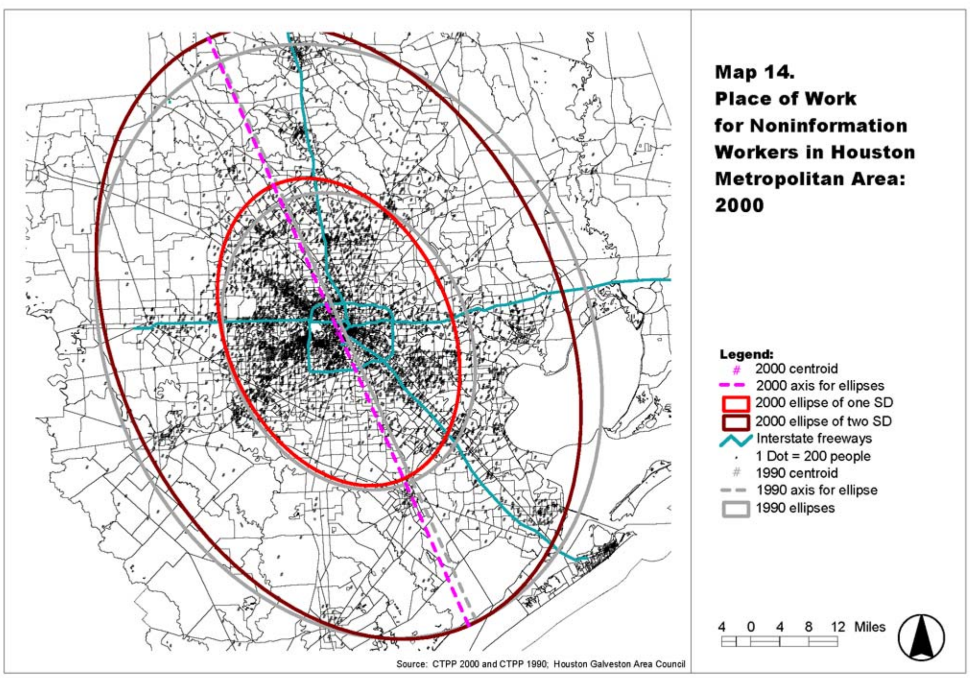

Map 9.
Place of Work
for Information
Workers in Denver
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; DRCOG

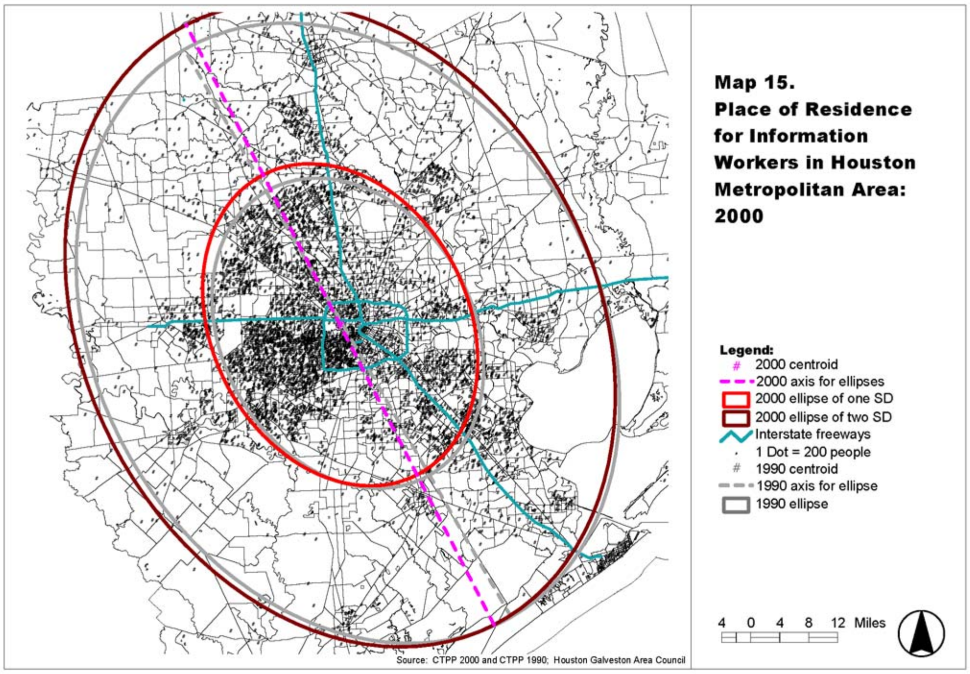

Map 10.
Place of Work for
Non-Information
Workers in Denver
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; DRCOG

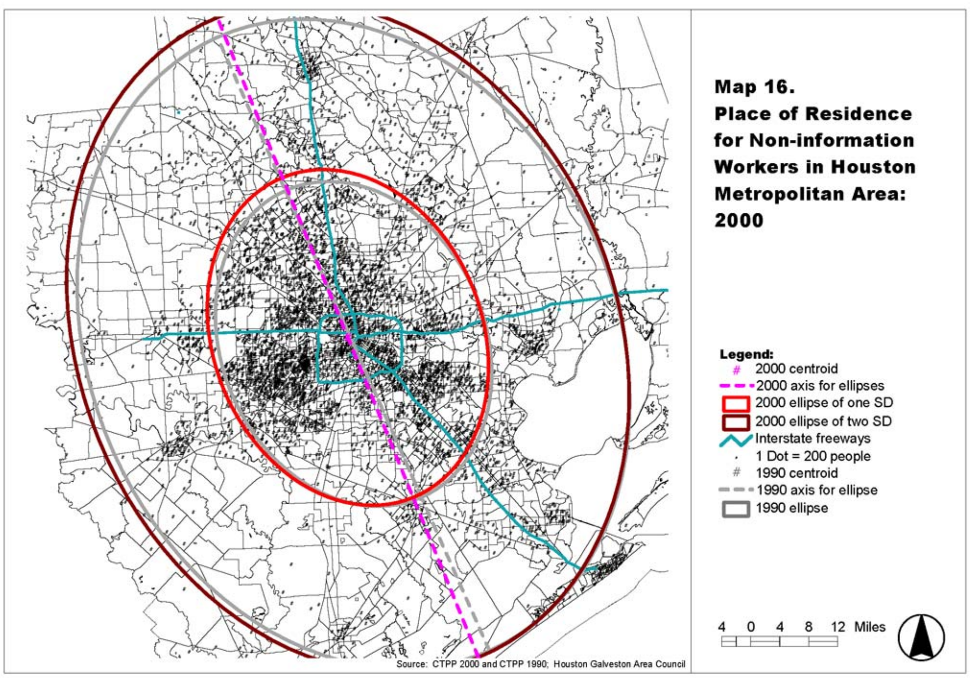

Map11.
Place of Residence
for Information
Workers in Denver
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; DRCOG

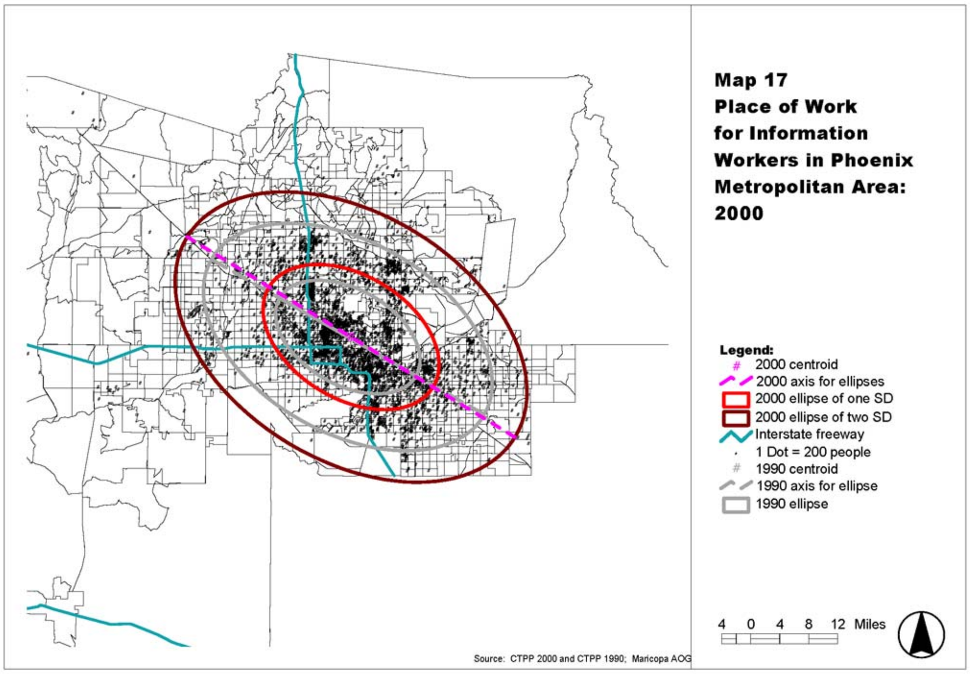

Map 12.
Place of Residence
for Non-Information
Workers in Denver
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Major highways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; DRCOG

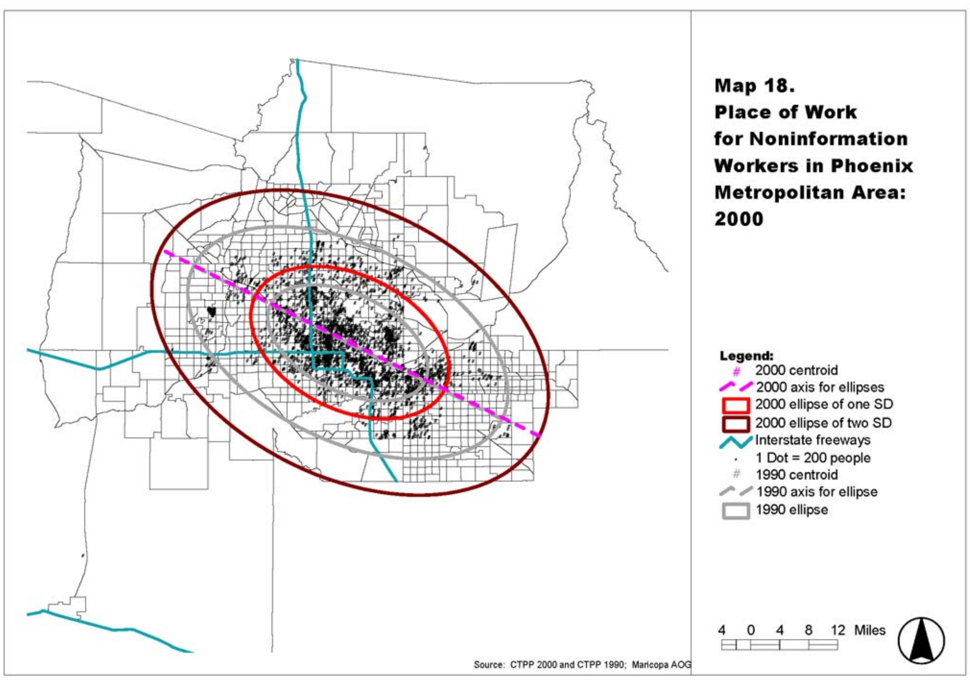

Map 13.
Place of Work
for Information
Workers in Houston
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Houston Galveston Area Council

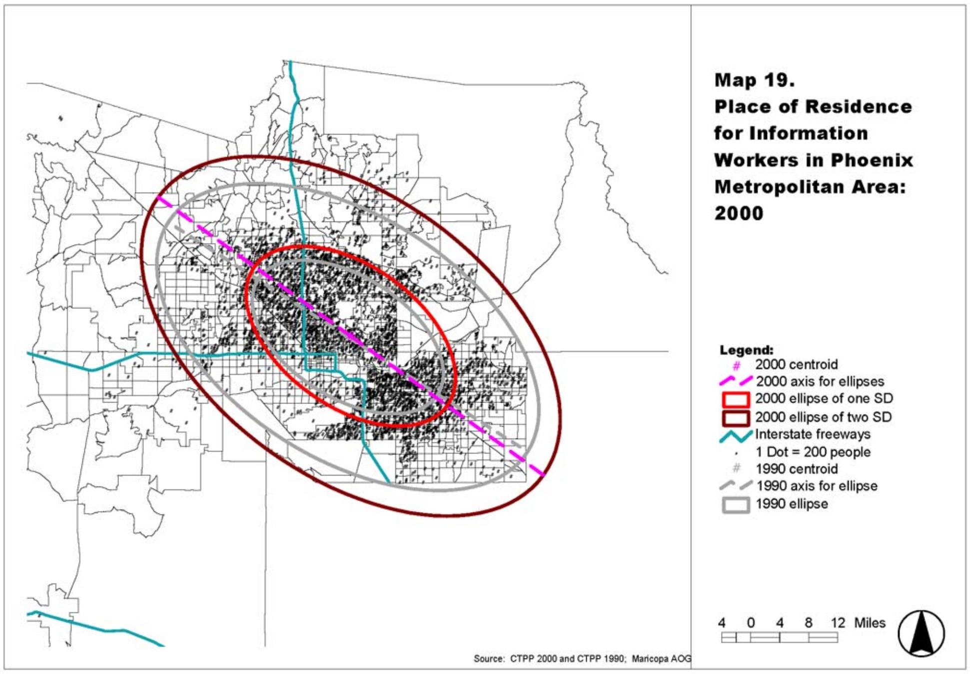

Map 14.
Place of Work
for Noninformation
Workers in Houston
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Houston Galveston Area Council

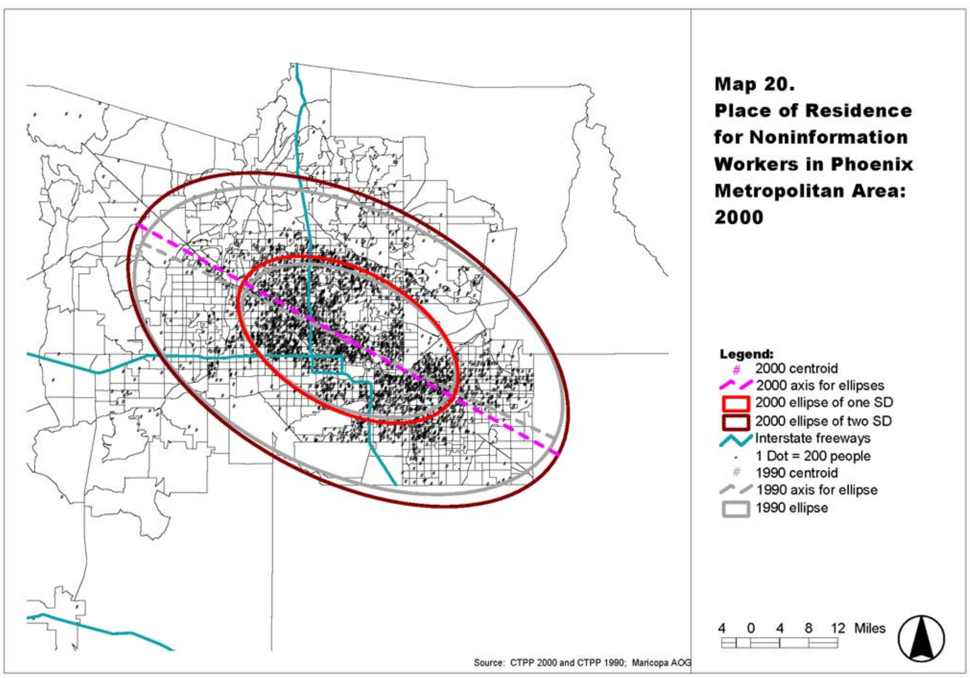

Map 15.
Place of Residence
for Information
Workers in Houston
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Houston Galveston Area Council

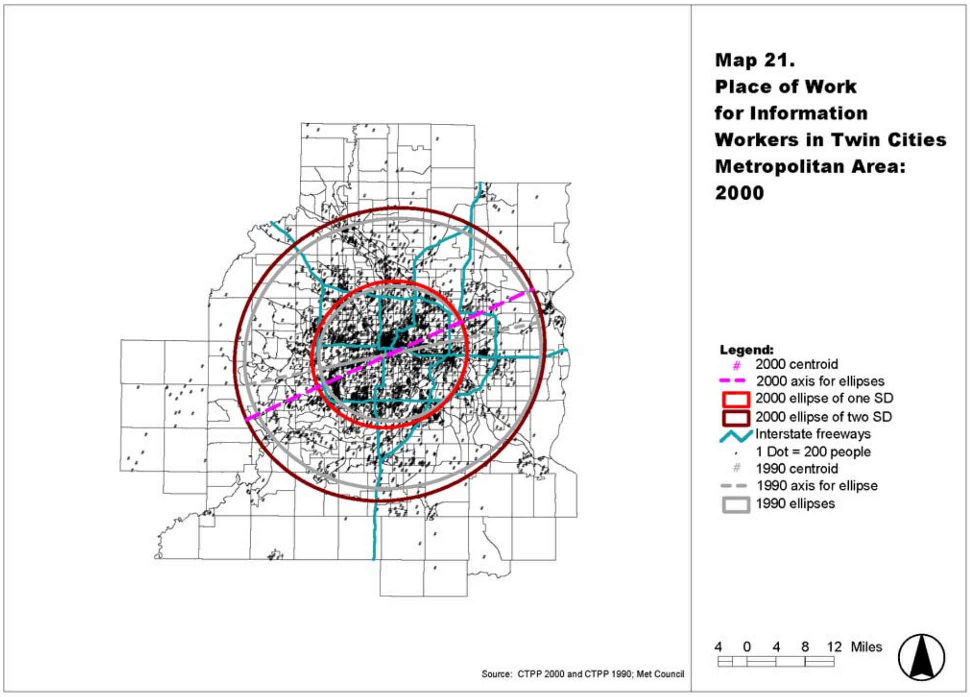

Map 16.
Place of Residence
for Non-information
Workers in Houston
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Houston Galveston Area Council

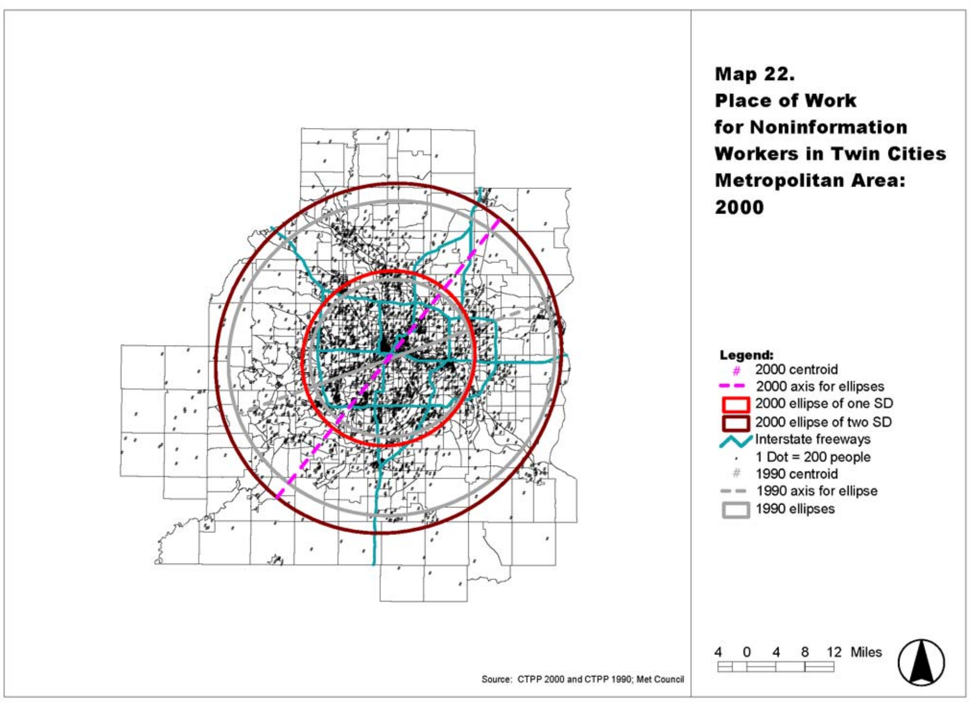

Map 17
Place of Work
for Information
Workers in Phoenix
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeway
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Maricopa AOG

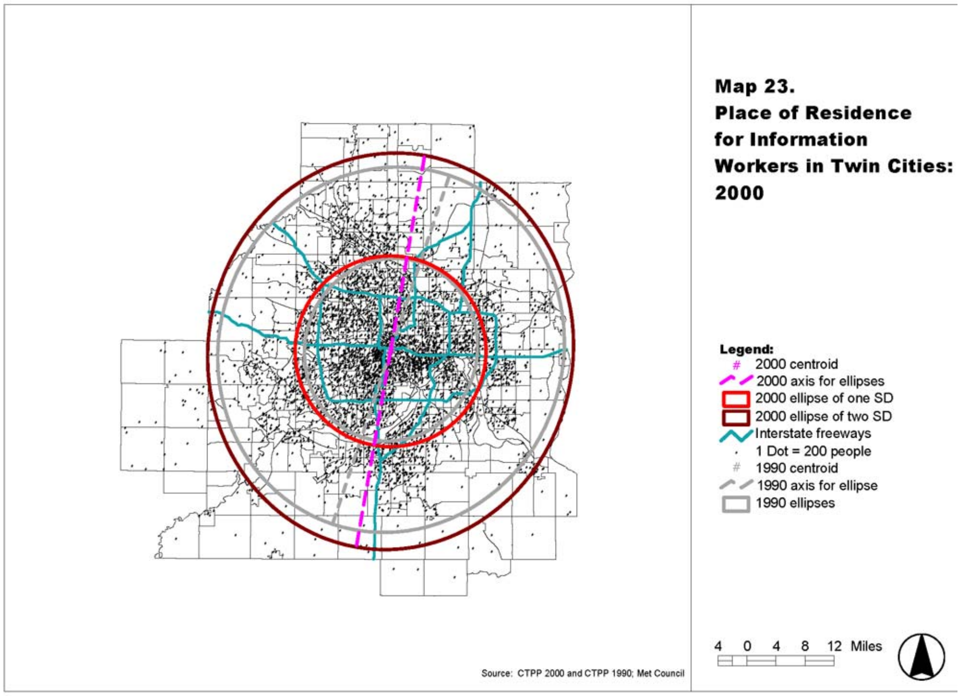

Map 18.
Place of Work
for Noninformation
Workers in Phoenix
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Maricopa AOG


Map 19.
Place of Residence
for Information
Workers in Phoenix
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Maricopa AOG

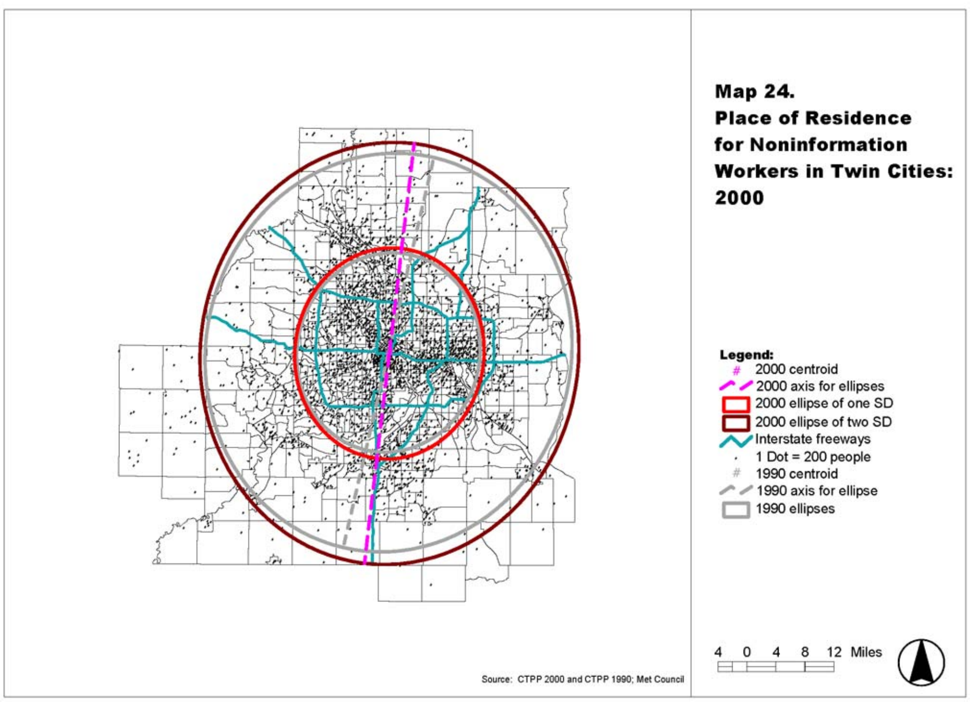

Map 20.
Place of Residence
for Noninformation
Workers in Phoenix
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipse
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Maricopa AOG


Map 21.
Place of Work
for Information
Workers in Twin Cities
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Met Council

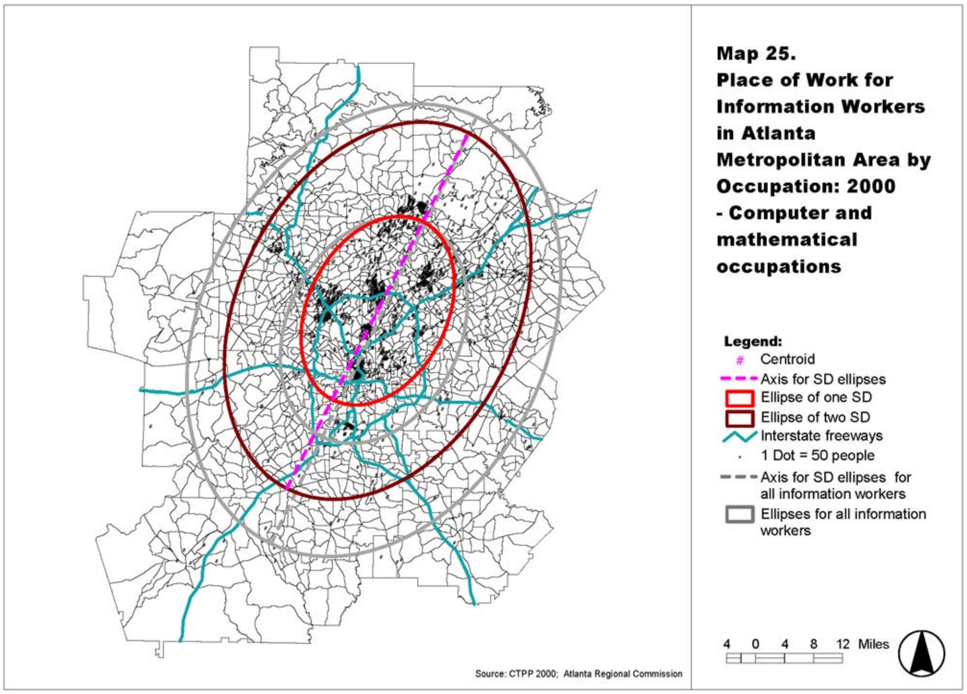

Map 22.
Place of Work
for Noninformation
Workers in Twin Cities
Metropolitan Area:
2000
Legend:
2000 centroid
2000 axis for ellipses
2000 ellipse of one SD
2000 ellipse of two SD
Interstate freeways
1 Dot = 200 people
1990 centroid
1990 axis for ellipse
1990 ellipses
4 0 4 8 12 Miles
Source: CTPP 2000 and CTPP 1990; Met Council



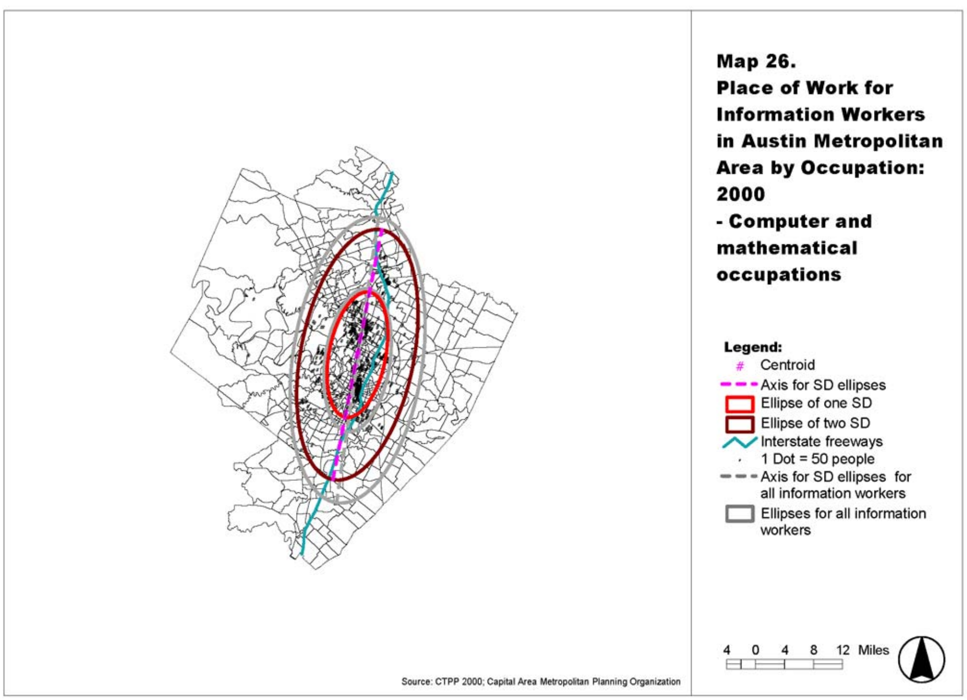
Map 23. Place of Residence for Information Workers in Twin Cities: 2000



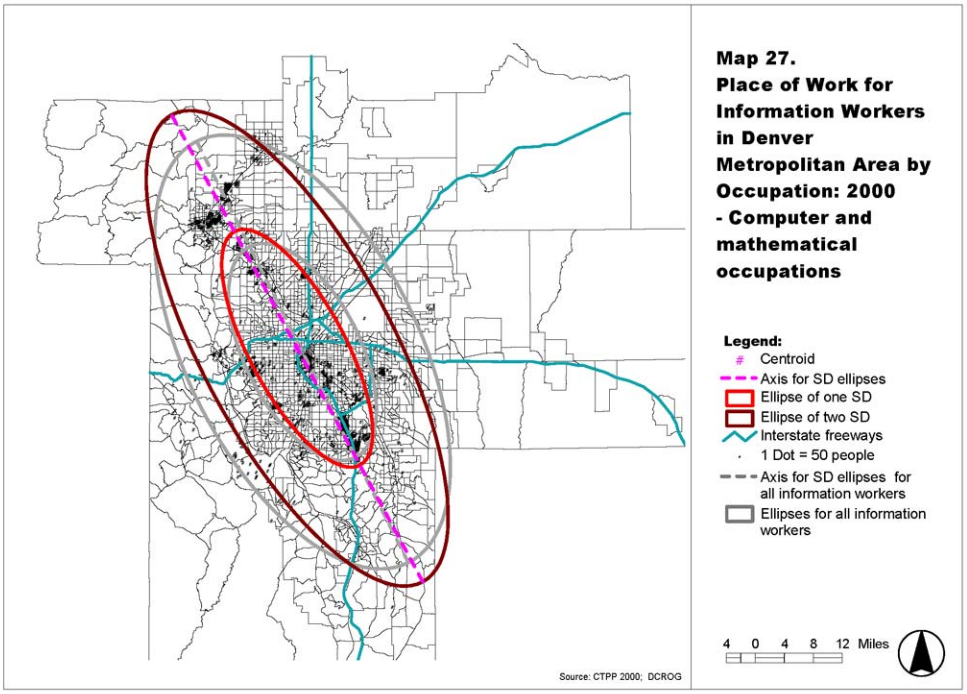
Map 24. Place of Residence for Noninformation Workers in Twin Cities: 2000



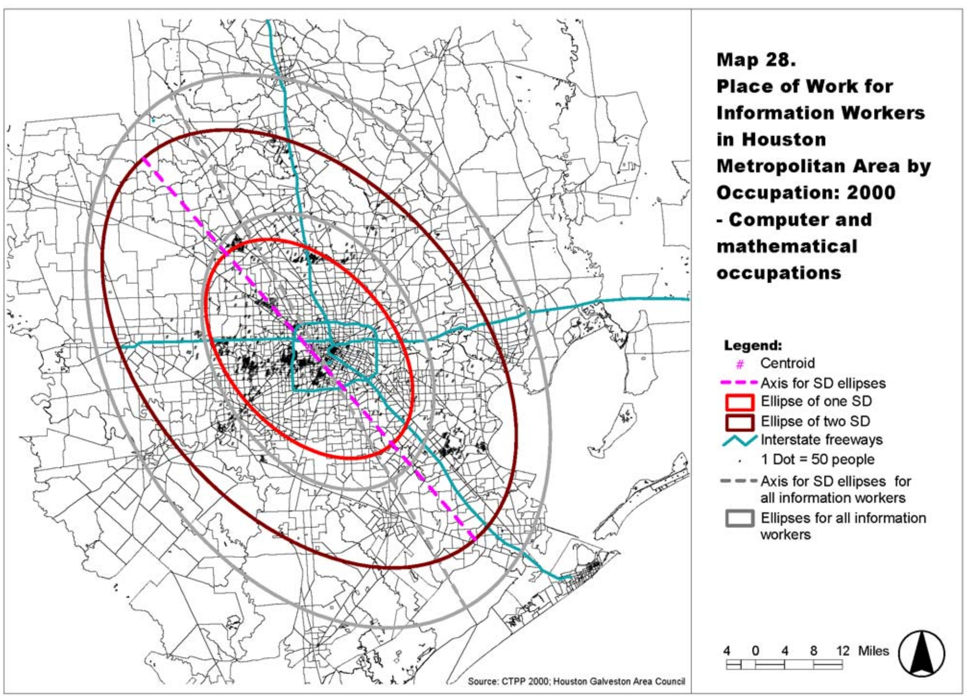
Map 25.
Place of Work for Information Workers in Atlanta Metropolitan Area by Occupation: 2000
- Computer and mathematical occupations



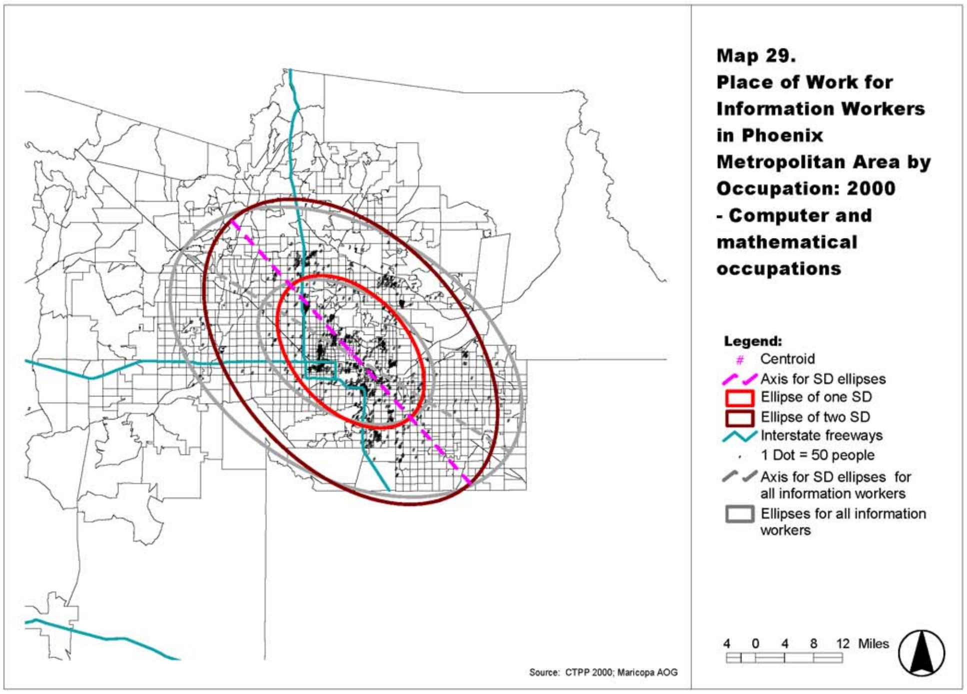
Map 26.
Place of Work for Information Workers in Austin Metropolitan Area by Occupation: 2000
- Computer and mathematical occupations

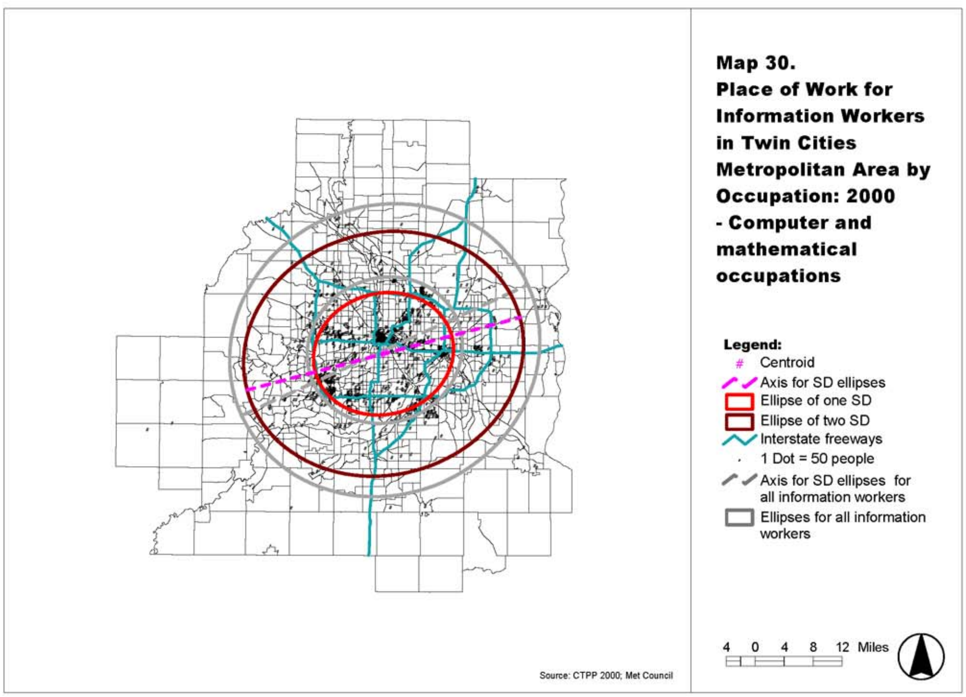


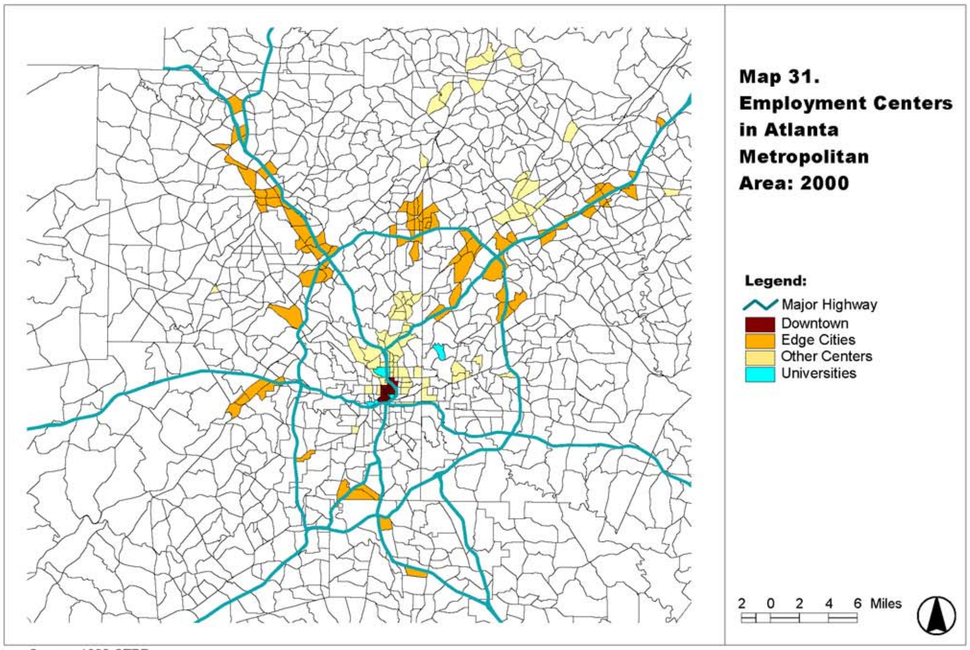
Map 27. Place of Work for Information Workers in Denver Metropolitan Area by Occupation: 2000 - Computer and mathematical occupations



Map 28. Place of Work for Information Workers in Houston Metropolitan Area by Occupation: 2000 - Computer and mathematical occupations

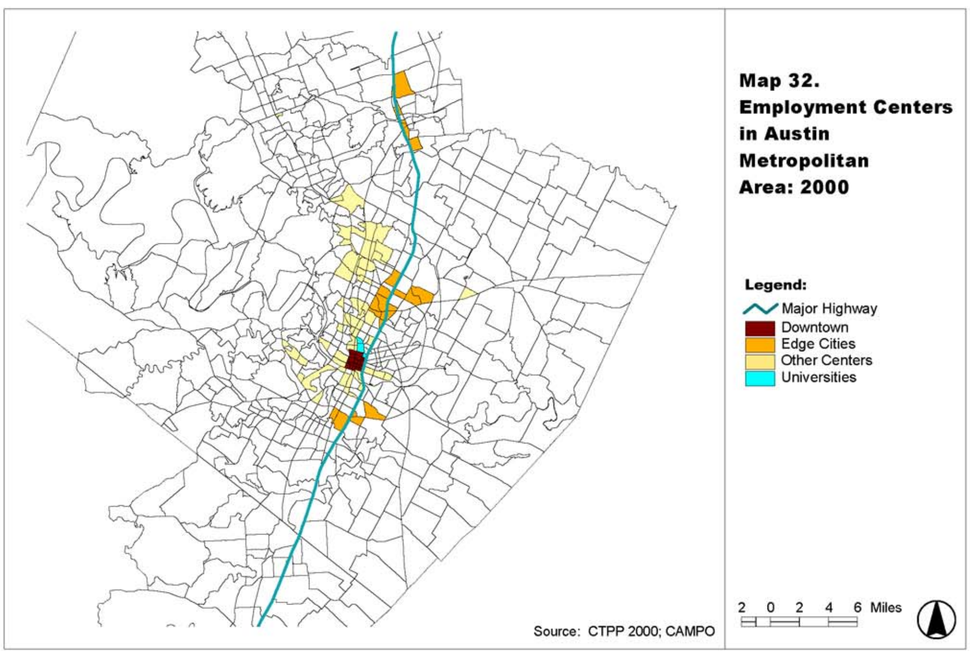

Map 29.
Place of Work for Information Workers in Phoenix Metropolitan Area by Occupation: 2000 - Computer and mathematical occupations
Legend:
Centroid
Axis for SD ellipses
Ellipse of one SD
Ellipse of two SD
Interstate freeways
1 Dot = 50 people
Axis for SD ellipses for all information workers
Ellipses for all information workers
4 0 4 8 12 Miles
Source: CTPP 2000; Maricopa AOG


Map 30.
Place of Work for Information Workers in Twin Cities Metropolitan Area by Occupation: 2000 - Computer and mathematical occupations
Legend:
Centroid
Axis for SD ellipses
Ellipse of one SD
Ellipse of two SD
Interstate freeways
1 Dot = 50 people
Axis for SD ellipses for all information workers
Ellipses for all information workers
4 0 4 8 12 Miles
Source: CTPP 2000; Met Council



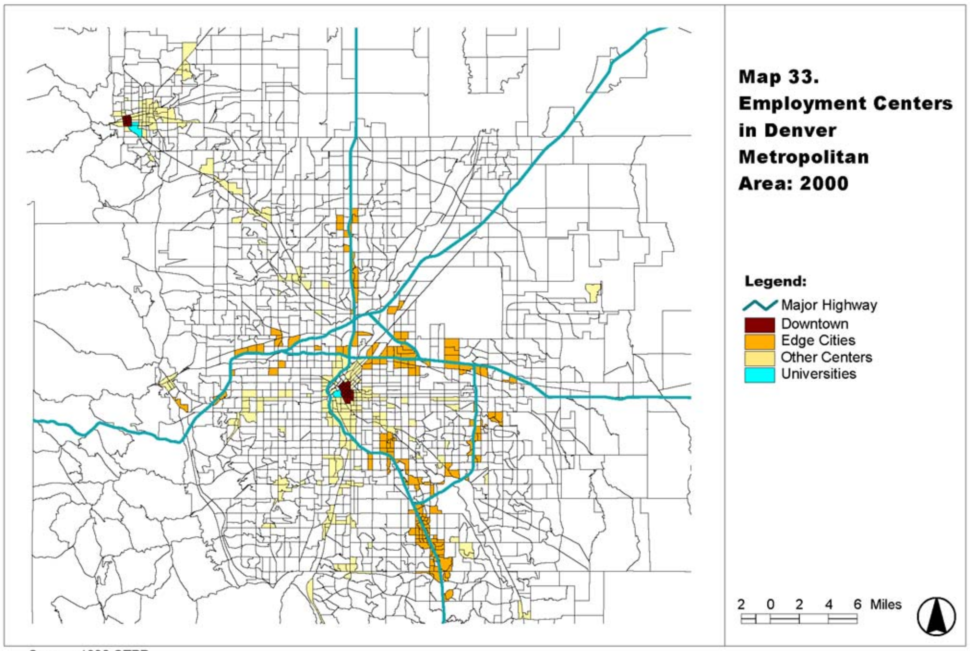
Map 31. Employment Centers in Atlanta Metropolitan Area: 2000

Source: 1990 CTPP

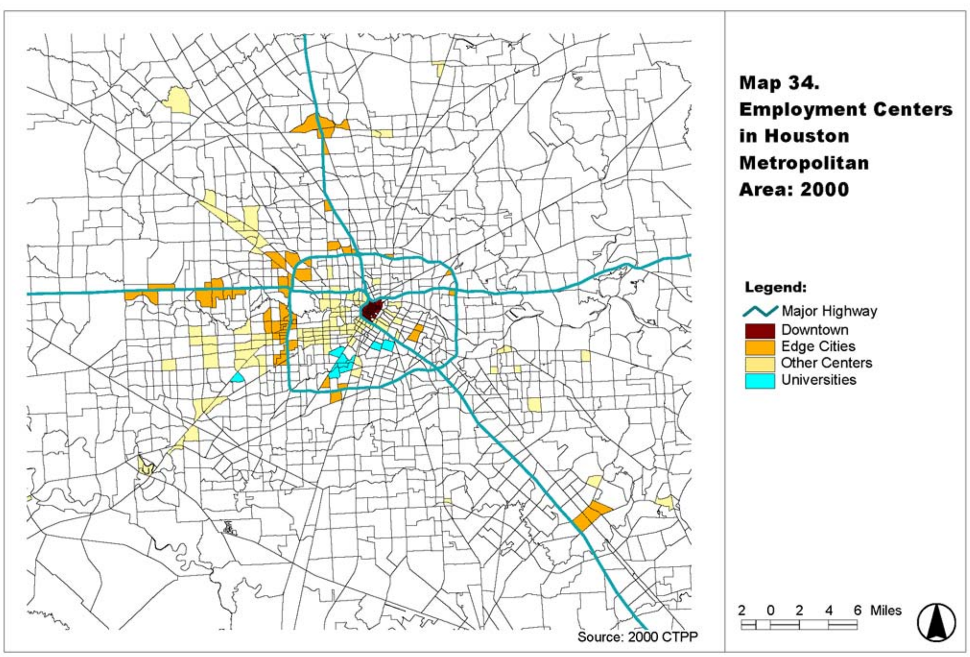


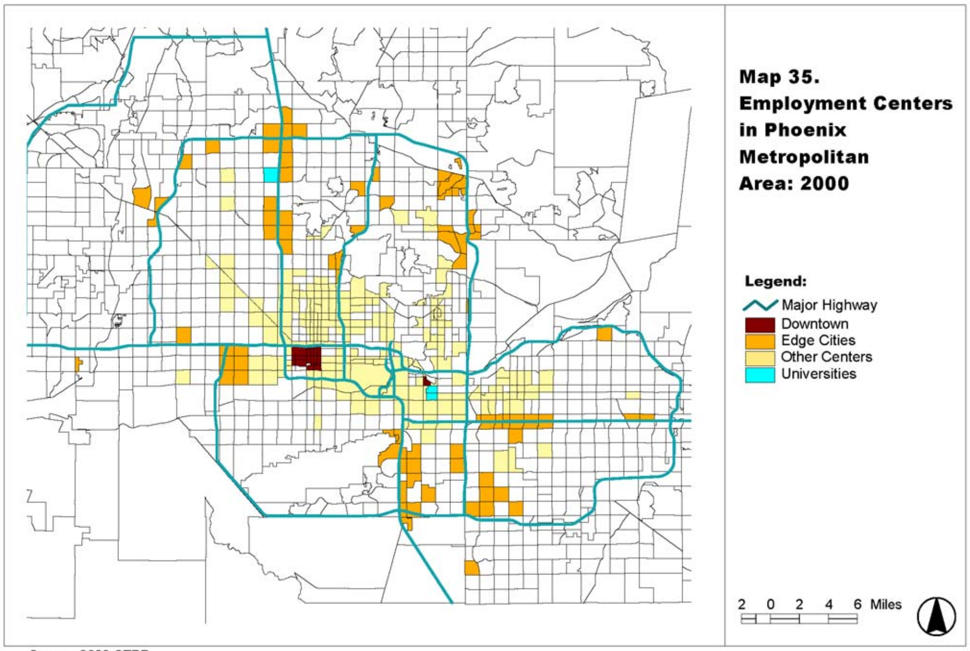
Map 32. Employment Centers in Austin Metropolitan Area: 2000

Source: CTPP 2000; CAMPO

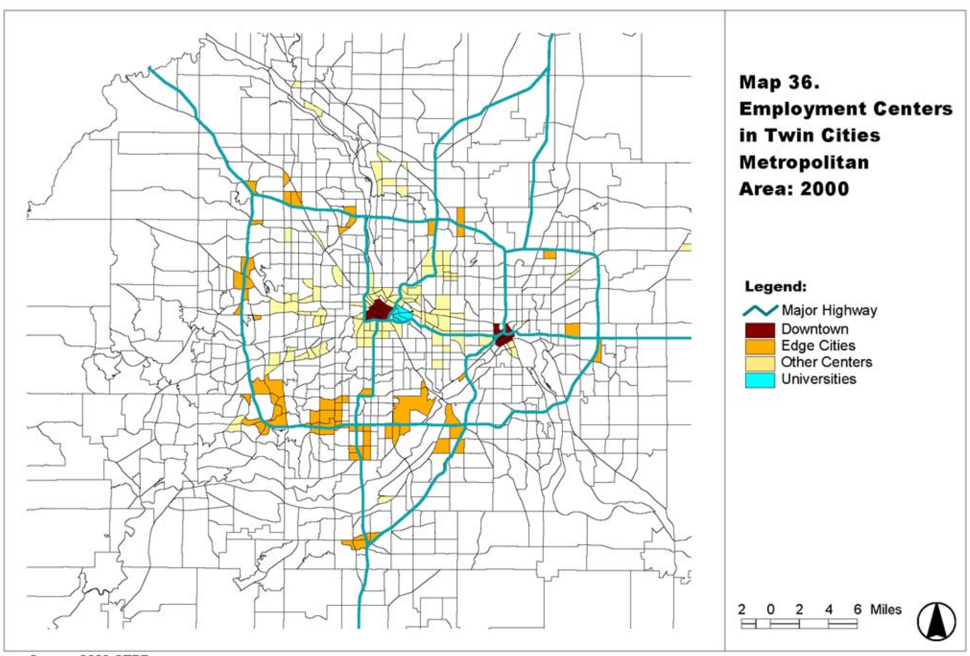


Source: 1990 CTPP

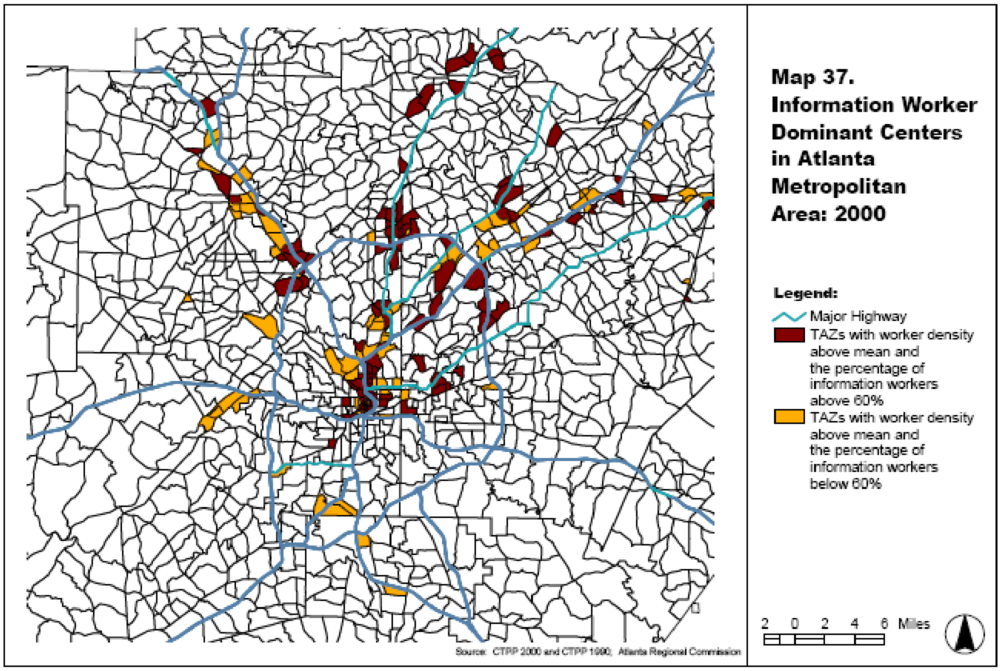




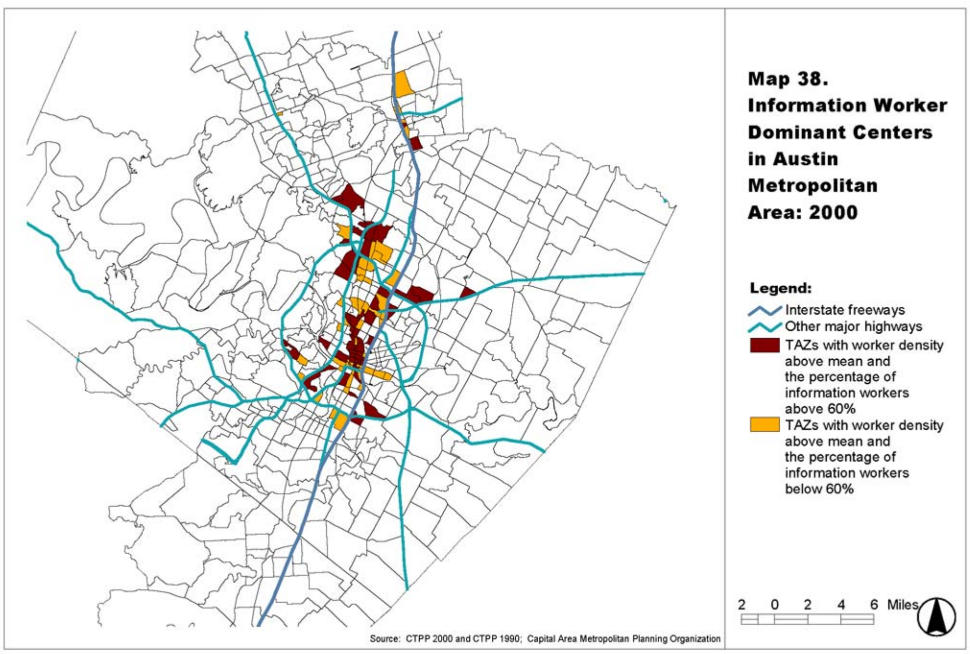
Source: 2000 CTPP



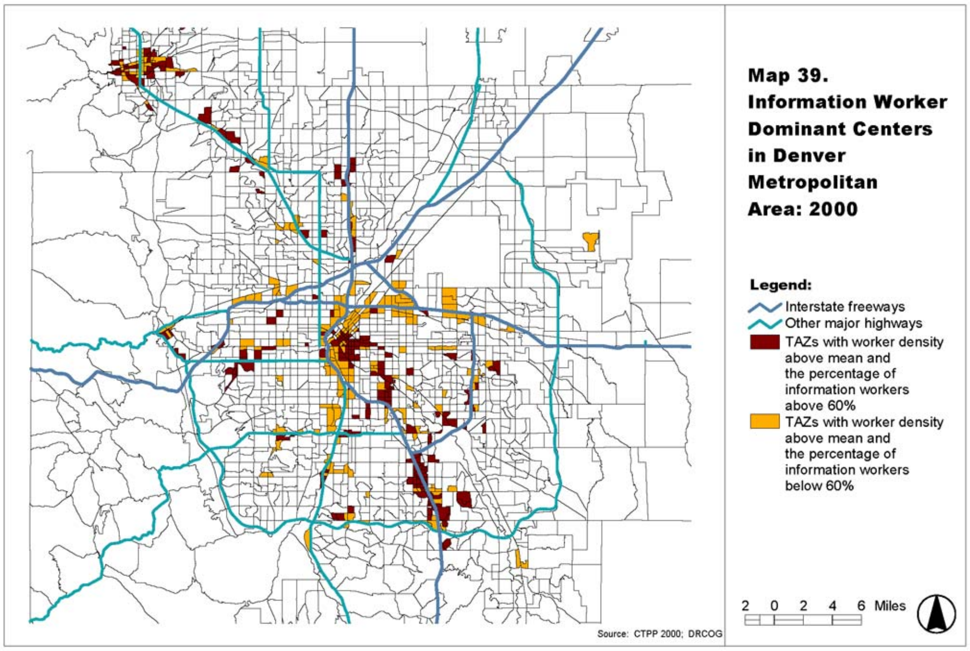
Source: 2000 CTPP



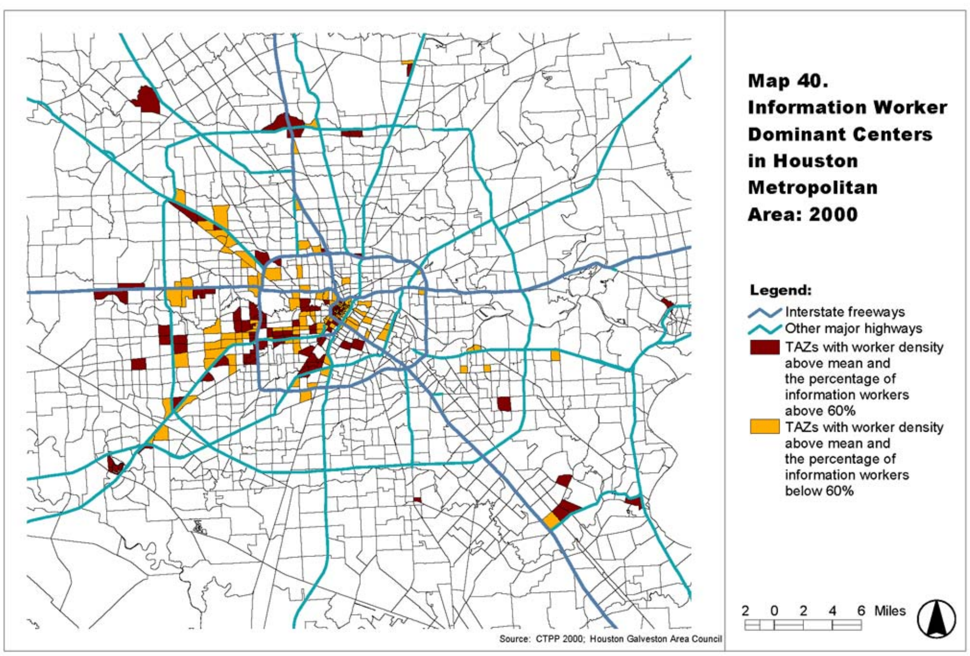
Map 37. Information Worker Dominant Centers in Atlanta Metropolitan Area: 2000



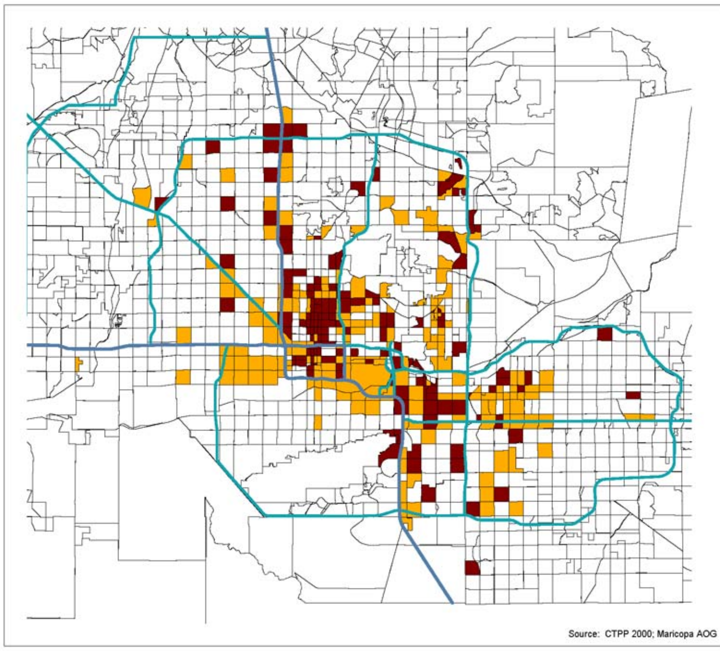
Map 38. Information Worker Dominant Centers in Austin Metropolitan Area: 2000



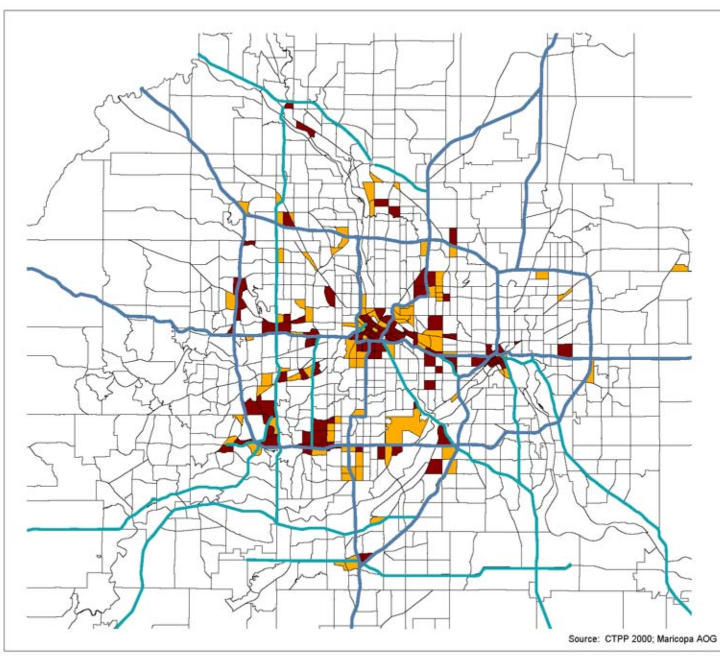
Map 39. Information Worker Dominant Centers in Denver Metropolitan Area: 2000



Map 40. Information Worker Dominant Centers in Houston Metropolitan Area: 2000

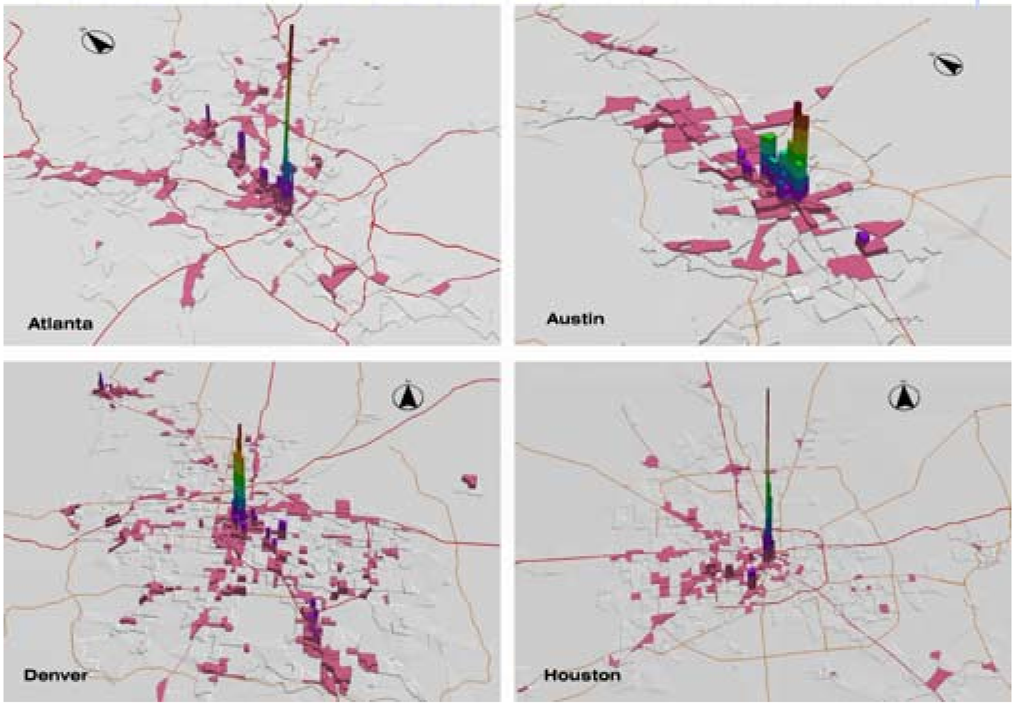


Map 41. Information Worker Dominant Centers in Phoenix Metropolitan Area: 2000



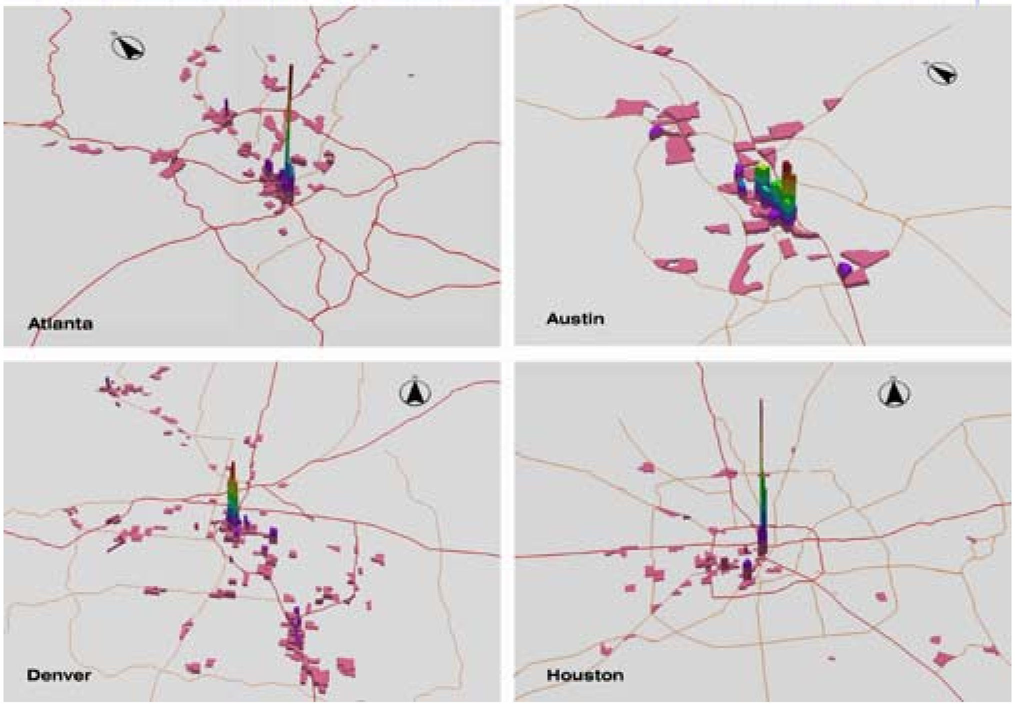
Map 42. Information Worker Dominant Centers in Twin Cities Metropolitan Area: 2000

MAP 43:All Worker Densities at All Centers: 2000

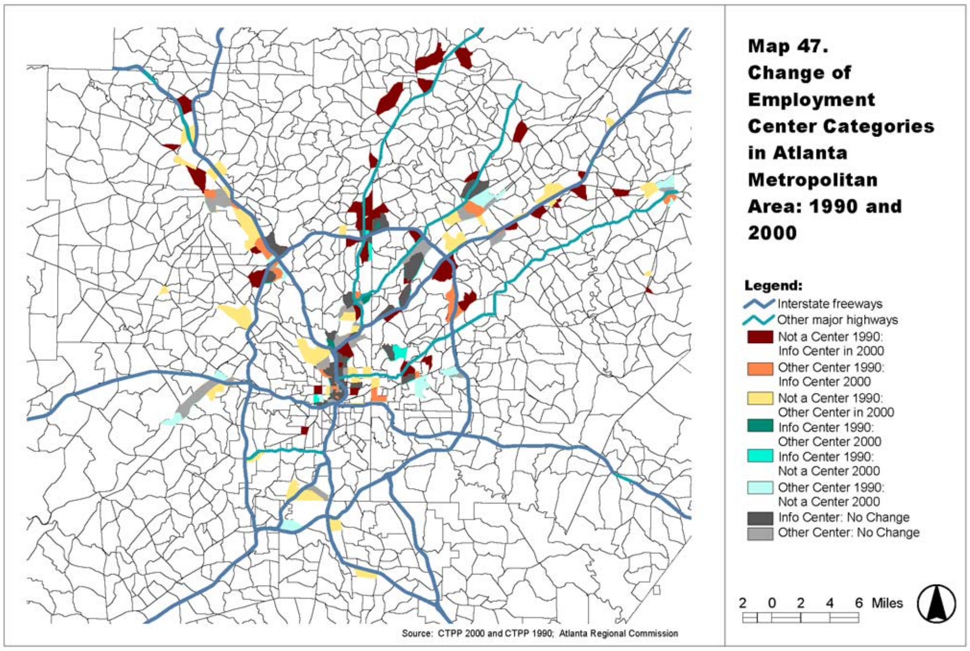


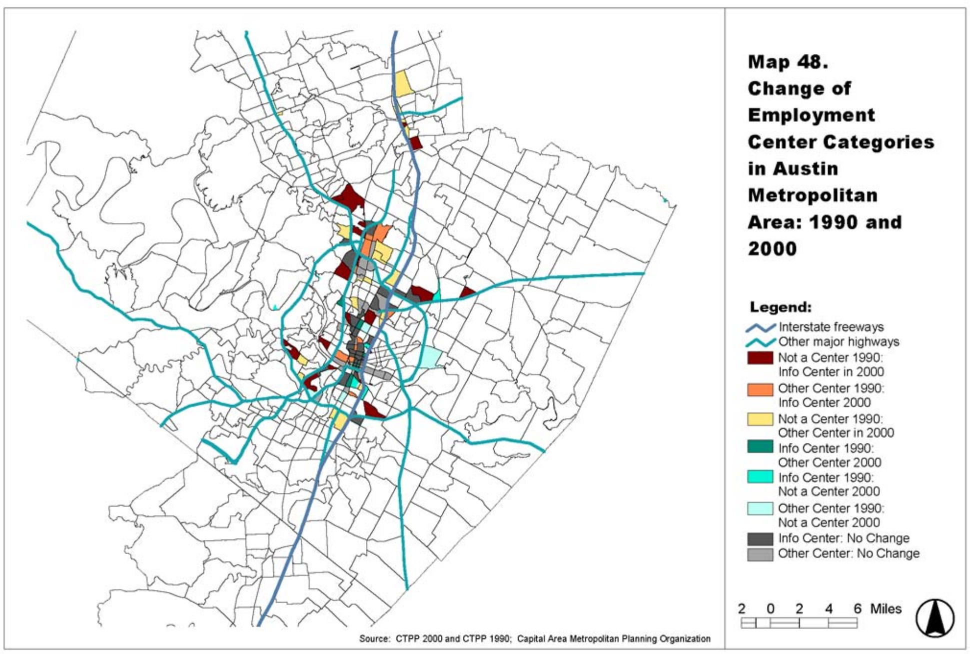
MAP 44: All Worker Densities at All Centers: 2000

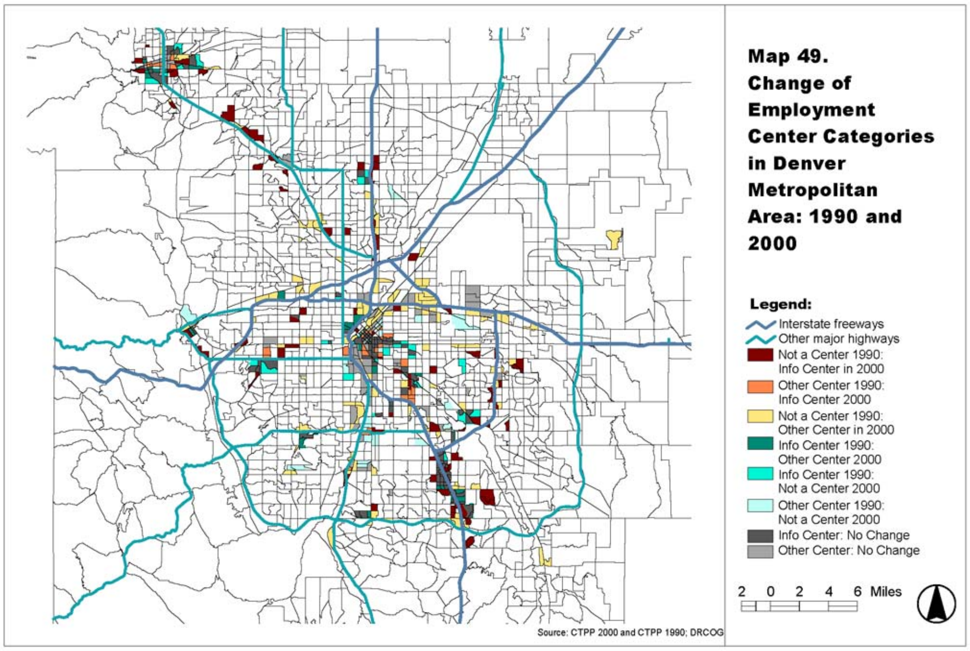
MAP 45:Information Worker Densities at Information Worker Dominant Centers: 2000



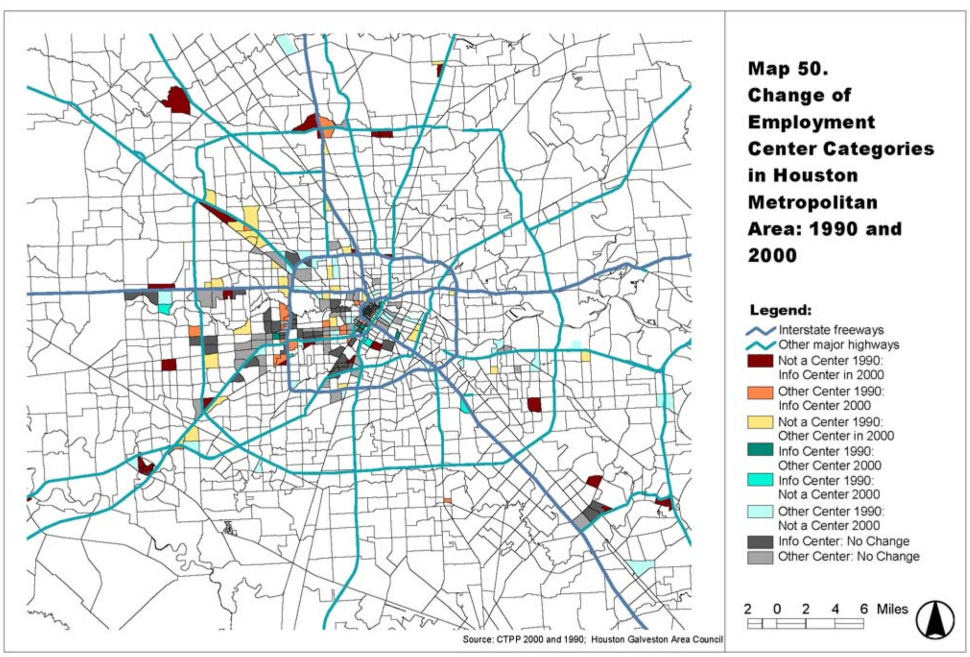
MAP 46:Information Worker Densities at Information Worker Dominant Centers: 2000


Map 47.
Change of Employment Center Categories in Atlanta Metropolitan Area: 1990 and 2000
Legend:
Interstate freeways
Other major highways
Not a Center 1990: Info Center in 2000
Other Center 1990: Info Center 2000
Not a Center 1990: Other Center in 2000
Info Center 1990: Other Center 2000
Info Center 1990: Not a Center 2000
Other Center 1990: Not a Center 2000
Info Center: No Change
Other Center: No Change
2 0 2 4 6 Miles
Source: CTPP 2000 and CTPP 1990; Atlanta Regional Commission

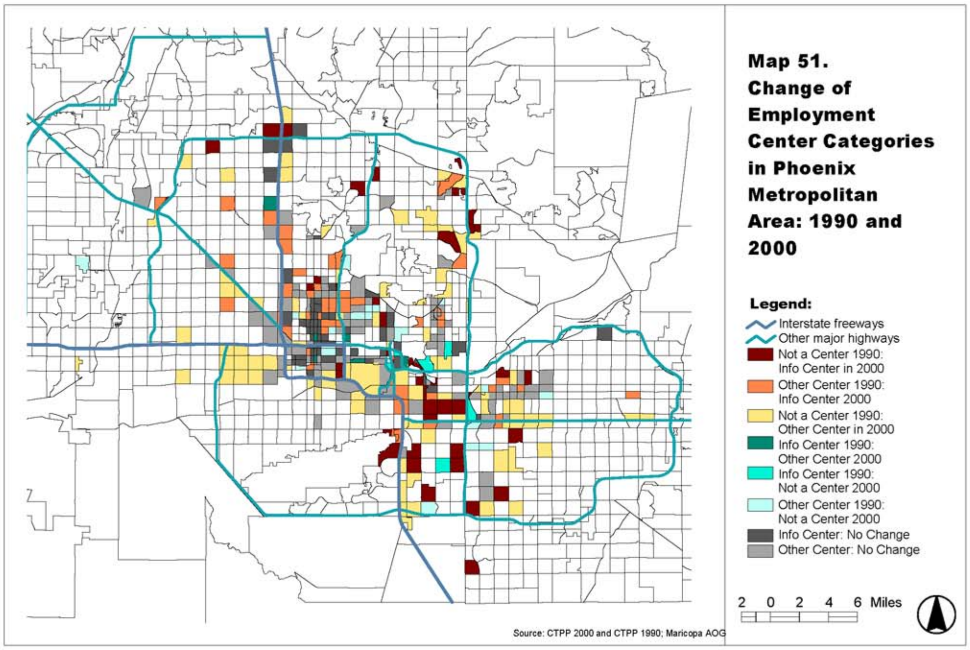

Map 48.
Change of Employment Center Categories in Austin Metropolitan Area: 1990 and 2000
Legend:
Interstate freeways
Other major highways
Not a Center 1990: Info Center in 2000
Other Center 1990: Info Center 2000
Not a Center 1990: Other Center in 2000
Info Center 1990: Other Center 2000
Info Center 1990: Not a Center 2000
Other Center 1990: Not a Center 2000
Info Center: No Change
Other Center: No Change
2 0 2 4 6 Miles
Source: CTPP 2000 and CTPP 1990; Capital Area Metropolitan Planning Organization



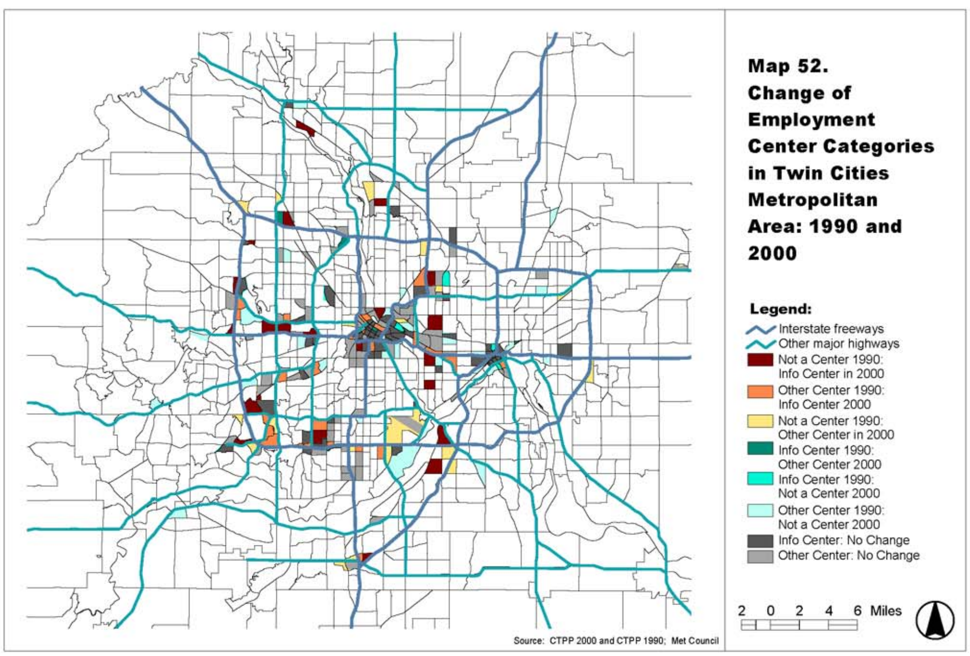
Map 49. Change of Employment Center Categories in Denver Metropolitan Area: 1990 and 2000



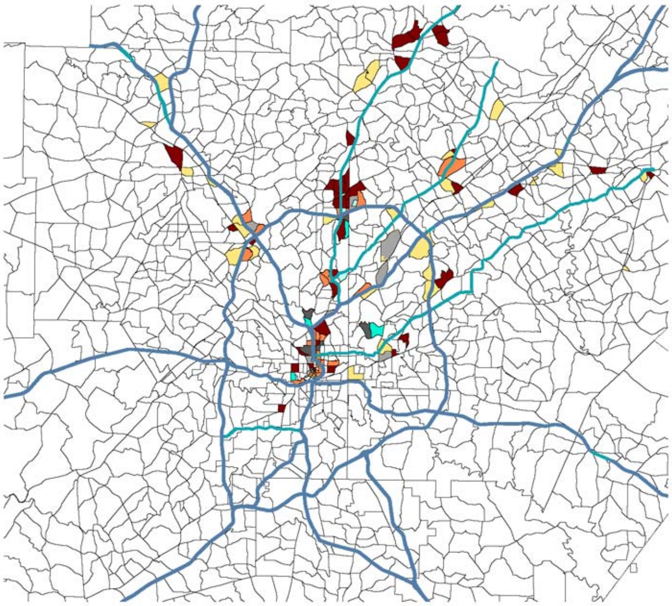
Map 50. Change of Employment Center Categories in Houston Metropolitan Area: 1990 and 2000



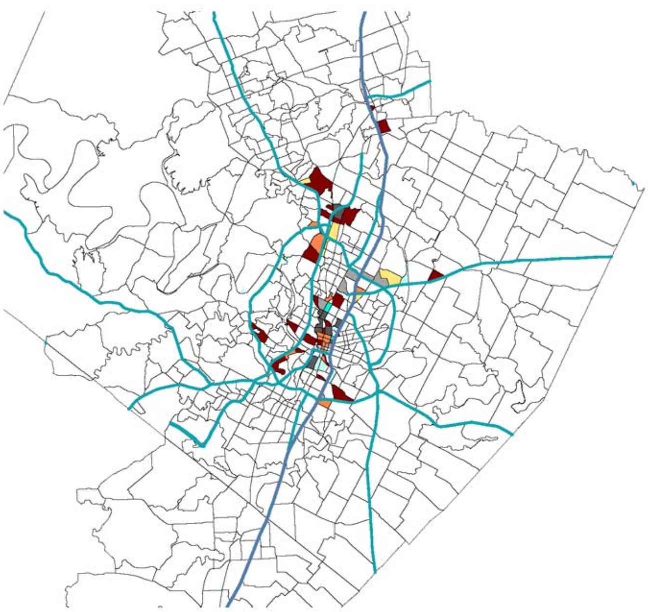
Map 51.
Change of
Employment
Center Categories
in Phoenix
Metropolitan
Area: 1990 and
2000
Legend:
Interstate freeways
Other major highways
Not a Center 1990: Info Center in 2000
Other Center 1990: Info Center 2000
Not a Center 1990: Other Center in 2000
Info Center 1990: Other Center 2000
Info Center 1990: Not a Center 2000
Other Center 1990: Not a Center 2000
Info Center: No Change
Other Center: No Change
2 0 2 4 6 Miles
Source: CTPP 2000 and CTPP 1990; Maricopa AOG



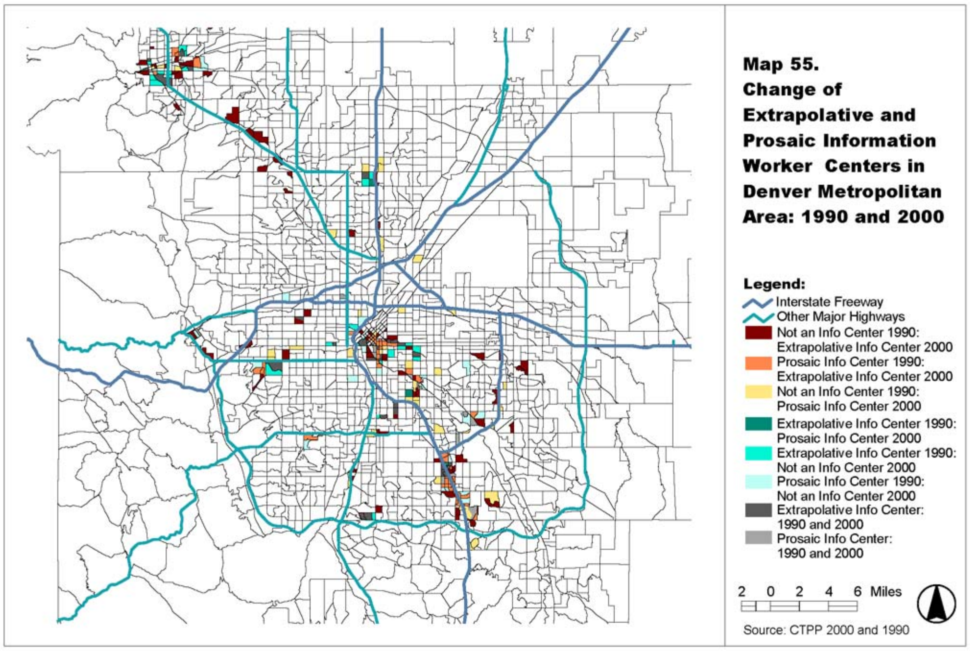
Map 52.
Change of
Employment
Center Categories
in Twin Cities
Metropolitan
Area: 1990 and
2000
Legend:
Interstate freeways
Other major highways
Not a Center 1990: Info Center in 2000
Other Center 1990: Info Center 2000
Not a Center 1990: Other Center in 2000
Info Center 1990: Other Center 2000
Info Center 1990: Not a Center 2000
Other Center 1990: Not a Center 2000
Info Center: No Change
Other Center: No Change
2 0 2 4 6 Miles
Source: CTPP 2000 and CTPP 1990; Met Council



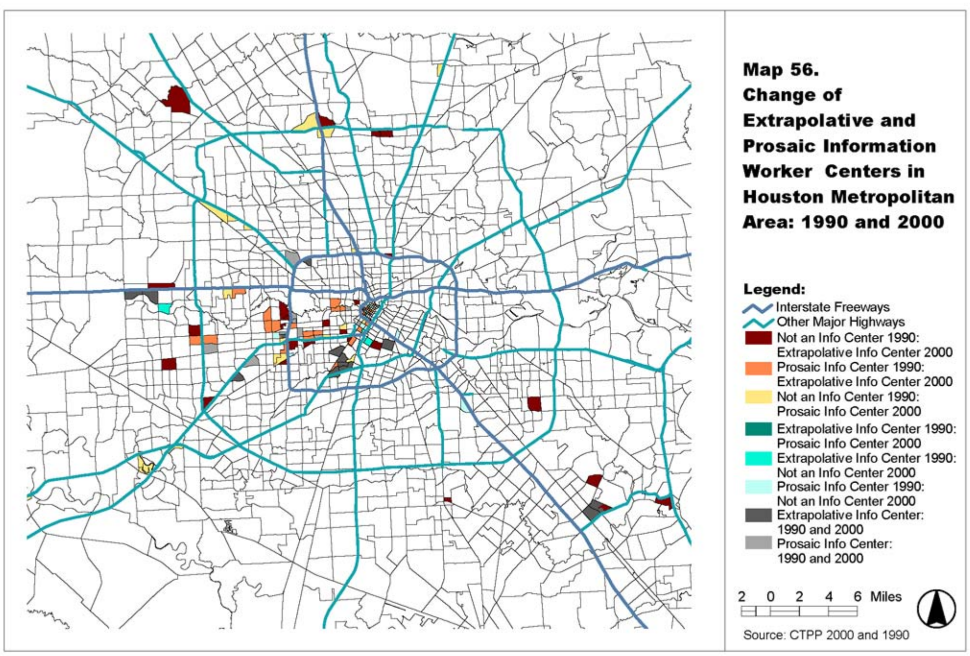
Map 53. Change of Extrapolative and Prosaic Information Worker Centers in Atlanta Metropolitan Area: 1990 and 2000



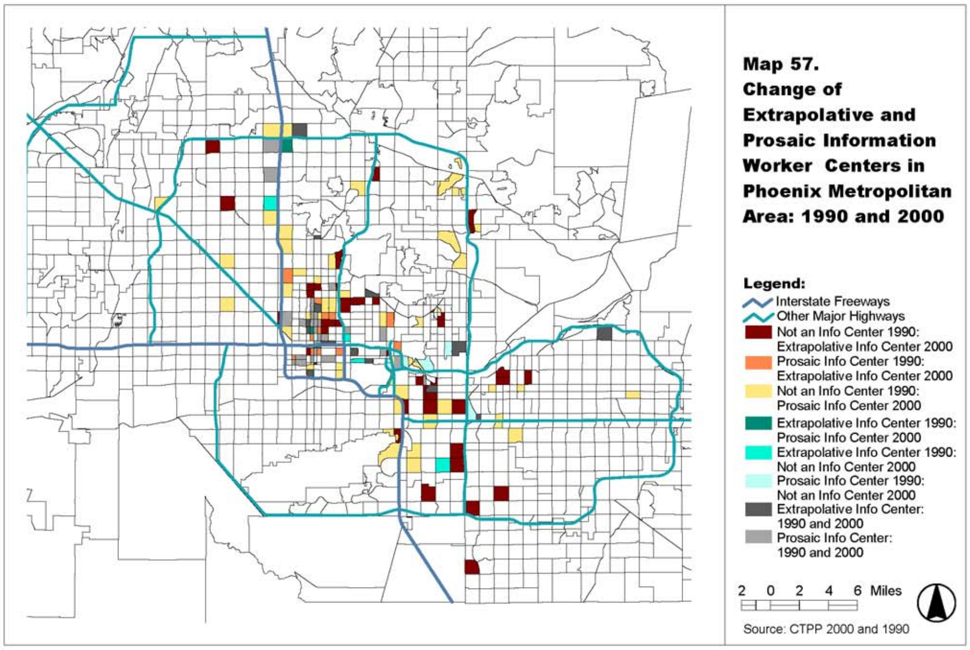
Map 54. Change of Extrapolative and Prosaic Information Worker Centers in Austin Metropolitan Area: 1990 and 2000



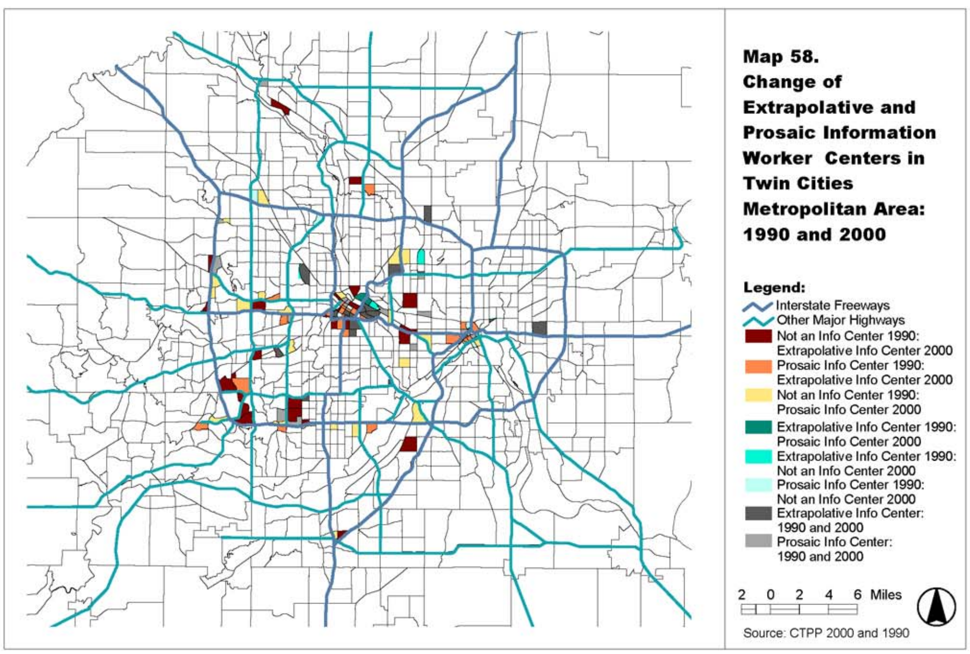
Map 55. Change of Extrapolative and Prosaic Information Worker Centers in Denver Metropolitan Area: 1990 and 2000



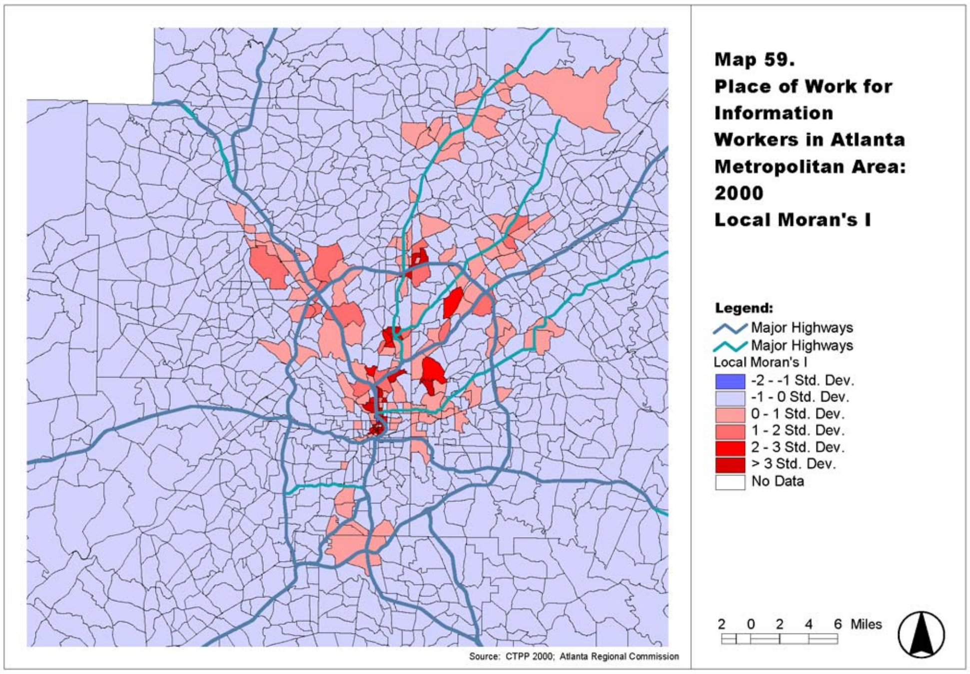
Map 56. Change of Extrapolative and Prosaic Information Worker Centers in Houston Metropolitan Area: 1990 and 2000



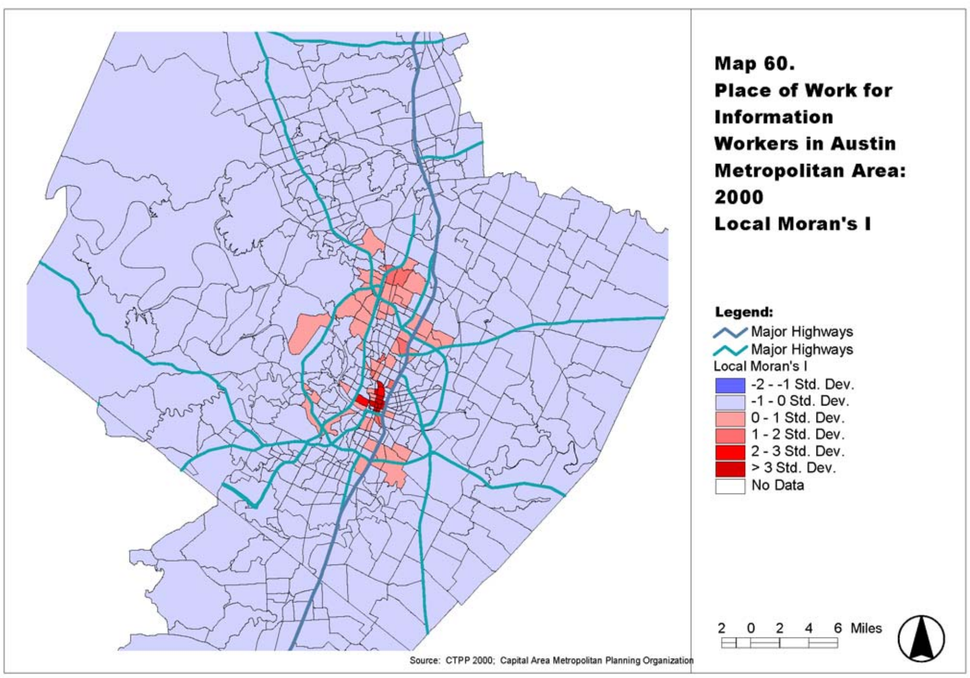
Map 57. Change of Extrapolative and Prosaic Information Worker Centers in Phoenix Metropolitan Area: 1990 and 2000



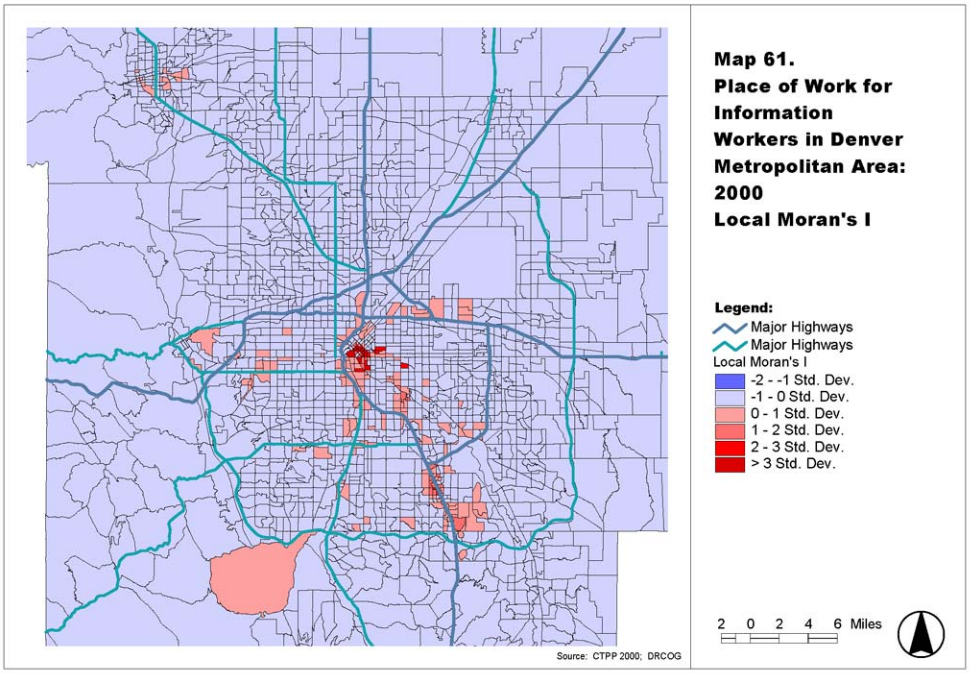
Map 58. Change of Extrapolative and Prosaic Information Worker Centers in Twin Cities Metropolitan Area: 1990 and 2000



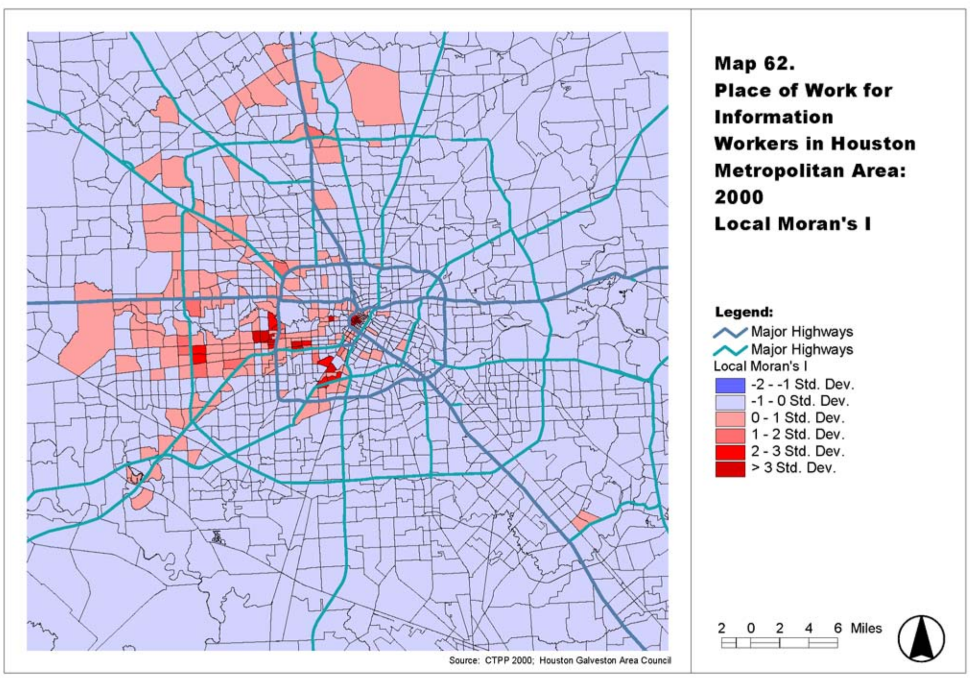
Map 59. Place of Work for Information Workers in Atlanta Metropolitan Area: 2000 Local Moran's I



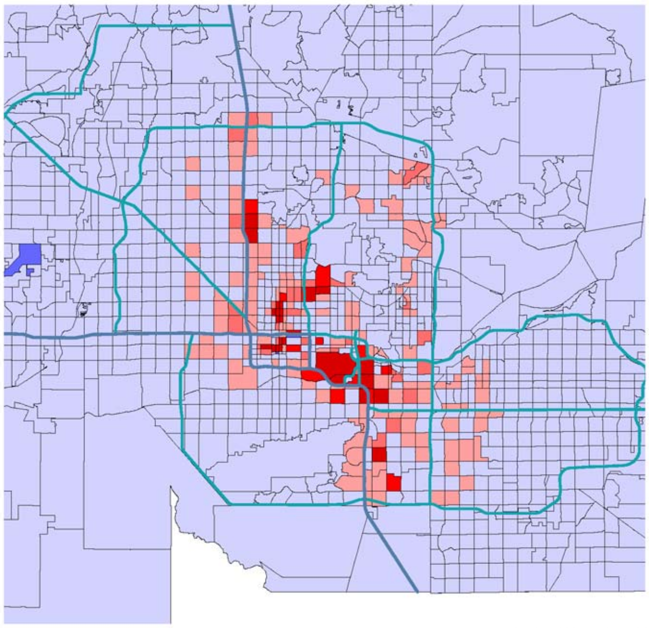
Map 60. Place of Work for Information Workers in Austin Metropolitan Area: 2000 Local Moran's I



**Map 61.
Place of Work for Information Workers in Denver Metropolitan Area: 2000
Local Moran's I**



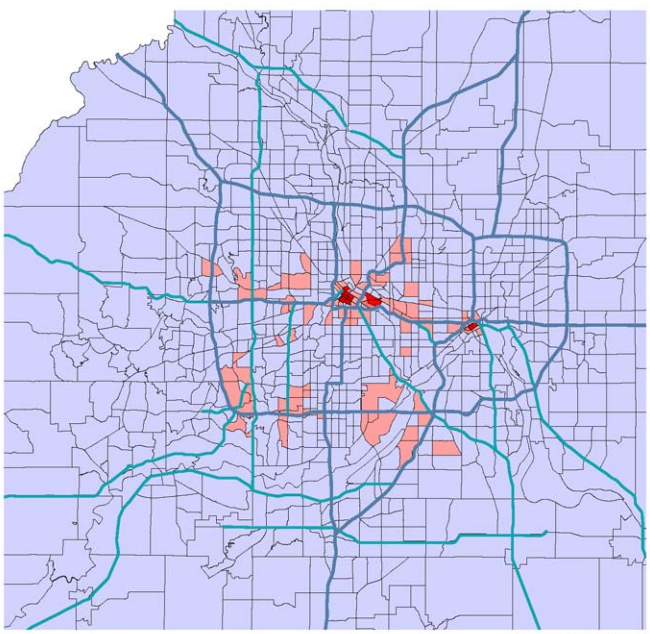
**Map 62.
Place of Work for Information Workers in Houston Metropolitan Area: 2000
Local Moran's I**



Map 63.
Place of Work for Information Workers in Pheonix Metropolitan Area: 2000
Local Moran's I

Legend:
Major Highways
Major Highways
Local Moran's I
-2 - -1 Std. Dev.
-1 - 0 Std. Dev.
0 - 1 Std. Dev.
1 - 2 Std. Dev.
2 - 3 Std. Dev.
> 3 Std. Dev.

2 0 2 4 6 Miles



Map 64.
Place of Work for Information Workers in Twin Cities Metropolitan Area: 2000
Local Moran's I

Legend:
Major Highways
Major Highways
Local Moran's I
-2 - -1 Std. Dev.
-1 - 0 Std. Dev.
0 - 1 Std. Dev.
1 - 2 Std. Dev.
2 - 3 Std. Dev.
> 3 Std. Dev.

2 0 2 4 6 Miles

Appendix C

User Perspectives on the Minnesota Interorganizational Mayday Information System

Study Participants and Types of Semi-structured Interview Questions Asked

Participant Organization	Position Title
Minnesota State Patrol	Communications Center Supervisor Emergency Communications Dispatcher Trooper Colonel
Department of Transportation, Office of Traffic Safety	Project Manager
Mayo Clinic Emergency Department	Emergency Physician, Director Emergency Physician Clinical Director Hospital Administrator
Mayo Medical Transport	Manager Paramedic Communications Specialist
Mayo Communications Dispatch Center	Supervisor Dispatch Operator
Software Development Consultant	Consultant
GM OnStar	Communications Center Manager Technician
Total Expert Participants: 17 Round discussions: 2	

Example Semi-Structured Interview Questions

Operational Dynamics

In regards to the Mayday system:

- Describe the end-to-end EMS service process?
- Describe your communication processes with organization A.
    - Who do you communicate with?
- When, in what circumstances (or service leg) does it take place?
- What information do you communicate (via voice or data)?
- Describe the data/information collection process.
- What information technologies does your organization use? (relative to incident informationexchange).
- What emergency related information does your organization not send/receive that you think should be?
- What specific service performance information do you obtain and use relative to your role in EMS?
    - What performance does your organization track?
    - What performance reports does your organization produce?

Organizational and Governance Dynamics

In regards to the Mayday system:

- What conditions inhibit or prohibit information sharing?
- What challenges does your organization face in terms of information sharing (with organization A)?
- What are the benefits of information sharing (a type of information with organization A)? Who decides what information you will share? How it will be shared? Who has oversight? Why?
- What role has IT played in enhancing/degrading information sharing?
- Why does your organization not send/receive (a type of) information?

Appendix D

Induced Demand: A Microscopic Perspective

Table D.1 Lag linear difference model for interstates

	Two-year lag $n = 8736$			Four-year lag $n = 6552$			Six-year lag $n = 4368$			Eight-year lag $n = 2184$		
Variable	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$
Delvkt	− 0.224	− 18.200	0.000	− 0.047	− 3.040	0.002	0.121	6.280	0.000	− 0.231	− 8.090	0.000
Delparalvkt	− 0.007	− 0.400	0.689	− 0.013	− 0.650	0.513	− 0.050	− 1.860	0.062	0.140	3.520	0.000
Dellanes	− 897.475	− 1.910	0.056	− 1004.312	− 2.110	0.035	575.491	1.080	0.078	2156.607	3.180	0.002
Delparalanes	− 344.935	− 0.750	0.451	− 682.426	− 1.450	0.148	995.676	1.920	0.055	3227.110	4.670	0.000
Delmcdpop	− 0.118	− 3.650	0.000	− 0.116	− 3.640	0.000	0.148	3.600	0.000	0.245	5.070	0.000
Deladjpop	− 0.014	− 0.850	0.393	0.015	1.140	0.256	− 0.030	− 2.210	0.027	− 0.051	− 3.260	0.001
Basevkt	0.097	41.740	0.000	0.170	39.160	0.000	0.270	37.590	0.000	0.454	43.220	0.000
Baseparalvkt	− 0.020	− 6.800	0.000	− 0.037	− 7.130	0.000	− 0.067	− 7.560	0.000	− 0.141	− 9.420	0.000
Baselanes	40.592	0.450	0.654	− 129.329	− 0.930	0.353	− 300.717	− 1.390	0.163	− 1414.214	− 4.270	0.000
Baseparalanes	− 45.514	− 0.430	0.665	− 145.272	− 0.900	0.370	− 226.041	− 0.910	0.364	− 460.144	− 1.200	0.231
Basemcdpop	− 0.001	− 1.360	0.173	− 0.004	− 2.990	0.003	− 0.006	− 3.100	0.002	− 0.009	− 3.200	0.001
Baseadjpop	− 0.003	− 4.660	0.000	− 0.008	− 8.680	0.000	− 0.013	− 9.400	0.000	− 0.012	− 6.210	0.000
Baseyrdum96	− 2008.057	− 6.550	0.000									
Baseyrdum94	− 3712.164	− 12.200	0.000	121.475	0.310	0.757						
Baseyrdum92	− 1436.683	− 4.650	0.000	741.420	1.740	0.083	517.305	0.940	0.346			
Baseyrdum90	− 2042.880	− 5.710	0.000	2059.803	4.910	0.000	− 284.327	− 0.540	0.587	449.419	0.800	0.424
Baseyrdum88	− 2909.350	− 9.500	0.000	180.690	0.460	0.644	655.772	1.330	0.183			
Baseyrdum86	− 2238.686	− 7.380	0.000	925.815	2.370	0.018						
Baseyrdum84	− 2805.922	− 8.990	0.000									
Constant	4055.013	12.000	0.000	3816.970	7.450	0.000	6856.693	8.890	0.000	8532.126	7.400	0.000
Inverse Mills ratio	0.102	6.730	0.000	1.322	2.170	0.030	0.562	0.440	0.662	− 4.310	− 1.910	0.056
R^2	0.2192			0.3552			0.4811			0.5485		

Table D.2 Lag linear difference model for trunk highways

	Two-year lag $n = 15\,914$			Four-year lag $n = 11\,924$			Six-year lag $n = 7936$			Eight-year lag $n = 3956$		
Variable	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$
Delvkt	−0.206	−23.730	0.000	−0.053	−5.120	0.000	0.039	2.480	0.013	0.012	0.480	0.634
Delparalvkt	−0.007	−1.100	0.270	−0.022	−2.740	0.006	−0.049	−3.970	0.000	−0.061	−2.640	0.008
Dellanes	−1159.346	−2.120	0.034	6195.932	9.400	0.000	2800.214	3.250	0.001	1843.694	1.620	0.104
Delparalanes	555.918	1.570	0.116	658.576	1.660	0.097	−954.399	−1.960	0.050	−1788.911	−2.670	0.008
Delmcdpop	0.010	0.700	0.485	0.039	2.810	0.005	0.007	0.420	0.674	−0.027	−1.230	0.218
Deladjpop	−0.015	−2.020	0.044	0.006	0.960	0.339	0.020	2.590	0.010	0.025	2.390	0.017
Basevkt	0.063	35.070	0.000	0.110	35.630	0.000	0.137	24.920	0.000	0.197	19.110	0.000
Baseparalvkt	0.003	2.160	0.031	0.008	2.780	0.005	0.016	3.480	0.000	0.022	2.540	0.011
Baselanes	242.579	4.890	0.000	347.363	4.530	0.000	483.508	3.950	0.000	775.658	3.700	0.000
Baseparalanes	−110.203	−2.520	0.012	−271.275	−4.000	0.000	−426.389	−3.940	0.000	−600.702	−3.210	0.001
Basemcdpop	−0.003	−8.480	0.000	−0.004	−7.790	0.000	−0.005	−6.110	0.000	−0.008	−5.480	0.000
Baseadjpop	0.000	−0.520	0.600	−0.001	−2.120	0.034	−0.002	−4.280	0.000	−0.003	−3.750	0.000
Baseyrdum96	−352.627	−3.140	0.002									
Baseyrdum94	−475.994	−4.260	0.000	−508.873	−3.470	0.000						
Baseyrdum92	−473.951	−4.190	0.000	−834.509	−5.260	0.000	−655.624	−3.190	0.001			
Baseyrdum90	−10.839	−0.080	0.934	−446.172	−2.820	0.005	−358.391	−1.790	0.074	18.473	0.080	0.940
Baseyrdum88	−419.623	−3.730	0.000	−247.842	−1.710	0.087	−279.090	−1.470	0.140			
Baseyrdum86	−224.498	−2.040	0.041	−455.896	−3.140	0.002						
Baseyrdum84	124.531	1.110	0.268									
Inverse Mills ratio	0.022	0.230	0.819	−0.296	−1.140	0.254	−0.169	−0.320	0.752	−0.070	−0.100	0.917
Constant	519.039	4.250	0.000	996.957	5.560	0.000	1630.599	6.000	0.000	1805.730	4.050	0.000
R^2	0.1035			0.1672			0.1982			0.2421		

Table D.3 Lag linear difference model for Hennepin county highways

	Two-year lag $n = 12\,962$			Four-year lag $n = 9\,716$			Six-year lag $n = 6\,470$			Eight-year lag $n = 3\,222$		
Variable	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$	Coefficient	t	$P > \|t\|$
Delvkt	− 0.330	− 38.250	0.000	− 0.354	− 30.680	0.000	− 0.367	− 24.880	0.000	− 0.145	− 8.010	0.000
Delparalvkt	− 0.006	− 1.960	0.050	− 0.023	− 6.740	0.000	− 0.023	− 5.290	0.000	− 0.012	− 2.320	0.021
Dellanes	585.068	2.820	0.005	999.932	3.320	0.001	715.083	1.590	0.111	1959.681	3.200	0.001
Delparalanes	325.361	2.000	0.045	641.810	3.480	0.000	1229.518	5.480	0.000	1002.940	4.410	0.000
Delmcdpop	0.041	7.220	0.000	0.077	14.910	0.000	0.075	12.390	0.000	0.037	5.820	0.000
Deladjpop	0.005	1.350	0.178	− 0.004	− 1.340	0.180	− 0.001	− 0.340	0.732	0.000	0.020	0.982
Basevkt	− 0.025	− 7.360	0.000	− 0.021	− 4.240	0.000	− 0.035	− 4.810	0.000	− 0.015	− 1.680	0.094
Baseparalvkt	0.002	3.520	0.000	0.007	6.370	0.000	0.010	5.830	0.000	0.009	3.380	0.001
Baselanes	− 63.749	− 2.140	0.032	− 157.335	− 3.940	0.000	− 311.187	− 5.440	0.000	− 615.939	− 8.630	0.000
Baseparalanes	− 85.106	− 3.650	0.000	− 147.551	− 4.450	0.000	− 206.292	− 4.340	0.000	− 153.224	− 2.570	0.010
Basemcdpop	0.000	1.300	0.195	0.001	2.390	0.017	0.002	5.690	0.000	0.004	9.800	0.000
Baseadjpop	− 0.001	− 7.360	0.000	− 0.002	− 10.080	0.000	− 0.003	− 9.940	0.000	− 0.004	− 9.980	0.000
Baseyrdum96	3.447	0.050	0.957									
Baseyrdum94	2.716	0.040	0.966	− 277.487	− 3.660	0.000						
Baseyrdum92	− 243.483	− 3.700	0.000	− 737.233	− 8.750	0.000	− 393.272	− 3.920	0.000			
Baseyrdum90	− 476.174	− 6.130	0.000	− 913.941	− 10.920	0.000	− 304.026	− 3.170	0.002	− 176.182	− 2.010	0.045
Baseyrdum88	− 113.921	− 1.710	0.087	− 282.218	− 3.720	0.000	− 16.985	− 0.190	0.848			
Baseyrdum86	− 72.249	− 1.140	0.255	− 224.704	− 2.970	0.003						
Baseyrdum84	276.217	4.180	0.000									
Inverse Mills ratio	0.003	0.370	0.711	− 1.895	− 0.220	0.822	66.490	2.580	0.010	− 20.778	− 0.440	0.662
Constant	1012.319	14.120	0.000	1976.988	19.910	0.000	2400.046	17.220	0.000	2681.014	15.540	0.000
R^2	0.1370			0.1478			0.1697			0.1237		

Table D.4 Cost model to estimate cost of link expansion

Description of the Variable	Coefficient	t	$P > \|t\|$
Lane-km of construction	0.28	4.1	0.00**
Dummy for new constructions	0.40	2.2	0.03**
Dummy for interstate roads	1.43	4.9	0.00**
Dummy for state roads	0.52	2.3	0.03**
Log of year—1979	0.76	7.1	0.00**
Log of duration of construction	0.36	2.7	0.01**
Distance from nearest downtown	− 0.03	− 2.1	0.04**
Constant	5.45	18.5	0.00**
Number of observations 102			
Adjusted R^2 0.65			

**indicates significant at the 5 per cent level.
Source of model: Levinson and Karamalaputi (2002).

Table D.5 Erroneous links in the network, which were removed

	Number of links
Links in built network initially	8281
Remove CSAH links in all other counties except Hennepin County	2592
Remove links with lanes > 0 but $aadt = 0$	630
Remove links with lanes $= 0$ and $aadt > 0$	70
New total	4989