Volume 6 Number 1, 2003
ISSN 1094-8848

BUREAU OF TRANSPORTATION STATISTICS UNITED STATES DEPARTMENT OF TRANSPORTATION

JOURNAL OF TRANSPORTATION AND STATISTICS

JOURNAL OF TRANSPORTATION AND STATISTICS

JOHN V. WELLS Editor-in-Chief
PEG YOUNG Associate Editor
MARSHA FENN Managing Editor
JENNIFER BRADY Assistant to the Editor-in-Chief
DORINDA EDMONDSON Desktop Publisher
ALPHA GLASS Editorial Assistant
LORISA SMITH Desktop Publisher
MARTHA COURTNEY Editor
DARCY HERMAN Editor

EDITORIAL BOARD

KENNETH BUTTON George Mason University
TIMOTHY COBURN Abilene Christian University
STEPHEN FIENBERG Carnegie Mellon University
GENEVIEVE GIULIANO University of Southern California
JOSE GOMEZ-IBANEZ Harvard University
DAVID GREENE Oak Ridge National Laboratory
KINGSLEY HAYNES George Mason University
DAVID HENSHER University of Sydney
PATRICIA HU Oak Ridge National Laboratory
RICHARD JOHN Volpe National Transportation Systems Center, USDOT
T.R. LAKSHMANAN Boston University
TIMOTHY LOMAX Texas Transportation Institute
PETER NIJKAMP Free University
KEITH ORD Georgetown University
ALAN PISARSKI Consultant
ROBERT RAESIDE Napier University
JEROME SACKS National Institute of Statistical Sciences
TERRY SHELTON U.S. Department of Transportation
KUMARES SINHA Purdue University
ROBERT SKINNER Transportation Research Board
CLIFFORD SPIEGELMAN Texas A&M University
DARREN TIMOTHY U.S. Department of Transportation
PIYUSHIMITA (VONU) THAKURIAH University of Illinois at Chicago
MARTIN WACHS University of California at Berkeley
C. MICHAEL WALTON The University of Texas at Austin

The views presented in the articles in this journal are those of the authors and not necessarily the views of the Bureau of Transportation Statistics. All material contained in this journal is in the public domain and may be used and reprinted without special permission; citation as to source is required.

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 6 Number 1, 2003
ISSN 1094-8848

BUREAU OF TRANSPORTATION STATISTICS UNITED STATES DEPARTMENT OF TRANSPORTATION

U.S. DEPARTMENT OF TRANSPORTATION

NORMAN Y. MINETA
Secretary

BUREAU OF TRANSPORTATION STATISTICS

RICK KOWALEWSKI
Deputy Director

WILLIAM J. CHANG
Associate Director for
Information Systems

JOHN V. WELLS
Chief Economist

JEREMY WU
Acting Chief Statistician

Bureau of Transportation Statistics

Our mission: To lead in developing transportation data and information of high quality and to advance their effective use in both public and private transportation decisionmaking.

Our vision for the future: Data and information of high quality supporting every significant transportation policy decision, thus advancing the quality of life and the economic well-being of all Americans.

The *Journal of Transportation and Statistics* releases three numbered issues a year and is published by the

Bureau of Transportation Statistics
U.S. Department of Transportation
Room 7412
400 7th Street SW
Washington, DC 20590
USA
journal@bts.gov

Subscription information

To receive a complimentary subscription or a BTS Product Catalog:

mail Product Orders
Bureau of Transportation Statistics
U.S. Department of Transportation
Room 7412
400 7th Street SW
Washington, DC 20590
USA

phone 202.366.DATA
fax 202.366.3197
internet www.bts.gov

Information Service

email answers@bts.gov
phone 800.853.1351

Cover and text design Susan JZ Hoffmeyer
Cover photo © 2003/Getty Images

The Secretary of Transportation has determined that the publication of this periodical is necessary in the transaction of the public business required by law of this Department.

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 6 Number 1
2003

Contents

Truth in Transportation Planning

DONALD C. SHOUP
University of California, Los Angeles

ABSTRACT

Transportation engineers and urban planners often report uncertain estimates as precise numbers, and unwarranted trust in the accuracy of these precise numbers can lead to bad transportation and land-use policies. This paper presents data on parking and trip generation rates to illustrate the misuse of precise numbers to report statistically insignificant estimates. Beyond the problem of statistical insignificance, parking and trip generation rates typically report the parking demand and vehicle trips observed at suburban sites with ample free parking and no public transit. When decisionmakers use these parking and trip generation rates for city planning, they create a city where everyone drives to their destinations and parks free when they get there.

Beware of certainty where none exists.
DANIEL PATRICK MOYNIHAN

INTRODUCTION

How far is it from San Diego to San Francisco? An estimate of 632.125 miles is precise but not accurate. An estimate of somewhere between 400 and 500 miles is less precise but more accurate, because

KEYWORDS: parking, regression analysis, urban planning.

the correct answer is 460 miles.[1] Nevertheless, if you did not know the distance from San Diego to San Francisco, whom would you believe: someone who confidently says 632.125 miles or someone who tentatively says somewhere between 400 and 500 miles? You would probably believe the one who says 632.125 miles, because precision creates the impression of accuracy.

Although reporting estimates with extreme precision suggests confidence in their accuracy, transportation engineers and urban planners often use precise numbers to report uncertain estimates. As examples of this practice, I will use two manuals published by the Institute of Transportation Engineers (ITE): *Parking Generation* (ITE 1987a) and *Trip Generation* (ITE 1987b, 1991, 1997). These manuals have enormous practical consequences for transportation and land use. Urban planners rely on parking generation rates to establish off-street parking requirements, and transportation planners rely on trip generation rates to predict the traffic impacts of development proposals. Yet a close look at the parking and trip generation data shows that placing unwarranted trust in these precise but uncertain estimates of travel behavior leads to bad transportation and land-use policies.

TRIP GENERATION

Trip Generation reports the number of vehicle trips as a function of land use. Transportation engineers survey the number of vehicle trips to and from a variety of locations, and for each land use the ITE reports a trip generation rate that relates the number of vehicle trips to a characteristic of the land use, such as the floor area or number of employees at a site. The sixth (and most recent) edition of *Trip Generation* (ITE 1997, vol. 3, pp. ix and 1) describes the data used to estimate trip generation rates as follows:

> This document is based on more than 3,750 trip generation studies submitted to the Institute by public agencies, developers, consulting firms, and associations. . . . Data were primarily collected at suburban localities with little or no transit service, nearby pedestrian amenities, or travel demand management programs.

ITE says nothing about the price of parking at the study sites, but since parking is free for 99% of vehicle trips in the United States, most of the study sites probably offer free parking.[2] *Trip Generation* uses these 3,750 studies to estimate 1,515 trip generation rates, one for each type of land use. Half the 1,515 reported trip generation rates are based on five or fewer studies, and 23% are based on a single study.[3] The trip generation rates thus typically measure the number of vehicle trips observed at a few suburban sites with free parking but little or no public transit service, pedestrian amenities, or travel demand management (TDM) programs. Urban planners who rely on these trip generation rates as guides to design the transportation system are therefore planning an automobile-dependent city.

Figure 1 shows a typical page from the fourth edition of *Trip Generation* (ITE 1987b).[4] It reports the number of vehicle trips to and from fast food restaurants on a weekday. Each point in the figure represents one of the eight studies and shows the number of vehicle trips per day and the floor area at a restaurant. Dividing the number of vehicle trips by the floor area at that restaurant gives the trip generation rate at that restaurant. A glance at the figure suggests that vehicle trips are unrelated to floor area in this sample. The extremely low R^2 of 0.069 for the fitted curve (regression) equation confirms this

[1] The airline distance between San Diego and San Francisco is calculated from the latitudes and longitudes of the two cities. See "How far is it?" at http://www.indo.com/distance/. "Accurate" implies fidelity to fact and freedom from error, while "precise" implies exactness.

[2] The U.S. Department of Transportation's 1990 Nationwide Personal Transportation Survey (NPTS) asked respondents, "Did you pay for parking during any part of this trip?" for all automobile trips made on the previous day. Of the responses to this question, 99% were "no." The NPTS asked the "did you pay for parking" question for all vehicle trips *except* trips that ended at the respondents' homes, thus free parking at home does not explain this high percentage.

[3] This refers to the sixth edition of *Trip Generation* (ITE 1997). The ITE *Trip Generation Handbook* (ITE 2001, p. 10) notes that the warning "Caution—Use Carefully—Small Sample Size" is placed on each trip generation report if the sample includes five or fewer sites. At most sites, vehicle trips are observed during the course of only one day.

[4] The fourth edition (ITE 1987b) is shown because this is the date of the most recent edition of *Parking Generation,* to which *Trip Generation* will be compared. Vehicle trips were surveyed at McDonald's, Dunkin Donuts, Burger Chef, and similar fast food restaurants.

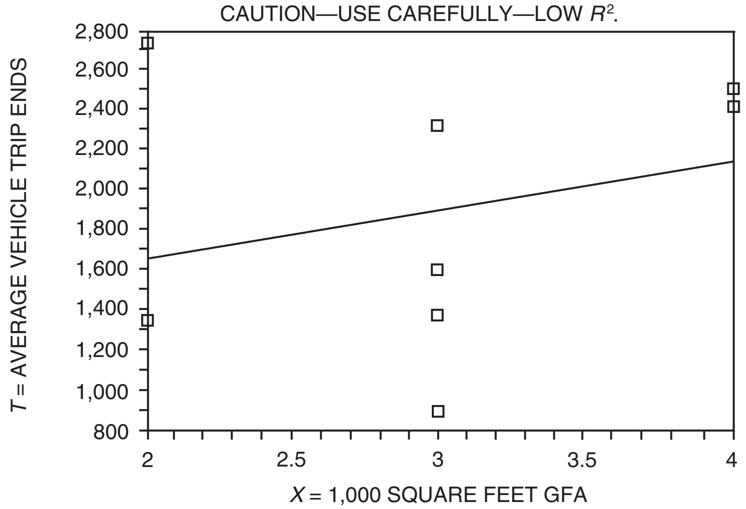
FIGURE 1 Fast Food Restaurant with Drive-Through Window (Land Use 834)
Average Vehicle Trip Ends vs: 1,000 Square Feet Gross Floor Area
On a: weekday

TRIP GENERATION RATES

Average Weekday Vehicle Trip Ends per 1,000 Square Feet Gross Floor Area				
Average Trip Rate	*Range of Rates*	*Standard Deviation*	*Number of Studies*	*Average 1,000 Square Feet GFA*
632.125	284.00–1,359.00	*	8	3.0

DATA PLOT AND EQUATION

CAUTION—USE CAREFULLY—LOW R^2.



☐ ACTUAL DATA POINTS —— FITTED CURVE

Fitted Curve Equation: $T = 242.75\ (X) + 1{,}168.0$
$R^2 = 0.069$

DIRECTIONAL DISTRIBUTION: Not available.

Institute of Transportation Engineers, *Trip Generation*, 4th edition (Washington, DC: 1987), p. 1,199.

impression.[5] Nevertheless, ITE reports the sample's *average* trip generation rate—which urban planners normally interpret as the significant relationship between floor area and vehicle trips—as precisely 632.125 trips per day per 1,000 square feet of floor area.[6] The trip generation rate looks accurate because it is so precise, but the precision is misleading. Few transportation or land-use decisions would be changed if the ITE reported the trip generation rate as 632 rather than 632.125 trips per 1,000 square feet, so the three-decimal-point precision serves no purpose other than to give the impression of accuracy.

The equation at the bottom of figure 1 suggests that a fast food restaurant generates 1,168 trips (the intercept) plus 242.75 trips per 1,000 square feet of floor area (the coefficient), but the 95% confidence interval around the floor area coefficient ranges from –650 to +1,141 trips per 1,000 square feet.[7] Since this confidence interval contains zero, the data

[5] "The coefficient of determination [R^2] is defined as the percent of the variance in the number of trips associated with the variance in the size of the independent variable" (ITE 1997, vol. 3, p. 19). An R^2 of zero shows complete lack of correlation between the two variables, and one would expect some correlation in a sample by chance. The significance test for the regression equation shows there is a 53% chance of getting an R^2 of 0.069 or higher even if there were no relationship between floor area and vehicle trips.

[6] ITE (1987b, p. 9) divides the sum of all vehicle trips by the sum of all floor areas to calculate the weighted average trip generation rate.

do not show that vehicle trips are related to floor area. Reporting the average trip generation rate implies that larger restaurants generate more vehicle trips, but the figure shows that the smallest restaurant generated the most trips, and a mid-sized restaurant generated the fewest. The data plot contains the warning "Caution—Use Carefully—Low R^2," which is good advice, but how can we carefully use a trip generation rate derived from data that show no relationship between vehicle trips and floor area? Despite its precision, the *average* trip generation rate (623.125 vehicle trips per day per 1,000 square feet) is far too uncertain to use for transportation planning.

PARKING GENERATION

Parking generation rates, which report peak parking occupancy as a function of land use, suffer from similar uncertainty. ITE's second, and most recent, edition of *Parking Generation* (ITE 1987a, p. vii–xv[8]) describes the data used to estimate parking generation rates.

> A vast majority of the data . . . is derived from suburban developments with little or no significant transit ridership. . . . The ideal site for obtaining reliable parking generation data would . . . contain ample, convenient parking facilities for the exclusive use of the traffic generated by the site. . . . The objective of the survey is to count the number of vehicles parked at the time of peak parking demand.

Half the 101 parking generation rates are based on 4 or fewer studies, and 22% are based on 1 study. The parking generation rates thus typically measure the peak parking demand observed at a few suburban sites with ample free parking but little or no transit ridership. Urban planners who use these parking generation rates to set minimum parking requirements therefore shape a city where everyone will drive wherever they go and park free when they get there.

Figure 2 shows the page for fast food restaurants from the most recent edition of *Parking Generation* (ITE 1987a). Each point in the plot represents one study (based on the observations at one site on one day). For example, if parking occupancy was observed at one restaurant for five days, this was counted as five studies.[9] Dividing the peak parking occupancy observed in a study by the floor area at the restaurant gives the parking generation rate for the study. The parking generation rates in the 18 studies range between 3.55 and 15.92 spaces per 1,000 square feet of leasable floor area. The largest restaurant in the sample generated one of the lowest peak parking occupancies, while a mid-sized restaurant generated the highest. The R^2 of 0.038 for the equation at the bottom of the figure confirms the visual impression that parking demand is unrelated to floor area in this sample. Nevertheless, ITE reports the *average* parking generation rate for a fast food restaurant as *precisely* 9.95 parking spaces per 1,000 square feet of floor area.[10]

Again, the precision is misleading. The fitted curve equation at the bottom of figure 2 suggests that a fast food restaurant generates a peak parking demand of 20 spaces plus 1.95 spaces per 1,000 square feet of floor area, but the 95% confidence interval around the floor area coefficient ranges from –3 to +7 spaces per 1,000 square feet. Since this confidence interval contains zero, the data do

---

[7] The confidence interval around the coefficient of floor area was calculated by re-estimating the regression equation from the eight observations in the data plot.

[8] ITE expects to publish a new edition of *Parking Generation* in 2003.

[9] It appears that eight restaurants were observed for one day, one restaurant was observed for two days, and two restaurants were observed for four days. We are not told the hour(s), the weekday, or the month when parking occupancy was observed. The 18 studies of parking occupancy at fast food restaurants are an unusually large sample. In contrast, consider the report on Technical Colleges (Land Use 541). Parking occupancy was observed for one hour on one day at one site, and on this basis the parking generation rate for a technical college is reported as 0.82 parking spaces per student (ITE 1987a, p. 88). Parking occupancy was observed for only one or two hours for many of the studies in *Parking Generation*. Because only the *peak* occupancy at a site is needed to calculate a parking generation rate, the observer's main concern is to report the peak number of cars parked during the hour(s) of expected peak demand.

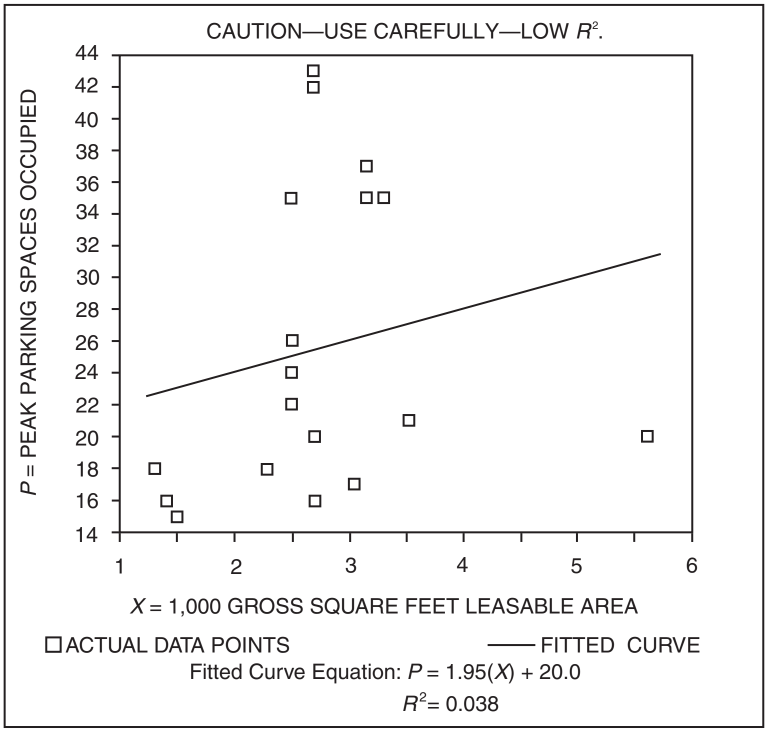
[10] The significance test for the regression equation shows there is a 42% chance of getting an R^2 of 0.038 or higher even if there were no relationship between floor area and parking occupancy. ITE (1987a, p. viii) divides the sum of all parking generation rates by the number of studies to calculate the unweighted average parking generation rate.

FIGURE 2 Fast Food Restaurant with Drive-In Window (Land Use 836)
Peak Parking Spaces Occupied vs:
1,000 Gross Square Feet Leasable Area
On a: weekday

PARKING GENERATION RATES

Average Rate	*Range of Rates*	*Standard Deviation*	*Number of Studies*	*Average 1,000 GSF Leasable Area*
9.95	3.55–15.92	3.41	18	3

DATA PLOT AND EQUATION



Institute of Transportation Engineers, *Parking Generation*, 2nd edition (Washington, DC: 1987), p. 146.

not show that parking demand is related to floor area.[11] The *average* parking generation rate of 9.95 spaces per 1,000 square feet is due mainly to the intercept, which is independent of floor area.[12] Predicting a parking demand of 26 spaces for every restaurant in this sample—regardless of restaurant size—produces about the same average error as predicting a parking demand of 9.95 spaces per 1,000 square feet.[13]

We cannot say much about how floor area affects either vehicle trips or parking demand, because the 95% confidence interval around the floor area coefficient includes zero in both cases.[14] This is not to say that vehicle trips and parking demand are unrelated to a restaurant's size, because common sense suggests some correlation. Nevertheless, factors other

[11] The confidence interval around the coefficient of floor area was calculated by re-estimating the regression equation from the 18 observations in the data plot.

[12] Because the intercept is 20 spaces and the average floor area is 3,000 square feet, the average parking generation rate would be 6.7 spaces per 1,000 square feet even if the coefficient of floor area were 0.

[13] The average peak parking occupancy for the 8 studies was 26 spaces.

[14] Statistical insignificance does not imply that floor area has no effect on parking demand or vehicle trips; rather, it means that floor area does not reliably predict either variable.

than the floor area explain most of the variation in vehicle trips and peak parking occupancy at these restaurants. Size does not matter much in these two samples of parking and trip generation, and it is misleading to publish precise *average* parking and trip generation rates based on floor area.

Parking generation rates are hardly scientific, but the authority inherent in ITE publications often means that planners automatically regard ITE rates as scientifically valid and do not examine them further. ITE offers a precise number without raising difficult public policy questions, although it does warn, "Users of this report should exercise extreme caution when utilizing data that is based on a small number of studies" (ITE 1987a, p. vii). Nevertheless, many planners recommend parking generation rates as minimum parking requirements. For example, the median parking requirement for fast food restaurants in the United States is 10 spaces per 1,000 square feet—almost identical to ITE's reported parking generation rate.[15]

Peak Parking Occupancy vs. Parking Demand

A big difference exists between "parking occupancy" and "parking demand." Transportation engineers define the former as the number of parked cars. Economists define the latter as the functional relationship between the price of parking and the number of parked cars, and they define the actual number of parked cars at any time as the quantity of parking demanded at a specific price. Economists call the peak parking occupancy observed at a site that offers free parking the quantity of parking demanded at a zero price at the time of peak parking demand. These differing definitions show the confusion that can result when ITE's parking generation rates are loosely referred to as parking demand.

STATISTICAL SIGNIFICANCE

The combination of extreme precision and statistical insignificance for the parking and trip generation rates for a fast food restaurant raises an important question: how many of the parking and trip generation rates for other land uses are statistically significant? The fourth edition of *Trip Generation* (ITE 1987b) does not state a policy on statistical significance, but it does show the plots and equations for most land uses with more than two data points. Nevertheless, it fails to show the plots and equations for some land uses with more than 10 data points. For example, consider the report of trip generation at recreational land uses. ITE presents 14 studies of trip generation at recreational land uses but says "No Plot or Equation Available—Insufficient Data." The trip generation rates in the 14 studies range from a high of 296 to a low of 0.066 trips per acre on a weekday: a ratio of 4,500 to 1. Given this wide range, reporting the *average* trip generation rate as *precisely* 3.635 trips per acre is clearly misleading.[16]

ITE first stated a policy regarding statistical significance in the fifth edition of *Trip Generation* (ITE 1991, p. I-8):

> Best fit curves are shown in this report only when each of the following three conditions are met:
>
> - The R^2 is greater than or equal to 0.25.
> - The sample size is greater than or equal to 4.
> - The number of trips increases as the size of the independent variable increases.[17]

The third criterion is notably unscientific. For example, suppose the R^2 is greater than 0.25 and the sample size is greater than four, but vehicle trips *decrease* as floor area increases (i.e., the first two criteria are met but the third is not). In this case, ITE would report the *average* trip generation rate (which implies that vehicle trips *increase* as floor area increases), but not the regression equation that would cast doubt on this rate. The stated policy, therefore, omits evidence that would contradict the presumed relationship.

Figure 3 from the fifth edition of *Trip Generation* (ITE 1991) shows how these three criteria affect the report of trip generation at a fast food restaurant. It shows the same eight data points from the fourth edition, but it omits the regression equation, the R^2, and the warning "Caution—Use Carefully—Low R^2." The omitted R^2 remains 0.069 because the data are

[15] The Planning Advisory Service (1991) surveyed the parking requirements in 127 cities. The median of 10 spaces per 1,000 square feet applies to cities that base their requirements for fast food restaurants on gross floor area.

[16] In the fourth edition of *Trip Generation,* Land Use 400 (Recreational) includes bowling alleys, zoos, sea worlds, lakes, pools, and regional parks (ITE 1987b, p. 537).

[17] ITE gives no explanation for showing the regression equation and the R^2 only when all three criteria are met.

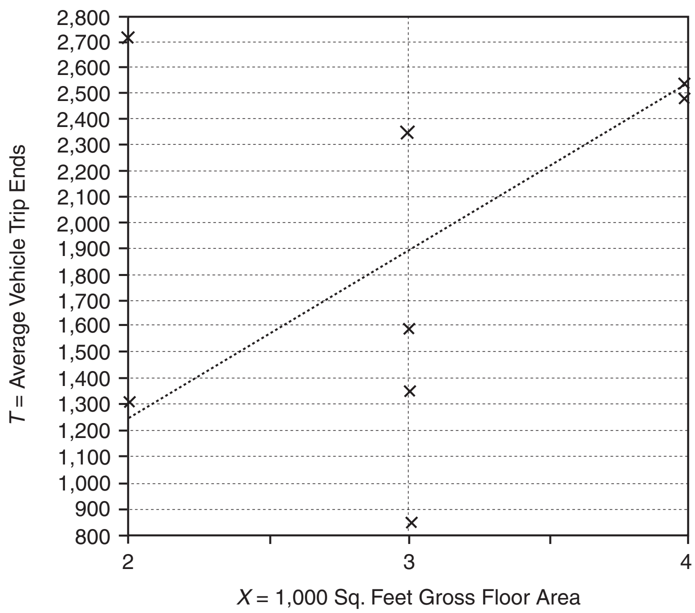
FIGURE 3 Fast Food Restaurant with Drive-Through Window (Land Use 834)
Average Vehicle Trip Ends vs:
1,000 Square Feet Gross Floor Area
On a: weekday

Number of Studies: 8
Average 1,000 Sq. Feet GFA: 3
Directional Distribution: 50% entering, 50% exiting

Trip Generation per 1,000 Sq. Feet Gross Floor Area

Average Rate	Range of Rates	Standard Deviation
632.12	284.00–1,359.50	266.29

Data Plot and Equation



Institute of Transportation Engineers, *Trip Generation*, 5th edition (Washington, DC: 1991), p. 1,308.

unchanged from the fourth edition, but the fifth edition is more cautious about needless precision; it truncates the average trip generation rate from 632.125 to 632.12 trips per 1,000 square feet.[18]

ITE revised its reporting policy in the sixth (most recent) edition of *Trip Generation* (ITE 1997, p. 19). Regression equations are shown only if the R^2 is greater than or equal to 0.5, while the other two criteria remain the same (the sample size is four or more, and vehicle trips increase as the independent variable increases). Figure 4 shows the sixth edition's report of trip generation at a fast food restaurant. The number of studies increased to 21, and the average trip generation rate fell to 496.12 trips per 1,000 square feet. The R^2 is below 0.5, but we are not told what it is. Since the fifth edition's rate was 632.12 trips per 1,000 square feet, anyone comparing the two editions might conclude that vehicle trips at fast food restaurants declined 22% between 1991 and 1997. But since both the previous rate (632.12) and the new one (496.12) were derived from data

[18] Figure 3 (from the fifth edition) also differs from figure 1 (from the fourth edition) in two other respects. First, the directional distribution of vehicle trips was "not available" in 1987, but for the same data became "50% entering, 50% exiting" in 1991. Second, the standard deviation was not reported in 1987 but was reported as 266.29 in 1991.

FIGURE 4 Fast Food Restaurant with Drive-Through Window (Land Use 834)

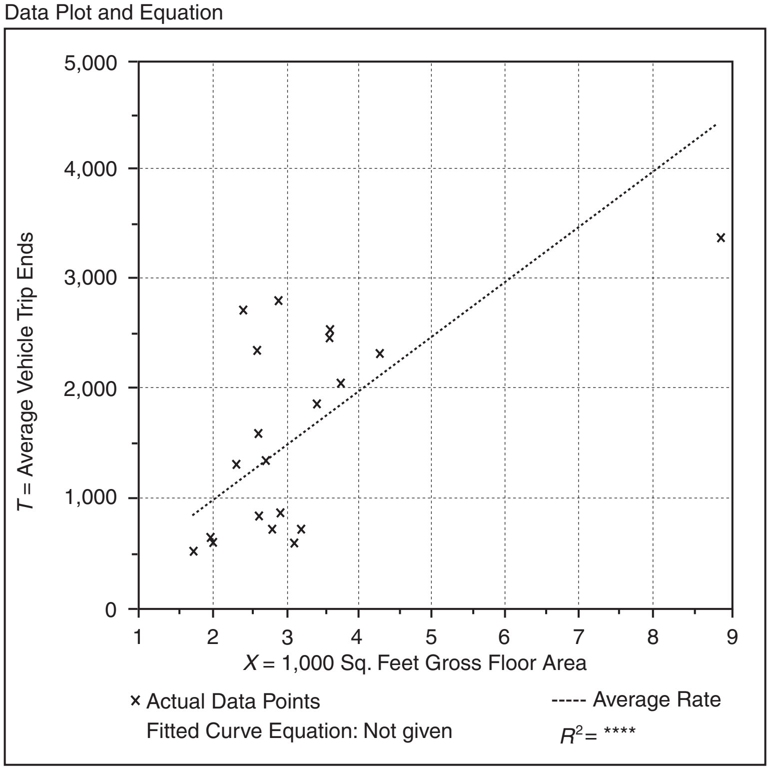
Average Vehicle Trip Ends vs: 1,000 Square Feet Gross Floor Area
On a: weekday

Number of Studies: 21
Average 1,000 Sq. Feet GFA: 3
Directional Distribution: 50% entering, 50% exiting

Trip Generation per 1,000 Sq. Feet Gross Floor Area

Average Rate	Range of Rates	Standard Deviation
496.12	195.98–1,132.92	242.52

Data Plot and Equation



Institute of Transportation Engineers, *Trip Generation*, 6th edition (Washington, DC: 1997), p. 1,401.

that show almost no relationship between floor area and vehicle trips, this decline seems unlikely.[19]

The 1997 edition shows regression equations for only 34% of the trip generation rates, which means that 66% of the 1,515 trip generation rates fail to meet at least one of the three criteria. This statistical insignificance is not surprising given that circumstances vary enormously among different sites for the same land use (e.g., a fast food restaurant). Floor area is only one among many factors that influence vehicle trips at a site, and we should not expect floor area or any other single variable to accurately predict the number of vehicle trips at any site or land use.[20]

[19] If the 8 studies from the fourth (ITE 1987b) and fifth (ITE 1991) editions are included among the 21 studies reported in the sixth (ITE 1997) edition, the average trip generation rate for the 13 new studies must be well below 496.12 in order to reduce the average rate for the 21 studies to 496.12. All of the 8 study sites in the fourth and fifth editions were exactly 2,000, 3,000, or 4,000 square feet, but none of the 21 study sites in the sixth edition matched these sizes.

[20] Trip generation rates are a stripped-down version of the gravity model for travel forecasting. The gravity model predicts aggregate traffic between origin and destination zones as a function of zone sizes and generalized travel cost, while trip generation rates predict traffic to and from one site as a function of floor area (or another variable) at that site, without reference to cost.

Although 66% of the trip generation rates fail to meet ITE's significance criteria, ITE nevertheless publishes a precise trip generation rate for every land use. For example, a report of trip generation at truck terminals (figure 5) presents two sites, with the larger site generating fewer vehicle trips. Nevertheless, ITE reports the *average* trip generation rate as precisely 81.90 vehicle trips per acre on a weekday and plots a line that suggests larger sites generate more vehicle trips.

Reporting statistically insignificant estimates with misleading precision creates serious problems, because many people rely on the ITE manuals to predict how urban development will affect parking and traffic. When estimating the traffic impacts of development, for example, developers and cities often debate over whether a precise trip generation rate is correct. Some cities even base zoning categories on trip generation rates. Consider this zoning ordinance in Beverly Hills, California:

> The intensity of use shall not exceed either sixteen (16) vehicle trips per hour, or 200 vehicle trips per day for each 1,000 gross square feet of floor area for uses as specified in the most recent edition of the Institute of Traffic Engineers' publication entitled *Trip Generation*.[21]

The precise but uncertain ITE data thus govern which land uses the city will allow.

Parking and trip generation rates are difficult to challenge once they are incorporated into municipal codes. Planning is an inherently uncertain activity, but the legal system of land-use regulation makes it difficult to acknowledge uncertainty in planning regulations. Calling attention to the flaws in the reporting of the parking and trip generation rates would expose land-use decisions to countless lawsuits from developers, neighborhood groups, and property rights advocates, all of whom could rightly question the legitimacy of the reasoning used to establish off-street parking requirements and to argue for either more or less parking. This desire for the appearance of certainty explains why transportation engineers, urban planners, developers, and elected officials rely on precise point estimates—rather than ranges—to report the highly uncertain parking and trip generation rates.

[21] Section 10-3.162(5) of the Beverly Hills Municipal Code. (ITE changed its name from the Institute of Traffic Engineers to the Institute of Transportation Engineers in 1976.)

PLANNING FOR FREE PARKING

ITE's parking and trip generation rates can create serious problems when they are used for urban planning. Most ITE samples are too small to draw statistically significant conclusions, and ITE's method of collecting data skews observations toward sites with high parking and trip generation rates. Larger samples might solve the problem of statistical insignificance, but a basic problem with parking and trip generation rates would remain: they measure the peak parking demand and the number of vehicle trips *at suburban sites with ample free parking*. This situation is troubling, because ITE rates greatly influence the outcome of transportation and land-use planning, ultimately contributing to decisions that result in more traffic, lower density, and more urban sprawl.

To explain how ITE's parking and trip generation rates influence transportation and land-use planning, consider what appears in practice to be the six-step process of planning for free parking in the United States.

- **Step 1.** Transportation engineers survey the peak parking demand at a few suburban sites with ample free parking but no transit service, and ITE publishes the results in *Parking Generation* with misleading precision.
- **Step 2.** Urban planners consult *Parking Generation* to set minimum parking requirements. The maximum observed parking demand thus becomes the minimum required parking supply.
- **Step 3.** Developers provide all the parking that planners require, and the ample supply of parking drives the price of most parking to zero, which increases vehicle travel.
- **Step 4.** Transportation engineers survey vehicle trips to and from suburban sites with ample free parking but no transit service, and ITE publishes the results in *Trip Generation* with misleading precision.

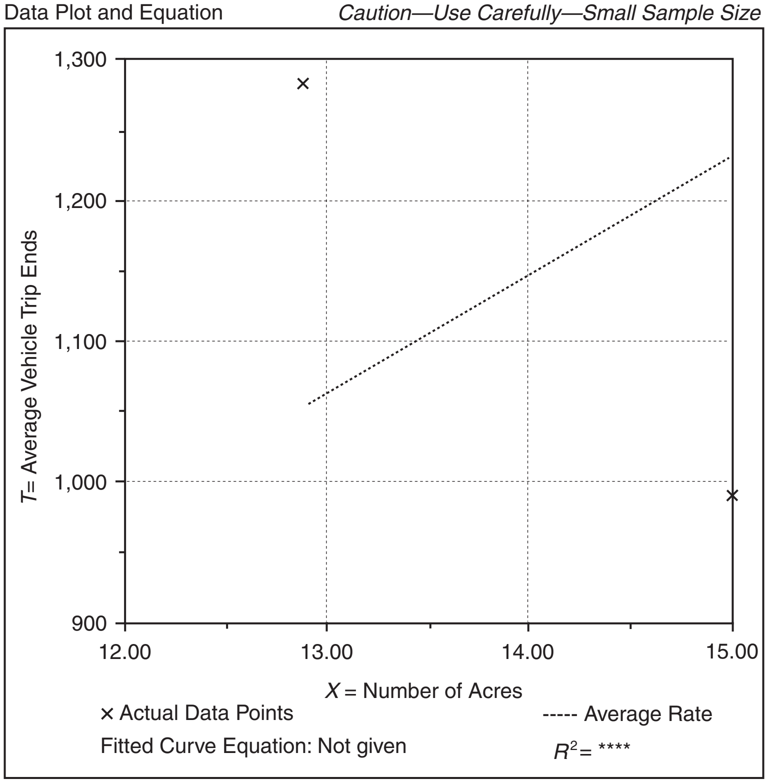
FIGURE 5 Truck Terminal
(Land Use 030)
Average Vehicle Trip Ends vs: Acres
On a: weekday

Number of Studies: 2
Average Number of Acres: 14
Directional Distribution: 50% entering, 50% exiting

Trip Generation per Acre

Average Rate	Range of Rates	Standard Deviation
81.90	66.27–100.08	*

Data Plot and Equation *Caution—Use Carefully—Small Sample Size*



Institute of Transportation Engineers, *Trip Generation*, 6th edition (Washington, DC: 1997), p. 66.

- **Step 5.** Transportation planners consult *Trip Generation* as a guide to design the transportation system with adequate capacity to bring cars to the free parking.[22]
- **Step 6.** Urban planners limit density so that development with ample free parking will not generate more vehicle trips than nearby roads can carry. This lower density spreads activities farther apart, further increasing both vehicle travel and parking demand.

We come full circle when transportation engineers again survey peak parking demand at suburban sites that offer free parking but no transit service and find that more parking spaces are "needed." Misusing precise numbers to report uncertain data gives a veneer of rigor to this elaborate but unscientific practice, and the circular logic explains why planning for transportation and land use has contributed to increased traffic and sprawl.

[22] Transportation planners often use the Urban Transportation Modeling System (UTMS) to predict modal flows on links between zones in a network, and the first of the four major steps in the UTMS model is "trip generation." The four-step UTMS model is thus used to carry out step 5 of the six-step process of planning for free parking. Meyer and Miller (2001) explain the UTMS model.

The ITE manuals do not cause this circular process, which started long before ITE began collecting data on parking and trip generation. In 1965, economist Edgar M. Hoover described the circular planning process in words that still apply today:

> In practice, the separation of highway-building programs from parking programs (they are in different and quite independent bureaucracies or authorities) introduces a still further pernicious element. We know the story of the man who took another piece of bread in order to finish his butter, then another piece of butter in order to finish his bread, and so on till he burst. Similarly, every provision of new freeways into a congested area heightens the observed demand and the public pressure for more parking facilities; every additional downtown parking garage heightens the demand for more new freeways to bring people to it; and so on back and forth indefinitely. Each of the two independent public authorities involved can argue persuasively that it is merely trying to keep up with an undeniably strong and growing demand. (Hoover 1965, pp. 188–189)

The main change that has occurred since 1965 is that engineers and planners now have precise parking and trip generation data to quantify the "undeniably strong and growing demand" for parking and highways. The interaction between transportation engineers and urban planners in gathering and interpreting these data helps to explain why planning for parking in the United States is essentially planning for free parking. Urban planners set parking requirements without taking into account the price of parking, the cost of parking spaces, the local context, or the wider consequences for transportation, land use, the economy, and the environment.

ITE warns users to be careful when the R^2 is low (although it removed this warning from the plots of trip generation rates in the two most recent editions of *Trip Generation*). ITE also advises users to modify trip generation rates in response to special circumstances.

> At specific sites, the user may want to modify the trip generation rates presented in this document to reflect the presence of public transportation service, ridesharing or other TDM measures, enhanced pedestrian and bicycle trip-making opportunities, or other special characteristics of the site or surrounding area. (ITE 1997, vol. 3, p. 1)

Nevertheless, ITE does not suggest how a user might modify the rates in response to any special characteristics of a site or its surrounding area, and the price of parking is prominently not on the list of special characteristics that might affect trip generation.

Data users should always ask themselves whether the data are appropriate for the intended purpose. Only users can misuse data, but ITE invites misuse when it reports statistically insignificant estimates as precise numbers. This spurious precision has helped to establish ITE parking requirements and trip generation rates as unquestionably authoritative in the planning profession.

CONCLUSION: LESS PRECISION AND MORE TRUTH

Estimates of parking and trip generation respond to a real demand for essential information. Citizens want to know how development will affect parking demand and traffic congestion in their neighborhood. Developers want to know how many parking spaces they should provide for employees and customers. Planners want to regulate development to prevent problems with parking and traffic. Politicians want to avoid complaints from unhappy parkers. These are all valid concerns, but reporting parking and trip generation rates with needless precision creates false confidence in the data. To unsophisticated users, these precise rates appear to carry the rigor of scientific constants.

When planners set parking requirements and design the transportation system, they treat parking and trip generation like established laws and ITE estimates like scientific observations. But parking and trip generation are poorly understood phenomena, and they both depend on the price of parking, an element not addressed by ITE in the two reports discussed. Demand is a function of price, not a fixed number, and this does not cease to be true merely because transportation engineers and urban planners ignore it. Most cities are planned on the unstated assumption that parking should be free—no matter how high the cost or how small the benefit.

American motor vehicles consume one-eighth of the world's total oil production, and ubiquitous free parking contributes to our automobile dependency.[23] What can be done to improve this situation? Here are four recommendations:

1. ITE should state in the report for each parking and trip generation rate that this rate refers only to suburban sites with ample free parking but no public transit, pedestrian amenities, or TDM programs.
2. ITE should show the regression equation and the R^2 for each parking and trip generation report and state whether the coefficient of floor area (or other independent variable) in the equation is significantly different from zero.
3. ITE should report the parking and trip generation rates as ranges, not as precise point estimates.
4. Urban planners should recognize that even if the ITE data were accurate, using them to set parking requirements would dictate an automobile-dependent urban form with free parking everywhere.

Both transportation engineers and urban planners should ponder this warning from Lewis Mumford: "The right to have access to every building in the city by private motorcar, in an age when everyone possesses such a vehicle, is actually the right to destroy the city." (Mumford 1981)

Parking and trip generation rates illustrate a familiar problem with statistics used in transportation planning, and placing unwarranted trust in the accuracy of these precise but uncertain data leads to bad transportation and land-use policies. Being roughly right is better than being precisely wrong. We need less precision—and more truth—in transportation planning.

[23] Transportation accounted for 66.4% of U.S. oil consumption in 1996, and highway transportation accounted for 78.3% of U.S. oil consumption for transportation. Therefore, highway transportation accounted for 52.0% of U.S. oil consumption (66.4% x 78.3%). The United States also consumed 25.7% of the world's oil production in 1996. Thus, U.S. highway transportation consumed 13.4% (slightly more than one-eighth) of the world's total oil production (52.0% x 25.7%). Highway transportation refers to travel by cars, trucks, motorcycles, and buses. See Davis (2000, tables 1.3, 2.10, and 2.7) for the data on energy consumption for transportation in the United States.

ACKNOWLEDGMENTS

I am grateful to the University of California Transportation Center for financial support. Douglas Kolozsvari provided superb research assistance. I am also grateful for excellent advice from Jeffrey Brown, Leland Burns, Daniel Chatman, Randall Crane, Melanie Curry, T.H. Culhane, Simon Fraser, Daniel Hess, Mimi Holt, Hiro Iseki, Joshua Kirshner, Robin Liggett, Bravishwar Mallavarapu, Jeremy Nelson, Don Pickrell, Thomas Rice, Michael Sabel, Lisa Schweitzer, Charles Sciammas, Patricia Shoup, Charanjeet Singh, Alexander Smith, Manual Soto, Brian Taylor, Florian Urban, Melvin Webber, Richard Willson, and two anonymous reviewers. Earlier versions of this paper were presented at the 2001 Annual Meeting of the Transportation Research Board in Washington, DC, the 2001 World Parking Symposium in St. Andrews, Scotland, and the 2002 Annual Meeting of the Western Regional Science Association in Monterey, California.

REFERENCES

Davis, S. 2000. *Transportation Energy Data Book: Edition 20,* ORNL-6959. Oak Ridge, TN: Oak Ridge National Laboratory.

Hoover, E. 1965. Motor Metropolis: Some Observations on Urban Transportation in America. *Journal of Industrial Economics* 13(3):17–192.

Institute of Transportation Engineers (ITE). 1987a. *Parking Generation,* 2nd edition. Washington, DC.

______. 1987b. *Trip Generation,* 4th edition. Washington, DC.

______.1991. *Trip Generation,* 5th edition. Washington, DC.

______. 1997. *Trip Generation,* 6th edition. Washington, DC.

______. 2001. *Trip Generation Handbook: An ITE Recommended Practice.* Washington, DC.

Meyer, M. and E. Miller. 2001. *Urban Transportation Planning,* 2nd edition. New York, NY: McGraw Hill.

Mumford, L. 1981. *The Highway and the City.* Westport, CT: Greenwood Press.

Planning Advisory Service. 1991. *Off-Street Parking Requirements: A National Review of Standards,* Planning Advisory Service Report Number 432. Chicago, IL: American Planning Association.

Author address: Department of Urban Planning, University of California, Los Angeles, CA 90095-1656. Email: shoup@ucla.edu.

Discussion

CARL H. BUTTKE
Consulting Transportation Engineer
EUGENE D. ARNOLD, JR.
Virginia Transportation Research Council

Mr. Shoup's article, "Truth in Transportation Planning," tends to view the Institute of Transportation Engineers' (ITE) *Trip Generation,* 6th edition and *Parking Generation,* 2nd edition reports as manuals to be followed step by step rather than as informational reports to be used to help guide transportation planning and development decisions. The intended purpose of the documents is stated in the reports. For example, page ix of the *Trip Generation User's Guide* contains the following:

> ITE Informational Reports are prepared for informational purposes only and do not include ITE recommendations on the best course of action or the preferred application of the data.

It is important to note that *Trip Generation* does not represent a quick fix for transportation problems or a shortcut to planning procedures; rather, it serves as a foundation on which the professional engineer can build his or her own knowledge and experience and apply this knowledge to any given transportation-related situation. The intended users who estimate vehicle trip generation or parking demand are transportation professionals trained in mathematics, statistics, traffic engineering, and planning fundamentals and who possess engineering judgment.

ITE's reports provide a compilation of available data collected from numerous sources. In the sixth edition of *Trip Generation,* data are combined from more than 3,750 individual trip generation studies. This information is by no means all inclusive; however, it represents the best information available at the time of publication. ITE's *Trip Generation* report is updated regularly to include supplemental information as it becomes available.

Some of Shoup's commentary, examples, and assertions are directed to the fourth and fifth editions of *Trip Generation.* While many of these references are used to make a point, some of the discussion is not relevant as the data, assumptions, and reporting techniques are updated and improved from edition to edition. Further, we expect that transportation professionals will use the latest edition to obtain the most recent knowledge and data available.

In his article, Shoup correctly points out that reporting statistics with "extreme precision may suggest confidence in their accuracy." He also rightfully acknowledges that generation rates such as 623.12 could be reported as 623 and not affect the accuracy of the calculation. However, there are also many instances in *Trip Generation* where rates presented with two decimal places are appropriate at that level of precision (e.g., as a rate of 0.57 pm peak-hour trips per occupied room of a business hotel, or 7.27 weekday trips per occupied room). When developing the first edition of *Trip Generation,* the Trip Generation Committee wrestled with this issue of decimal placement and decided to be consistent in reporting all rates with two decimal places.

Shoup also notes that, from a statistical standpoint, some of the independent variables used are simply not related to trips (e.g., he points to an extremely low R^2 value). This may be a valid point; however, in many instances the particular independent variable is chosen because it is the only information available in the early stages of development when these analyses are often undertaken. To that end, the *Trip Generation User's Guide* (vol. 3, p. 21) notes that: "Selecting an appropriate method for estimating trips requires use of engineering judgment and a thorough understanding of the three methodologies…."

In reference to Shoup's remarks regarding figure 4, the only independent variables available for this land use for measuring weekday trips were gross square feet and seats. We acknowledge that it is the customers and employees who make the trips, but these data were not available when the measurements were made and are rarely known when estimating proposed traffic impacts. Page 14 of the *User's Guide* addresses the variation in the statistics:

> These variations may be due to the small sample size, the individual marketing of the site, economic conditions of the business market, the geographic location of sites studied, or the unique character of the specific site. Accordingly, judgment must be exercised in the use of the statistics in this report.

Shoup continues with a dialogue regarding ITE's advice to users to modify trip rates in response to special situations, such as the presence of public transportation service, ridesharing, and enhanced pedestrian facilities. We feel it is appropriate for ITE to point out potential cautions with the use of data without necessarily providing a solution if it cannot be supported by current research.

In Shoup's conclusion, he recommends that *Trip Generation* data be reported as ranges and not as precise point estimates. Current editions of *Trip Generation* and *Parking Generation* do provide ranges, average rates, and a data plot. This diversity in data presentation provides the user with a more comprehensive look at the data. Additionally, page 18 of the *User's Guide* provides a detailed description of a sample data page.

To produce resources supporting *Trip Generation,* ITE relies on the voluntary submittal of data from the transportation community. Calls for the submission of data have been ongoing over the years, with the intent to provide additional data to assist transportation professionals. ITE's openness about the availability of data can be seen on page one of the *User's Guide:*

> In some cases, limited data were available; thus, the statistics presented may not be truly representative of the trip generation characteristics of a particular land use.

Such cautionary statements run throughout both the *Trip Generation* and the *Parking Generation* informational reports.

Trip Generation, 7th edition, and *Parking Generation,* 3rd edition, are slated for release in 2003. Data collected from various sources, as well as comments, including those provided by Shoup, are reviewed and taken into consideration during the revision process. ITE's intent is to provide a helpful resource that will guide transportation professionals in their decisionmaking.

Editor-in-Chief's Note: The discussants were chosen by the Institute of Transportation Engineers.

Author Addresses: Corresponding author—Eugene Arnold, Senior Research Scientist, Virginia Transportation Research Council, 530 Edgemont Road, Charlottesville, VA 22903. Email: Gene.Arnold@VirginiaDOT.org.

Carl Buttke, Consulting Transportation Engineer, PO Box 2740, Hailey, ID 83333. Email: buttke@tripgeneration.com.

DONALD C. SHOUP

Carl Buttke and Eugene Arnold argue that nothing is wrong with the Institute of Transportation Engineers' (ITE) *Trip Generation* and *Parking Generation*. In part, their confidence may derive from their assumption that "the intended users . . . are transportation professionals trained in mathematics, statistics, traffic engineering, and planning fundamentals and who possess engineering judgment." But the actual users are a much broader and more diverse group. The ITE itself says, "*Trip Generation* is an educational tool for planners, transportation professionals, zoning boards, and others who are interested in estimating the number of vehicle trips generated by a proposed development" (ITE 1997, vol. 3, p. ix). Many of these people are *not* trained in mathematics, statistics, and traffic engineering. Zoning boards are rarely trained in anything—they are elected or appointed to their positions, perform their duties as volunteers, and rely heavily on references such as *Parking Generation* and *Trip Generation.* They will not realize that the reported rates are often statistically insignificant and refer only to suburban sites with ample free parking and no public transit.

I would like to address three issues that Buttke and Arnold raise, and make a recommendation.

SIGNIFICANT DIGITS

ITE's convention of rounding every parking and trip generation rate to two digits after the decimal point blurs the distinction between precision and accuracy. Buttke and Arnold agree that the two-digits-after-the-decimal-point convention leads to inappropriate precision in some instances, but then say,

> There are also many instances in *Trip Generation* where rates presented with two decimal places are appropriate at that level of precision (e.g., as a rate of 0.57 pm peak-hour trips per occupied room of a business hotel, or 7.27 weekday trips per occupied room).

But *Trip Generation's* estimate of 7.27 weekday trips per occupied room of a business hotel is based on only one observation.[1] It illustrates perfectly the statistical insignificance and inappropriate precision of many parking and trip generation rates.

An estimate always has some associated uncertainty. The number of significant digits used to express an estimate should reflect this uncertainty. The least significant digit in a number is the one farthest to the right, and the accuracy of any number is usually assumed to be ±1 of the least significant digit, unless stated otherwise. In a typical engineering context, one would assume that an estimate expressed with five significant digits had been measured more accurately than an estimate expressed with only two significant digits. Because the number of significant digits used to express an estimate should be related to the uncertainty surrounding the estimate, the ITE's automatic two-digits-after-the-decimal-point convention is inappropriate and unscientific.

Buttke and Arnold note that the Trip Generation Committee wrestled with the issue of decimal placement in preparing the first edition of *Trip Generation* in 1976, and decided to be consistent in reporting all rates with two digits after the decimal point.[2] Accuracy is more important than digits-after-the-decimal-point consistency, however, and one should not use more (or less) precision than is warranted simply for the sake of uniformity. Precision refers to the number of significant digits, not to the number of digits after the decimal point.

---

[1] ITE (1997, vol. 1, p. 543). The estimate of 0.57 pm peak-hour trips per occupied room is based on only four studies.

[2] The first (1976), second (1979), and third (1983) editions of *Trip Generation* report some rates with no digits after the decimal point and other rates with one or two digits after the decimal point. The fourth (1987) edition reports all rates with three digits after the decimal point. The fifth (1991) and sixth (1997) editions report all rates with two digits after the decimal point.

MISUSE

Statistically sophisticated users understand the extreme uncertainty of trip generation rates and can ignore the false precision. But many users are *not* statistically sophisticated. To them, ITE's trip generation rates are *the* relationship between transportation and land use. Some zoning codes explicitly specify ITE's trip generation rates as the basis for making land-use decisions and as the basis for assessing traffic impact fees, regardless of the sample size or statistical significance of the rates.

In Signal Hill, California, for example, the traffic impact fee is $66 per daily vehicle trip generated by a development project. The number of trips is calculated by multiplying the size of the project times its trip generation rate "as set forth in the most recent edition of the Traffic [*sic*] Generation manual of the Institute of Transportation Engineers."[3] The sixth edition's trip generation rate for a fast food restaurant is 496.12 trips per 1,000 square feet, so Signal Hill's traffic impact fee is $32.74 per square foot of restaurant space. The uncertain trip generation rates thus determine cities' tax rates.

FREE PARKING

Buttke and Arnold conclude that "ITE's intent is to provide a helpful resource that will guide transportation professionals in their decisionmaking." Spurious precision is not a real impediment for this purpose, although it is misleading.[4] The real problem with *Parking Generation* and *Trip Generation* is that they measure the peak parking demand and the number of vehicle trips *at suburban sites with ample free parking and no public transit.* Using these precise but poorly understood parking and trip generation rates as a guide to planning leads to bad transportation and land-use decisions. *Parking Generation* and *Trip Generation* are helpful resources in designing cities where everyone will drive everywhere they go and park free when they get there.

RECOMMENDATION

What can be done to make the ITE reports more reliable? The British counterpart to *Trip Generation* suggests some possible improvements. The "Trip Rate Information Computer System" (TRICS) gives full information about the characteristics of every surveyed site and its surroundings.[5] Users can thus estimate a trip generation rate based on sites comparable to the one under consideration. In addition to counts of vehicles, TRICS also includes counts of all the people (pedestrians, cyclists, public transport users, and car occupants) who arrive at and depart from a site. By including more than vehicle trips, TRICS takes a broader view of transportation. When all modes are included, the *person* trip rates are often much higher than the *vehicle* trip rates.

With its narrow focus on counting cars at suburban sites with free parking, *Trip Generation* presents a precise but uncertain, skewed, and incomplete measure of the relationship between transportation and land use in the United States. Fortunately, the ITE's Parking and Trip Generation Committees seek to improve each successive edition of *Parking Generation* and *Trip Generation.* In future editions, they should settle for less precision, and strive for more accuracy.

---

[3] Section 21.48.020 of the Signal Hill Municipal Code. The code is available online at http://www.ci.signal-hill.ca.us/homepage.php.

[4] Even if everyone who refers to *Parking Generation* and *Trip Generation* were an engineer or statistician, that does not excuse unjustified precision. Journalists do not casually break grammar and spelling rules just because intelligent readers might be able to figure out what they mean anyway. The burden of clarity and accuracy falls on the writer—it cannot be shifted to the reader, no matter who one supposes the reader to be.

[5] The TRICS database is available online at http://www.trics.org/.

Modeling Trip Duration for Mobile Source Emissions Forecasting

HARIKESH S. NAIR
University of Chicago

CHANDRA R. BHAT
The University of Texas at Austin

ABSTRACT

Metropolitan area trip duration distributions are an important input for estimating area-wide running loss emissions, operating mode fractions, and vehicle-miles traveled accumulated on local roads in the region. This paper discusses the formulation and implementation of a methodology for modeling trip durations. The approach develops a log-linear regression model of trip duration as a function of trip purpose, time of day of the trip start, and other land-use and socio-demographic characteristics of the zone of trip start, using vehicle trip data from household travel surveys and supplementary zonal demographic and land-use data. A distinguishing characteristic of the methodology is the straightforward manner in which model parameters estimated from vehicle trip data can be applied to obtain zonal-level trip duration distributions. The modeling framework is applied to develop trip duration distributions for the Dallas-Fort Worth area of Texas.

BACKGROUND AND SIGNIFICANCE OF WORK

The Intermodal Surface Transportation Efficiency Act of 1991 and the Clean Air Act Amendments of 1990 require localities not meeting the National

KEYWORDS: local-road VMT, MOBILE emissions factor model, mobile source emissions, operating mode fractions, running loss emissions, trip duration.

Ambient Air Quality Standards set by the Environmental Protection Agency (EPA) to demonstrate area-wide conformity with mobile source emissions budgets established in their respective state implementation plans. For such conformity analyses, the Mobile Source Emissions Factor Model (MOBILE) is generally used.[1]

The emissions factor models use traffic-related data as inputs. Regional vehicle trip duration distribution data are one type and are important for several reasons. First, the trip duration distribution provides information for developing trip duration activity parameters used by the MOBILE emissions factor model to estimate running loss emissions. Running loss emissions are evaporative emissions that escape from a vehicle while the engine is operating (from spots where the vehicle's evaporative/purge system has become inoperative). Due to greater heating of the engine fuel and evaporative system on longer trips, running loss emissions continue to increase as a function of trip duration until the emissions reach a plateau at a trip duration of about 50 to 60 minutes (Glover and Brzezinski 2001). Second, the trip duration distribution enables the estimation of operating mode fractions, which are needed by MOBILE5 to calculate emissions rates. Third, the trip duration distribution can be used to predict the vehicle-miles traveled (VMT) on local roads in the region.

Trip duration is likely to depend on various factors such as the trip purpose, the time of day the trip began, and other land-use and socio-demographic characteristics of the zone of the trip start. In this paper, we formulate and implement a methodology for modeling trip durations using vehicle trip data from household travel surveys and supplementary zonal demographic and land-use data. The implementation is demonstrated in the context of mobile source emissions analysis for the Dallas-Fort Worth area in Texas.

The next section reviews earlier studies relevant to our subject and the motivation for our research. We then discuss the development of the model estimation and application framework. The following section focuses on data sources and data assembly procedures, after which we present the empirical results. The next section discusses issues related to integrating the trip duration model with travel demand models, and we conclude the paper with an evaluation of the model.

LITERATURE REVIEW AND MOTIVATION FOR THE STUDY

Running Loss Emissions

The methodology to estimate running loss emissions differs from MOBILE5 to MOBILE6. In MOBILE5, running loss emissions are modeled as a direct function of the input temperature, fuel volatility, average speed, and trip duration. The procedure for calculating the running loss emissions entails partitioning the vehicle trip duration into 6 time duration bins (less than 10 minutes, 11 to 20 minutes, 21 to 30 minutes, 31 to 40 minutes, 41 to 50 minutes, and 51 minutes and longer) and obtaining the proportion of VMT accumulated by trips that fall into each time duration bin (these proportions are referred to as the trip duration activity parameters).

Within MOBILE5, the running loss emissions value of an average vehicle trip is calculated as the sum of the product of the emissions factors associated with each time duration bin (embedded within MOBILE5) and the corresponding trip duration activity parameter. The product of these average running loss emissions and the number of trips per day represent the running loss emissions level. The user can then accept default daily running loss emissions values available within MOBILE5 (developed using default trip-time distributions representing national average conditions) or develop region-specific estimates by using a local set of trip duration activity parameters. The MOBILE5 manual suggests using area-specific trip duration activity parameters to more accurately estimate running loss emissions.

MOBILE6 advances the state-of-the-art and practice by providing activity parameters for each of 14 time periods in a day and by distinguishing between weekdays and weekends. The default MOBILE6 hourly activity estimates are based on an EPA survey of 168 vehicles and are invariant across

[1] Some metropolitan planning organizations (MPOs) still use the MOBILE5b version of this model, although MOBILE6 is now available. EPA has given MPOs two years to transition to MOBILE6 from MOBILE5.

geographic regions and trip purpose categories. Thus, as in MOBILE5, EPA recommends the use of locally estimated trip duration activity parameters whenever possible. However, to our knowledge, there has been no earlier attempt in the literature to develop such locally estimated trip duration parameters.

In summary, using trip duration activity parameters developed from local data for estimating running loss emissions constitutes an important improvement over using the default values embedded in the MOBILE emissions factor model. In this paper, we present a methodology to develop zone-specific trip duration activity parameters that vary by time-of-day and trip purpose using a trip duration model estimated from local data.

Operating Mode Fractions

Operating mode fractions are an important input to MOBILE5 for estimating mobile source emissions. There are two dimensions associated with operating mode fractions; one is the *start mode* of vehicle trips (cold versus hot) and the second is the *running mode* of vehicle trips (transient versus stabilized). In an earlier paper, we focused on the start mode of trips (Nair et al. 2002). Trip duration modeling, the focus of this paper, affects the latter dimension of the operating mode, that is, the running mode of trips. To the extent that running mode fractions can be more accurately estimated using a trip duration model calculated from local data, such a model can contribute toward improved mobile source emissions forecasting.

EPA defines the *transient* mode of operation as all vehicle operations before 505 seconds after the start of a trip and the *stabilized* mode as all operations after 505 seconds of a trip. EPA recommends the following default values for running mode fractions: transient (47.9%) and stabilized (52.1%). Most metropolitan planning organizations (MPOs) accept these default running mode fractions. However, these default values were developed over 20 years ago and earlier research (USEPA 1993) suggests that these values may no longer adequately represent overall vehicle emissions control performance. In addition, the default fractions do not vary by trip purpose, time of day, or regional land-use and socio-demographic characteristics.

Few studies have attempted to develop locally estimated running mode fractions of trips. Brodtmen and Fuce (1984) used field data obtained by direct on-road measurement of engine conditions to develop running mode fractions in New Jersey. Ellis et al. (1978) analyzed origin-destination data from travel surveys in Alabama to develop aggregate measures of running mode fractions. Frank et al. (2000) developed transient and stabilized mode fractions based on vehicle trip times, using the Puget Sound Panel Survey. Chatterjee et al. (1996) and Venigalla et al. (1999) used a network-based approach for modeling running modes, in which they traced the elapsed time of vehicles from trip origins during the assignment of trips to the highway network. Allen and Davies (1993) have similarly used the ASSIGN module of MINUTP, a commercially available planning model, to determine trips operating in the transient mode for the southern New Jersey area.

A limitation of the above studies is that they compute a single set of running mode fractions for an entire state (or for aggregate regions within a state) and for various times of day and trip purposes. In this paper, we estimate a trip duration model using local data from a metropolitan region and present a methodology to use this estimated model to develop running mode fractions that vary by zone within the region, time of day, and trip purpose. In addition, our methodology allows for the estimation of running mode fractions for travel on local roads.

VMT on Local Roads

Local roads are usually not included in the travel demand model networks used by most MPOs; thus, the travel speeds and volumes required to calculate the VMT on local links are unavailable. Many MPOs simply calculate the VMT as a percentage (typically about 10%) of the VMT on all other roads and use it in developing their emissions inventories. This method is rather ad hoc in nature and can result in VMT estimates quite different from the actual values. A few MPOs calculate the VMT on local roads as a product of the total intrazonal trips for each zone (obtained from the origin-destination trip-interchange matrices at the end of trip distribution) and an average intrazonal trip length parameter. The average intrazonal trip length parameter is typically calculated as a function of the total area of the zone. While this method is a substantial improvement over using a percentage of VMT on nonlocal roads, it is still limited by the restrictive nature of variation of the intrazonal trip length parameter. In particular, in this method, the intrazonal trip length (and, therefore, local VMT) does not vary by trip purpose, time of day, and zonal spatial attributes (other than zonal area).

Our study develops the intrazonal trip length as a function of time of day, purpose, and zonal attributes. We accomplish this by estimating a trip duration model and then multiplying the predicted intrazonal trip duration by an estimate of average speed on local links (it is more straightforward to develop a direct model of intrazonal trip length, but most household surveys collect data only on trip duration and not trip length).

MODEL FRAMEWORK

Our modeling approach uses vehicle trip data from household travel surveys and zonal demographic and land-use data from supplementary sources. The approach develops the distribution of the duration of trips using a log-linear regression model. The use of a log-linear form for trip duration guarantees the nonnegativity of trip time in application of the model.

The application step of the model predicts the trip duration distributions for each traffic analysis zone in a metropolitan region and for each combination of time of day and trip purpose. An important characteristic of the proposed method is the ease with which the estimated models using vehicle trip data can be immediately applied to obtain zonal-level trip-time distributions.

Model Estimation

Let q be the index for vehicle trip, t be the index for time of day, i be the index for activity purpose (i may be defined as a function of the activity purposes at both ends of the trip q), and z be the index for zone. Define ω_{qti} to be a dummy variable taking the value 1 if vehicle trip q occurs in time period t with trip purpose i, and 0 otherwise; define δ_{qz} as another dummy variable taking the value 1 if vehicle trip q is produced from zone z, and 0 otherwise. Define I_q to be a variable that takes the value 1 if vehicle trip q is intrazonal, and 0 otherwise. Let x_z be a vector of zonal attributes.

We assume the trip duration to be log-normally distributed in the population of trips, and develop a linear regression model for the duration as a function of trip purpose, time of day, and land-use and socio-demographic characteristics of the zone of trip production. Let d_q be the duration of vehicle trip q. Then, we write the log-linear regression equation for the trip duration as

$$\begin{aligned} ln(d_q) &= \eta + \sum_{t,\,i} \alpha_{ti}\omega_{qti} + \lambda\left(\sum_z \delta_{qz}x_z\right) \\ &\quad + I_q\left(\chi + \sum_{t,\,i} \zeta_{ti}\omega_{qti} + \rho\left(\sum_z \delta_{qz}x_z\right)\right) + \varepsilon_q, \\ &\quad \varepsilon_q \sim N\left[0,\ \sigma^2\right] \end{aligned} \qquad (1)$$

where

η is the generic constant to be estimated;

the α_{ti}'s (t = 1, 2,…,T; i = 1,2,…,I) are scalars capturing the effects of time of day and activity purpose on trip duration (these scalars are to be estimated);

λ is a vector of parameters representing the effects of the characteristics of the zone of trip production (the vector λ is also to be estimated).

χ, ζ_{ti}, and ρ are similar to η, α_{ti}, and λ, respectively, but are introduced as specific to intrazonal trips (note that I_q takes the value 1 if vehicle trip q is an intrazonal trip, and 0 otherwise). ε_q is a normally distributed random error term introduced to complete the statistical specification.

In equation (1) above, we have not allowed interactions between zonal attributes x_z and time of day/trip purpose combinations ω_{qti}; however, this is purely for notational convenience and for ease in presentation of the model application step. Interactions between x_z and ω_{qti} can be included within the model structure without any additional conceptual or estimation complexity. Similarly, the notation structure implies full interactions of time and trip purpose (as defined by the dummy variable ω_{qti}), though more restrictive structures such as single dimensional effects without interaction can be imposed by appropriately constraining the α_{ti} and

ζ_{ti} scalars across the different time/trip purpose combinations.

The reader will note that the inclusion of the intrazonal dummy variable, and interactions of this variable with exogenous variables, allows us to accommodate separate trip duration distributions for intrazonal vehicle trips and interzonal vehicle trips. The model from equation (1) can be estimated using any commercially available software with a linear regression module.

Model Application

Trip Duration Activity Parameters for Running Loss Emissions

The trip duration distribution for any zone in the study area by time period and trip purpose can be predicted in a straightforward manner after estimation of equation (1). The (log) trip duration distribution of interzonal vehicle trips in time t for trip purpose i produced from zone z may be written as

$$ln\left(d^{a}_{tiz}\right) \sim N\left[\eta + \alpha_{ti} + \lambda x_{z},\ \sigma^{2}\right] = N\left[\Delta^{a}_{tiz},\ \sigma^{2}\right]. \quad (2)$$

The superscript a in the above equation is used to denote interzonal trips. The mean Δ^{a}_{tiz} and variance σ^2 of this distribution can be estimated from the parameter estimates obtained in the estimation stage. The corresponding distribution of intrazonal vehicle trips in time t for trip purpose i in zone z may be written as

$$ln\left(d^{l}_{tiz}\right) \sim N\left[\eta + \alpha_{ti} + \lambda x_{z} + \chi + \zeta_{ti} + \rho x_{z},\ \sigma^{2}\right] = N\left[\Delta^{l}_{tiz},\ \sigma^{2}\right] \quad (3)$$

where the superscript l is used to denote intrazonal trips.

The objective in our effort is to obtain the fraction of VMT accrued by trips in each of six trip duration bins (as needed by MOBILE; see the Running Loss Emissions section earlier in this paper) for each zone and for each trip purpose and time-of-day combination. Let k be an index for time-bin (k = 1,2,...,6), and let k be bounded by the continuous trip duration value of m_{k-1} to the left and by m_k to the right. Let V^k be the average speed of trips in time-bin k and let ϑ_z be the fraction of trips originating in zone z that are intrazonal.[2] Then, the fraction of VMT accrued by interzonal trips in time-bin k, during time of day t, for trip purpose i, produced from zone z ($FVMT^{ka}_{tiz}$) can be obtained as (the derivation of the expression is available from the authors)

$$FVMT^{ka}_{tiz} = \frac{L^{ka}_{tiz} * \Omega^{ka}_{tiz} * V^{k}}{VMT^{a}_{tiz}} \quad (4)$$

where

$$L^{ka}_{tiz} = \Phi\left(\frac{ln\left(m^{k} - \Delta^{a}_{tiz}\right)}{\sigma}\right) - \Phi\left(\frac{ln\left(m^{k-1} - \Delta^{a}_{tiz}\right)}{\sigma}\right) \quad (5)$$

$$\Omega^{ka}_{tiz} = \exp\left[\Delta^{a}_{tiz} + \sigma\frac{\phi\left(\frac{ln\left(m_{k-1}\right) - \Delta^{a}_{tiz}}{\sigma}\right) - \phi\left(\frac{ln\left(m_{k}\right) - \Delta^{a}_{tiz}}{\sigma}\right)}{\Phi\left(\frac{ln\left(m_{k}\right) - \Delta^{a}_{tiz}}{\sigma}\right) - \Phi\left(\frac{ln\left(m_{k-1}\right) - \Delta^{a}_{tiz}}{\sigma}\right)}\right] \quad (6)$$

$$VMT^{a}_{tiz} = \sum_{k} L^{ka}_{tiz} * \Omega^{ka}_{tiz} * V^{k} \quad (7)$$

In the above equation structure, L^{ka}_{tiz} represents the proportion of interzonal trips in time period t, for trip purpose i, produced from zone z that fall in trip duration bin k. Ω^{ka}_{tiz} represents the mean trip duration of interzonal trips in time period t, for trip

[2] V^k may be obtained from local metropolitan area data or using the following national default values obtained from the 1995 Nationwide Personal Transportation Survey data: 18.96 mph (for trips of duration 0–10 minutes), 20.80 mph (for trips of duration 11–20 minutes), 26.40 mph (for trips of duration 21–30 minutes), 29.14 mph (for trips of duration 31–40 minutes), 33.60 mph (for trips of duration 41–50 minutes), and 45.30 mph (for trips of duration greater than 51 minutes). ϑ_z represents the fraction of intrazonal trips originating from zone z and can be obtained from the sample used for estimation. If the sample data do not support evaluation of ϑ_z for all zones, ϑ_z can be determined from the zone-to-zone production-attraction trip interchanges matrices obtained at the end of the trip distribution step in the travel demand modeling process.

purpose i, produced from zone z that fall in trip duration bin k. The product of L_{tiz}^{ka} and Ω_{tiz}^{ka} with V^k represents the VMT accrued by interzonal trips in time period t, for trip purpose i, produced from zone z that fall in trip duration bin k. VMT_{tiz}^{a} represents the total VMT accrued by interzonal trips in time period t, for trip purpose i, produced from zone z, and is obtained by summing the VMT across all trip duration bins.

The fraction of VMT accrued by intrazonal trips in time-bin k, in time t, for trip purpose i, produced from zone z ($FVMT_{tiz}^{kl}$) can be obtained by substituting Δ_{tiz}^{l} instead of Δ_{tiz}^{a} in equations (4) through (7).

Finally, the fraction of VMT accrued by all trips in each time-bin k, for trip purpose i, from zone z, during time t may be written as

$$FVMT_{tiz}^{k} = \vartheta_z * FVMT_{tiz}^{kl} + \left(1-\vartheta_z\right) * FVMT_{tiz}^{ka} \quad (8)$$

Running Mode Fractions for MOBILE5

This section presents the method to obtain the proportion of transient and stabilized trips required as an input to MOBILE5. We begin by discussing the approach for interzonal trips; the approach is identical for intrazonal trips, with appropriate replacements to reflect the mean and variance of intrazonal trips.

Let the assumed speed of vehicles be v mph. Let the mean of the distribution of trips of duration less than 8.42 minutes (505 seconds) occurring in time period t, with trip purpose i, produced from zone z be μ_{tiz}^{1a} and let the corresponding mean of the distribution of trips of duration greater than 8.42 minutes be μ_{tiz}^{2a} (μ_{tiz}^{1a} and μ_{tiz}^{2a} represent the means of the right- and left-truncated normal distributions of trip duration, respectively).

We obtain the analytical expression for μ_{tiz}^{1a} (see Greene 1997) as

$$\mu_{tiz}^{1a} = \exp\left\{ \Delta_{tiz}^{a} - \sigma \frac{\phi\left(\frac{ln(8.42) - \Delta_{tiz}^{a}}{\sigma}\right)}{\Phi\left(\frac{ln(8.42) - \Delta_{tiz}^{a}}{\sigma}\right)} \right\} \quad (9)$$

The transient mode VMT accumulated by trips of duration less than (or equal to) 8.42 minutes, is given by

$\left(\mu_{tiz}^{1a} / 60\right) * (v) *$[Number of trips of duration ≤8.42 min]. Trips of duration greater than 8.42 minutes are in the transient mode for the first 8.42 minutes of their operation. The transient mode VMT accumulated by such trips is given by $(8.42/60) * (v) *$[Number of trips of duration > 8.42 min]. Therefore, the total *transient* mode VMT in time period t, of purpose i, due to trips produced from zone z is given by the following expression

$$VMT_{tiz,\ transient}^{a} = \frac{1}{60}\left[\left(\mu_{tiz}^{1a}\right) * (v) * \Phi\left(\frac{ln(8.42) - \Delta_{tiz}^{a}}{\sigma}\right) + (8.42) * (v) * \left[1 - \Phi\left(\frac{ln(8.42) - \Delta_{tiz}^{a}}{\sigma}\right)\right]\right] * \text{[Total number of trips]} \quad (10)$$

The mean duration of trips greater than 8.42 minutes, μ_{tiz}^{2a}, is given by

$$\mu_{tiz}^{2a} = \exp\left\{ \Delta_{tiz}^{a} + \sigma \frac{\phi\left(\frac{ln(8.42) - \Delta_{tiz}^{a}}{\sigma}\right)}{1 - \Phi\left(\frac{ln(8.42) - \Delta_{tiz}^{a}}{\sigma}\right)} \right\} \quad (11)$$

The VMT in the stabilized mode in time t, for trip purpose i, originating in zone z can be obtained as

$$\left[\frac{\mu_{tiz}^{2a} - (8.42)}{60}\right] * (v) * \text{[Number of trips of duration > 8.42 min].}$$

Therefore, the expression for the VMT accumulated in the *stabilized* mode is

$$VMT_{tiz,\ stabilized}^{a} = \left[\frac{\mu_{tiz}^{2a} - (8.42)}{60}\right] * (v) * \left[1 - \Phi\left(\frac{ln(8.42) - \Delta_{tiz}^{a}}{\sigma}\right)\right] * \text{[Total number of trips]} \quad (12)$$

Finally, the fraction of VMT in transient versus stabilized modes in zone z, during time period t, and trip purpose i can be obtained from equations (10) through (12).

The reader will also note that the running mode fractions for intrazonal trips may be readily obtained after replacing Δ_{tiz}^{a} with Δ_{tiz}^{l}, μ_{tiz}^{1a} with μ_{tiz}^{1l}, and μ_{tiz}^{2a} with μ_{tiz}^{2l}, respectively, in equations (10) through (12).

VMT on Local Roads

As noted in the model estimation section, the intrazonal nature of a trip is captured through the interaction effects of I_q with exogenous determinants of trip duration. The logarithm of the trip duration of intrazonal trips in time t, with trip purpose i, in zone z is normally distributed, as shown in equation (3). It follows from this that the trip duration distribution of intrazonal vehicle trips in time t, with trip purpose i, in zone z is log-normally distributed with a mean θ_{tiz}^{l} and variance λ_{tiz}^{l} given by the following expressions (see Johnson and Kotz 1970):

$$\vartheta_{tiz}^{l} = \exp\left(\Delta_{tiz}^{l} + \sigma^2 / \ 2\right) \tag{13}$$

$$\lambda_{tiz}^{l} = \exp\left(2 * \Delta_{tiz}^{l} + \sigma^2\right)\left[\exp\left(\sigma^2\right) - 1\right] \tag{14}$$

The mean trip length of intrazonal trips is the product of θ_{tiz}^{l} and the average speed on local roads (which many MPOs assume to be around 20 mph; if data on the variation of average speeds with zone, time period, and/or trip purpose are available, the corresponding average speed may be used). The total VMT on local roads can next be estimated as the product of the mean intrazonal trip length and the total intrazonal vehicle trips (obtained from the trip distribution step in the travel demand modeling process). Our methodology accommodates the variation in VMT on local roads with time of day, trip purpose, and zonal socio-demographic and land-use characteristics through the variation of the average intrazonal trip duration with these characteristics.

To summarize this section, we presented the formulation for a model of trip duration as a function of trip purpose, time of day, and zonal/trip attributes. We proposed methods that can be implemented after estimation of the trip duration model to predict 1) running loss emissions, 2) running mode fractions, and 3) local road VMT. The outputs from the application of the model are the above three mobile source-related parameters for each time of day and trip purpose combination and for each zone in a planning area. The model framework can be integrated within a broader travel demand-air quality forecasting procedure in a straightforward fashion, as discussed later in this paper.

DATA PREPARATION

Data Sources

The data used in the empirical analysis were drawn from two sources: the 1996 Activity Survey conducted in the Dallas-Fort Worth (D-FW) area and the zonal land-use and demographics characteristics file for the D-FW area. These data sources were obtained from the North Central Texas Council of Governments (NCTCOG).

Sample Formation

Several data assembly steps were involved in developing the sample. First, we converted the raw composite (travel and nontravel) activity file into a corresponding person trip file. Second, we identified person trips that were pursued using a motorized vehicle owned by the household. Third, we translated the person trip file into a corresponding vehicle trip file, which provided the sequence of trips made by each vehicle in the household. In this process, we extracted and retained information on the time of day of each vehicle trip start, traffic analysis process (TAP) zone of trip production, and the purpose of the activity being pursued at the attraction-end of the trip. Fourth, we aggregated the traffic survey zone (TSZ) level land-use and demographic characteristics to the TAP level and appended this information to each vehicle trip start based on the TAP in which the trip start occurs (there are about 5,000 TSZs in the D-FW planning area). Finally, we conducted several screening and consistency checks on the resulting dataset from the previous steps.[3] As part of this screening process, we eliminated observations that

[3] A flow chart of this screening process is available from the authors.

had missing data on departure times, activity purposes, and/or on the TAP location of the vehicle trip start.

The final sample used for analysis includes 19,455 vehicle trip observations. Of these, 2,940 trips (15.1%) are intrazonal.

EMPIRICAL ANALYSIS

Sample Description

The dependent variable of interest in our analysis is the time duration of trips. The mean trip duration for interzonal trips is about 21 minutes with a standard deviation of 24 minutes, while the mean trip duration for intrazonal trips is about 11 minutes with a standard deviation of 18 minutes (the standard deviation is higher than the mean because of the substantial scatter of trip duration values, particularly in the higher end of the duration spectrum).

Three types of variables were considered to explain trip duration. These were: 1) trip purpose variables indicating the purpose of the trip, 2) time-of-day variables identifying the time of trip start, and 3) zonal and trip attributes. Interactions among these three sets of variables were also considered. In the description below, we briefly highlight some of the characteristics of these sets of variables.

Trip purpose was characterized by two dimensions: whether or not the trip was produced at home (home-based versus nonhome-based trips) and the purpose at the attraction end of the trip (i.e., whether the attraction end activity is work, school, social or recreational, shopping, personal business, or other). Of the 19,455 trips, 14,294 (73.4%) were home-based, and this percentage is independent of the intrazonal or interzonal nature of trips (see table 1). However, the percentage of interzonal trips with work as the attraction end is higher than the percentage of intrazonal trips with work as the attractor end.

The time-of-day variables were associated with one of the following six time periods: morning (midnight–6:30 a.m.), a.m. peak (6:30 a.m. –9:00 a.m.), a.m. off peak (9:00 a.m.–noon), p.m. off peak (noon–4:00 p.m.), p.m. peak (4:00 p.m.–6:30 p.m.), and evening (6:30 p.m.–midnight). The time periods for the a.m. and p.m. peaks were based on the peak period definitions employed by the NCTCOG transportation department in the D-FW area. The times for the offpeak periods were determined by splitting the remaining blocks of time at noon and midnight. The distribution of intrazonal and interzonal trips by time of day is presented in table 2. In general, the distributions by time of day are rather similar across intrazonal and interzonal trips.

We considered several zonal (TAP-level) land-use and demographic characteristics in our analysis. Of these, the following zonal attributes were significant determinants of trip duration: total zonal area, zonal household density, acreage in retail facilities, acreage in office space, number of people in service employment, acreage in institutional facilities (e.g., hospitals and churches), acreage in manufacturing and warehousing facilities, zonal median income, and presence of airports or airport-related infrastructure in the zone. The trip-related attribute included in the model was an indicator variable for whether or not the trip was intrazonal.

TABLE 1 Distribution of Trips by Trip Purpose

Trip purpose	Percentage distribution for: Intrazonal	Interzonal
Production end		
Home	73.4%	73.5%
Nonhome	26.6%	26.5%
Attraction end		
Work	12.0%	22.3%
School	3.2%	2.4%
Social/recreational	10.2%	11.3%
Shopping	6.9%	6.3%
Personal business	40.0%	41.5%
Other	27.8%	16.2%

TABLE 2 Distribution of Trips by Time of Day

Time of day of trip start	Percentage distribution for: Intrazonal	Interzonal
Morning	1.5%	4.0%
a.m. peak	22.4%	21.0%
a.m. offpeak	13.5%	12.8%
p.m. offpeak	28.4%	23.5%
p.m. peak	19.0%	22.6%
Evening	15.1%	16.1%

We obtained the final model specification of trip duration by systematically eliminating statistically insignificant variables and combining those found to have similar and comparable effects in terms of magnitude and significance.

Results of the Trip Duration Model

The empirical results for the log-linear regression model are presented in table 3, which provides the estimated values of η, α_{ti}, λ, χ, and ζ_{ti} (t = 1,2,…,T; i = 1,2,…,I) in equation (1). All the coefficients are statistically significant at a level of 0.02 or lower (except for the school attraction end coefficient, which is significant only at about the 0.15 level). The R^2 value (bottom of the table) is about 0.2, which indicates that the model does better than a naive model that predicts the average (log) duration value for all trips. However, the low R^2 value also suggests room for further specification improvement in future studies.

The trip purpose variables were included with nonhome-based trips as the base category (for home-based vs. nonhome-based trips) and with work as the base attraction end activity. The time-of-day variables were introduced with the evening period as the base category. The morning period is combined with the a.m. peak period because of the very small fraction of trips in the morning period (as can be observed in table 1).

The positive coefficient on the home-based trip variable under the trip purpose category in table 3 indicates that home-based trips tend to be significantly longer than nonhome-based trips. The coefficients on the attraction end variables under the trip purpose category need to be interpreted jointly with the time-of-day and trip purpose interaction effects. The results show that work trips are of the longest duration across all times of the day and purposes, except for school trips pursued during the midday/evening periods and social-recreational trips pursued in the late evening. Shopping trips are the shortest across all times of the day. The coefficient on the time-of-day variables, when considered jointly with the time-purpose interaction effects, indicate: 1) peak period trips are longer than non-peak period trips, and this is particularly the case for work trips; and 2) work and nonwork trips undertaken during the evening period are shorter than trips taken at other times of the day (except for social-recreational trips, which are longer in the evening than earlier times of the day).

Several zonal and other trip attributes have a statistically significant effect on trip duration. We classify these attributes into three categories: zonal size-related variables, zonal nonsize-related variables, and trip-related variables.

Among the size-related variables, a larger total area of a zone, in general, increases the duration of trips produced from that zone. This is particularly the case if the zone has a high acreage in office space. Similarly, trips produced from zones with a high number of people in service employment and with large acreage in manufacturing facilities also have longer durations. These may reflect congestion effects. On the other hand, acreage in retail and institutional facilities has a negative effect on trip duration, possibly due to greater accessibility to shopping and service-related activities in these zones.

The zonal nonsize-related variables indicate shorter trip durations in zones with high household density and with high household income. However, trips produced from zones with an airport have a longer duration. This latter effect may be caused by increased congestion on roadways in zones with airport-related infrastructure or may be due to airports occupying a large area in the zone and thereby reducing the number of activity opportunities in the zone. Of course, other reasons may be equally plausible.

Finally, intrazonal trips are significantly shorter in duration than interzonal trips, especially during the p.m. peak, although the magnitude of this effect is less for shopping and social-recreational trip purposes.

INTEGRATION WITH TRAVEL DEMAND MODELS

Existing travel demand models may be based on an activity approach or a trip approach. Activity-based travel demand models focus on the activities that people pursue, as a function of the locations and attributes of potential destinations, the state of the transportation network, and the personal and household characteristics of individuals (Ettema and

TABLE 3 Empirical Results for Trip Duration Model

Variable	Coefficient	*t*-statistic
Constant	2.504	76.81
Trip purpose		
Production end (nonhome-based trip is base)		
Home-based	0.213	16.16
Attraction end (work purpose is base)		
School	0.041	1.56
Social-recreational	0.125	5.11
Shopping	–0.299	–16.07
Other	–0.215	–13.07
Time-of-day variables ("evening" period is base)		
morning–a.m. peak/p.m. peak	0.445	15.57
a.m. offpeak/p.m. offpeak	0.176	6.75
Time of day and trip purpose interaction effects		
morning–a.m. peak/p.m. peak x nonwork	–0.155	–11.23
a.m. offpeak/p.m. offpeak x social-recreational	–0.261	–8.94
Zonal and trip-related attributes		
Zonal size-related variables		
Zonal area x 10^{-5}	1.243	2.56
Zonal acreage in office space x 10^{-3}	2.363	3.85
Number of people in service employment x 10^{-5}	2.544	9.75
Zonal acreage in manufacturing facilities x 10^{-4}	6.917	5.19
Zonal acreage in retail facilities x 10^{-3}	–2.266	–4.82
Zonal acreage in institutional facilities x 10^{-3}	–1.020	–2.54
Zonal nonsize-related variables		
Zonal household density x 10^{-3}	–1.612	–4.42
Median income of zone x 10^{-6}	–2.576	–6.55
Presence of an airport or airport-related infrastructure x 10^{-2}	5.342	2.56
Trip-related variables		
Intrazonal trip	–0.777	–33.95
Intrazonal p.m. peak trip	–0.189	–5.08
Intrazonal shopping/social-recreational trip	0.158	6.12
Number of observations	19,455	
Regression sums of squares	2,737.09	
Residual sums of squares	11,058.14	
Standard error of estimate	0.754	
R^2	0.198	
Adjusted R^2	0.198	

Timmermans 1997). If such an approach is adopted in travel analysis, the activity stops made by individuals are explicitly modeled as a function of origin and destination activity categories, time of day, and zone of origin. Thus, information on trip purpose, time of trip start, and attributes of the zone of trip origin are readily available for all trips. Integration of the trip duration model developed in this paper within this framework is straightforward.

If a trip-based travel demand modeling framework is used, the trip duration model can be directly applied if the MPO develops zone-to-zone production-attraction interchanges for the disaggregate trip purpose and time of day categories identified in this paper. However, most MPOs use more aggregate classes of trip purpose and time periods (typically home-based work, home-based other, and nonhome-based trip purposes, and peak versus offpeak time periods). In this situation, the trip duration model can be used after post-processing the aggregate production-attraction trip interchanges matrix to reflect the disaggregate classifications employed here. Factors obtained from travel surveys can be applied to achieve this post-classification. Tables 4, 5, and 6 present such factors developed for the D-FW region.

MODEL EVALUATION

In this section we conduct two evaluations of our proposed model. First, we evaluate our assumption of normality of the distribution of (log) trip durations in our regression (equation 1). For this purpose, we utilize normal probability plots and also a formal statistical test of normality. Second, we compare the performance of our proposed model for trip duration activity parameters with the "default" MOBILE model parameters that remain fixed for all zones and for all time of day and trip purpose categories (this is the state of the practice in the D-FW and other metropolitan areas).

Testing the Normality Assumption

We use two methods to evaluate the assumption of normality of log trip durations in our regression model (equation 1). We first develop an "eyeball" evaluation of the distribution of trip durations in

TABLE 4 Split of Home-Based Work Trips by Time of Day

	Time of day of trip start					
Trip purpose	**morning**	**a.m. peak**	**a.m. offpeak**	**p.m. offpeak**	**p.m. peak**	**evening**
Home-based work	9.02%	34.97%	6.33%	13.57%	26.57%	9.54%

TABLE 5 Split of Home-Based Other Trips by Disaggregate Trip Purpose and Time of Day

	Time of day of trip start					
Trip purpose	**morning**	**a.m. peak**	**a.m. offpeak**	**p.m. offpeak**	**p.m. peak**	**evening**
Home-based school	0.66%	1.15%	2.42%	3.69%	4.57%	9.67%
Home-based social-recreational	0.12%	0.71%	3.07%	5.15%	4.93%	4.73%
Home-based shopping	0.29%	1.89%	4.41%	4.84%	3.78%	2.53%
Home-based personal business	0.64%	10.22%	2.20%	6.73%	6.09%	5.69%
Home-based other	0.06%	3.35%	0.94%	2.90%	1.66%	0.90%

TABLE 6 Split of Nonhome-Based Trips by Disaggregate Trip Purpose and Time of Day

	Time of day of trip start					
Trip purpose	**morning**	**a.m. peak**	**a.m. offpeak**	**p.m. offpeak**	**p.m. peak**	**evening**
Nonhome-based work	0.02%	1.10%	3.53%	4.05%	1.05%	0.27%
Nonhome-based school	0.00%	0.12%	0.21%	0.48%	0.41%	0.04%
Nonhome-based social-recreational	0.16%	0.76%	4.36%	10.13%	2.34%	2.56%
Nonhome-based shopping	0.06%	0.41%	2.07%	4.84%	3.60%	2.19%
Nonhome-based personal business	0.14%	1.72%	6.67%	10.11%	5.27%	2.13%
Nonhome-based other	0.41%	8.35%	3.08%	7.63%	6.68%	3.04%

our sample using normal probability plots. Next, we apply a rigorous statistical test for examining the normality assumption. These evaluations are discussed in the next two sections.

The "Eyeball" Evaluation

The eyeball evaluation typically entails two probability plots. The first, the normal Q-Q plot (i.e., the normal quantile-quantile plot), is the plot of the ordered data values against the associated quantiles of the normal distribution. For data from a normal distribution, the points of the plot should lie close to a straight line. The procedure to produce a Q-Q plot involves the following two steps:

1. Sort the n observed data points in ascending order so that
   $x_1 \leq x_2 \leq . \leq x_n$,
   and plot these observed values against one axis of the graph; and
2. Plot
   $F^{-1}((i - r_{adj}) / (n + n_{adj}))$
   on the other axis

where i is the rank of the respective observation on the ascending scale,

r_{adj} and n_{adj} are adjustment factors (≤ 0.5), and

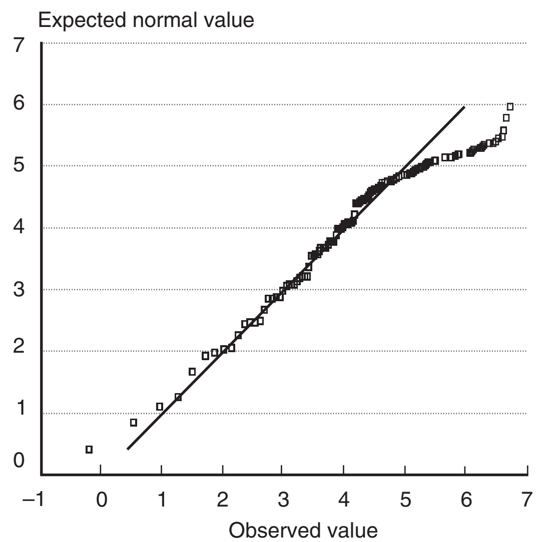
F^{-1} denotes the inverse of the cumulative standard normal distribution function.

The resulting Q-Q plot is thus a scatterplot of the observed values against the (standardized) expected values, given the normal distribution.

The second plot, the P-P plot (i.e., the probability-probability plot) is similar to the Q-Q plot except that the observed cumulative distribution function is plotted against the theoretical cumulative distribution function. As in the Q-Q plot, the values of the respective variable are first sorted into ascending order. The i^{th} observation is then plotted on one axis as i/n (i.e., the observed cumulative distribution function), and $F(x_{(i)})$ is plotted on the other axis, where $F(x_{(i)})$ represents the value of the cumulative normal distribution function for the respective observation $x_{(i)}$. If the normal cumulative distribution approximates the observed distribution well, then all points in this plot should fall onto a diagonal line.

In figures 1 and 2, we present the Q-Q and P-P plots for the logarithm of the trip duration values in our sample. The Q-Q plot is fairly linear except for large values of trip duration. This is not uncommon, because the Q-Q plot will amplify small differences between the model and sample probabilities when they are both close to one. The P-P plot is also fairly linear, except in the middle; this effect is also expected, because the P-P plot amplifies small differences in the "middle" of the model and sample probabilities.[4] The linearity of the two probability plots provides some justification for the use of a normal distribution for the logarithm of trip durations.

FIGURE 1 Normal Q-Q Plot of Log Trip Duration

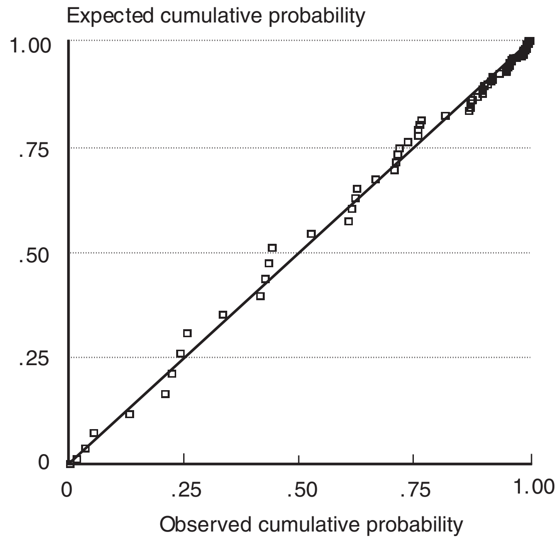


Formal Statistical Test

We next conduct a formal statistical test of normality on the logarithm of trip duration data in our sample. In the statistical literature, several goodness-of-fit statistics have been proposed to test the (null) hypothesis that the sample observations are independent draws from a normal distribution. These tests work well for small to medium sample sizes; however, they will almost always reject the null when n is large (as is the case in our sample; see Gibbons 1985, p. 76). To circumvent this problem, we randomly chose a subsample of 1,000 vehicle trips from our data, and conducted the Kolmogorov-Smirnov (with Lilliefors correction) test for normality on this subsample (see Lilliefors 1967 for a descrip-

[4] See Law and Kelton (1991, pp. 374–380) for a description of the P-P and Q-Q probability plots.

FIGURE 2 Normal P-P Plot of Log Trip Duration



tion of this test). The Kolmogorov-Smirnov statistic for the subsample was 0.028, which is less than the critical value of 0.033 (computed as $=1.031/\sqrt{N}$, at the 0.01 level of significance). Therefore, based on this test, we cannot reject the null hypothesis of normality at the 0.01 level of significance.

From the informal plots discussed in the previous section and the statistical test discussed above, we conclude that our normality assumption for the logarithm of trip duration is reasonable.

Model Comparison

In this section, we compare the performance of our proposed model with that of the "default" MOBILE model. The default model uses the same trip duration activity parameters for all zones in the region, and these parameters do not vary by time of day, and trip purpose.

For model evaluation, we observed data on the proportion of trip duration activity (equivalent to VMT) in each of the six time bins for each zone, each time of day, and each trip purpose.[5] The corresponding trip duration activity parameters predicted by both our model and the default MOBILE model are available for analysis (note again that the default MOBILE model-predicted trip duration activity parameters do not vary by zone, time of day, or trip purpose).

The evaluation of the proximity of estimated and actual trip duration activity parameters can then be based on a pseudo-R^2 value computed as shown below

$$R^2 = \frac{\sum_z \sum_t \sum_i \sum_{k=1}^{6} (\hat{y}_{tiz}^k - \bar{y}_{ti}^k)^2}{\sum_z \sum_t \sum_i \sum_{k=1}^{6} (y_{tiz}^k - \bar{y}_{ti}^k)^2} \tag{15}$$

where

y_{tiz}^k is the actual trip duration activity parameter (i.e., the proportion of VMT accrued) for trip duration bin k, trip purpose i, time- period t, and zone z,

$\hat{y}_{tiz}^k$ is the model-predicted trip duration activity parameter,

$\bar{y}_{ti}^k$ is the area-wide average parameter for trip duration bin k for trip-purpose i in time-period t.

The denominator is the total variation in the actual trip duration activity parameter values around the mean area-wide value, summed across all trip duration bins for all zones, time periods, and trip purposes. The numerator represents the variation explained by the model.

The denominator in the above equation cannot be computed with the sample used in this paper, because we do not have adequate observations in each zone to obtain meaningful averages of VMT accrued in each time duration bin k for each zone and for each time-of-day and activity purpose combination. However, since the denominator remains the same for our proposed model and the default model, a comparison of the two alternative models can be made by taking the ratio of the values of the numerators of the proposed and default models. This statistic can be viewed as a "net performance measure" that represents an index of the total variation in the trip duration activity parameters explained by the proposed model as compared with the default model. A value of the net performance measure that is greater than unity will reveal that our proposed model is superior.

[5] This information is not available in our dataset, because we observed only a sample of trips made by households in each zone that did not span all of the time of day, trip purpose, and trip duration bin categories.

The computation of the net performance index as discussed earlier is still tedious. To simplify, we computed the measure for a restricted version of our proposed model vis-à-vis the default model. The restricted version ignores variations in the trip duration activity parameters across zones. Thus, we get the activity parameter predictions from our proposed model for a single representative zone (with characteristics representing the average of all the zones in the sample) for each time duration bin k and for each activity purpose and time-of-day combination. We also obtained the corresponding trip duration activity parameters from the default model, which assigns the same values of the proportion of VMT accrued in each time bin k across all activity purposes and times of the day (table 7). The net performance measure can be computed by evaluating the closeness of these predictions to the average values for each time of day and activity purpose obtained from the sample.

The value of the net performance measure is 3.89. This shows that our proposed model is performing about four times better than the default model in explaining the variation in trip duration activity parameters across zones, time of day, and trip purpose. Additionally, we also computed performance measures for each of three broad time periods (peak, offpeak, and evening) and for each of the 12 trip purpose combinations. Table 8 presents these performance measures. As can be seen, the proposed model out-performs the default model in all the categories, revealing its efficiency in capturing the trip duration activity parameters relative to the default model.

In addition to our performance measure, we compared the two models through a statistical test. For this, we note that the disaggregate version of the default MOBILE model is equivalent to the "constant-only" specification for the log (trip duration) in equation 1. An F-test of the null hypothesis that all the parameters (except the constant) are equal to zero would therefore test our model (the "unconstrained" specification) against the default model (the "constrained" specification). The corresponding F-statistic is 229.05. This very strongly rejects the null hypothesis that the zonal characteristics, time of day, and trip purpose do not affect trip duration (the relevant critical F-value at the 99% significance level is 2.36).

CONCLUSIONS

The modeling of trip durations in a metropolitan area is important for three reasons. First, trip duration activity parameters used by the MOBILE emissions factor model to estimate running loss emissions can be developed from the trip duration distribution. Second, the trip duration distribution provides information for estimating operating mode fractions, which are needed by MOBILE5 to estimate emissions rates. Third, the trip duration distribution can be used to predict the VMT accumulated on local roads in the region.

Trip duration is likely to depend on various factors such as trip purpose, time of day of the trip start, and other land-use and socio-demographic characteristics of the zone of trip production. In this paper, we formulated and implemented a methodology for modeling trip durations as a function of these characteristics, using vehicle trip data from household travel surveys and supplementary zonal, demographic, and land-use data. The approach involves developing the distribution of the duration of trips using a log-linear regression model. The modeling framework is implemented in the context of mobile source emissions analysis for the Dallas-Fort Worth area of Texas. Model evaluation indicates that the proposed model is superior to the default model in explaining the variation in trip duration activity parameters across zones, times of day, and trip purposes.

TABLE 7 EPA-Recommended Default Trip Duration Proportions

Trip duration (minutes)	0–10	11–20	21–30	31–40	41–50	>50
Proportion	6.7%	18.5%	16.8%	13.2%	8.3%	36.5%

TABLE 8 Performance Measures of Proposed vs. Default Models for Each Trip Purpose and Time Period

	Time period		
Trip purpose	Peak	Offpeak	Evening
Home-based work	7.74	5.65	6.07
Home-based school	4.99	5.97	4.85
Home-based social-recreational	3.75	4.19	4.12
Home-based shopping	3.24	3.66	2.56
Home-based personal-business	6.68	4.66	4.82
Home-based other	3.47	3.19	3.82
Nonhome-based work	6.15	5.62	3.78
Nonhome-based school	5.52	4.85	1.83
Nonhome-based social-recreational	4.73	3.12	4.09
Nonhome-based shopping	4.53	3.59	3.07
Nonhome-based personal-business	4.01	3.44	3.13
Nonhome-based other	4.60	4.21	4.44
All trip purposes/time periods		3.89	

The proposed model contributes significantly toward improved mobile source emissions forecasting by systematically developing area-specific estimates of running loss emissions, running mode fractions, and VMT on local roads. A distinguishing characteristic of the methodology is the straightforward manner in which model parameters estimated from vehicle trip data can be applied to obtain zonal-level trip duration distributions. The model can be integrated easily within various travel demand-air quality modeling frameworks.

Future work in this area could include developing a generalized posterior distribution for trip duration using a Bayesian framework that infers the nature of the distribution of trip durations from the sample. The trip duration activity parameters could then be computed by numerically evaluating the VMT proportions from this posterior distribution. Further, rather than parameterize the trip duration distribution, a flexible semi-nonparametric model could be developed, which could capture (possible) nonlinear responses in trip durations to changes in the exogenous variables.

ACKNOWLEDGMENTS

This study was funded by the Texas Department of Transportation project "Transportation Control Measure Effectiveness Evaluation in Ozone Non-Attainment Areas." The authors would like to thank Bill Knowles, Carol Nixon, Wayne Young, and George Reeves for their valuable input throughout this research effort. Thanks are also due to Ken Cervenka and Mahmoud Ahmadi of the North Central Texas Council of Governments (NCTCOG) for providing, and clarifying data issues related to, the 1996 activity survey and the zonal land-use characteristics file for the Dallas-Fort Worth area. Ken Kirkpatrick and Christopher Klaus helped clarify current NCTCOG transportation air quality procedures. The authors would also like to thank three anonymous reviewers for their constructive suggestions on an earlier version of the paper. Finally, Lisa Weyant helped with typing and formatting.

REFERENCES

Allen, W.G. and G.W. Davies. 1993. A New Method for Estimating Cold-Start VMT. *Compendium of Technical Papers from ITE 63rd Annual Meeting,* The Hague, The Netherlands, September 22.

Brodtmen, K.J and T.A. Fuce. 1984. *Determination of Hot and Cold Start Percentages in New Jersey,* Report FHWA/NJ-84/001. Prepared for the U.S. Department of Transportation, Federal Highway Administration by the New Jersey Department of Transportation. July.

Chatterjee, A., P.M. Reddy, M. Venigalla, and T.L. Miller. 1996. Operating Mode Fractions on Urban Roads Derived by Traffic Assignment. *Transportation Research Record* 1520:97–103.

Ellis, G.W., W.T. Camps, and A. Treadway. 1978. The Determination of Vehicular Cold and Hot Operating Mode Fractions for Estimating Highway Emissions. Alabama Highway Department.

Ettema, D. and H. Timmermans, eds. 1997. *Activity-Based Approaches to Travel Analysis*. Pergamon Books.

Frank, L.D., B. Stone, Jr., and W. Bachman. 2000. Linking Land-Use with Household Vehicle Emissions in the Central Puget Sound: Methodological Framework and Findings. *Transportation Research D* 5:173–196.

Gibbons, J.D. 1985. *Nonparametric Methods for Quantitative Analysis,* 2nd Edition. Columbus, OH: American Sciences Press.

Glover, E.L. and D.J. Brzezinski. 2001. *Trip Length Activity Factors for Running Loss and Exhaust Running Emissions,* Report M6.FLT.005. U.S. Environmental Protection Agency, Assessment and Standards Division. Available at http://www. epa.gov/otaq/models/mobile6/m6tech.htm, as of June 9, 2003.

Greene, W.H. 1997. *Econometric Analysis*. Upper Saddle River, NJ: Prentice-Hall International. pp. 948–952.

Johnson, N., and S. Kotz. 1970. *Distributions in Statistics—Continuous Univariate Distributions*. New York, NY: Houghton Mifflin.

Law, A.M, and W.D. Kelton. 1991. *Simulation Modeling and Analysis,* 2nd Edition. New York, NY: McGraw Hill.

Lilliefors, H.W. 1967. On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown. *Journal of the American Statistical Association* 62, June.

Nair, H.S., C.R. Bhat, and R.J. Kelly. 2002. Modeling Soak-Time Distribution of Trips for Mobile Source Emissions Forecasting: Techniques and Applications. *Transportation Research Record* 1750:24–31.

U.S. Environmental Protection Agency (USEPA), Office of Mobile Sources. 1993. *Federal Test Procedure Review Project: Preliminary Technical Report,* EPA 420-R-93-007. Available at http://www.epa.gov/oms/regs/ld-hwy/ftp-rev/ftp-summ.txt, as of June 9, 2003.

Venigalla, M., A. Chatterjee, and M.S. Bronzini. 1999. A Specialized Equilibrium Assignment Algorithm for Air Quality Modeling. *Transportation Research D* 4:29– 44.

Authors' addresses: Harikesh S. Nair, 5532 S. Kenwood Ave., Apt. 108, Chicago, IL 60637. Email: harikesh.nair@gsb.uchicago.edu.

Corresponding author: Chandra R. Bhat, Department of Civil Engineering, The University of Texas at Austin, 1 College Station C1761, Austin, TX 78712. Email: bhat@mail.utexas.edu.

Roadway Traffic Crash Mapping: A Space-Time Modeling Approach

SHAW-PIN MIAOU
Texas Transportation Institute

JOON JIN SONG
BANI K. MALLICK
Texas A&M University

ABSTRACT

Mapping transforms spatial data into a visual form, enhancing the ability of users to observe, conceptualize, validate, and communicate information. Research efforts in the visualization of traffic safety data, which are usually stored in large and complex databases, are quite limited at this time. This paper shows how hierarchical Bayes models, which are being vigorously researched for use in disease mapping, can also be used to build model-based risk maps for area-based traffic crashes. County-level vehicle crash records and roadway data from Texas are used to illustrate the method. A potential extension that uses hierarchical models to develop network-based risk maps is also discussed.

INTRODUCTION

Transportation-related deaths and injuries constitute a major public health problem in the United States. Injuries and fatalities occur in all transportation modes, but crashes involving motor vehicles account for almost 95% of all transportation fatalities and most injuries. Despite the progress made in roadway safety in the past several decades, tens of thousands of people are still killed and millions of people are injured in motor vehicle crashes each

KEYWORDS: Bayes models, risk, space-time models, traffic safety.

year. For example, in 1999 nearly 42,000 people were killed in traffic crashes and over 3.2 million more were injured.

Motor vehicle fatalities are the leading cause of unintentional injury deaths, followed by falls, poisonings, and drownings (about 16,000, 10,000, and 4,400 deaths per year, respectively) (NSC 2002). They are also responsible for as many pre-retirement years of life lost as cancer and heart disease, about 1.2 million years annually. In fact, motor vehicle crashes are the leading cause of death for people aged 1 to 33. Societal economic losses from these crashes are huge, estimated by the National Highway Traffic Safety Administration to exceed $230 billion in 2000. Thus, much work remains to be done to develop a better understanding of the causes of vehicle crashes—their chains of events and operating environments—and to develop countermeasures to reduce the frequency and severity of these crashes (USDOT 1996–1999).

Safety is one of the U.S. Department of Transportation's (USDOT's) five current strategic goals, and Rodney Slater, a former Transportation Secretary stated: "Safety is a promise we keep together." Indeed, roadway safety intersects with all five core functional areas within conventional highway engineering (planning, design, construction, operation, and maintenance) and crosscuts the boundaries of other engineering (vehicle and material) and non-engineering areas (human factors, public health, law enforcement, education, and other social sciences). Thus, research in roadway safety requires interdisciplinary skills and essential cooperation from various engineering and social science fields.

In 2002, a series of conferences was hosted by the Bureau of Transportation Statistics under the general title of "Safety in Numbers: Using Statistics to Make the Transportation System Safer." These conferences supported the top strategic safety goal of promoting public health and safety "by working toward the elimination of transportation-related deaths, injuries, and property damage" (USDOT 2002).

Contributing Factors, Countermeasures, and Resources

Motor vehicle crashes are complex events involving the interactions of five major factors: drivers, traffic, roads, vehicles, and the environment (e.g., weather and lighting conditions) (e.g., Miaou 1996). Among these factors, driver error has been identified as the main contributing factor to a great percentage of vehicle crashes, and many research efforts are being undertaken to better understand human and other synergistic factors that cause or facilitate crashes. These factors include operator impairment due to the use of alcohol and drugs, medical conditions, or human fatigue and the operator's interaction with new technologies used on the vehicle.

Countermeasures to reduce the number and severity of vehicle crashes are being sought vigorously through various types of community, education, and law enforcement programs and improved roadway design and vehicle safety technology. However, many of these programs have limited resources and need better tools for risk assessment, prioritization, and resource scheduling and allocation.

Recognizing that "to err is human" and that driver behavior is affected by virtually all elements of the roadway environment, highway engineers are constantly redesigning and rebuilding roadways to meet higher safety standards. This includes designing and building roadways and roadsides that are more "forgiving" when an error is made, more conforming to the physical and operational demands of the vehicle, and that better meet drivers' perceptions and expectations in order to reduce the frequency of human errors (TRB 1987). The relatively low fatality rate on the Interstate Highway System (about half the fatality rate of the remainder of the nation's highways) is evidence of the impact of good design on highway safety (Evans 1991).

Many impediments keep highway engineers from achieving their design and operational goals, including a lack of resources and a vast highway system that needs to be built, operated, maintained, audited, and improved. They must make incremental improvements over time and make difficult decisions on the tradeoffs among cost, safety, and other operational objectives. Consequently, knowing where to improve and how to prioritize and schedule improvements is as important as knowing which roadway and roadside features and elements to add or improve. Tools for identifying, auditing, ranking, and clinically evaluating problem sites; developing countermeasures; and allocating resources are essential for highway engineers who make these decisions.

Disease Mapping and Methods

In recent years, considerable progress has been made in developing methodology for disease mapping and ecological analysis, particularly in the application of hierarchical Bayes models with spatial-temporal effects. This model-based development has led to a dramatic gain in the number and scope of applications in public health policy studies of risks from diseases such as leukemia, pediatric asthma, and lung cancer (Carlin and Louis 1996; Knorr-Held and Besag 1997; Xia et al. 1997; Ghosh et al. 1999; Lawson et al. 1999; Zhu and Carlin 1999; Dey et al. 2000; Sun et al. 2000; Lawson 2001; Green and Richardson 2001). A special issue of *Statistics in Medicine* entitled "Disease Mapping with a Focus on Evaluation" was also recently published to report this development (vol. 19, Issues 17–18, 2000).

Among other applications, disease maps have been used to

- describe the spatial variation in disease incidence for the formulation and validation of etiological hypotheses;
- identify and rank areas with potentially elevated risk and time trends so that action may be taken; and
- provide a quantitatively informative map of disease risk in a region to allow better risk assessment, prioritization, and resource allocation in public health.

Clearly, roadway traffic safety planning has similar requirements and can potentially benefit from these kinds of maps.

Studies have shown that risk estimation using hierarchical Bayes models has several advantages over estimation using classical methods. One important point that has been stressed by almost all of these studies is that individual incidences of diseases of concern are relatively rare for a typical analysis unit such as census tract or county. As a result, estimates based on simple aggregation techniques may be unreliable because of large variability from one analysis unit to another. This variability makes it difficult to distinguish chance variability from genuine differences in the estimates and is sometimes misleading for analysis units with a small population size. Hierarchical Bayes models, however, especially those Poisson-based generalized linear models with spatial random effects, have been shown to have the ability to account for the high variance of estimates in low population areas and at the same time clarify overall geographic trends and patterns (Ghosh et al. 1999; Sun et al. 2000).

Note that in the context of sample surveys the type of problem described above is commonly referred to as a small area, local area, or small domain estimation problem. Ghosh and Rao (1994) conducted a comprehensive review of hierarchical Bayes estimations and found them favorable for dealing with small area estimation problems when compared with other statistical methods. Hierarchical models are also gaining enormous popularity in fields such as education and sociology, in which data are often gathered in a nested or hierarchical fashion: for example, as students within classrooms within schools (Goldstein 1999). In these fields, hierarchical models are often called multilevel models, variance component models, or random coefficients models.

The overall strength of the Bayesian approach is its ability to structure complicated models, inferential goals, and analyses. Among the hierarchical Bayes methods, three are most popular in disease mapping studies: empirical Bayes (EB), linear Bayes (LB), and full Bayes methods. These methods offer different levels of flexibility in specifying model structures and complexity in computations. As suggested by Lawson (2001): "While EB and LB methods can be implemented more easily, the use of full Bayesian methods has many advantages, not least of which is the ability to specify a variety of components and prior distributions in the model set-up."

To many statistical practitioners, it is fair to say that the challenges they face dealing with real-world problems come more often from the difficulties of handling nonsampling errors and unobserved heterogeneity (because of the multitude of factors that can produce them) than from handling sampling errors and heterogeneity due to observed covariates. One potential advantage of using the full Bayes model is the flexibility that it can provide in dealing with and adjusting for the unobserved heterogeneity in space and time, whether it is structured or unstructured.

Objectives and Significance of Work

Mapping transforms spatial data into a visual form, enhancing the ability of users to observe, conceptualize, validate, and communicate information. Research efforts in the visualization of traffic safety data, which are usually stored in large and complex databases, are quite limited at this time because of data and methodological constraints (Smith et al. 2001). As a result, it is common for engineers and other traffic safety officials to analyze roadway safety data and make recommendations without actually "seeing" the spatial distribution of the data. This is not an optimal situation.

To the best of our knowledge, unlike the public health community, which has developed models for disease mapping, the roadway safety research community has not done much to develop model-based maps for traffic crash data. One of the objectives of the study presented here was to initiate development of model-based mapping for roadway traffic crashes. Vehicle crash records and roadway inventory data from Texas were used to illustrate the nature of the data, the structure of models, and results from the modeling.

Overall, TxDOT maintains nearly 80,000 centerline-miles of paved roadways, serving about 400 million vehicle-miles per day. Over 63% of the centerline-miles are rural two-lane roads that, on average, carry fewer than 2,000 vehicles per day. These low volume rural roadways carry only about 8% of the total vehicle-miles on state-maintained (or on-system) highways and have less than 7% of the total reported on-system vehicle crashes. Due to the low volume and relatively low crash frequency on these roads, it is often not deemed cost-effective to upgrade these roads to the preferred design standards. However, vehicles on these roadways generally travel at high speeds and thus tend to have relatively more severe injuries when vehicle crashes occur. For example, in 1999, about 26% of the Texas on-system crashes were fatal (K), incapacitating injury (A), and nonincapacitating injury (B) (or KAB) crashes, compared with over 40% of the crashes on rural, two-lane, low volume on-system roads (Fitzpatrick et al. 2001). As a result, we have chosen to focus this study on crashes occurring on rural, two-lane, low-volume, on-system roads.

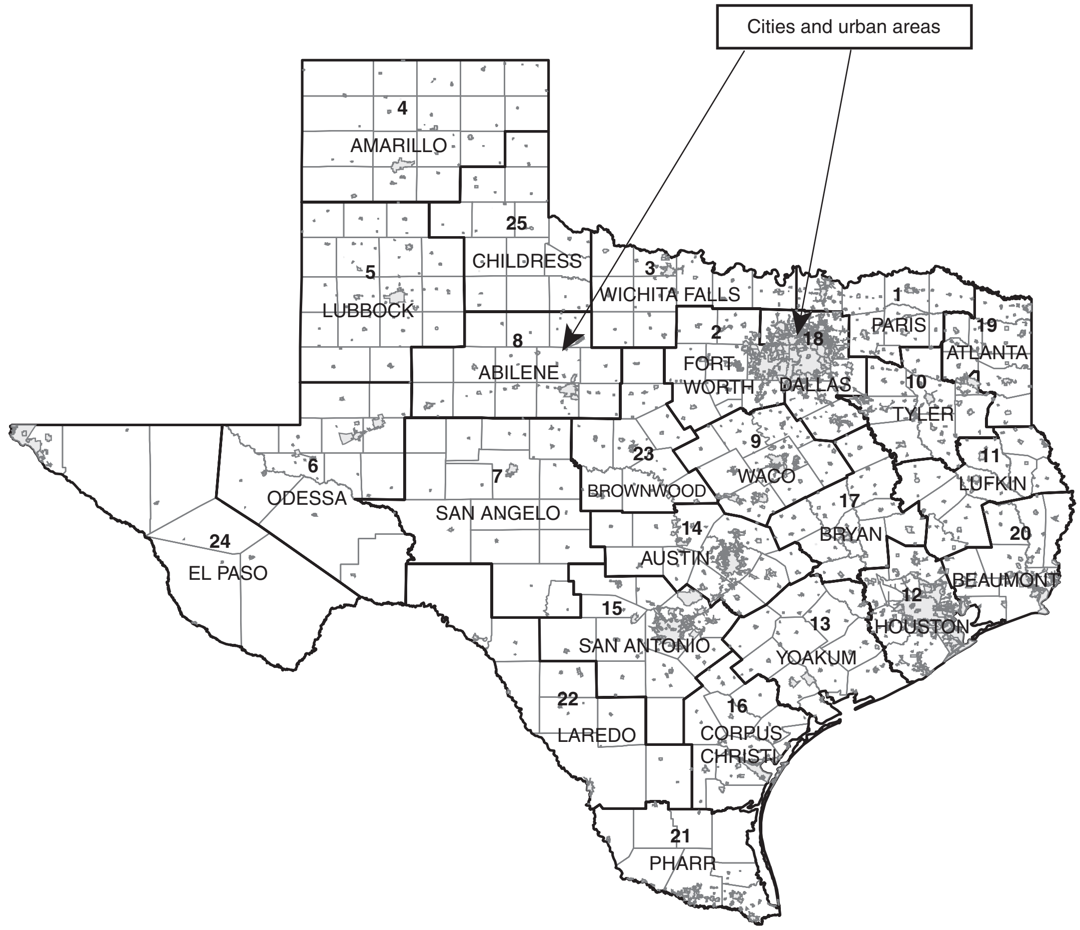
This paper is organized as follows: the next section briefly describes the sources and nature of the data analyzed in this study, followed by a quick review of modeling and computational techniques and a discussion of Poisson-based hierarchical Bayes model with space-time effects and possible variants. Results from models of various levels of complexities are then presented and compared, and we conclude with a discussion of future work.

DATA

The Texas Department of Transportation (TxDOT) currently has 25 geographic districts that are responsible for highway development. The state's 254 counties are divided among the districts (figure 1). Each district includes 6 to 17 counties. District offices divide their work into area offices and area offices into local maintenance offices. The variety of climates and soil conditions in Texas places differing demands on its highways, so design and maintenance, right-of-way acquisition, construction oversight, and transportation planning are primarily administered and accomplished locally.

Annual KAB crash frequencies for rural, two-lane, low volume on-system roads at the county level from 1992 to 1999 were used for modeling in this study. Figure 2 shows the number of reported KAB crashes by county in 1999, while figure 3 shows total vehicle-miles incurred for the same year

FIGURE 1 Geographic Districts, Counties, and Urbanized Areas in Texas



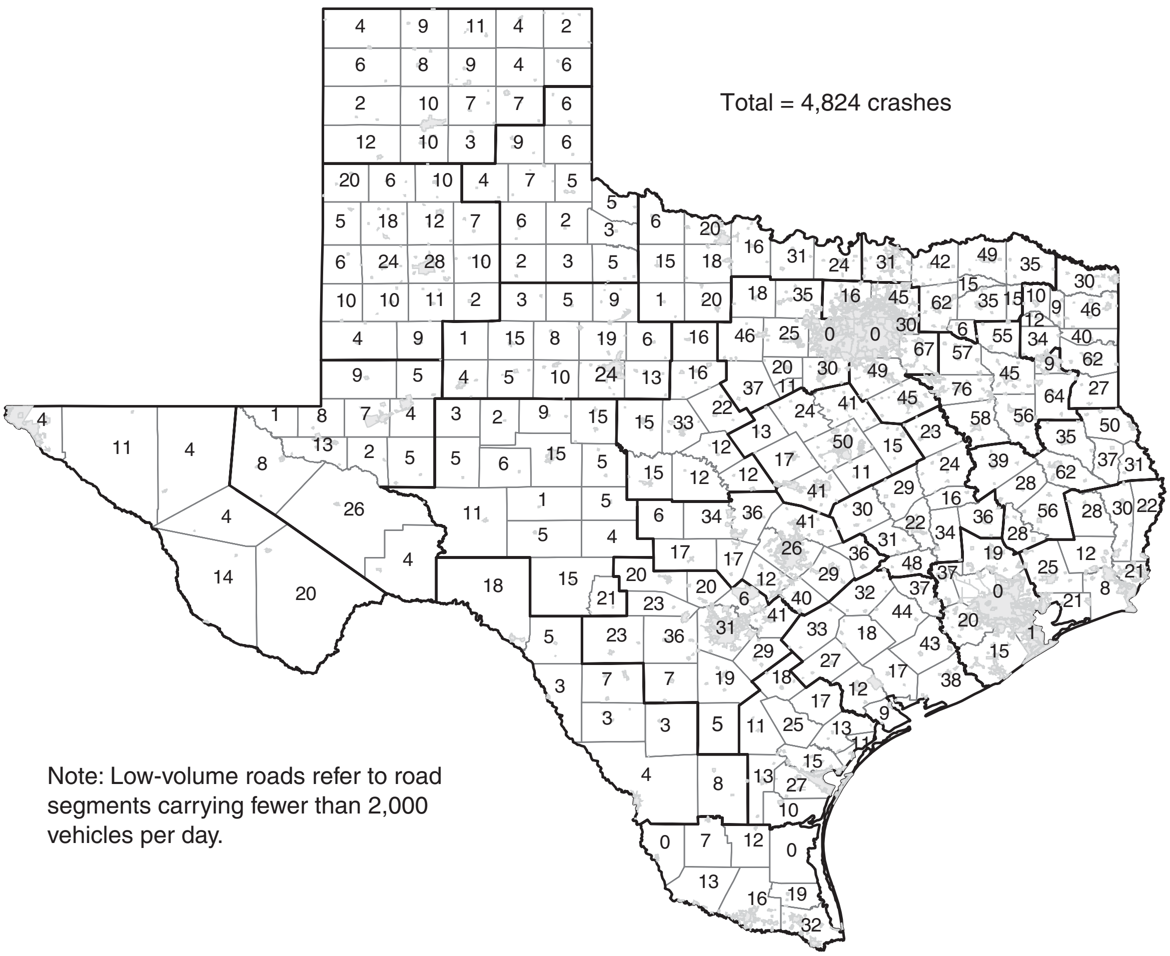
Note: The Texas Department of Transportation has 25 geographic districts; each district contains 6 to 17 counties, and gray areas show urbanized area contained in each district.

(in millions of vehicle-miles traveled, or MVMT). In a bubble plot, figure 4 shows the highest, lowest, and average of the "raw" annual KAB crash rates by county (in number of crashes per MVMT). Note that two of the urban counties and one rural county were removed from the analysis for having no (or almost no) rural two-lane roads with the level of traffic volume of interest, i.e., fewer than 2,000 vehicles per day on average.

As shown in figure 4, crash rates in most counties were stable over the eight-year period, while several counties exhibited marked differences between the high and the low. There is a clear east-west divide in terms of the KAB crash rates, with eastern counties on average showing considerably higher rates. Rural roadways in the eastern counties are limited by the rolling terrain and tend to have less driver-friendly characteristics, with more horizontal and vertical curves (figures 5 and 6), restricted sight distance, and less forgiving roadside development (e.g., trees closer to the travelway and steeper side slopes). In addition, with more and larger urbanized areas in the east, rural roads tend to have higher roadside development scores, higher access density, and narrower lanes and/or shoulders (Fitzpatrick et al. 2001). In general, northern and eastern counties have higher proportions of wet-weather-related crashes (figure 7). Also, on average, rural roads in eastern counties were found to have more crashes at intersections than western counties (figure 8).

The National Highway System Designation Act of 1995 repealed the national maximum speed limit and returned authority to set speed limits to the states. In early 1996, speed limits on many Texas highways during daylight hours were raised from 55 mph to 70 mph for passenger vehicles and 60

FIGURE 2 Number of KAB Crashes on Rural, 2-Lane, Low-Volume, On-System Roads in Each Texas County: 1999



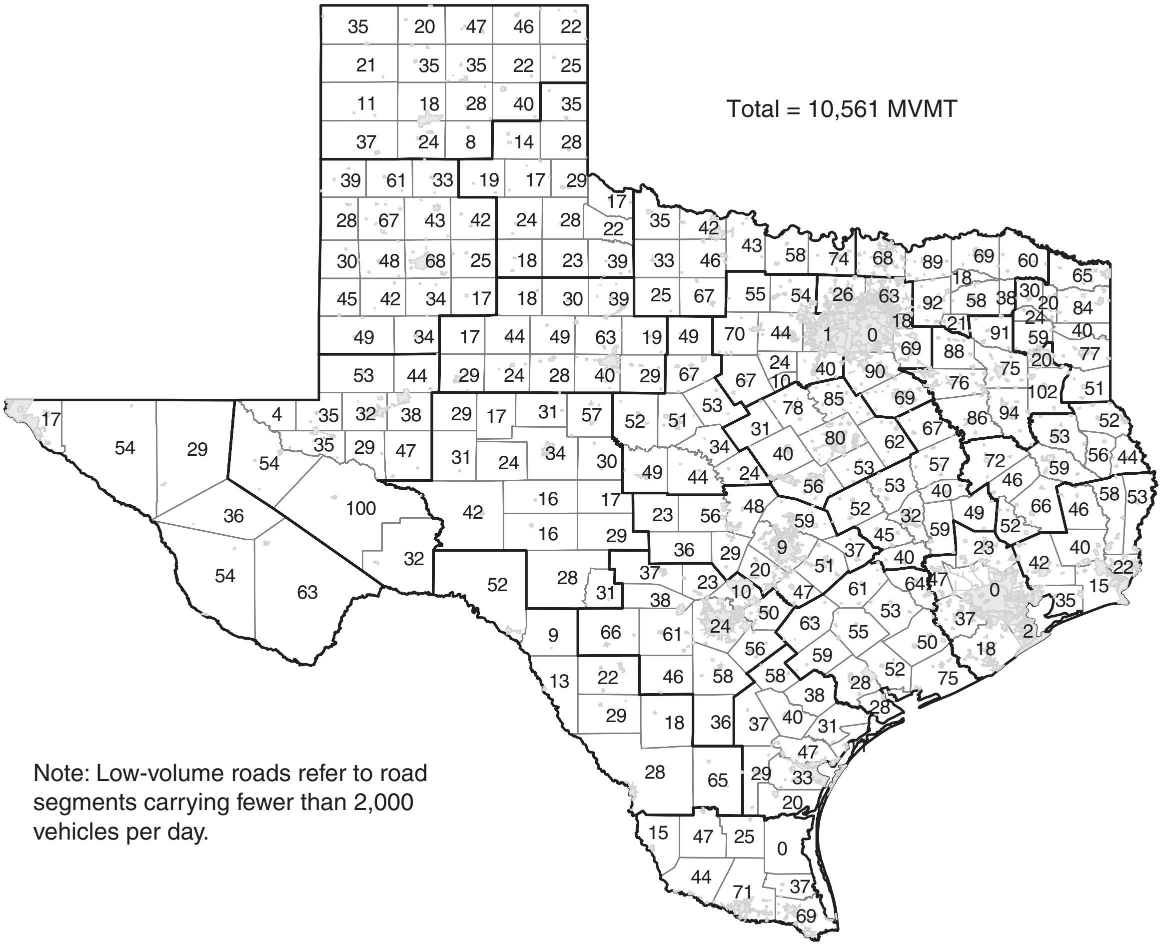
mph for trucks. In a study using monthly time series data from January 1991 to March 1997, it was shown that for those roads on which speed limits were raised, the number of KAB crashes increased in five out of the six highway categories studied during the post-intervention periods (Griffin et al. 1998). The speed limit increase also coincided with a 14% jump in speed-related fatalities, from 1,230 in 1995 to 1,403 in 1996. The number of speed-related injuries increased during that period as well: 3.3% for incapacitating injuries and 7.0% for non-incapacitating injuries. Thus, for the low volume roads considered by this study, we expected to see a change in KAB crash rates in 1996.

MODEL

As part of our modeling efforts, we developed a Poisson hierarchical Bayes model for traffic crash risk mapping at the county level for state-maintained rural, two-lane, low volume roads (fewer than 2,000 vehicles per day) in Texas. In general, the model consists of six components:

1. an offset term (i.e., a covariate with a fixed regression coefficient equal to 1), representing the amount of travel occurring on these roads;
2. a fixed TxDOT district effect;
3. a fixed or random covariate effect component modeling the spatial variation in crash risk due to spatial differences in number of wet days, number of sharp horizontal curves, and degrees of roadside hazards;
4. one random spatial effect component using the inverse of the Great Circle distance between the centroid of counties as the weights for determining spatial correlations;
5. a fixed or random time effect component representing year-to-year changes; and

FIGURE 3 Vehicle-Miles Traveled on Rural, 2-Lane, Low-Volume, On-System Roads in Each Texas County: 1999
In millions of vehicle-miles traveled or MVMT



Note: Low-volume roads refer to road segments carrying fewer than 2,000 vehicles per day.

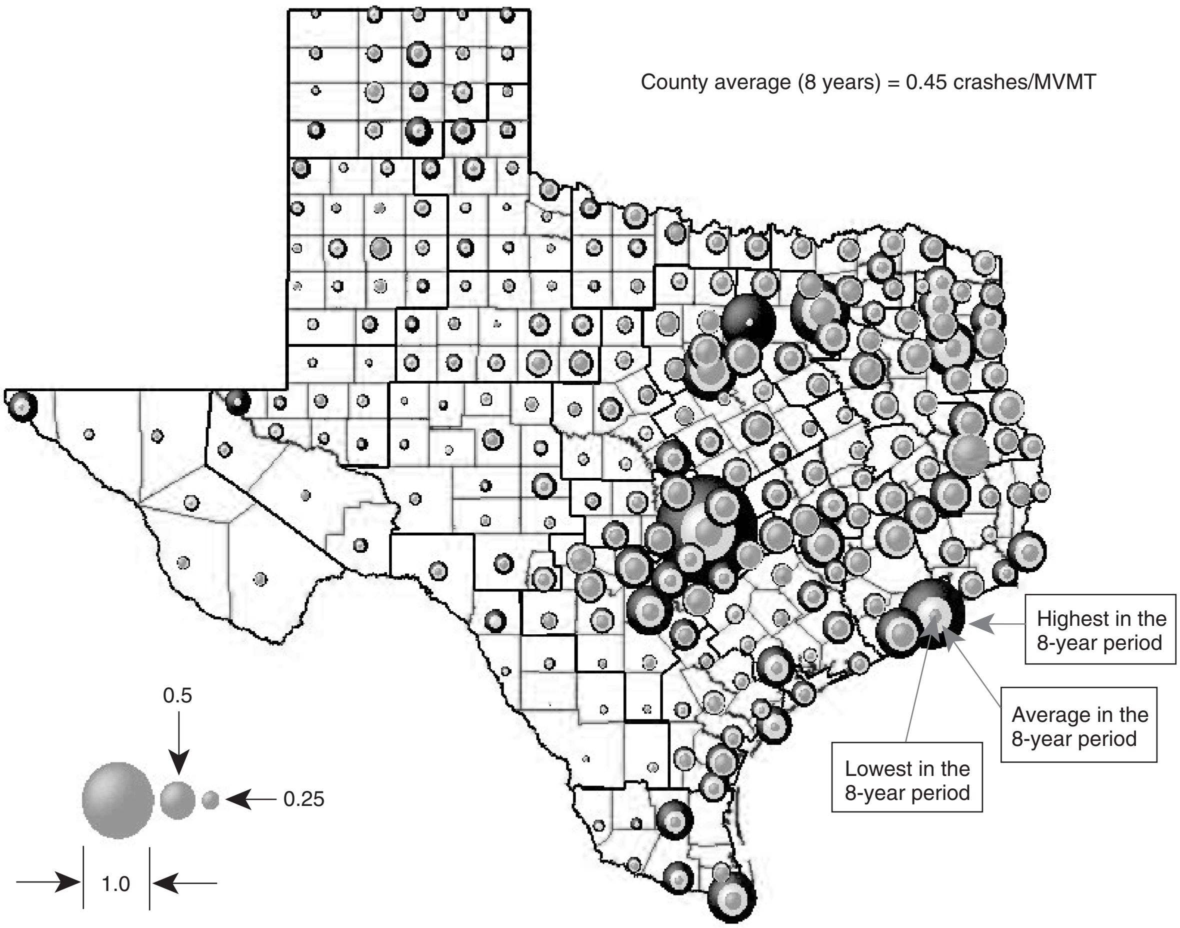
6. an exchangeable random effect term, which, for the purpose of this study, can be deemed as a pure independent random local space-time variation that is independent of all other components in the model.

In this paper, we consider a fixed effect as an effect that is subject only to the uncertainty associated with an unstructured noninformative prior distribution with no unknown parameters and the sampling variation.[1] A fixed effect can, however, vary by individual districts, counties, and time periods (see the discussion of model hierarchy). Note also that unlike the traditional traffic crash prediction models (Maher and Summersgill 1996; Miaou 1996; and Hauer 1997), which were concerned principally with modeling the fixed effects for individual sites (e.g., road segments or intersections), this study focuses more on exploring the structure of the random component of the model for area-based data.

The rediscovery by statisticians in the last 15+ years of the Markov chain Monte Carlo (MCMC) methods and new developments, including convergence diagnostic statistics, are revolutionizing the entire statistical field (Besag et al. 1995; Gilks et al. 1996; Carlin and Louis 1996; Roberts and Rosenthal 1998; Robert and Casella 1999). At the same time, improved computer processing speed and lower data-collection and storage costs are allowing more complex statistical models to be put into practice. These complex models are often hierarchical and high dimensional in their probabilistic and functional structures. Furthermore, many

[1] *Bayesian Data Analysis* (Gelman et al. 1995) provides a Bayesian interpretation of fixed and random effects.

FIGURE 4 “Raw” Annual KAB Crash Rates in Crashes per MVMT by County: Highest, Lowest, and Average Rates, 1992–1999
On rural, 2-lane, low-volume, on-system roads



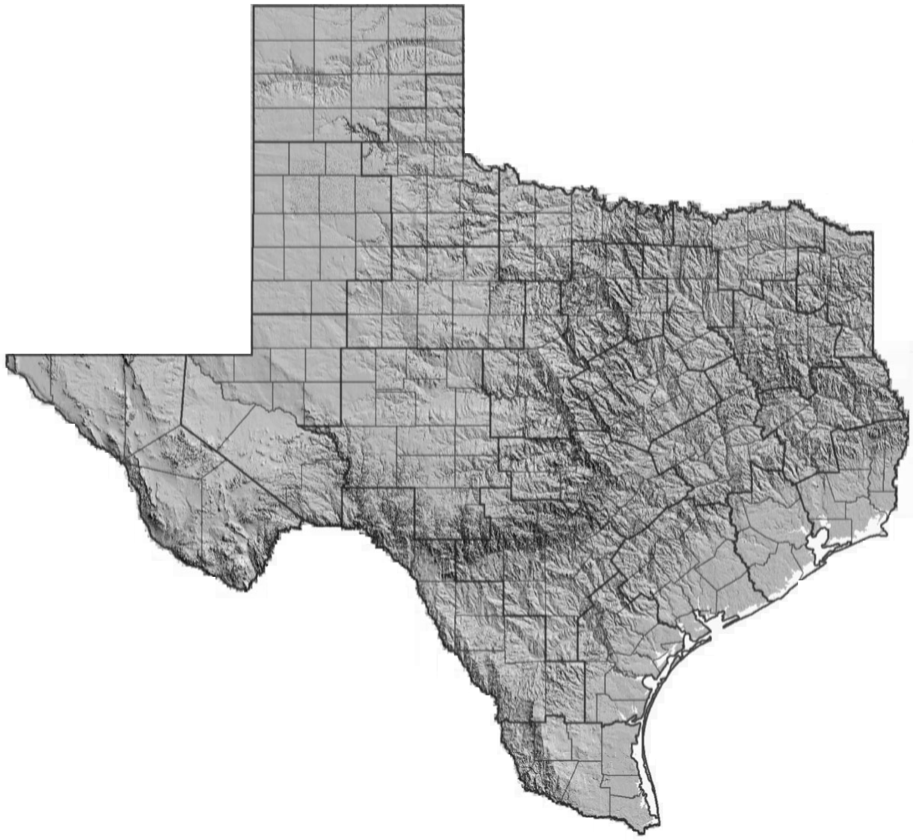
Notes: Crash rate is expressed in terms of the diameter of the ball. The three balls on the lower left corner indicate 1.0, 0.5, and 0.25 crashes per MVMT, respectively. Low-volume roads refer to road segments carrying fewer than 2,000 vehicles per day.

models also need to include dynamics of unobserved and unobservable (or latent) variables; deal with data distributions that are heavily tailed, highly overdispersed, or multimodal; and work with datasets with missing data points. MCMC provides a unified framework within which model identification and specification, parameter estimation, performance evaluation, inference, prediction, and communication of complex models can be conducted in a consistent and coherent manner.

With today’s desktop computing power, it is relatively easy to sample the posterior distributions using MCMC methods that are needed in full Bayes methods. The advantage of full Bayesian treatment is that it takes into account the uncertainty associated with the estimates of the random-effect parameters and can provide exact measures of uncertainty. Maximum likelihood methods, on the other hand, tend to overestimate precision, because they ignore this uncertainty. This advantage is especially important when the sample size is small. Other estimation methods for hierarchical models are also available, e.g., iterative generalized least squares (IGLS), expected generalized least squares (EGLS), and generalized estimating equations (GEE). These estimation procedures tend to focus on obtaining a consistent estimate of the fixed effect rather than exploring the structure of the random component of the model (Goldstein 1999).

For some problems, existing software packages such as WinBUGS (Spiegelhalter et al. 2000) and MLwiN (Yang et al. 1999) can provide Gibbs and

FIGURE 5 Topographical Map of Texas



other MCMC sampling for a variety of hierarchical Bayes models. For the models presented in this paper, we relied solely on the WinBUGS codes. At present, however, the type of spatial and temporal models available in WinBUGS is somewhat limited and will be discussed later.

Notations

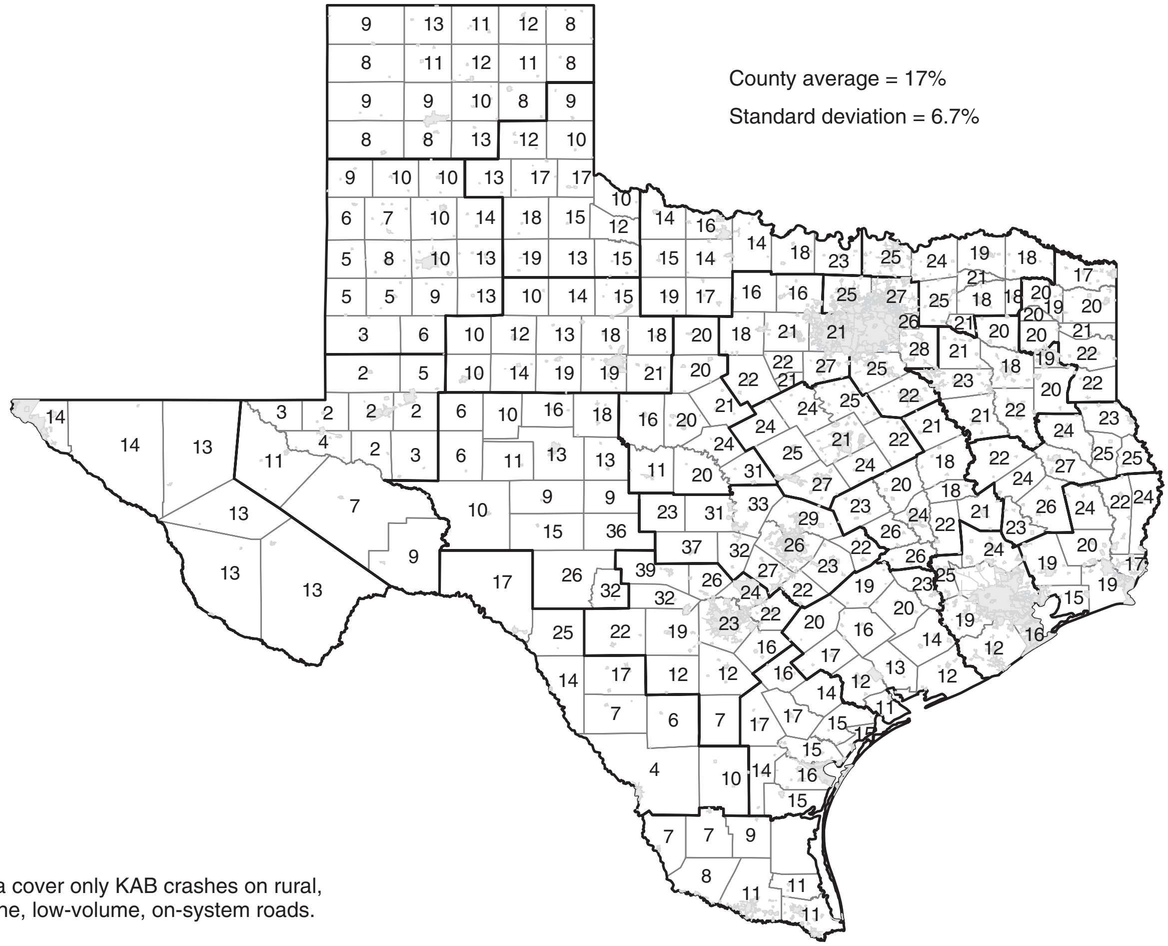
We let the indices i, j, and t represent county, TxDOT district, and time period, respectively,
where $i = 1,2,...,I$; $j = 1,2,...,J$; and $t = 1,2,...,T$.
For the data analyzed, we have 251 counties, divided among 25 districts, and 8 years of annual data (i.e., $I = 251$, $J = 25$, and $T = 8$). As indicated earlier, each district may include 6 to 17 counties, which will be represented by county set D_j, where $j = 1,2,...,25$. That is, D_j is a set of indices representing counties administered by TxDOT district j.

We define variable Y_{it} as the total number of reported KAB crashes on the rural road of interest in county i and year t. We also define v_{it} as the observed total vehicle-miles traveled (VMT) in county i and year t for the roads in discussion, representing the size of the population at risk. In addition, we define x_{itk} as the kth covariate associated with county i and year t. Three covariates were considered.

Covariates

The first covariate x_{it1} is a surrogate variable intended to represent the percentage of time that the road surface is wet due to rain, snow, etc. Not having detailed weather data, we chose to use the proportion of KAB crashes that occurred under wet pavement conditions as a surrogate variable. In addition, we do not expect general weather characteristics to vary much between neighboring counties. Therefore, the proportion for each county is

FIGURE 6 Proportion of KAB Crashes[1] that Occured on Sharp Horizontal Curves in Each County
In percent; averaged over the 1992–1999 period and 6 neighboring counties



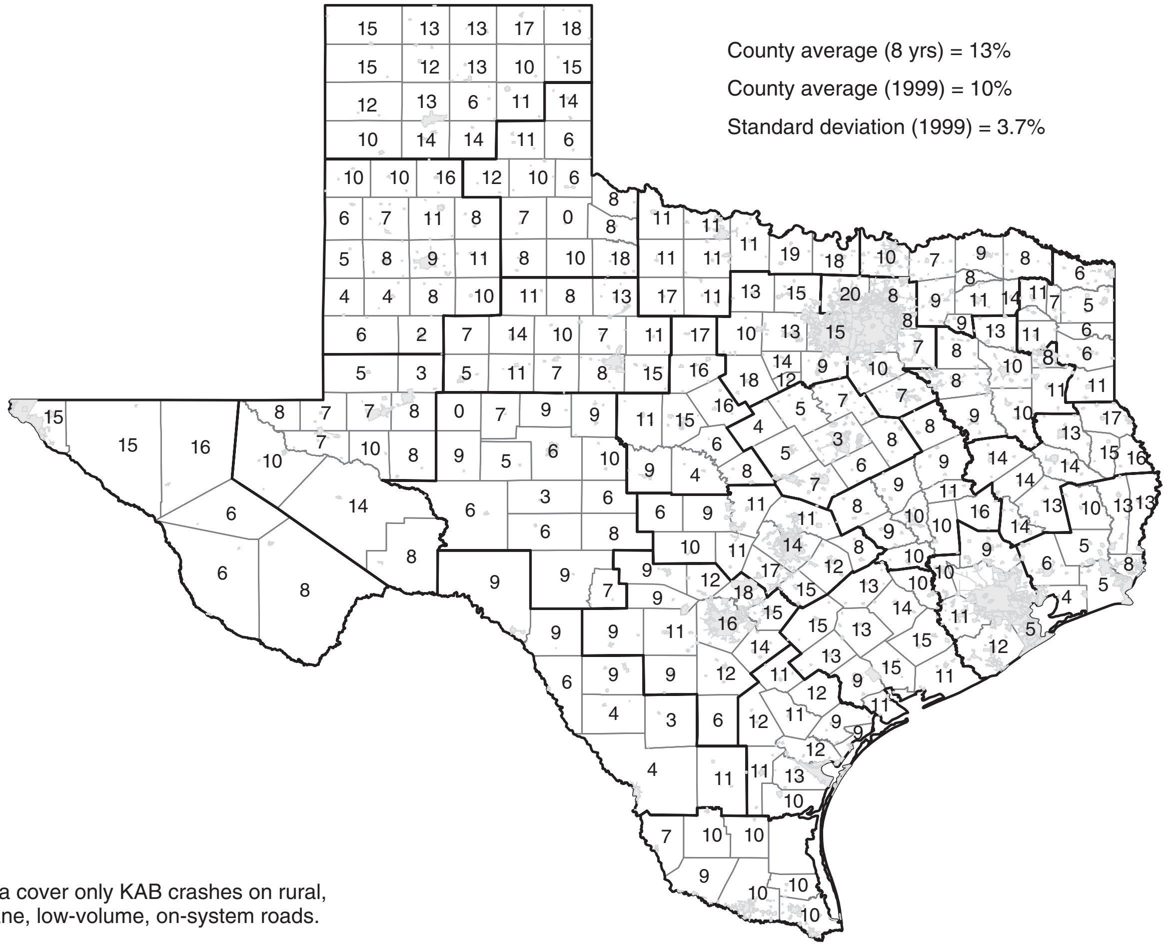
[1] Data cover only KAB crashes on rural, 2-lane, low-volume, on-system roads.

computed as the average of this and six other neighboring counties that are close to the county in terms of their Great Circle distances. We do, however, expect weather conditions to vary significantly from year to year. Thus, for each county *i*, we have x_{it1} change with *t*.

The second covariate x_{it2} is intended to represent spatial differences in the number of sharp horizontal curves in different counties. The actual inventory of horizontal curves on the highway network is not currently available. However, when a traffic crash occurs, site characteristics including the horizontal curvature are coded in the traffic crash database. We chose to use the proportion of KAB crashes that occurred on sharp horizontal curves in each county as a surrogate variable, and we define a sharp horizontal curve as any road segment having a horizontal curvature of 4 or higher degrees per 100-foot arc. Given that this roadway characteristic is mainly driven by terrain variations, we do not expect this characteristic to vary much between neighboring counties. Therefore, as in the first covariate, the proportion for each county is computed as the average of this and six other neighboring counties that are close to the county in terms of their Great Circle distances. Furthermore, for this type of road, we did not expect the proportion to vary in any significant way over the eight-year period in consideration. Thus, the average proportion from 1992 to 1999 was actually used for all *t*. In other words, for each county *i*, x_{it2} are the same for all *t*.

The third covariate x_{it3} is a surrogate variable intended to represent degrees of roadside hazards. As in the second covariate, the actual inventory of hazards (ditches, trees, and utility poles), available clear zones, and geometry and surface type of roadsides are not available. Similar to the first covariate, a surrogate variable was devised to indicate the

FIGURE 7 Proportion of KAB Crashes[1] that Occurred Under Wet Pavement Conditions for Each County: 1999
In percent; averaged over 6 neighboring counties



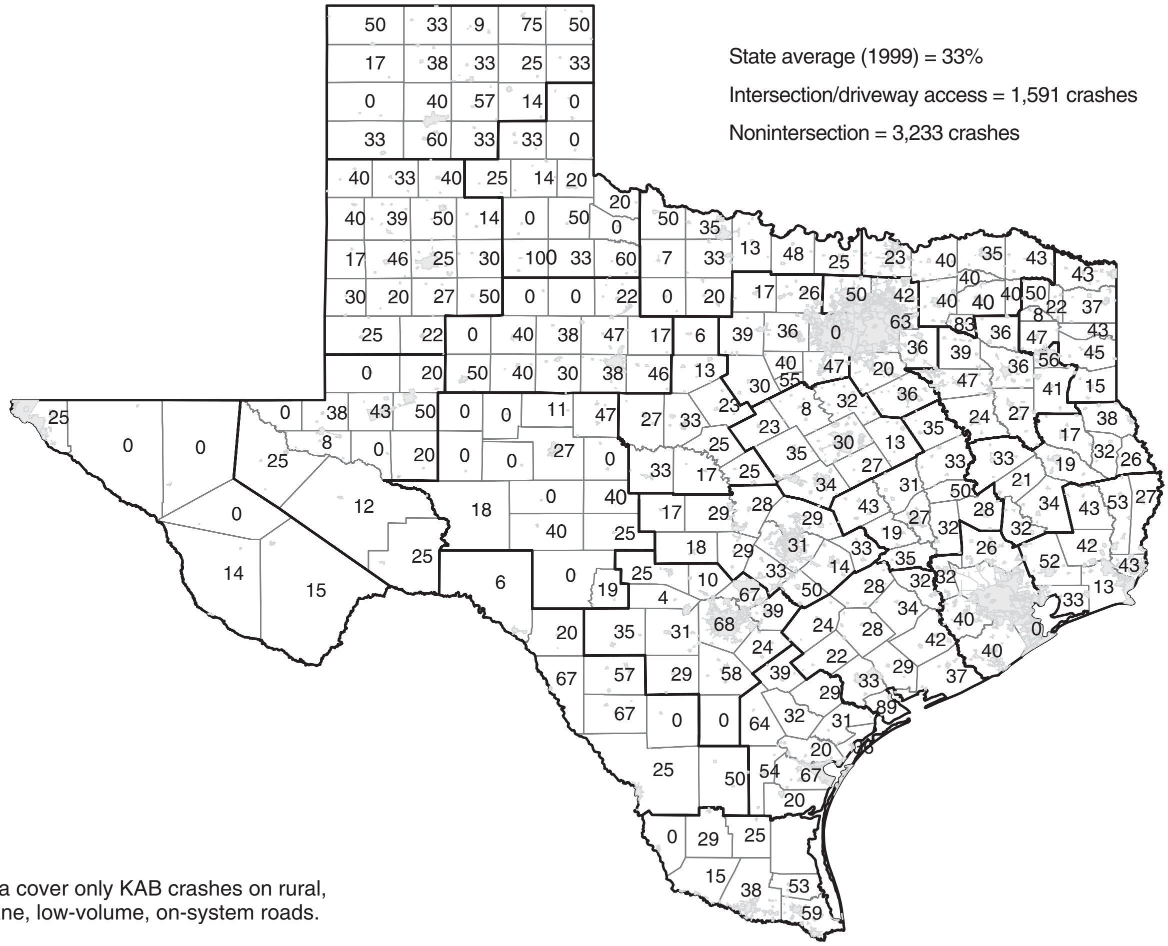
[1] Data cover only KAB crashes on rural, 2-lane, low-volume, on-system roads.

proportion of KAB crashes that ran off roads and hit fixed objects on the roadside. We also do not expect this characteristic to vary much between neighboring counties over the eight-year period in consideration. Again, the average proportion from 1992 to 1999 was used for all *t*, i.e., for each county *i*, x_{it3} are the same for all *t*. Figure 9 shows the spatial distribution of this variable.

The use of these surrogate variables is purely data driven (as opposed to theory driven) and empirical in nature. We use the proportion of wet crashes (x_{it1}) as an example to explain the use and limitation of such surrogate measures in practice. First, variables such as "percentage of wet crashes" and "wet crashes to dry crashes ratio" are commonly used in wet-weather accident studies. Examples in the literature include Coster (1987), Ivey and Griffin (1990), and Henry (2000). These authors reviewed various wet-weather accident studies, and the relationships between 1) skid numbers (or friction values) of pavement and percentage of wet weather accidents, and 2) skid numbers and wet/dry pavement surfaces were quite well documented. Although they were conducted with limited data, these wet weather accident studies also suggest that crash rates are higher during wet surface conditions than under dry surface conditions, and some indicate that traffic volumes are reduced by about 10% to 20% during wet weather in rural areas (no significant reduction was found in urban areas).

Second, the use of percentage of wet crashes as a surrogate variable in this study to explain the variation of crash rates by county mixes several possible relationships and has limited explanatory power. A positive correlation of percentage of wet crashes and crash rate mixes has at least two possible

FIGURE 8 Proportion of KAB Crashes[1] that were Intersection, Intersection Related, or Driveway Access Related for Each County: 1999
In percent



[1]Data cover only KAB crashes on rural, 2-lane, low-volume, on-system roads.

relationships: 1) the effect of wet surface conditions on crash rates, and 2) the effect of rainfall (or other precipitation) on traffic volumes. Everything else being equal, if the wet surface crash rate is the same as the dry surface crash rate, then we do not expect this positive correlation to be statistically significant in the model regardless of the relative traffic volumes during wet or dry surface conditions. We interpret a positive correlation as an indication that a higher crash rate is indeed experienced during wet surface conditions than during dry conditions. However, because of the lack of data on traffic volumes by wet and dry surface conditions, we are not able to quantify the difference in crash rates under the two surface conditions. This is the main limitation in using such a surrogate measure.

Probabilistic and Functional Structures

The space-time models considered in this study are similar to the hierarchical Bayes generalized linear model used in several disease mapping studies cited earlier. At the first level of hierarchy, conditional on mean μ_{it}, Y_{it} values are assumed to be mutually independent and Poisson distributed as

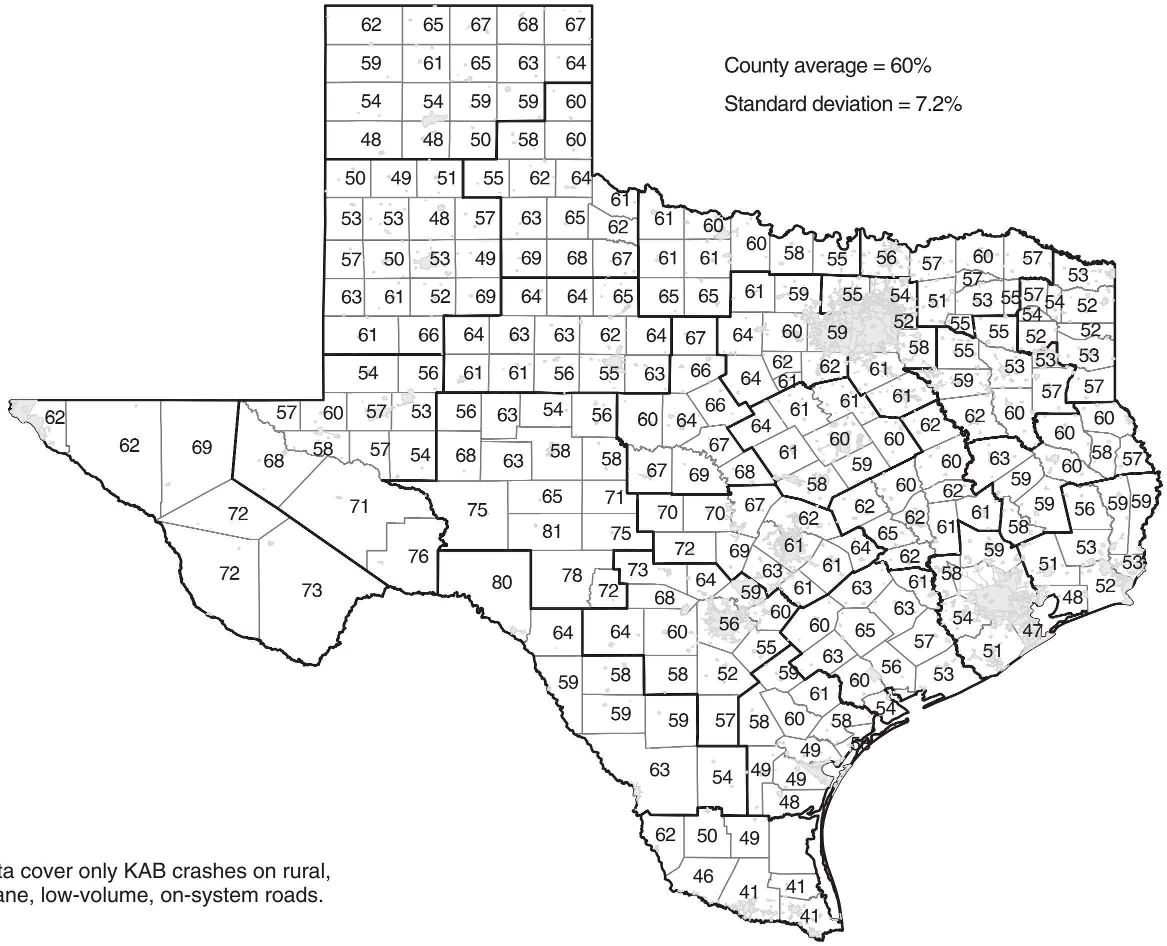
$$Y_{it} \overset{ind}{\sim} Po(\mu_{it}) \tag{1}$$

The mean of the Poisson is modeled as

$$\mu_{it} = v_{it}\lambda_{it} \tag{2}$$

FIGURE 9 Proportion of KAB Crashes[1] Involving Vehicles that Ran Off Roads and Hit Fixed Objects on the Roadside for Each County

In percent; averaged over the 1992–1999 period and 6 neighboring counties



[1] Data cover only KAB crashes on rural, 2-lane, low-volume, on-system roads.

where total VMT v_{it} is treated as an offset and λ_{it} is the KAB crash rate. The rate, which has to be nonnegative, is further structured as

$$\log(\lambda_{it}) = \sum_{t=1}^{T} \sum_{j=1}^{J} \alpha_{jt} I(i \in D_j) + \sum_{k} \beta_k x_{itk} + \delta_t + \phi_i + e_{it} \tag{3}$$

where log is the natural logarithm, *I(S)* is the indicator function of the set *S* defined as

$$I(i \in D_j) = \begin{cases} 1 & \text{if } i \in D_j \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

This makes the first term on the right hand side of equation 3 the intercept representing district effects at different years; x_{itk} are covariates discussed earlier and their interactions; δ_t represents year-to-year time effects due, e.g., to speed limit, weather, and socioeconomic changes; ϕ_i is a random spatial effect; e_{it} is an exchangeable, unstructured, space-time random effect; and α_{jt} and β_k are regression parameters to be estimated from the data. As defined earlier, D_j is a set of indices representing counties administered by TxDOT district *j*.

Many possible variations of equation 3 were and could potentially be considered in this study. For each component that was assumed to have a fixed effect, the second level of hierarchy was chosen to be an appropriate noninformative prior. On the other hand, for each component that was assumed to have a random effect, the second level of hierarchy was a prior with certain probabilistic structure that contained unknown parameters. The priors for these unknown parameters (called hyperpriors) constitute the third level of the hierarchy. What follows are discussions of the variation of models considered by this study, some limitations of the WinBUGS software, and possible extensions of the models considered.

The intercept term, which represents the district effect over time, was assumed to have fixed effects with noninformative normal priors. For the covariates x_{itk}, we considered both fixed and random effects. That is, β_k was assumed to be either a fixed value or random variable. The three covariates discussed earlier and three of their interactive terms, $x_{it4} = x_{it1}x_{it2}$, $x_{it5} = x_{it1}x_{it3}$, and $x_{it6} = x_{it2}x_{it3}$, were included in the model. It is important to note that the values of these covariates were *centered* for better numerical performance. Noninformative normal priors were also assumed for fixed-effect models. For the random-effect model, β_k, $k = 1,2,\ldots,6$, are assumed to be independent and normally distributed with mean μ_{β_k} and variance $\sigma^2_{\beta_k}$, expressed as $N(\mu_{\beta_k}, \sigma^2_{\beta_k})$. Noninformative normal and inverse gamma priors (or more precisely hyperpriors) were assumed for μ_{β_k} and $\sigma^2_{\beta_k}$, respectively.

With 251 counties and 8 years of data, the data are considered to be quite rich spatially but rather limited temporally, as are data in many disease mapping studies. Because of this limitation, we only considered two simple temporal effects for δ_t: fixed effects varying by t (or a year-wise fixed-effect model) and an order-one autoregressive model (AR(1)) with the same coefficient for all t. Again, noninformative priors were used for both models. For the model to be identifiable, in the fixed-effect model, δ_1 was set to zero, and in the AR(1) model, δ_1 was set to be an unknown fixed constant. From the fixed effect, we expected to see a change in δ_t at $t = 5$ (1996), due in part to the speed limit increase in that year.

Recent disease mapping research has focused on developing more flexible, yet parsimonious, spatial models that have attractive statistical properties. Based on the Markov random field (MRF) theory, Besag's conditional autoregressive (CAR) model (Besag 1974 and 1975) and its variants are by far the most popular ones adopted in disease mapping. We considered several Gaussian CAR models, all of which have the following general form

$$p(\phi_i|\phi_{-i}) \propto \eta^{1/2} \exp\left(-\frac{\eta}{2} \sum_{i^* \in C_i} w_{ii^*}(\phi_i - \phi_{i^*})^2 \right) \qquad (5)$$

where $p(\phi_i|\phi_{-i})$ is the conditional probability of ϕ_i given ϕ_{-i};
ϕ_{-i} represents all ϕ except ϕ_i,
$\propto$ stands for "proportional to,"
C_i is a set of counties representing "neighbors" of county i,
η is a fixed-effect parameter across all i, and
w_{ii^*} is a positive weighting factor associated with the county pair (i, i^*).

This equation is shown to be equivalent to

$$p(\phi_i|\phi_{-i}) \sim N\left(\mu_{\phi_i}, \sigma^2_{\phi_i}\right)$$

where $\mu_{\phi_i} = \sum_{i^* \in C_i} \left(w_{ii^*} / w_{i+} \right) \phi_{i^*}$,

$\sigma^2_{\phi_i} = 1/(\eta w_{i+})$, and

$w_{i+} = \sum_{i^* \in C_i} w_{ii^*}$.

In our study, we had $w_{ii^*} = 1/d^c_{ii^*}$,
where d_{ii^*} is the Great Circle distance between the centroid of county i and i^*, and c is a constant parameter equal to 1 or 2 (note that d_{ii^*} ranges roughly from 30 to 700 miles.)

With regard to the number of neighbors, we adopted a more generous definition by allowing every other county i^* $(\neq i)$ to be a neighbor of county i.

In theory, we could treat the constant c as an unknown parameter and estimate it from the data. However, in the current version of WinBUGS, the weights of the built-in CAR spatial model do not allow unknown parameters (Spiegelhalter et al. 2000), which we found to be a limitation for our application. In a separate attempt to find a good range of the decay constant for the inverse distance weight in the CAR model, we adopted a simpler model that included only the offset, the yearwise time effect, and the Gaussian CAR components. We estimated the same model with different c values between 0 and 4 and found that model performance was best achieved when the decay constant was set between 1 and 2 (based on the deviance information criterion to be discussed shortly). Weights with an exponential form $w_{ii^*} = \exp(-cd_{ii^*})$ were also examined but are not reported in this paper.

We also explored the L-1 CAR models of the following form:

$$p(\phi_i|\phi_{-i}) \propto \eta \exp\left(-\eta \sum_{i^* \in C_i} w_{ii^*}|\phi_i - \phi_{i^*}|\right) \tag{6}$$

where η is a fixed-effect parameter the same for all *i*. Weights with the same *c* as in the Gaussian CAR models were considered. WinBUGS constrains the sum of ϕ_i to zero to make both the Gaussian CAR and L-1 CAR spatial models identifiable. A noninformative gamma distribution was used as hyperpriors for η in equations 5 and 6.

The spatial correlation structure represented by equations 5 and 6 is considered global in the sense that the distribution functions and associated parameters (c and η) do not change by *i*. More sophisticated models allowing spatial correlation structure to be adaptive or location specific are being actively researched (e.g., Lawson 2000; Green and Richardson 2001). Still, computational challenges seem to be keeping researchers from exploring more flexible, yet parsimonious, space-time interactive effects, and more research in this area needs to be encouraged (Sun et al. 2000).

For the exchangeable random effects, we considered two commonly used distributions. One distribution assumed e_{it} to be independent and identically distributed (*iid*) as

$$e_{it} \overset{iid}{\sim} N(0, \sigma_e^2) \tag{7}$$

Another distribution assumed an *iid* one-parameter gamma distribution as

$$\exp(e_{it}) \overset{iid}{\sim} Gam(\psi, \psi) \tag{8}$$

which has a mean equal to 1 and a variance $1/\psi$. The use of a one-parameter gamma distribution (instead of a two-parameter gamma) ensures that all model parameters are identifiable. Again, noninformative inverse gamma and gamma distributions were used as hyperpriors for σ_e^2 and ψ, respectively.

Deviance Information Criterion and Variants

The deviance information criterion (DIC) has been proposed to compare the fit and complexity (measured by the effective number of parameters) of hierarchical models in which the number of parameters is not clearly defined (Spiegelhalter et al. 1998; Spiegelhalter et al. 2002). DIC is a generalization of the well-known Akaike Information Criterion (AIC) and is based on the posterior distribution of the deviance statistic

$$D(\theta) = -2\log[p(y|\theta)] + 2\log[f(y)]$$

where $p(y|\theta)$ is the likelihood function for the observed data vector *y* given the parameter vector θ, and

$f(y)$ is some standardizing function of the data alone. For the Poisson model, $f(y)$ is usually set as the saturated likelihood, i.e., $f(y) = p(y|\mu = y)$

where μ is a vector of the statistical means of vector *y*.

DIC is defined as a classical estimate of fit plus twice the effective number of parameters, which gives

$$DIC = D(\bar{\theta}) + 2p_D = \bar{D} + p_D \tag{9}$$

where $D(\bar{\theta})$ is the deviance evaluated at $\bar{\theta}$, the posterior means of the parameters of interest;

p_D is the effective number of parameters for the model; and

$\bar{D}$ is the posterior mean of the deviance statistic $D(\theta)$.

As with AIC, models with lower DIC values are preferred. From equation 9, we can see that the effective number of parameters p_D is defined as the difference between the posterior mean of the deviance $\bar{D}$ and the deviance at the posterior means of the parameters of interest $D(\bar{\theta})$. It was shown that in nonhierarchical models (or models with negligible prior information) DIC is approximately equivalent to AIC. It has also been emphasized that the quantity of p_D can be trivially obtained from an MCMC analysis by monitoring both θ and $D(\theta)$ during the simulation. For the random-effect model considered in equations 1 through 3, the parameter vector θ should include α_{jt}, β_k, δ_t, ϕ_i and e_{it} for all *i*, *j*, *k*, and *t*.

In addition to DIC values and associated quantities $\overline{D}$, $D(\bar{\theta})$, and p_D, we also used some goodness-of-fit measures that attempted to standardize DIC in some fashion. This includes DIC divided by sample size *n* and R^2_{DIC}, which is defined as

$$R^2_{DIC} = 1 - \frac{DIC_{model} - DIC_{ref}}{DIC_{max} - DIC_{ref}} \tag{10}$$

where DIC_{model} is the DIC value for the model under evaluation;

DIC_{max} is the maximum DIC value under a fixed one-parameter model; and

DIC_{ref} is a DIC value from a reference model that, ideally, represents some expected lower bound of the Poisson hierarchical model for a given dataset.

Clearly, R^2_{DIC} is devised in the spirit of the traditional R^2 goodness-of-fit measure for regression models. Through simulations, Miaou (1996) evaluated several similar measures using AIC for overdispersed Poisson models. Since DIC is known to be noninvariant with respect to the scale of the data (Spiegelhalter et al. 1998; Spiegelhalter et al. 2002), an analytical development of DIC_{ref} is difficult. However, we know that for a model with a good fit, $\overline{D}$ should be close to sample size *n* (Spiegelhalter et al. 2002). We, therefore, chose $DIC_{ref} = n$ as a conservative measure for computing R^2_{DIC}; that is, the effective number of parameters was essentially ignored.

Another goodness-of-fit indicator considered is $1/\psi$, which is the variance of $\exp(e_{it})$ under the gamma model, indicating the extent of overdispersion due to exchangeable random effects. In theory, this value could go to zero when such effects vanish. Thus, similar to R^2_{DIC}, we can devise the following measure:

$$R^2_{\psi} = 1 - (1/\psi)_{\text{model}} / (1/\psi)_{\text{max}}$$

where $(1/\psi)_{\text{model}}$ is the variance of $\exp(e_{it})$ for the model under consideration, and

$(1/\psi)_{\text{max}}$ is the amount of overdispersion under the simplest model.

In essence, $1/\psi_{\text{ref}}$, the expected lower bound, is set to zero.

RESULTS

Table 1 lists 42 models of various complexities examined by this study. These models include simplified versions of the general model presented in equations 2 and 3, as well as models for reference purposes, e.g., models 1 to 3. Model 1 is a saturated model, in which the estimates of the Poisson means $\hat{\mu}_{it}$ are equal to y_{it}. Model 2, expressed as Alpha0, is a one-parameter Poisson model without the offset, and model 3 is another one-parameter model with the offset. Essentially, model 2 focuses on traffic crash frequency and model 3 on traffic crash rate.

In table 1, the following symbols are used:

- Alpha(*j*) stands for fixed district effects.
- Beta.Fix and Beta.N respectively represent fixed covariate effects and random covariate effects with independent normal priors.
- Time.Fix and Time.AR1 respectively stand for fixed time and AR(1) time effects.
- For the random spatial effects, Space.CAR.N1 and Space.CAR.L1, represent the Gaussian and L-1 CAR models shown in equations 5 and 6, respectively, and both have a decay constant *c* equal to 1.
- Space.CAR.N2 and Space.CAR.L2 represent similar spatial models with a decay constant *c* equal to 2.
- The components e.N and e.Gam represent exchangeable random effects as presented in equations 7 and 8, respectively.

We experienced some computational difficulties for the models that included the Beta.N component when we tried to include all six main and interactive effects. Therefore, for all models with the Beta.N component, we only included the three main effects.

In computing R^2_{DIC}, DIC_{max} is defined as the maximum DIC value under a fixed one-parameter model, which is model 2 in the table when crash frequency is the focus and model 3 when crash rate is the focus. Similarly, in computing R^2_{ψ}, $(1/\psi)_{max}$ is set as the amount of overdispersion under the simplest model with an e.Gam error component, which is model 11 for models focusing on the crash rate.

TABLE 1 Deviance Information Criterion and Related Performance Measures for Models of Various Complexities

Model no.	Model components (equations 2 and 3)	$\overline{D}$	$D(\bar{\theta})$	P_D	*DIC*	*DIC/n*	R^2_{DIC} *(Freq)*	R^2_{DIC} *(Rate)*	$1/\psi$	R^2_{ψ} *(Rate)*
1	Alpha(*i,t*) [Saturated Model]	2026	61	1965	3991	1.99	0.91	0.77		
2	Alpha0 [Constant Frequency Model]	23416	23414	1	23417	11.66	0.00	-----		
3	Offset+ Alpha0 [Constant Rate Model]	10706	10701	5	10710	5.33	0.59	0.00		
4	Offset+Alpha0+Beta.Fix	6716	6713	2	6718	3.35	0.78	0.46		
5	Offset+Alpha0+Beta.Fix+Time.Fix	6686	6676	10	6695	3.33	0.78	0.46		
6	Offset+Alpha(*j*)+Beta.Fix	5126	5090	36	5161	2.57	0.85	0.64		
7	Offset+Alpha(*j,t*)	5316	5113	202	5518	2.75	0.84	0.60		
8	Offset+Alpha(*j,t*)+Beta.Fix	4816	4608	208	5024	2.50	0.86	0.65		
9	Offset+Alpha(*j,t*)+Beta.Fix+Time.Fix	4816	4623	193	5009	2.49	0.86	0.66		
10	Alpha0+e.N	2116	350	1765	3881	1.93	0.91	0.78		
11	Offset+Alpha0+e.Gam	2086	582	1504	3590	1.79	0.93	0.82	0.263	0.00
12	Offset+Alpha0+e.N	2146	611	1534	3680	1.83	0.92	0.81		
13	Offset+Alpha0+Space.CAR.N1	2846	2473	373	3218	1.60	0.94	0.86		
14	Offset+Alpha(*j*)+e.N	2156	988	1168	3324	1.66	0.94	0.85		
15	Offset+Alpha(*j*)+e.Gam	2136	978	1158	3294	1.64	0.94	0.85	0.103	0.61
16	Offset+Alpha(*j*)+Space.CAR.N1	2846	2483	362	3208	1.60	0.94	0.86		
17	Offset+Alpha(*j*)+Beta.Fix+e.N	2186	1085	1101	3287	1.64	0.94	0.85		
18	Offset+Alpha(*j*)+Beta.Fix+e.Gam	2156	1046	1110	3266	1.63	0.94	0.86	0.089	0.66
19	Offset+Alpha(*j*)+Beta.Fix+Time.AR1+e.N	2176	1080	1096	3272	1.63	0.94	0.85		
20	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.N1+e.N	2066	1376	689	2755	1.37	0.97	0.91		
21	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.L1+e.N	2066	1375	691	2757	1.37	0.97	0.91		
22	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.N1+Time.AR1+e.N	2066	1380	686	2751	1.37	0.97	0.91		
23	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.N1+Time.Fix+e.N	2056	1377	679	2735	1.36	0.97	0.92		
24	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.N1+Time.Fix+e.Gam	2026	1342	684	2709	1.35	0.97	0.92	0.022	0.92
25	Offset+Alpha($_j$)+Beta.N+Space.CAR.N1+Time.Fix+e.Gam	2036	1327	709	2744	1.37	0.97	0.92	0.020	0.92
26	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.N2+Time.Fix+e.N	2046	1360	686	2731	1.36	0.97	0.92		

(*continues on next page*)

TABLE 1 Deviance Information Criterion and Related Performance Measures for Models of Various Complexities (continued)

Model no.	Model components (equations 2 and 3)	$\bar{D}$	$D(\bar{\theta})$	P_D	*DIC*	*DIC/n*	R^2_{DIC} *(Freq)*	R^2_{DIC} *(Rate)*	$1/\psi$	R^2_{ψ} *(Rate)*
27	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.N2+Time.Fix+e.Gam	2016	1329	687	2703	1.35	0.97	0.92	0.021	0.92
28	Offset+Alpha(*j*)+Beta.N+Space.CAR.N2+Time.Fix+e.Gam	2026	1318	708	2733	1.36	0.97	0.92	0.020	0.92
29	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.L1+Time.Fix+e.N	2056	1375	681	2736	1.36	0.97	0.92		
30	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.L1+Time.Fix+e.Gam	2026	1339	687	2712	1.35	0.97	0.92	0.022	0.92
31	Offset+Alpha(*j*)+Beta.N+Space.CAR.L1+Time.Fix+e.Gam	2036	1330	706	2742	1.37	0.97	0.92	0.020	0.92
32	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.L2+Time.Fix+e.N	1966	1181	785	2751	1.37	0.97	0.91		
33	Offset+Alpha(*j*)+Beta.Fix+Space.CAR.L2+Time.Fix+e.Gam	2016	1327	689	2704	1.35	0.97	0.92		
34	Offset+Alpha(*j*)+Beta.N+Space.CAR.L2+Time.Fix+e.Gam	2026	1317	709	2735	1.36	0.97	0.92	0.020	0.92
35	Offset+Alpha(*j,t*)+e.N	2166	917	1249	3415	1.70	0.93	0.84		
36	Offset+Alpha(*j,t*)+e.Gam	2126	886	1240	3365	1.68	0.94	0.84	0.108	0.59
37	Offset+Alpha(*j,t*)+Beta.Fix+e.N	2156	965	1191	3347	1.67	0.94	0.85		
38	Offset+Alpha(*j,t*)+Beta.Fix+e.Gam	2136	962	1174	3310	1.65	0.94	0.85	0.089	0.66
39	Offset+Alpha(*j,t*)+Beta.Fix+Space.CAR.N1+e.N	1956	1103	853	2808	1.40	0.96	0.91		
40	Offset+Alpha(*j,t*)+Beta.Fix+Space.CAR.N1+e.Gam	1936	1091	845	2781	1.38	0.96	0.91	0.027	0.90
41	Offset+Alpha(*j,t*)+Beta.Fix+Space.CAR.N2+e.N	1946	1086	860	2806	1.40	0.96	0.91		
42	Offset+Alpha(*j,t*)+Beta.Fix+Space.CAR.N2+e.Gam	1926	1072	854	2780	1.38	0.96	0.91	0.027	0.90

Notes: i = county, j = district, and t = time period; n = sample size = 2,008; and the saturated part of the deviance statistics = 2log[$f(y)$] = –8,894.

As a rule, in our development we started with simpler models, and the posterior means of the estimated parameters of these simple models were then used to produce initial values for the MCMC runs of more complex models. In general, the models presented in the table are ordered by increasing complexity: intercepts only, intercepts + covariate effect, intercepts + covariate effect + exchangeable effect, intercepts + covariate effect + exchangeable effect + spatial/temporal effects, and so on. Models 7 to 9 and the last eight models include a more complex fixed-effect intercept term. The models are presented in the table in line with the order in which they were estimated with the WinBUGS codes.

The MCMC simulations usually reached convergence quite quickly. Depending on the complexity of the models, for typical runs, we performed 10,000 to 20,000 iterations of simulations and removed the first 2,000 to 5,000 iterations as burn ins. As in other iterative parameter estimation approaches, good initial estimates are always the key to convergence. For some of the models, we have hundreds of parameters and MCMC monitoring plots based on the Gelman-Rubin statistics (which are part of the output from the WinBUGS codes). Because estimated parameters usually converge rather quickly, their convergence plots, which are not particularly interesting to show, are not presented here. Table 2 shows some statistics of the estimated posterior density of a selected number of parameters for model 27, which was one of the best models in terms of the DIC value and other performance measures discussed above. Also, figure 10 presents estimated posterior mean crash rates, as well as their 2.5 and 97.5 percentiles, in a bubble plot for 1999 by county.

From table 2, one can see that the fixed-time effect δ_t jumps from about 0 in previous years to about 0.05 in $t = 4$ (1995) and has another increase to about 0.09 at $t = 5$ (1996). The value comes down somewhat (about 0.06) in 1998 ($t = 7$) and 1999 ($t = 8$) but is still significantly higher than those in the preintervention periods. It has been suggested that the jump in 1995 was perhaps due to higher driving speeds by drivers in anticipation of a speed limit increase, and higher crash rates in 1996 were due in part to the speed limit increase and less favorable winter weather (Griffin et al. 1998). Lower δ_t values in 1998 and 1999 may suggest that drivers had adjusted themselves and become more adapted to driving at higher speeds.

From the same model (model 27), estimates of α_j, i.e., district effects, range from about –0.5 to –1.5, indicating significant district-level variations in crash risk. The covariate effects β_k indicate that the horizontal curve variable is the most influential and statistically significant variable in explaining the crash rate variations over space. Wet pavement condition is the second-most significant variable. The ran-off-road fixed-object variable is not a statistically significant variable, which suggests that ran-off-road fixed-object crash risk is correlated with and perhaps exacerbated by the presence of sharp horizontal curves and wet pavement conditions.

From DIC and other performance measures in table 1, several observations can be made:

- For the exchangeable random effect, models with a gamma assumption (equation 8) are preferred over those with a normal assumption (equation 7). This is observed by comparing the performance of, e.g., model 15 with model 14, model 18 with model 17, and model 27 with model 26.
- Models with fixed covariate effects are favored over their random-effect counterparts. This is seen by comparing, e.g., model 25 with model 24 and model 33 with model 34.
- Models with fixed time effects (e.g., model 23) performed better than those with AR(1) time effects (e.g., model 22).
- Models with separate district and time effects (α_j and δ_t) are preferred over those with joint district time effects (α_{jt}). For example, we can compare the performance of model 27 with model 42 and model 40 with model 24.
- For comparable model structures, adding a spatial component decreases the DIC value quite significantly, which indicates the importance of the spatial component in the model. As an example, we can compare model 17 with

TABLE 2 Example MCMC Simulation Output for Model 27: Some Statistics of the Estimated Posterior Density for a Selected Number of Parameters

Parameter	Mean	Standard error	2.5%	Median	97.5%
α_1	–0.963	0.154	–1.269	–0.964	–0.662
α_2	–0.639	0.148	–0.929	–0.635	–0.356
α_3	–1.131	0.162	–1.450	–1.128	–0.823
α_4	–1.240	0.183	–1.595	–1.237	–0.882
α_5	–1.288	0.155	–1.595	–1.283	–0.993
α_6	–1.427	0.182	–1.768	–1.429	–1.066
α_7	–1.376	0.128	–1.629	–1.375	–1.127
α_8	–1.218	0.130	–1.479	–1.215	–0.978
α_9	–0.984	0.158	–1.283	–0.986	–0.666
α_{10}	–0.582	0.162	–0.889	–0.584	–0.260
α_{11}	–0.610	0.156	–0.924	–0.602	–0.321
α_{12}	–0.498	0.208	–0.918	–0.489	–0.097
α_{13}	–0.919	0.149	–1.232	–0.914	–0.634
α_{14}	–0.668	0.137	–0.943	–0.668	–0.398
α_{15}	–0.770	0.139	–1.045	–0.772	–0.503
α_{16}	–0.893	0.165	–1.216	–0.891	–0.551
α_{17}	–0.754	0.139	–1.030	–0.756	–0.495
α_{18}	–0.630	0.170	–0.966	–0.621	–0.294
α_{19}	–0.649	0.171	–0.975	–0.645	–0.326
α_{20}	–0.877	0.208	–1.282	–0.880	–0.459
α_{21}	–1.005	0.308	–1.656	–0.981	–0.442
α_{22}	–1.561	0.224	–1.980	–1.566	–1.114
α_{23}	–1.189	0.147	–1.483	–1.187	–0.901

(*continues on next page*)

TABLE 2 Example MCMC Simulation Output for Model 27: Some Statistics of the Estimated Posterior Density for a Selected Number of Parameters (continued)

Parameter	Mean	Standard error	2.5%	Median	97.5%
α_{24}	–1.114	0.378	–1.831	–1.127	–0.379
α_{25}	–1.401	0.156	–1.712	–1.396	–1.094
δ_1 (set to 0)					
δ_2	0.0132	0.026	–0.0380	0.0129	0.0645
δ_3	–0.0156	0.027	–0.0677	–0.0156	0.0376
δ_4	0.0508	0.027	–0.0009	0.0508	0.1034
δ_5	0.0929	0.027	0.0418	0.0926	0.1453
δ_6	0.0886	0.027	0.0365	0.0886	0.1408
δ_7	0.0632	0.027	0.0111	0.0631	0.1155
δ_8	0.0603	0.026	0.0089	0.0601	0.1123
β_1	0.00286	0.0018	–0.00079	0.00290	0.00648
β_2	0.00723	0.0019	0.00350	0.00721	0.01103
β_3	–0.00057	0.0014	–0.00346	–0.00057	0.00229
β_4	–0.00004	0.0002	–0.00050	–0.00004	0.00040
β_5	0.00009	0.0002	–0.00028	–0.00010	0.00048
β_6	–0.00015	0.0002	–0.00045	–0.00015	0.00014
ψ	46.52	5.04	37.83	46.18	57.41
$1/\eta$	0.0023	0.0002	0.0019	0.0023	0.0028

model 20. Except for the spatial component, these two models have the same structures (in intercept terms, covariate effects, and the error component). Model 17 does not have any spatial component, while model 20 includes a normal CAR model. The DIC value drops from 3,287 for model 17 to 2,755 for model 20, a very significant reduction when compared with the differences in DIC values for various models presented in table 1. Other comparisons that would give the same conclusion include model 19 vs. model 22 or model 38 with models 40 and 42.

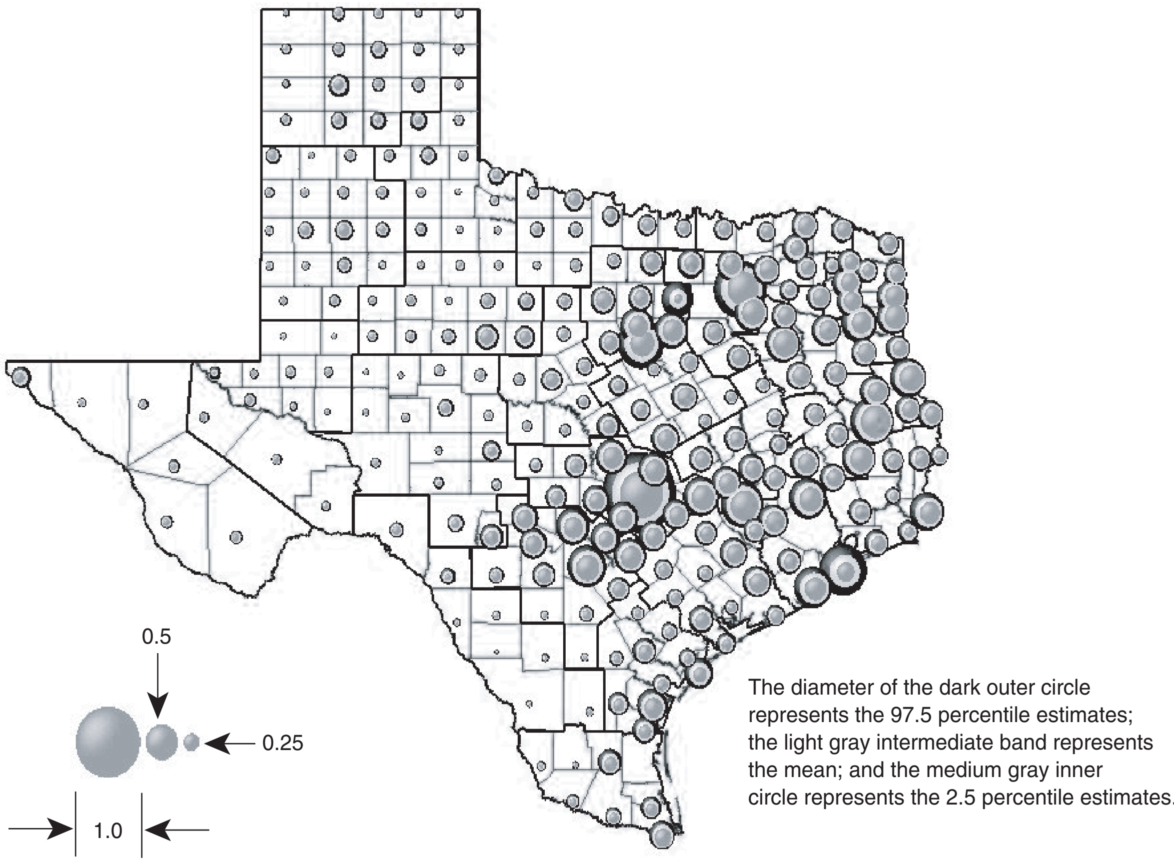
- No particular spatial CAR models considered by this study, i.e., CAR.N1, CAR.L1, CAR.N2, or CAR.L2, were clearly favored over other CAR models.
- Despite the empirical nature of the two goodness-of-fit measures R^2_{DIC} and R^2_{ψ}, seeing some of the better models that have values exceeding 0.9 provides some comfort as to the general explanatory capability of these models.

DISCUSSION

Most of the methodologies developed in disease mapping were intended for area-based data, e.g., number of cancer cases in a county or census tract during a study period. While we demonstrate the use of some of these methodologies for roadway traffic crashes at the county level, we recognize that, fundamentally, traffic crashes are network-based data, whether they are intersection, intersection-related, driveway access-related, or nonintersection crashes. Figure 11 gives an example of the locations of KAB crashes on the state-maintained highway network of a Texas county in 1999.

Thus, an obvious extension of the current study is to develop risk maps for traffic crashes on road

FIGURE 10 Estimated KAB Crash Rates[1] in Crashes per MVMT by County from Model 27: 1999
97.5 percentile, mean, and 2.5 percentile of the posterior density



[1] Data cover only KAB crashes on rural, 2-lane, low-volume, on-system roads.

networks. The problem is essentially one of developing hierarchical models for Poisson events on a network (or a graph). We expect that, in different applications, these maps may need to be developed by roadway functional classes, vehicle configurations, types of crashes (e.g., those involving drunk drivers), and crash severity types (e.g., fatal, injury, and noninjury crashes). We also expect these network-based maps to be useful for roadway safety planners and engineers to 1) estimate the cost and benefit of improving or upgrading various design and operational features of the roadway, 2) identify and rank potential problem roadway locations (or hotspots) that require immediate inspection and remedial action, and 3) monitor and evaluate the safety performance of improvement projects after the construction is completed. Such maps need to be constructed from quality accident-, traffic-, and roadway-related databases and with scientifically grounded data visualization and modeling tools.

Modeling and mapping of traffic crash risk need to face all the challenges just as in the field of disease mapping, i.e., multilevel data and functional structures, small areas of occurrence of studied events at each analysis unit, and strong unobserved heterogeneity. The hierarchical nature of the data can be described as follows: In a typical roadway network, other than the fact that roadway networks are connected or configured in specific ways, individual road entities are classified by key geometric charac-

FIGURE 11 Locations of KAB Crashes on the State-Maintained Highway Network of a Texas County in 1999

teristics (e.g., segments, intersections, and ramps), nested within roadway functional or design classifications, further nested within operational and geographical units, and subsequently nested within various administrative and planning organizations. Strong unobserved heterogeneity is expected because of the unobserved driver behaviors at individual roadway entities that are responsible for a large percentage of crash events.

Every state maintains databases on vehicle crash records and roadway inventory data. We hope that the results of our study using Texas data will motivate the development of similar studies in other states. We also envision that the network-based hierarchical models we propose can potentially be utilized in other transportation modes and in computer and communication network studies to further the exploration and interpretation of incidence data. Furthermore, the hierarchical Bayes models with spatial random effects described in this paper can be used to develop more efficient sampling surveys in transportation that alleviate multilevel and small-area problems. Finally, the models have been shown to have the ability to account for the high variance of estimates in low-population areas and at the same time clarify overall geographic trends and patterns, which make them good tools for addressing some of the equity issues required by the Transportation Equity Act for the 21st Century.

ACKNOWLEDGMENTS

This research was supported in part by the Bureau of Transportation Statistics, U.S. Department of Transportation, via transportation statistics research grant number DTTS-00-G-B005-TX. The authors are, however, solely responsible for the contents and views expressed in this paper.

REFERENCES

Besag, J. 1974. Spatial Interaction and the Statistical Analysis of Lattice Systems (with Discussion). *Journal of the Royal Statistical Society. Series B* 36:192–236.

Besag, J. 1975. Statistical Analysis of Non-Lattice Data. *Statistician* 24(3):179–195.

Besag, J., P. Green, D. Higdon, and K. Mengersen. 1995. Bayesian Computation and Stochastic Systems (with Discussion). *Statistical Science* 10:3–66.

Carlin, B.P. and T.A. Louis. 1996. *Bayes and Empirical Bayes Methods for Data Analysis.* London, England: Chapman and Hall/CRC.

Coster, J.A. 1987. Literature Survey of Investigations Performed to Determine the Skid Resistance/Accident Relationship, Technical Report RP/37. National Institute for Transport and Road Research, South Africa.

Dey, D., S. Ghosh, and B.K. Mallick. 2000. *Bayesian Generalized Linear Model.* New York, NY: Marcel Dekker.

Evans, L. 1991. *Traffic Safety and the Driver.* New York, NY: Van Nostrand Reinhold.

Fitzpatrick, K., A.H., Parham, M.A. Brewer, and S.P. Miaou. 2001. Characteristics of and Potential Treatments for Crashes on Low-Volume, Rural Two-Lane Highways in Texas, Report Number 4048-1. Texas Transportation Institute, College Station, TX.

Gelman, A., J.B. Carlin, H.S. Stern, and D.B. Rubin. 1995. *Bayesian Data Analysis,* 1st ed. Boca Raton, FL: CRC Press.

Ghosh, M., and J.N.K. Rao. 1994. Small Area Estimation: An Appraisal. *Statistical Science* 9(1):55–76.

Ghosh, M., K. Natarajan, L.A. Waller, and D. Kim. 1999. Hierarchical Bayes for the Analysis of Spatial Data: An Application to Disease Mapping. Journal of Statistical Planning and Inference 75:305–318.

Gilks, W.R., S. Richardson, and D.J. Spiegelhalter. 1996. *Markov Chain Monte Carlo in Practice*. London, England: Chapman and Hall.

Goldstein, H. 1999. *Multilevel Statistical Models,* 1st Internet edition. Available at http://www.ioe.ac.uk/multilevel/.

Green, P.J. and S. Richardson. 2001. Hidden Markov Models and Disease Mapping, Working Paper. Department of Mathematics, University of Bristol, United Kingdom.

Griffin, L.I., O. Pendleton, and D.E. Morris. 1998. An Evaluation of the Safety Consequences of Raising the Speed Limit on Texas Highways to 70 Miles per Hour, Technical Report. Texas Transportation Institute, College Station, TX.

Hauer, E. 1997. *Observational Before-After Studies in Road Safety.* New York, NY: Pergamon Press.

Henry, J.J. 2000. *Design and Testing of Pavement Friction Characteristics,* NCHRP Project 20-5, Synthesis of Highway Practice, Topic 30-11. Washington, DC: Transportation Research Board.

Ivey, D.L. and L.I. Griffin, III. 1990. *Proposed Program to Reduce Skid Initiated Accidents in Texas.* College Station, TX: Texas Transportation Institute.

Knorr-Held, L. and J. Besag. 1997. Modelling Risk from a Disease in Time and Space, Technical Report. Department of Statistics, University of Washington.

Lawson, A.B., A. Biggeri, D. Bohning, E. Lessafre, J.F. Viel, and R. Bertollini, eds. 1999. *Disease Mapping and Risk Assessment for Public Health.* Chichester, UK: Wiley.

Lawson, A.B. 2000. Cluster Modelling of Disease Incidence via RJMCMC Methods: A Comparative Evaluation. *Statistics in Medicine* 19:2361–2375.

____. 2001. Tutorial in Biostatistics: Disease Map Reconstruction. *Statistics in Medicine* 20:2183–2204.

Maher, M.J. and L. Summersgill. 1996. A Comprehensive Methodology for the Fitting of Predictive Accident Models. *Accident Analysis & Prevention* 28(3):281–296.

Miaou, S.P. 1996. Measuring the Goodness-of-Fit of Accident Prediction Models, FHWA-RD-96-040. Prepared for the Federal Highway Administration, U.S. Department of Transportation.

National Safety Council (NSC). 2002. *Report on Injuries in America, 2001.* Available at http://www.nsc.org/library/rept2000.htm, as of July 2003.

Robert, C.P. and G. Casella. 1999. *Monte Carlo Statistical Methods.* New York, NY: Springer-Verlag.

Roberts, G.O. and J.S. Rosenthal. 1998. Markov Chain Monte Carlo: Some Practical Implications of Theoretical Results. *Canadian Journal of Statistics* 26:5–31.

Smith, R.C., D.L. Harkey, and B. Harris. 2001. Implementation of GIS-Based Highway Safety Analyses: Bridging the Gap, FHWA-RD-01-039. Prepared for the Federal Highway Administration, U.S. Department of Transportation.

Spiegelhalter, D.J., N. Best, and B.P. Carlin. 1998. Bayesian Deviance, the Effective Number of Parameters, and the Comparison of Arbitrarily Complex Models, Research Report 98-009. Division of Biostatistics, University of Minnesota.

Spiegelhalter, D.J., A. Thomas, and N.G. Best. 2000. *WinBUGS Version 1.3 User Manual.* Cambridge, UK: MRC Biostatistics Unit. Available at http://www.mrc-cam.ac.uk/bugs.

Spiegelhalter, D.J., N. Best, B.P. Carlin, and A. Linde. 2002. Bayesian Measure of Model Complexity and Fit. *Journal of Royal Statistical Society* 64(3):1–34.

Sun, D., R.K. Tsutakawa, H. Kim, and Z. He. 2000. Spatio-Temporal Interaction with Disease Mapping. *Statistics in Medicine* 19:2015–2035.

Transportation Research Board (TRB). 1987. *Designing Safer Roads: Practices for Resurfacing, Restoration, and Rehabilitation,* Special Report 214. Washington, DC: National Research Council.

U.S. Department of Transportation (USDOT), Bureau of Transportation Statistics (BTS). 1996–1999. *Transportation Statistics Annual Report.* Washington, DC.

____. 2002. Safety in Numbers Conferences. Available at http://www.bts.gov/sdi/conferences, as of August 2003.

Xia, H., B.P. Carlin, and L.A. Waller. 1997. Hierarchical Models for Mapping Ohio Lung Cancer Rates. *Environmetrics* 8:107–120.

Yang, M., J. Rasbash, H. Goldstein, and M. Barbosa. 1999. MLwiN Macros for Advanced Multilevel Modelling, Version 2.0: Multilevel Models Project. Institute of Education, University of London. Available at http://www.ioe.ac.uk/multilevel/, as of December 1999.

Zhu, L. and B.P. Carlin. 1999. *Comparing Hierarchical Models for Spatio-Temporally Misaligned Data Using the DIC Criterion,* Research Report 99-006. Division of Biostatistics, University of Minnesota.

Author Addresses: Corresponding author: Shaw-Pin Miaou, Research Scientist, Texas Transportation Institute, Texas A&M University System, 3135 TAMU, College Station, TX 77843-3135. Email: s-miaou@tamu.edu.

Joon Jin Song, Research Assistant, Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143. Email: j-song@ttimail.tamu.edu.

Bani K. Mallick, Professor, Department of Statistics, Texas A&M University, 3143 TAMU, College Station, TX 77843-3143. Email: bmallick@stat.tamu.edu.

Development of a Random Sampling Procedure for Local Road Traffic Count Locations

SARAH T. BOWLING
JWA/HMB Indiana, LLC

LISA AULTMAN-HALL
University of Connecticut

ABSTRACT

Traffic counting programs traditionally designed for traffic management require some re-thinking in order to provide accurate estimates of daily vehicle-miles traveled (DVMT) by road class for air quality planning. Predicting DVMT usually involves traffic counting at random points along highways, but local/minor roads, despite their extensive mileage, are not routinely counted.

We present a procedure to determine random count locations on functionally local roads. We used a geographic information system (GIS)-generated grid to cut roads into point-like sections from which we drew a random sample. The advantages of this procedure are that it overcomes GIS local road database limitations, uses standard GIS functions, and generates output that can be directly mapped for field crews. Cutting roads into various sizes and shapes introduced some bias during this process. A weighting procedure based on 750 local road counts in Kentucky measured the effect of the bias (which was deemed minimal and is therefore not needed in the application). Our experience using the sampling procedures allows us to recommend grid sizes that take into account computer processing time and file size limitations while limiting bias and ensuring acceptable randomness.

KEYWORDS: geographic information systems, sampling, traffic counting, vehicle-miles traveled.

INTRODUCTION

Traditionally, transportation agencies conduct routine traffic volume counts on higher volume highway corridors. However, local roads[1] are also important and unique, because they account for a considerable amount of total roadway mileage. For example, local roads make up 67% of the total roadway mileage in Kentucky, the study area for this project (CKTC 1997). Because traffic counts are typically conducted on local roads only for events such as road improvement projects and specific developments, the counts are not random and thus cannot provide accurate estimates of total travel on this class of road.

In September 1998, the need for estimates of the overall travel on local roads was bolstered by a U.S. Environmental Protection Agency (EPA) mandate requiring 22 states and the District of Columbia to submit state implementation plans relating to the transport of ozone across state lines (USEPA 1998). Oxides of nitrogen (NO_x) form ozone, or smog, which can negatively affect the environment and human health (e.g., damaged vegetation, water quality deterioration, acid rain, and respiratory and heart disease). Sources of NO_x emissions include motor vehicles and electric utilities. EPA requires state agencies to provide daily vehicle-miles traveled (DVMT) by land-use classification, road type, and vehicle type in order to estimate the amount of vehicle emissions produced on the county level.

DVMT is most commonly estimated from average 24-hour traffic counts at points along roads or a subset of roads. To obtain the DVMT, the traffic count is adjusted for daily and seasonal factors and then multiplied by the length of the road section. For example, if 1,000 vehicles a day travel a 2-mile section of road, the DVMT is estimated to be 2,000 vehicle-miles. Likewise, if there are a total of 100 miles of a particular road class in a county and the mean of a number of random traffic counts is 40,000 vehicles per day, then the countywide DVMT estimate is 4 million vehicle-miles for that class of roads. DVMT estimated from the existing nonrandom local road counts and total mileage would overestimate DVMT given that the more heavily traveled local roads are the ones more often counted.

When air quality, as opposed to traffic management, is the focus of DVMT and traffic count efforts, random locations must be chosen. One common source of random traffic counts is the Highway Performance Monitoring System (HPMS) established in 1978 by the Federal Highway Administration (FHWA). Data in the HPMS provide current statistics on the condition, use, operating characteristics, and performance of the nation's major highways. This travel information is routinely available for major highway systems, statewide and nationally, and is useful for the estimation of DVMT.

In Kentucky and elsewhere, HPMS data are used for estimating the total DVMT for the entire arterial and collector road systems, even though the sample is not completely random. In order to get the HPMS sample, each state had to break the arterial and collector routes into logical roadway sections. Rural section lengths were to range from 3 to 10 miles while attempting to ensure homogeneous traffic sections. Similarly, urban access-controlled facility sections were not to exceed five miles. All other urban sections were to be between one and three miles. A random sample[2] was then taken from this total set of road sections, but did not include the National Highway System and major arterial roads that in theory have complete coverage (USDOT 2000). What made the sample nonrandom was the various section lengths and the fact that there were no instructions for selecting the point on the section to take the traffic count. Some agencies may have counted at the busiest point or at the center. Although some states count local roads as part of the HPMS, most do not.

It might seem easy to produce a spatially random sample by dividing the local roads into segments of a particular length (one-tenth of a mile is common for a number of purposes) and selecting a random sample from this database. However, local road

[1] In this paper, local roads are all public roads in the state of Kentucky classified as "functionally local" by the Kentucky Transportation Cabinet. These roads may be paved or unpaved but nearly all in this study area are paved. All local roads, regardless of the responsible jurisdiction, were included in this study (i.e., city- and county-maintained roads are included).

[2] Unless otherwise noted, "random sample" refers to a simple random sample as opposed to any sampling technique involving weights or resampling.

geographic information systems (GIS) databases from which sample locations would be drawn are less developed than those for more major roadways. Given a one-tenth of a mile section it would be necessary to attribute starting points, ending points, and mile point locations to every road segment in the database in order to produce maps of the count locations for field workers. Additionally, many local roads, especially in urban areas, are shorter than the segment length into which roads are normally divided. This makes discretizing the routes complicated. Roads shorter than the segment length would always be a single segment and would have a higher chance per unit length of being selected. If shorter or longer roads have systematically higher or lower traffic volumes, respectively, this would bias the DVMT estimate if the shorter length roads were more likely to be selected.

It would be useful to have a method for selecting random points on the roads directly or graphically using simple random procedures rather than depending on weighted or proportional sampling. In this case, the spatial procedure is analogous to throwing a dart at a map blindfolded and counting at the road location that the dart hit. Moving from counting traffic on homogeneous traffic road sections to counting traffic at random points represents a fundamental change in philosophy and is consistent with the idea that traffic volume changes from point to point at driveways and intersections. Because of the variety of land uses on local roads the nonhomogeneity of traffic is particularly problematic.

The objective of this study is to develop a GIS-based random sampling procedure to determine count locations as random points on functionally local roads. A total of 750 24-hour local road counts were taken during this study in order to evaluate the sample properties resulting from the procedure. The large sample allowed us to analyze the bias issues resulting from 1) the grid-based nature of the procedure, 2) the shorter length of some local roads, and 3) the various directions or curves of individual roads. In application, much smaller sample sizes are likely to be used.

The following section describes other efforts to estimate DVMT on local roads. The remainder of the paper describes the GIS grid-based procedure and the evaluation of the bias it creates. The results of the bias analysis are presented along with a description of a procedure to correct for the sampling bias. However, the sampling bias was considered small enough to recommend use of the straightforward sampling procedure without the more complicated bias correction procedure.

OTHER EFFORTS TO ESTIMATE LOCAL ROAD DVMT

Programs in several states estimate overall travel on local roads through random samples. For example, Tennessee takes counts on local roads for specific highway projects, railroad crossing studies, and intersection analysis, although the count locations are not typically selected randomly. Because of this, the Tennessee Department of Transportation (TDOT) sought other methods to get a random sample of count locations (Crouch et al. 2001). Their study analyzed a program that collects traffic count information for all bridges in the state with a span length of 24 feet or greater.

Crouch et al. (2001) proposed a method to measure the randomness of these bridge counts for DVMT estimation on rural local roads. The traffic counts at bridge locations were compared with a random sample of traffic counts at nonbridge locations on local roads in eight counties. The researchers developed the procedure used to collect the random sample for nonbridge locations. Each of the eight counties was divided into four-square-mile grids (the width and length were two miles), and a process of repeated systematic sampling was used.

First, the grids throughout the county were sampled. Then, within each grid, the location of the actual count was chosen by randomly selecting x and y coordinates. Each grid cell consisted of a 10 by 10 matrix. From the randomly selected coordinates, the closest local road location was selected, and at this location, a traffic count was collected by TDOT. This is indeed a random procedure with one possible bias: shorter roads may be less likely to be closest to the 0.2 mile by 0.2 mile grid selected. When working with a large number of counties, the process could be labor intensive and time consuming. Using the random counts generated in this manner, the researchers found the bridge counts were not a representative sample of all rural local roads in each county.

In a California study (Niemeier et al. 1999), vehicle-miles traveled on dead-end unpaved roads were estimated from a random sample. Traffic counts were collected at random unpaved local road access points to paved roads. Because counting was conducted at the access points to prevent trespassing on the private roads, researchers did not have to deal with the issue of selecting the point along a road and, thus, a random sample of whole roads was taken. The count locations were mapped using GIS so the sites could be easily found. The count provided an estimate of the number of trips generated on the unpaved road, which was converted into DVMT by assuming there was a single destination on the road and each vehicle entering or exiting the road traveled half the length of the segment. However, the assumption that the vehicle is traveling to or from the midpoint of the road may cause the DVMT to be incorrectly estimated. For example, dead-end, unpaved local roads could have one origin/destination point at the end of the road. This method is random, but it is only suitable for local roads that dead end and have very few origin/destination points.

As part of this research study, an email survey of 45 states was conducted using contact names provided by the FHWA division office. The 29 replies indicated various methods for obtaining local road volume counts and sample locations. In Oregon, locations are picked from a select group of local roads that a computer program indicates are undersampled. The most recent counts from the local roads that are frequently sampled are then added to the counts of the sampled roads. The total sample may be nonrandom because the frequently sampled local roads are usually selected based on where road improvement projects will be located, developments built, or traffic problems exist. These are historically the more highly traveled areas. The random sample of the undersampled road segments is built by aggregating the full dataset as if it were one continuous road. Microsoft Excel then randomly picks a mile point along the road segments, and each pick becomes a location for a traffic count. The urban sample segments are 0.1 miles in length, while the rural sample segments are 1 mile. The count is taken at the center of the segment.

Other states provided less detailed input in the email survey. Vermont, for instance, selects the most "important" local roads for the counts. This, of course, is not random. West Virginia does not sample roads that have an average daily traffic value of less than 50 vehicles per day. This nonrandom method would certainly cause the DVMT to be inflated if total road length were used for the estimate. In Wisconsin, local roads are counted for special reasons, such as a traffic problem or new development. Again, this is not a random sample and, therefore, the DVMT estimate for EPA purposes could be incorrect. Wisconsin proposed developing a random sample of locations on local roads, but costs were prohibitive.

Until recently, DVMT estimates were mainly used to determine if a road needed improvements or expansion. Now that EPA requires DVMT to predict total vehicle emissions for each county, accurate estimates are much more important. The formerly sufficient nonrandom sampling methods used by many states are no longer adequate. Clearly, the need exists for a random sampling procedure that is not extremely labor-intensive in order to count locations to be used for estimating the DVMT on all functionally local roads.

GIS GRID-BASED SAMPLING METHODOLOGY

Challenges of Finding a Methodology

Location and alignment information for roads in most jurisdictions is usually stored in GIS databases, and sampling from these databases is desirable. In addition, because maps are useful to direct field workers to count locations it is logical to proceed with a GIS-based method. Roadways stored in a GIS are usually divided into segments (and, therefore, individual GIS features) at all intersections and many other points, some unsystematic.

In the road databases for the three Kentucky counties in this study, local road segments ranged in length from a few feet to 10 miles. ArcView, a

Windows-based GIS produced by the Environmental Systems Research Institute (ESRI), has a built-in function that can select a random set of such features or, in this case, segments. However, a random sample taken from this form of road database would not be appropriate for several reasons. First, the exact location on the road must be chosen and more than one location on the same road segment must have the opportunity to be chosen. The reasoning for this is based on the nonuniform variation in traffic volume along a road segment, especially for longer local roads where different intersecting roads and land uses affect traffic levels. Another reason the sample could not be taken from this database is that short and long segments would have been weighted equally. If the sample were taken from the existing GIS line theme, the precise location on the selected segment would then have to be subsequently chosen, and thus an individual point on a short segment would have a greater opportunity of being selected than a point on a longer segment. Therefore, equal weighting is not desirable.

There are other reasons why weighting is not a good method in our process. First, it adds two extra steps to the sampling procedure, which is intended to be straightforward. The length of each section would have to be determined to be used as weights. This might require GIS spatial analysis with poorer quality GIS databases. After segments were selected, another sampling procedure would be required to choose the random point along the given road segment. Second, weighted random sampling cannot be undertaken with built-in functions in most GIS programs requiring data to be transferred between programs.

As discussed in the introduction, another logical approach to developing the random sample would involve picking a random mile point or distance measure along these roads and then mapping it for the people conducting the counts. Knowing the length of every local road in a particular county, a line or row in a spreadsheet program could represent each one-tenth of a mile section. Most spreadsheet programs are capable of taking a random sample from the whole set. However, once the sample is taken it is difficult to direct the people making the traffic counts to the count location. On local roads, there are typically no mile markers to indicate location as there are with more major or higher volume roads. Maps of count locations made in ArcView could solve this problem. However, limitations in the coding of local road databases present a further difficulty for this mapping.

Mapping a specific point on a road is very easy with GIS road databases with a feature called "dynamic segmentation." Using dynamic segmentation, every road segment has two "special" attributes. One indicates the beginning linear reference marker at the start of the segment and the second indicates the end reference. The GIS can then locate any mile point on the road segment based on this information. This allows the mile-point reference system to span across adjacent segments. For example, the system could span across an intersection. However, the available GIS databases for local roads rarely contain dynamic segmentation. Therefore, use of a sampling procedure that required start and end mile points to allow mapping would become very labor-intensive.

As an alternative to creating dynamic segmentation attributes in the database, each individual road segment (as opposed to the whole road) could be coded automatically with a starting mile point of zero and an ending mile point of its length. However, using discrete mile-point demarcations, such as one-tenth in the spreadsheet listing, and random sampling presents another problem for very short local roads, especially in urban areas. Selection of a random continuous number between zero and each segment's length would be necessary in a two-stage process like that used in Tennessee. In the first stage, a weighted random sample with replacement, with probability proportional to road segment length, would be taken. In the second stage, a point or points along the segment would be selected by random number generation. This procedure would require separate programming outside the GIS, and the results would necessitate subsequent transfer back into the GIS for mapping because mile points are not meaningful on a segment-by-segment basis or on local roads without field mile-point markers.

The new methodology proposed here is also two stage but uses standard built-in functions of the typical GIS: grid generation, database intersection, and random sampling from a feature table. The product is already a line feature in the database and is immediately mapped. In the first stage, a GIS grid is generated and used to cut road segments into sections. As the grid size becomes smaller, the sections become more point-like, enabling a new theme from which the random sample can be drawn using the direct built-in random sample command. This avoids the use of any weighting or resampling. The procedure ensures that the sample locations are spread randomly throughout the study area and that each point-like section along all roads has an equal chance of being in the sample regardless of the total length of the road.

Creating the Point-Like Sections for Three Study Areas

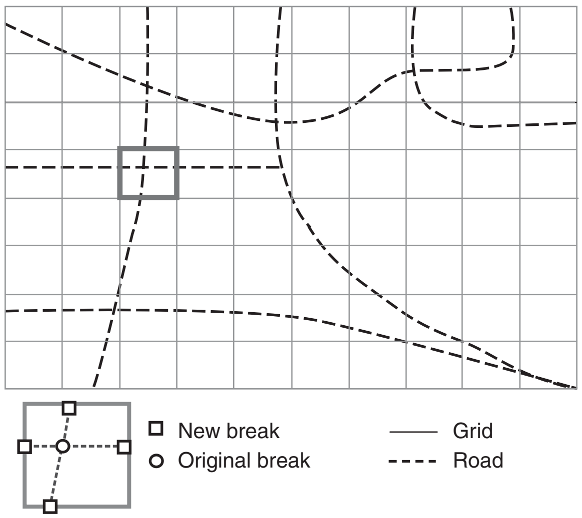
In this case, the primary GIS used was ArcView. We developed a procedure that cut the roads into small sections using a grid; thus, the shape and density of the local roads were considered potentially influencing and affected the selection of study areas. Because it was not feasible to include all 120 Kentucky counties, we chose three counties for this study: Henderson, Pike, and Fayette. In total, the Kentucky Transportation Cabinet agreed to count up to 750 locations in these 3 counties for analysis of the sampling strategy. The counts were performed by a state contractor using "tube style" Peek Automatic Data Recorders (ADR-1000) between fall 1999 and spring 2000. Counts were taken for 24 to 48 hours and adjusted for season and day of the week using factors developed with historic counts by the Kentucky Transportation Cabinet. No axle counts or adjustments for heavy vehicles were undertaken. This large number of counts was not expected to be routine but was undertaken to address the issue of variability in local road volumes in order to design future counting programs.

The three counties for the chosen sampling procedure are very different from one another. Henderson County (440 square miles or 1,140 km^2) is in the western part of the state where the flat plain topography results in gridlike roads (total of 601 miles or 968 km of local road). It includes the small city of Henderson, which has a population of approximately 27,000. Pike County (788 square miles or 2,041 km^2) is in the eastern, mountainous part of the state, has winding and curvy roads, and is considered a relatively rural county (total of 829 miles or 1,335 km of local roads). Fayette County (284 square miles or 736 km^2), with a population of approximately 250,000, represents an urban county with a dense road network (total of 734 miles or 1,182 km of local roads). The separate GIS themes for state-maintained, county-maintained, and city-maintained local roads were combined for the three test counties to obtain three local road GIS databases. All GIS local road databases were developed and maintained by the Kentucky Transportation Cabinet.

Unfortunately, ArcView does not have the capability to create a grid (a set of adjacent polygon squares covering a certain area or extent), so grids were created in ArcInfo (a compatible ESRI GIS) by specifying the extent of the area and the grid size. These grids can be used directly in ArcView. Using the intersection function in ArcView, a "cookie cutter" grid shows, for example, in the designated square in figure 1, that the roads in the square are now in four separate pieces or features. Each separate, tiny line feature in the output database has a record in the attribute table from which ArcView's sampling script draws the random sample. Note that the random point-like road segments are selected, not the squares. Therefore, there is no need to select the road segment within a given selected cell as was done in some past procedures.

One obstacle of the grid approach is that some bias can be introduced by virtue of the point-like segments not being of equal length, as illustrated in figure 1. The grid used to cut the roads into small sections was orthogonal, so the roads were cut at different angles. As a result, some of the sections were considerably longer than others. If you have two roads of equal length, one cut into several short pieces and the other cut into a few long pieces, then the road cut into several short pieces will have a greater chance of being selected in the random sample. Given that the local road traffic volume was found to correlate with the original road segment length and also with differences in rural and urban areas, in order to avoid bias, the number of segments into which a particular road was divided would have to be directly proportional to the length

FIGURE 1 A "Cookie Cutter" Grid on a Network of Roads



of that road. This means that a road with twice the length of another road should be divided into twice the number of sections.

Our objective then is to determine the size of the largest grid square that brings an acceptably low bias to the sample. As the grid size approaches zero, the point-like sections approach true points of zero length, which present absolutely no bias. The smaller the grid square size, the more computer space and time are needed for the spatial analysis that cuts the road segments. The three counties were analyzed with 0.2 mile, 0.15 mile, 0.1 mile, and 0.05 mile grid square sizes. Although space issues needed to be considered (the grid for one county at the 0.05 mile size was 148 MB) in choosing the final grid square size, the computing time and ability of a personal computer to do the intersection (cutting) without crashing were certainly critical issues.

CONSIDERATION OF BIAS IN THE POINT-LIKE SECTIONS

In order to compare grid sizes and determine if the straightforward sampling procedure could be used without a more complicated weighting procedure to correct for the bias, it was necessary to develop a method to measure the bias that would be present in an average traffic count from a sample drawn using this process. Once the road segments were cut by the grid, the length of the original road section and the number of point-like segments into which it was divided were available for use in measuring bias. Figure 2 illustrates these data for one 0.2 mile grid in Pike County (lines and equations on this figure are described below).

The first of several indicators of bias considered was the coefficient on the x^2 variable in the equation for the best-fit quadratic curve. This curve is not represented on the figure but has the form

$$y = a + bx + cx^2$$

where a, b and c are parameter coefficients,

x is the original road segment length,

y is the number of segments into which the road is cut.

The value of the coefficient on the x^2 variable is an indication of the curvature of the line, and increasing values of the coefficient would indicate bias. A negative value would indicate that the line curved downward, specifying that the longer roads were being cut into relatively fewer pieces and were therefore underrepresented in the sample. A positive value would denote the opposite: longer roads were overrepresented in the sample. The magnitude of the coefficient for the x^2 term also provides an indication of whether it is appropriate to proceed using a linear regression-based representation of the relationship between road length and number of point-like segments.

A bias analysis graph and equation such as that in figure 2 was generated for each county and grid size analyzed. The coefficients on the x^2 variable in the equation for the best-fit quadratic line as generated by Microsoft Excel are shown in table 1. Within an individual county, the value of the coefficient varies. This alone is not insightful; it is the comparison between counties that provides useful information. The magnitude of the coefficient is substantially greater for Fayette County than it is for Henderson and Pike Counties, showing that the grid process works better for rural roads than for urban roads because they are longer and less dense. We considered the low magnitude of these coefficients to be the justification to proceed with representing the relationship with a linear equation.

However, it is important to note that bias could still exist even in a linear relationship (x^2 coefficient = zero). Therefore, we undertook further consideration of the

FIGURE 2 Bias Analysis for Pike County
(0.2 mile grid)

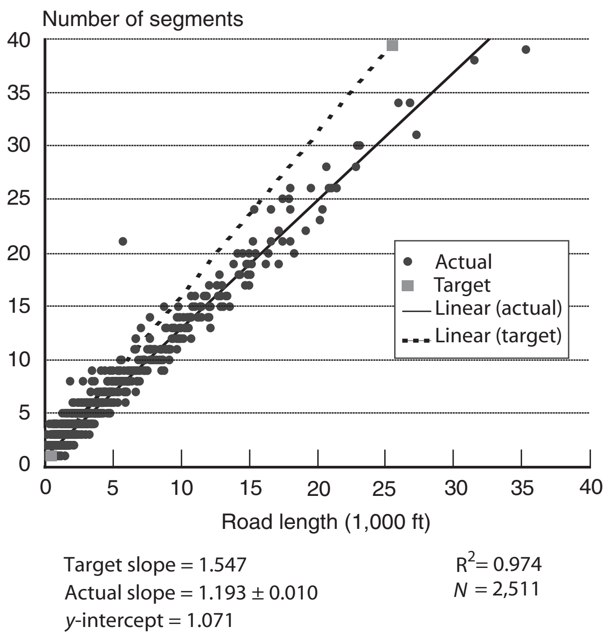


TABLE 1 y-Intercept, R^2, and x^2 Coefficient

	Coefficient on the x^2 variable	y-intercept (linear)	R^2 (linear)
Pike County			
0.20 mile grid	**–0.0007**	1.071	0.97
0.15 mile grid	–0.0005	1.066	0.98
0.10 mile grid	0.00001	1.023	0.99
0.05 mile grid	**–0.0016**	1.026	0.99
Henderson County			
0.20 mile grid	–0.0005	1.049	0.98
0.15 mile grid	**–0.0011**	1.110	0.98
0.10 mile grid	–0.0001	1.038	0.99
0.05 mile grid	–0.0013	1.036	0.99
Fayette County			
0.20 mile grid	**–0.0054**	1.005	0.74
0.15 mile grid	**–0.0120**	1.012	0.81
0.10 mile grid	**–0.0073**	1.016	0.90
0.05 mile grid	**–0.0123**	1.032	0.97

Note: Values in bold are statistically significant at the 0.05 level.

linear regression equation. One factor considered in measuring this bias was the y-intercept of the best-fit line. On one hand, this value would ideally seem to be zero, because a road of zero length should be divided into zero sections. However, a y-intercept of one would indicate that a road of very small length was divided into one section, meaning that very short roads will be automatically overrepresented in the sample. As evident in figure 2, some very short roads were divided into up to three or four segments. Table 1 shows that the y-intercept value did not vary significantly as the grid size changed. For all counties and grid sizes, the y-intercept hovered just above one, which is expected because very short segments would most often be cut into one piece or, at most, two pieces. This result illustrates that some bias will be present with all grid sizes given that short segments are overrepresented.

The line indicating no sampling bias due to road length would be expected to have a certain slope, referred to here as the "target slope." The target slope is obtained by dividing the total number of segments in a county by the total length of local roadway in that county. For example, if there are 5 million distance units of local road in a particular county, and a specific grid size cuts these roads into 7,000 segments, the segments should be on average 714.29 distance units (i.e., 5 million distance units/7,000 segments) long. The target slope is the inverse of this number (divided by 1,000 for the graph scale shown), and the line on figure 2 was derived by using this slope with a y-intercept of 1.

Comparison of the target slope to the actual slope first required consideration of the R^2 value. The R^2 values shown in table 1 indicate that both the sampling procedure and the weighting procedure described below, which is based on the linear slope, are better suited to non-urban areas. The variation in the number of segments decreases with the smaller grid square sizes, as expected. However, the relatively high overall R^2 values indicate that the best-fit line does indeed represent the data well, adding legitimacy to the comparison of the actual and target slopes described below.

Table 2 includes the target slope, the actual slope of the best-fit line, and the percentage difference between its slope and the target slope. The range included with the slope is the 95% confidence interval. The confidence interval was inspected for the inclusion of the target slope. None of the target slopes were included in the 95% confidence interval, indicating bias was present.

TABLE 2 Slope Comparison

	Target slope	Actual slope	Error (percent)
Pike County			
0.20 mile grid	1.547	1.193 ± 0.0100	22.9
0.15 mile grid	1.945	1.593 ± 0.0103	18.1
0.10 mile grid	2.752	2.414 ± 0.0116	12.3
0.05 mile grid	5.182	4.841 ± 0.0137	6.6
Henderson County			
0.20 mile grid	1.552	1.260 ± 0.0120	18.8
0.15 mile grid	1.977	1.668 ± 0.0143	15.6
0.10 mile grid	2.802	2.512 ± 0.0161	10.3
0.05 mile grid	5.296	5.009 ± 0.0245	5.4
Fayette County			
0.20 mile grid	3.029	1.228 ± 0.0170	59.5
0.15 mile grid	3.441	1.629 ± 0.0183	52.7
0.10 mile grid	4.258	2.441 ± 0.0190	42.7
0.05 mile grid	6.754	4.908 ± 0.0206	27.3

In each county the percent error between the target slope and the actual slope decreased as the grid square size approached zero, as expected. The target slopes are greater than the actual slopes, indicating that as road length increases the road becomes underrepresented in the sample. Fayette County had percent errors greater than that for the other two counties, again indicating that less dense roads are better suited to the grid process. Henderson County's grid-like roads had smaller errors than Pike County where roads are curvier. Therefore, it can be inferred that the grid procedure works best for grid-like roads and rural roads. The grid size is more crucial in urban areas.

In order to consider the impact of the bias due to road length and the grid procedure, weights were developed based on slope comparison; these weights were then applied to the traffic counts for these three counties. Counts were performed during calendar year 2000 at points selected using the 0.2 mile grid procedure (a worst-case scenario). The number of 24-hour counts performed in Henderson, Pike, and Fayette counties were 164, 243, and 337, respectively. These totals were designed so that the number of counts in each county were proportional to the length of local roads but also ensured a minimum number of rural and urban counts in each county (this constraint was imposed by the Transportation Cabinet). Counts were corrected for seasonal and weekly factors using constants developed in Kentucky based on counts on all functionally classed roads over many years.

The best-fit line and the target line were known for each county for the 0.2 mile grid size. In other words, for a road of a particular length, the number of segments into which it was divided and the number of segments into which it should have been divided were known. The weight was calculated as the ratio of the number of segments into which the road of a given length should have been divided if no bias by road length existed and the actual average number of segments into which the road was divided. This weight varied by road length as illustrated in figure 3 for Pike County for all grid sizes. We calculated a weighted average for the 24-hour traffic count, or average daily traffic (ADT) using the weights for the 0.2 mile grid size.

FIGURE 3 Pike County Weights

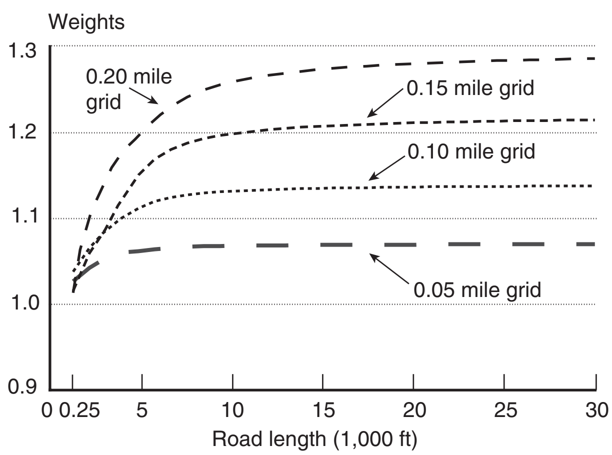


Table 3 presents the sampled and weighted ADT and the subsequent sampled and weighted DVMT estimate for local roads in each county based on the 0.2 mile grid process. The table demonstrates that without the weighted ADT, the DVMT estimate for each county would be slightly overestimated, with the greatest difference in Fayette County. This is further evidence that the weighting procedure is important for urban areas but is also a function of the greater number of shorter roads in those areas.

TABLE 3 Corrected and Uncorrected ADT and DVMT Values
(0.2 mile grid-based sample)

	Number of counts	Average[1] ADT (veh/day)	Average weighted[2] ADT (veh/day)	DVMT estimate (veh-miles)	Weighted DVMT estimate (veh-miles)
Pike County	243	454.87	453.18	377,232.79	375,831.24
Henderson County	164	386.27	367.59	232,105.78	220,881.16
Fayette County	337	747.06	719.56	548,177.69	527,998.74

[1] Straight arithmetic mean.
[2] Mean is weighted by the ratio of target and actual slopes from the regression analysis.
Key: ADT = average daily traffic; DVMT = daily vehicle-miles traveled; veh/day = vehicles per day; veh-miles = vehicle-miles.

However, the percentage difference due to the sampling bias is small and deemed acceptably low for modeling purposes for either the planning or air quality considerations described at the beginning of this paper. Based on the slope comparison the bias would be even less with the smaller grid sizes. It would not be useful to undertake the multistage weighting procedure calculations.

CONCLUSIONS

In summary, we developed and validated a straightforward sampling procedure that will allow random sampling of traffic count locations on extensive local road systems. Because built-in GIS commands can be used, sampling does not require time-intensive processes and the results can be directly mapped for field use. The procedure offers a means to determine not only a random road but also the point along the road where counting should occur. Furthermore, the procedure can handle very short local roads without greatly biasing the sample.

The analysis presented here provides guidance for determining a recommended grid size for use in sampling that takes into account computer capabilities in terms of file size and processing time while ensuring acceptable randomness of sampling. Attempts to use grid sizes below 0.05 miles were not successful in ArcView for the study areas used. Although individuals should select a grid square size based on their computer processing capabilities and the characteristics of the roads in their study, these results indicated that a larger grid size can be used for rural roads and grid-like roads. The grid square size needs to be smaller for urban counties due to the dense, short roads. Because it is very difficult to work with the 0.05 mile grid square size, the 0.1 mile size is recommended for urban counties. The recommendation for rural counties is to use the smallest grid square size feasible, but a 0.2 mile size would be sufficient, especially if roads are in a grid-like pattern.

Although not directly related to the main topic of this paper, several observations can be made regarding traffic counts on local roads and the estimation of accurate countywide DVMT. The state of Kentucky undertook a significant number of 24- to 48-hour local road traffic counts for this project, which is a very unusual and expensive undertaking, particularly for local roads. A total of 3,800 counts were obtained (including the 750 used in this sample procedure research). The counts had extraordinarily high standard deviations (386 for 2,702 counts in rural areas and 1,323 for 1,099 counts in urban areas), suggesting that sample sizes beyond those realistically possible would be necessary to obtain average counts with reasonable confidence intervals. Further disaggregation of roads beyond simple use of the functional classification system will be necessary before any reasonable traffic data-collection plan can be undertaken by states for EPA travel estimations. For this reason, we recommend that the next stage of research be to apply the sampling procedure to higher functional class roads where it might decrease the total number of counts required. If tests were conducted on the National Highway System where the HPMS provides near universal coverage, valuable sample size recommendations might be possible.

ACKNOWLEDGMENTS

The authors would like to thank the Kentucky Transportation Cabinet for funding this research as well as members of the University of Kentucky Transportation Center and Wilbur Smith Associates for their contributions.

REFERENCES

Commonwealth of Kentucky Transportation Cabinet (CKTC), Division of Transportation Planning. 1997. *Traffic Characteristics of Kentucky Highways*. Frankfort, KY.

Crouch, J.A., W.L. Seaver, and A. Chatterjee. 2001. Estimation of Traffic Volumes on Rural Local Roads in Tennessee, presented at the Transportation Research Board 80th Annual Meeting, Washington, DC.

Niemeier, D., J. Morey, J. Franklin, T. Limanond, and K. Lakshminarayanan. 1999. *An Exploratory Study: A New Methodology for Estimating Unpaved Road Miles and Vehicle Activity on Unpaved Roads,* RR-99-2. Davis, CA: Institute of Transportation Studies.

U.S. Department of Transportation (USDOT), Federal Highway Administration. 2000. *Highway Performance Monitoring System Field Manual,* Chapter 7: Sample Selection and Maintenance. Available at http://www.fhwa.dot.gov/ohim/hpmsmanl/hpms.htm, as of May 28, 2003.

U.S. Environmental Protection Agency (USEPA), Office of Air Quality Planning and Standards. 1998. *The Regional Transport of Ozone: New EPA Rulemaking on Nitrogen Oxide Emissions*. Research Triangle Park, NC.

Authors' addresses: Sarah T. Bowling, JWA/HMB Indiana, LLC, 624 W. Main St., Suite 300, Louisville, KY 40202.

Corresponding author: Lisa Aultman-Hall, Associate Professor, Department of Civil and Environmental Engineering, University of Connecticut, Unit 2037 261 Glenbrook Road, Storrs, CT 06269-2037. Email: aultman@engr.uconn.edu.

Containerized Cargo Shipper's Behavior in China: A Discrete Choice Analysis

PIYUSH TIWARI
Infrastructure Development
Finance Co., Ltd.

HIDEKAZU ITOH
MASAYUKI DOI
University of Tsukuba

ABSTRACT

Shippers choose ports for export or import of goods based on a number of variables, including location, preferences for particular shipping line services, and facilities offered. A huge port infrastructure investment is necessary to attract shippers, and ports compete with each other for business. This paper models the port and shipping line choice behavior of shippers in China, using a shipper-level database obtained from a 1998 survey of containerized cargo shippers. We used a discrete choice model where each shipper chooses among 10 shipping line and port combinations and makes decisions based on various shipper and port characteristics. This paper incorporates the shipping line choice behavior through model specification by nesting the choices in a hierarchical fashion where shippers choose from Chinese and non-Chinese shipping lines and then from ports or vice versa. The results indicate that the distance of the shipper from the port, the number of ship calls at the port, the efficiency of the port infrastructure, and the number of routes offered at the port strongly influence decisions to use a port.

KEYWORDS: China, discrete choice model, port choice, shipping.

INTRODUCTION

Shippers choose shipping lines and ports through which their goods can be moved reliably and economically, particularly for containerized high value-added cargo. In an earlier study (Tiwari et al. 2003), we modeled the choice of ports and shipping lines by light manufacturing industrial cargo shippers in China. Our study emphasized the importance of joint modeling the choice of ports and shipping lines.

Policy changes in China during the last two decades (e.g., economic liberalization and trade, the unification of mainland China and Hong Kong, and membership in the World Trade Organization) have strengthened the role of ports in the economy. Ports and shipping lines play a key role in the export and import of goods. Imports constituted 17.5% of the Chinese gross domestic product (GDP) and exports were 21.9% of the GDP (US$) in 1998.

China has 16 major coastal ports (SSB Annual). Figure 1 shows the location of the key ports. The ports of Shanghai and Qingdao are relatively close to each other and are more likely to compete with one another than more distant ports. The ports of Tianjin and Dalian compete against one another as they are close to Beijing. In 1998, the port of Shanghai, the largest containerized cargo handler in China, handled 28.4% of this cargo, measured in 20-foot equivalent units (TEUs). Port shares for Dalian, Tianjin, and Qingdao in total TEUs were 4.9%, 9.5%, and 11.3%, respectively. Other ports in China handled about 45.9% of the total cargo in TEUs (Year Book House of China 1999; SSB Annual).

Chinese ports have started to compete among themselves, including traditionally busy ports such as Hong Kong. Ports in South China that are closer to cargo origins are growing by more than 40%, while Hong Kong, which depends on South China for about 80% of its cargo, is seeing only single-digit increases (Mongeluzzo 2002).

China, an emerging market economy, is in the process of formulating policies intended to make its ports more efficient. The port sector in China has undergone many changes during the last decade and is poised for significant gains in the future. Key areas being considered include strengthening of port infrastructure and issues related to the market entry or exit of ports and related services and shipping lines. It is essential to quantify the impact that changes in key port infrastructure or shipping line variables would have on the demand for port and shipping line services.

FIGURE 1 Location of Chinese Ports

This paper attempts to fill the gap in the literature on this topic by formulating a model of port and shipping line choice decisions by shippers in China. Specifically, the model seeks to specify and empirically estimate the underlying factors that influence the port and shipping line selection behavior of containerized cargo shippers in China. The model estimates market elasticities that measure changes in market shares of various ports and shipping lines due to changes in key infrastructure variables.

This study is important for two reasons. First, it supplements the currently limited amount of literature available on port and shipping line choice behavior of shippers, particularly for China. Second, as nations become more "global" and their industries are affected by the pressures of international competition, services must be provided on an internationally competitive basis. Thus, port authorities must understand the necessity of improving their services in order to respond to greater competition among ports and the growing pressure from shippers for lower port and shipping charges.

Ports form a vital link in the overall trading chain and, consequently, their level of utilization determines to a large extent their domestic and international competitiveness. In order to maintain a competitive edge in these markets, port authorities must understand the underlying factors that affect their competitiveness.

The objectives of this paper are to 1) estimate the demand for ports and shipping lines in China, 2) determine what factors affect the demand for ports and shipping lines, and 3) estimate the market elasticity parameters of demand. This analysis covers shippers from most of the Chinese coastal regions but excludes Guangdong and Hainan provinces, which ship most of their cargo from the ports of Hong Kong and Yantian, because these provinces are not covered by our dataset.

The paper is structured as follows. The next section briefly reviews the literature, followed by a section describing our model. Next, we present key statistics based on the study data, and the final two sections discuss the results and conclusions.

LITERATURE REVIEW

Earlier literature in this area focused on analyzing the performance of shipping lines or factors influencing the choice of ports. Our earlier study (Tiwari et al. 2003) hypothesized that shippers choose shipping lines and ports simultaneously. This is important because most earlier studies are based on the assumption that shippers' deal with forwarders and base their decisions only on service factors. The port and shipping line choices would then be made by forwarders on behalf of shippers. This is not always the case in Asia. Shippers base their logistic decisions on an overall cost-minimizing strategy. They choose shipping lines and ports so that their goods can be shipped economically and efficiently to market locations. Although our earlier study emphasized the need for direct modeling of shipping line and port choice by shippers, our empirical results only weakly supported the hypothesis. In this analysis, we include all types of containerized cargo to and from China (extending the Tiwari et al. database) and test a hypothesis similar to that proposed in our earlier study.

Prior research analyzing factors responsible for port and shipping line efficiency guided our choice of variables for this study. Network and scale economies affect shippers' choice of ports. Studies estimating port performance indicate that the number of ship calls is an important factor in determining the performance of a port. The speed at which cargo moves through a port influences its value.

Slack (1985) analyzed port end users and freight forwarders engaged in trans-Atlantic container trade between the United States and Europe and found that the number of sailings was the most important criterion for port choice. Bird's (1988a and 1988b) results from perception analysis of European freight forwarders indicate that the frequency of ship service is the main reason for port choice. Tongzon (1995) also confirmed that time is critical in freight forwarding and the frequency of shipping service is the major determinant of time.

Investigations of what influences the choice of shipping line have identified three categories of factors: route (e.g., frequency, capacity, convenience, directness, flexibility, and transit time), cost (freight rate and other costs), and service (delays, reliability and urgency, avoidance of damage, loss and theft, fast response to problems, cooperation between shipper and carrier, and documentation and tracing capability) (Gilmour 1976; McGinnis 1979; Ogden and Rattray 1982; Brooks 1985; Wilson et al. 1986; and Meyrick and D'Este 1989). These studies found that shippers are generally risk averse in their choice of shipping lines and thus have limited options. Service factors, particularly service frequency, take precedence over price (Bayliss and Edwards 1970; Brooks 1984, 1985; Wilson et al. 1986; Meyrick and D'Este 1989; Pearson 1980).

On the other hand, Suthiwartnarueput's (1988) study on the efficiency of the shipping industry in Thailand suggests that the most important service attribute is cost, followed by punctuality, transit times, frequency of sailings, directness of sailings, as well as past loss and damage experience. Jamaluddin (1995), with reference to the Far East/Europe trade, reported that the six service factors to which *shippers* attach the most importance were freight rate, cargo care and handling, knowledgeability, punctuality, transit time, and service frequency. Furthermore, the six service attributes to which *carriers* attach the most importance are knowledgeability, freight rate, cargo care and handling, punctuality, transit time, and service frequency.

Chiu (1996) evaluated the logistics performance of liner shipping in Taiwan and found that the most important service attributes for *shippers* were prompt response by the carrier to problems, transit time reliability, documentation services, notice of delay, and assistance with loss or damage claims from the carrier. For *carriers,* the five most important service attributes of carriers were transit time reliability, prompt response by them to problems, knowing the needs of the shippers, their own reputation, and knowledgeability of sales personnel.

Previous studies concluded that service attributes significantly influence distribution and logistics activities. Services such as transit time, frequency of service, reliability of delivery, speed of claims response, on-time pickup and delivery, as well as other aspects of physical distribution, are perceived as crucial by shippers.

MODEL AND VARIABLES

This paper extends our earlier analysis, which looked at shippers' logistics decisions for light manufacturing industrial cargo in China. That analysis was based on a multinomial logit (MNL) model in which alternatives available to shippers were combinations of ports and shipping lines. The methodology is briefly discussed below.

Shippers require a shipping line and port to move goods. In our model, because of data limitation, we considered only two groups of shipping lines (Chinese and non-Chinese). We assumed (a rather strong assumption but we explain our rationale later) that all service aspects of shipping lines are subsumed by grouping shipping lines this way. Shippers maximize profits by minimizing their cost of transportation and while doing so they will choose a combination of shipping line and port that is the most cost-effective in terms of the overall chain of production. This paper classifies choices as combinations of five ports and two shipping lines. The total number of choices available to a shipper is 10 (2 shipping lines x 5 ports) (figure 2). We estimated the probabilities of choosing alternative i (i = 1,…,10) using a discrete choice modeling framework.

The simplest and most convenient functional form for a discrete choice probability of alternative i is the standard MNL form (McFadden 1981).

$$'(i|N,\ Z,\ \beta) = \frac{\exp(Z_i\beta)}{\sum\limits_{N = 1,\ \dots,\ n}\exp(Z_N\beta)},$$

where

$N = \{1,\dots,n\}$ denotes the set of n discrete port-shipping line choices,

Z_i = a vector of K attributes specific to choice i, and

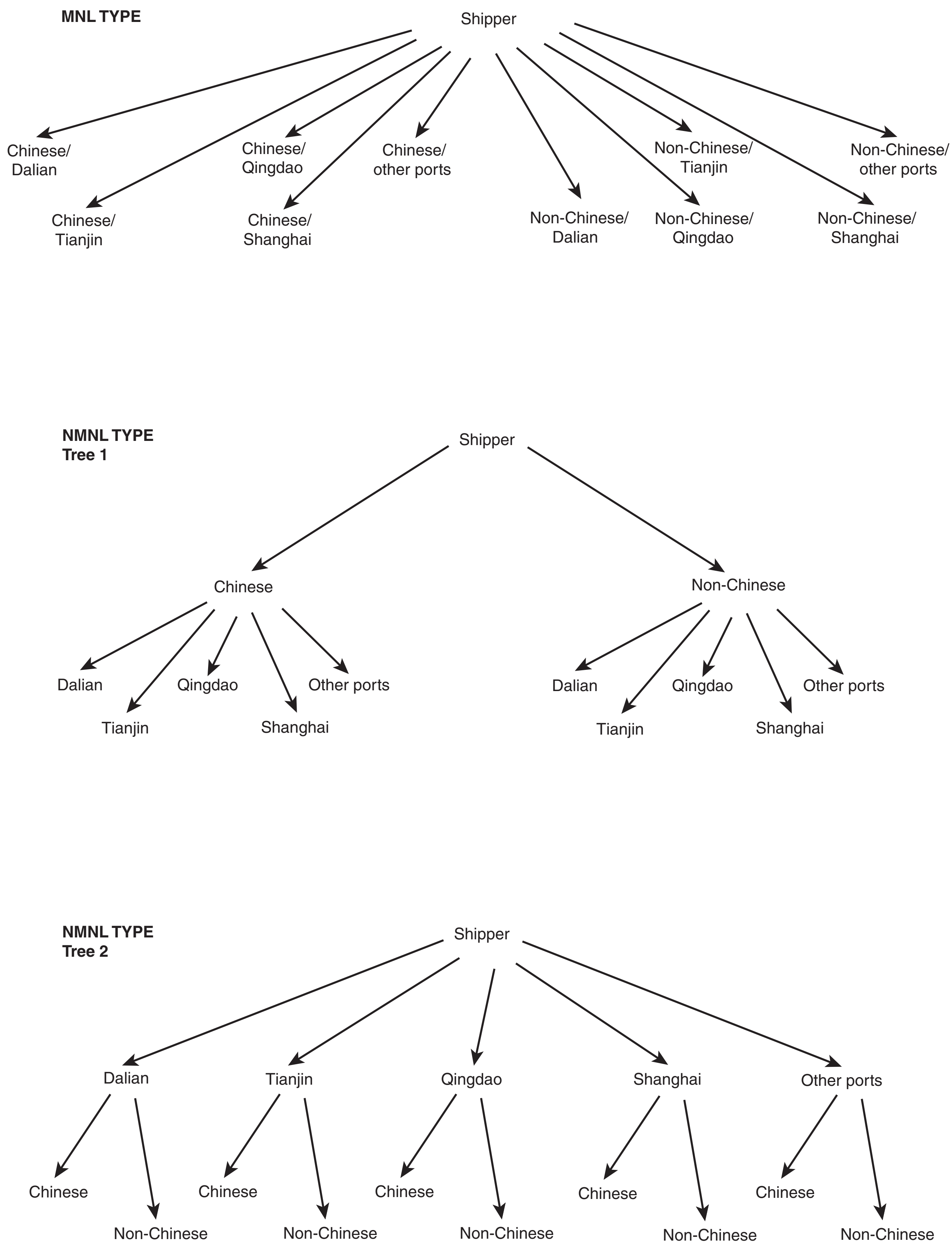
β = a vector of corresponding cost parameters.[1]

The simple MNL model is constrained by the Independence of Irrelevant Alternatives (IIA), which implies that the cross-elasticities of the probability shares must be equal (Boersch-Supan and Pitkin 1988).

McFadden (1981) generalized the MNL specification as the nested multinomial logit (NMNL) model based on a hierarchy that groups alternatives into subsets of similar choices. The choice within a cluster and the choices among the clusters within each nest are described by a conditional logit choice probability and conform to the IIA assumption. Following McFadden, this paper formulates an NMNL model with shipping lines and ports at dif-

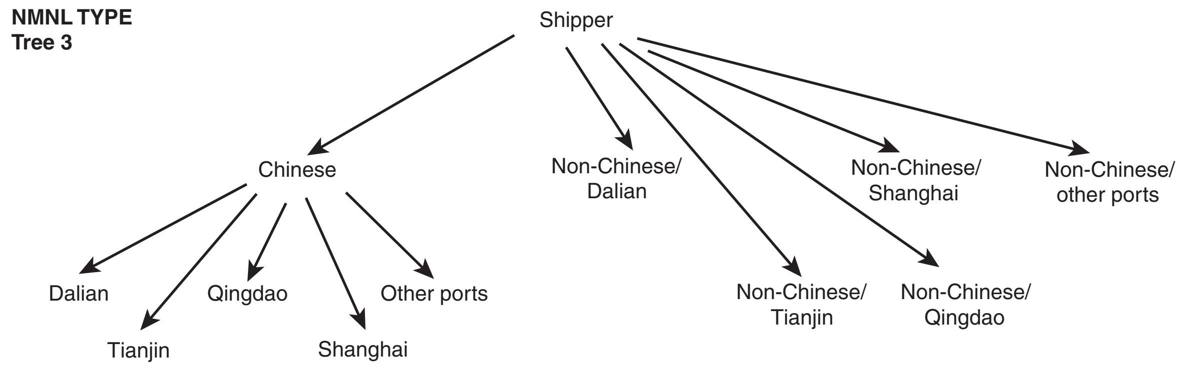
[1] The choice probability can be derived from cost minimization of shippers by defining $C_i^* = Z_i\beta + \varepsilon_i$ as the stochastic cost of port-shipping line choice i where ε_i follows a type I extreme value distribution (see McFadden 1981).

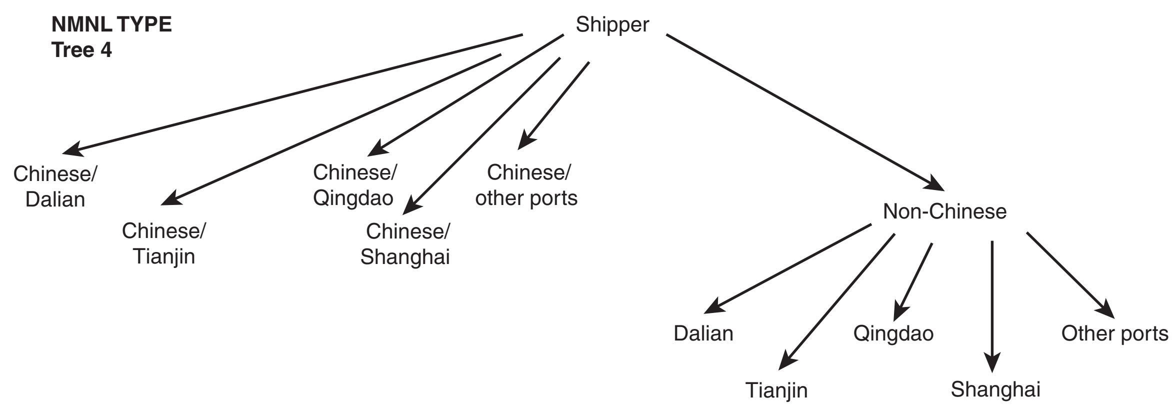
FIGURE 2 Alternatives in Shippers' Decision Process in China



(Figure 2 continues on next page)

FIGURE 2 Alternatives in Shipper's Decision Process in China (continued)





ferent levels of hierarchy. Figure 2 shows the various illustrative specifications of the NMNL models considered in this paper.

Mathematically, the probability of choosing alternative *ij* in the NMNL model is

$$P(ij) = P(i) \times P(j|i)$$

where

i = number of subsets (e.g., in figure 2, two subsets in NMNL Tree 1 and five subsets in NMNL Tree 2) and each subset has some or all of j = shipping lines-port combinations (some or all of the total of 10),

$P(i)$ = the marginal choice probability of subset i, and

$P(j|i)$ = the conditional probability of choosing alternative j from the alternatives included in subset i.

The conditional probabilities of choosing alternative j in subset i have the form of MNL choice probabilities

$$P(j|i) = \frac{\exp\{I(i) \times s(i)\}}{\sum_{j' = 1, \ldots, i} \exp\{I(j') \times s(j')\}}$$

The marginal choice probability is represented as

$$P(i) = \frac{\exp(Z\beta /\ s(i))}{\exp(I(i))}$$

where the "inclusive values" *I(i)* is defined by

$$I(i) = \log \sum_{j = 1}^{k-1} \exp(-Z\beta) /\ s(i)$$

are weighted by "similarity coefficients" $s(i)$. These similarity coefficients refer to their respective subsets and characterize the degree of substitutability among the alternatives in the subset. Values of similarity coefficients between zero and one are a measure of the importance of similarities and dissimilarities among choices. If, however, these values are equal to one, the trees reduce to simple MNL.

This paper estimates the probability of choosing alternative *i* based on the following variables.

Port Characteristics

Ship calls is an important variable. The sign for this variable is ambiguous, because, as explained earlier, an increase in the frequency of ship calls is an attractive variable for exporters and importers. A positive sign may be expected for this variable. However, if the ship call variable is already high for a port, increased congestion may result and have a negative effect.

Total TEUs handled at the port may represent the use of the port by shippers. However, a higher volume of TEUs may also represent congestion. We do not have any a priori notion about the sign of this variable.

A higher number of **TEUs per berth at the port** means that one berth caters to a large number of TEUs. As will be discussed in more detail in the Results section, we believe this may represent congestion at the berth and shippers are expected to react negatively to this variable.

TEUs of cargo per crane loaded and unloaded from a ship is also another time-related factor. As will be discussed in more detail in the Results section, we believe a higher value for this variable indicates efficiency in loading and unloading a ship's cargo. This variable is expected to influence the shippers' decisions positively.

The **usage factor** is defined as the handling volume (thousand tons) per length of quay. The expected sign for this factor is negative, because a higher value for this variable represents congestion.

The **number of routes offered** is another frequency of shipping-related variable. The greater the number of routes offered at a port, the quicker it is for shippers to move their goods to various destinations.

Port and loading charges is another variable that would affect the decisions of shippers. However, these charges are the same for all ports (Water Transportation Ministry 1997).

Shipping Line Characteristics

The literature in this area considers marketing to be a service factor. Advertising, frequency, and quality of service are crucial factors in determining the selection of a shipping line. Traditionally in China, forwarding and logistics have been tightly controlled. The entry of foreign companies has not been easy. Foreign companies can operate only if they have a Class A license from the Ministry of Foreign Trade and Economic Cooperation. The license strictly limits the services companies can offer and the cities or regions in which they can operate.

In a highly regulated environment, it is difficult to distinguish shipping lines on the basis of services or competitive factors. However, to capture some aspect of competition we classified shipping lines as Chinese and non-Chinese based on their flag of registration. Chinese shippers prefer Chinese shipping lines for two reasons: their business relationship has been built over a number of years, and the licensing system offers the most profitable routes to the Chinese shipping lines. Because foreign shipping lines are relatively new to the Chinese markets, cost and service data are limited. Thus, this paper tries to capture the differences in the attitudes of shippers toward shipping lines using the NMNL model.

We have good data for two characteristics related to shipping lines.

- **Total TEUs handled** during the year by shipping lines. We do not have an a priori expectation regarding its sign.
- **Number of vessels.** The greater the number of vessels operated by a shipping line, the lower the turnaround time for shippers' cargo. The number of vessels is highly related to frequency and timeliness and is an important service variable. We expect that it has a positive sign.

Shippers' Characteristics

The characteristics of the shippers themselves influence their decisions regarding port and shipping line choices. The variable we considered in our model is the distance of the shipper from the port.

We imputed the distance of shippers from ports based on our calculation from *China's Road Map Collection* (Measurement Publishing 1994) and *China's Route Map* (PTPH 2000). The farther the port is from the shipper's location, the less likely the port will be chosen. Distance is expected to be negatively related to port choice.

Cargo type also influences the choice of port and shipping line. This paper, however, considers only containerized cargo so the type of cargo is not a variable that will determine the choice of ports and shipping lines.

DATA

The data used in this paper are from the Survey of International Containerized Cargo Flows on the Yellow Sea Rim, conducted by the International Center for the Study of East Asian Development in Japan for 1998. The data do not include domestic cargo flows within China. The total sample size is 2,424 and covers shippers located in 9 provinces and 6 cities of China (Guangdong and Hainan provinces are not covered). The survey asked shippers about their choice of shipping line and port through which they export or import goods. The data report only the behavior of shippers. We supplemented these data with various port infrastructure-related variables such as number of berths and cranes, port charges, number of ship calls, routes offered, and so forth, from the *China Statistical Yearbook, China Shipping Development Annual Report 1998,* and the *Year Book of China Transportation & Communications.*

The data indicate that the most used port in China is at Shanghai, with 4.22 million total TEUs handled in 1998. Shanghai's better infrastructure accounts for this; it has 28 cranes and 19 berths. In 1998, 9,660 ships called at the port of Shanghai, traveling on 624 routes. Qingdao and Tianjin followed Shanghai with about 1.5 million and 1.3 million total TEUs, respectively.

The data used in this paper include some characteristics of shipping lines, for example, number of vessels and TEUs handled. However, due to the confidential nature of the data, we could not obtain the names of the shipping lines. This lack of information kept us from incorporating shipping line service-related variables. Moreover, the survey provided only quantitative variable data and no information on qualitative or behavioral aspects of shipping lines. As mentioned earlier, we classified shipping lines into two groups according to their country of registration: Chinese and non-Chinese. Of those surveyed, the Chinese shipping lines had the largest fleets, with an average of 128 vessels. The total TEUs handled by these shipping lines is shown in table 1.

TABLE 1 Descriptive Statistics

Port variables	Dalian	Tianjin	Qingdao	Shanghai	Average for other ports[1]
Berths (number)	5	7	5	19	4
Water depth (meters)	14	12	12	13	13
Cranes (number)	7	16	10	28	6
Total TEUs handled during the year (millions)	0.74	1.3	1.5	4.22	0.64
TEUs per crane	106	81	150	151	107
TEUs per berth	148	186	300	222	160
Ship calls (number)	4,216	2,498	2,549	9,660	2,445
Usage factor	5.61	6.25	8.02	8.35	4.58
Routes (number)	133	183	202	624	77
Shipping line variable		**Chinese**		**Non-Chinese**	
Average number of vessels		128.0		71.3	

[1] Other ports include Qinhuangdao, Yantai, Rizhaogang, Yingkou, Lianyungang, Ningbo, Xiamen, Yantiai, Guangzhou, and Shekou.

Note: Usage factor is handling volume/length of quay (thousand tons/meter).

RESULTS

As formulated in this paper, a shipper has a choice of 10 port-shipping line combinations. Based on the characteristics of shipping lines and ports, shippers choose the combinations that minimize costs. Based on model estimation trials, we retained only those variables that were significant at the 95% level, with the exception of the constant term. In our model, the constant term interacts with various alternatives to take advantage of nesting, and technically this increases the degrees of freedom. All other variables are alternative-specific.

The NMNL structure requires that we impute values not only for a combination that is chosen by the shipper but also for other choices. We imputed these values for all levels of combinations for each of the variables as averages for the variables. The results are presented in table 2.

We tried five nesting specifications (including MNL) of the model. Though there was little difference in the performance statistics (likelihood value, McFadden's Rho-squared, and ex-post correct predictions) of MNL and NMNL specifications, theoretical reasons suggest that NMNL specification is better than MNL, because unlike MNL it is not constrained by IIA (McFadden 1981).

Among NMNL specifications, Tree 3 (figure 2) performs slightly better than other specifications, and we computed elasticity of market shares based on this specification. The value of McFadden's Rho-squared of NMNL Tree 3 is 0.4042. The ex-post percentage of correct prediction is 58.75%. The dissimilarity coefficients of all NMNL trees differ

TABLE 2 Estimated Discrete Choice Functions

		NMNL			
Performance statistics	**MNL**	**Tree1**	**Tree2**	**Tree3**	**Tree4**
Log likelihood	–3,326.46	–3,326.10	–3,272.11	–3,325.52	–3,325.71
McFadden Rho-squared	0.4040	0.4041	0.4038	0.4042	0.4042
Correctly predicted (%)	58.746	58.746	58.705	58.746	58.746
Alternative specific variables					
Distance	–0.37(–37.0)	–0.80(–40.3)	–0.37(–36.8)	–0.35(–16.9)	–0.38(–31.8)
TEUs per berth	–2.05(–4.5)	–4.37(–9.4)	–1.55(–3.3)	–1.95(–4.2)	–2.09(–4.6)
TEUs per crane	4.01(6.5)	8.57(13.1)	3.70(5.5)	3.80(5.8)	4.08(6.5)
Ship calls	–0.91(–7.0)	–1.95(–14.0)	–0.72(–5.0)	–0.86(–6.2)	–0.93(–7.1)
Usage factor	–3.27(–3.7)	–7.09(–8.0)	–2.82(–3.2)	–3.09(–3.5)	–3.33(–3.8)
Routes	0.10(6.6)	0.22(13.2)	0.07(4.1)	0.10(5.9)	0.11(6.6)
Constant variables					
Chinese	1.13(23.9)	1.07(13.3)	11.33(12.8)	1.19(17.3)	1.18(17.7)
Dalian, Qingdao, other ports	0.74(0.8)	1.50(1.5)	0.69(0.7)	0.71(0.7)	0.76(0.8)
Tianjin, Qingdao	–1.41(–1.4)	–3.15(–3.2)	–1.23(–1.3)	–1.33(–1.4)	–1.44(–1.5)
Shanghai, other ports	3.75(4.4)	7.91(9.0)	3.30(3.7)	3.55(4.1)	3.81(4.4)
Dissimilarity parameters					
Chinese	1.00	2.16(49.7)		0.93(1.2)	1.00
Non-Chinese	1.00	2.08(18.2)		1.00	1.06(1.2)
Dalian			9.10(7.6)		
Tianjin			12.20(7.0)		
Qingdao			3.82(7.0)		
Shanghai			12.7(30.0)		
Other ports			12.56(7.6)		

Note: Numbers in parentheses are *t*-statistics.

from one for different levels of hierarchy and are greater than one at some or all levels for all except Tree 3. This suggests that trees other than Tree 3 are inconsistent (Boersch-Supan and Pitkin 1988). Because Tree 3 shows better performance, this suggests that shippers compare Chinese shipping lines and port combinations based on shipping line and port characteristics and the ability to choose one combination conditional on the presence of other alternatives. However, among non-Chinese combinations, choices are independent of the presence of other alternatives (like MNL structure). Results indicate that the distance of the port from shippers is an important determinant of port choice. If the port is far from the shipper's location, the probability of its being chosen by the shipper decreases. In fact, distance is so important in the overall decision process of shippers that many shippers have located closer to ports. An increase in the TEUs per berth, *ceteris paribus*, decreases the probability of choice of that port. The TEUs per crane at a port has a positive sign, as expected, capturing the efficiency.

The estimated coefficient for TEUs per berth is negative, while the coefficient of TEUs per crane is positive. This confirms our hypothesis that, for China, TEUs per berth represent port congestion, while TEUs per crane indicate handling efficiency. The following discussion explains our hypothesis and expectations about signs.

The port of Shanghai is the largest port in China, handling more than 4 million TEUs in 1998. Furthermore, Shanghai is a river port and it suffers from accumulated silt, making it unable to accommodate large container vessels. This historically important port is located near the largest economic center of China. As mentioned earlier, Shanghai handled 222,000 TEUs per berth in 1998.

Qingdao, whose economy is partly supported by trade and direct investment from Korea, is the second largest container port. In 1998, it handled 1.5 million TEUs; however, there are only 5 container berths at Qingdao (the same number as Dalian but fewer than Tianjin). As a result, Qingdao's reported container-handling volume was 300,000 TEUs per berth. The port authority at Qingdao developed new deep-water berths relatively late compared with Tianjin or Dalian and suffered from terminal area congestion for a long time. Berth undercapacity and terminal area congestion mean other vessels cannot choose their desired port call timing and may have to wait offshore until berths are vacant. Some vessels carrying cargoes requiring quick turnover avoid such congested ports and shift to other ports that can handle 150,000 to 180,000 TEUs per berth.

The ship size allowed by the water depth, the number of berths, and limited terminal areas determine the absolute container port capacity. Congestion due to the limited port capacity is a discouraging factor.

Two variables, the number of vessels and total TEUs handled by the shipping line during the year, were dropped from our estimated models because they were not significant. As mentioned earlier, this paper does not include variables related to cost and service characteristics due to limited data. We opine that classifying shipping lines as Chinese and non-Chinese would capture essential differences.

These results, when translated into market share elasticities, present an interesting picture. The proximity of shippers to ports plays an important role in determining their choices. Market share elasticities are useful for port planning and shipping line operations and, in turn, lead to economic development through advantages to shippers. Market share elasticities with respect to shipper-port distance differ depending on shipping lines and ports. Non-Chinese shipping line users tend to stay with the ports of their choice more than Chinese shipping line users, presumably because non-Chinese shipping line calls are limited to certain ports.

If a shipper is located far away from a port, *ceteris paribus,* preference for that port is low. The magnitude of decrease in market share relative to the shipper's distance from ports also depends on the chosen shipping line.

Table 3 presents distance elasticities and shows how the market share of ports changes if the shipper's distance from the port increases by 1%. For example, if the distance of a shipper from Dalian increases by 1% and if the shipper uses Chinese shipping lines (row 1), the market share of this combination is reduced by 6.2% (column 1), while the market shares of all other port-shipping line

TABLE 3 Distance Elasticity of Port-Shipping Line Market Shares

Port-shipping line combinations	1 Chinese/ Dalian	2 Chinese/ Tianjin	3 Chinese/ Qingdao	4 Chinese/ Shanghai	5 Chinese/ other ports	6 Non-Chinese/ Dalian	7 Non-Chinese/ Tianjin	8 Non-Chinese/ Qingdao	9 Non-Chinese/ Shanghai	10 Non-Chinese/ other ports
1. Chinese/Dalian	–6.206	0.283	0.283	0.283	0.283	0.256	0.256	0.256	0.256	0.256
2. Chinese/Tianjin	0.206	–3.542	0.206	0.206	0.206	0.186	0.186	0.186	0.186	0.186
3. Chinese/Qingdao	0.241	0.241	–4.405	0.241	0.241	0.218	0.218	0.218	0.218	0.218
4. Chinese/Shanghai	0.208	0.208	0.208	–4.122	0.208	0.189	0.189	0.189	0.189	0.189
5. Chinese/other ports	3.665	3.665	3.665	3.665	–15.433	3.318	3.318	3.318	3.318	3.318
6. Non-Chinese/Dalian	0.830	0.830	0.830	0.830	0.830	–5.930	0.830	0.830	0.830	0.830
7. Non-Chinese/Tianjin	0.640	0.640	0.640	0.640	0.640	0.640	–3.409	0.640	0.640	0.640
8. Non-Chinese/Qingdao	0.750	0.750	0.750	0.750	0.750	0.750	0.750	–4.231	0.750	0.750
9. Non-Chinese/Shanghai	0.620	0.620	0.620	0.620	0.620	0.620	0.620	0.620	–3.950	0.620
10. Non-Chinese/other ports	1.092	1.092	1.092	1.092	1.092	1.092	1.092	1.092	1.092	–16.604

Note: The table shows the elasticity of market shares of the port-shipping line combination shown in the column with respect to a 1% change in the distance of the shipper (using a particular type of shipping line) from the port shown in the row.

combinations (columns 2–5) increase by 0.28%; for other port-shipping line alternatives with non-Chinese shipping lines (columns 6–10) the increase is 0.26%. Moreover, if shippers using non-Chinese shipping lines to move goods through Dalian (row 6) change their location and the distance from this port increases by 1%, the market share of the Dalian-non-Chinese shipping line combination decreases by 5.9% (column 6), while the market shares of all other port-shipping lines combinations increase by 0.83% (columns 7–10).

Distance elasticities of market shares are high for new and secondary ports, and Dalian and Qingdao have the potential to increase their shares by improving their accessibility. On the other hand, Shanghai tends to retain its shippers due to economic maturity and related logistic functions achieved during its long port history.

As table 3 shows, ports most affected by distance increases are the "other ports," whose market share decreases by a large percentage. Of all ports surveyed, Tianjin is the least affected. Furthermore, Chinese shipping lines are the most affected. These elasticity estimates are smaller in magnitude than we estimated in our earlier study.

Table 4 presents market share elasticities for TEUs per berth. These numbers indicate the changes in the port-shipping line market shares shown in the column with respect to a 1% increase in TEUs per berth at a port. The elasticity values are highest for Qingdao, followed by Shanghai. Although Qingdao is the second largest container handling port, congestion is a serious problem. However, Qingdao's congestion relief efforts and efficiency improvement through port area development and other measures can be expected to enlarge its share.

The elasticity of market shares for TEUs per crane (table 5) indicates that with an increase of TEUs per crane, Qingdao's market share will increase the most followed by Shanghai.

Table 6 shows that with an increase in the number of routes offered at a port, Shanghai would gain the most. With an increase in ship calls, however, Shanghai loses the most, because, as mentioned earlier, Shanghai is served by the largest number of routes (table 7).

These results are interesting and indicate that after Shanghai, Qingdao is experiencing higher levels of use and now faces congestion problems. Qingdao's cargo handling services are efficient and an increase in the routes offered and improved infrastructure facilities would bring more cargo to this port.

Dalian has a new container terminal and the results indicate that its market share is less elastic than the other major ports with respect to congestion. Since its market share elasticity with respect to routes is second highest after Shanghai, Dalian is expected to experience increased demand due to greater economic activity and ship calls.

In sum, secondary and new ports other than Shanghai have the potential to increase their market shares through accessibility improvements, congestion relief, and route diversification. Addressing these challenges could make the Chinese port system more efficient and further increase inter-port competition.

CONCLUSION

China's acceptance into the World Trade Organization has generated significant interest in the forwarding and logistics services sector. Demand has grown for quality forwarding and logistics services, better services at ports, and upgraded port infrastructure.

Newer Chinese ports are starting to compete with older, more established ports. For example, newer ports on the southeastern coast of China have challenged the traditional monopoly of Hong Kong (80% of the cargo leaving Hong Kong comes from South China). Although this paper does not include Hainan and Guangdong provinces, where a lot of the Chinese cargo going to the port of Hong Kong originates, ports in these regions compete with Hong Kong.

China is a country of cargo originators, and the international competitiveness of its goods depends on cost-effective and efficient shipment. With a view to attract more cargo, Chinese ports are competing with each other.

TABLE 4 TEUs per Berth Elasticity of Port-Shipping Line Market Shares

Port-shipping line combinations	1	2	3	4	5	6	7	8	9	10
1. Chinese/Dalian	–26.738	4.329	4.329	4.329	4.329	3.919	3.919	3.919	3.919	3.919
2. Chinese/Tianjin	5.397	–33.584	5.397	5.397	5.397	4.884	4.884	4.884	4.884	4.884
3. Chinese/Qingdao	6.561	6.561	–56.413	6.561	6.561	5.936	5.936	5.936	5.936	5.936
4. Chinese/Shanghai	9.319	9.319	9.319	–37.303	9.319	8.437	8.437	8.437	8.437	8.437
5. Chinese/other ports	6.434	6.434	6.434	6.434	–27.090	5.824	5.824	5.824	5.824	5.824
6. Non-Chinese/Dalian	1.234	1.234	1.234	1.234	1.234	–27.553	1.234	1.234	1.234	1.234
7. Non-Chinese/Tianjin	1.601	1.601	1.601	1.601	1.601	1.601	–34.519	1.601	1.601	1.601
8. Non-Chinese/Qingdao	1.973	1.973	1.973	1.973	1.973	1.973	1.973	–56.379	1.973	1.973
9. Non-Chinese/Shanghai	2.655	2.655	2.655	2.655	2.655	2.655	2.655	2.655	–40.545	2.655
10. Non-Chinese/other ports	1.916	1.916	1.916	1.916	1.916	1.916	1.916	1.916	1.916	–29.146

TABLE 5 TEUs per Crane Elasticity of Port-Shipping Line Market Shares

Port-shipping line combinations	1	2	3	4	5	6	7	8	9	10
1. Chinese/Dalian	37.276	–6.035	–6.035	–6.035	–6.035	–5.463	–5.463	–5.463	–5.463	–5.463
2. Chinese/Tianjin	–4.612	28.701	–4.612	–4.612	–4.612	–4.173	–4.173	–4.173	–4.173	–4.173
3. Chinese/Qingdao	–6.403	–6.403	55.059	–6.403	–6.403	–5.794	–5.794	–5.794	–5.794	–5.794
4. Chinese/Shanghai	–12.343	–12.343	–12.343	49.406	–12.343	–11.174	–11.174	–11.174	–11.174	–11.174
5. Chinese/other ports	–8.375	–8.375	–8.375	–8.375	35.263	–7.582	–7.582	–7.582	–7.582	–7.582
6. Non-Chinese/Dalian	–1.720	–1.720	–1.720	–1.720	–1.720	38.412	–1.720	–1.720	–1.720	–1.720
7. Non-Chinese/Tianjin	–1.368	–1.368	–1.368	–1.368	–1.368	–1.368	29.499	–1.368	–1.368	–1.368
8. Non-Chinese/Qingdao	–1.925	–1.925	–1.925	–1.925	–1.925	–1.925	–1.925	55.025	–1.925	–1.925
9. Non-Chinese/Shanghai	–3.517	–3.517	–3.517	–3.517	–3.517	–3.517	–3.517	–3.517	53.700	–3.517
10. Non-Chinese/other ports	–2.494	–2.494	–2.494	–2.494	–2.494	–2.494	–2.494	–2.494	–2.494	37.941

TABLE 6 Number-of-Routes Elasticity of Port-Shipping Line Market Shares

Port-shipping line combinations	1	2	3	4	5	6	7	8	9	10
1. Chinese/Dalian	12.263	–1.985	–1.985	–1.985	–1.985	–1.797	–1.797	–1.797	–1.797	–1.797
2. Chinese/Tianjin	–2.174	16.891	–2.174	–2.174	–2.174	–2.456	–2.456	–2.456	–2.456	–2.456
3. Chinese/Qingdao	–2.255	–2.255	19.386	–2.255	–2.255	–2.040	–2.040	–2.040	–2.040	–2.040
4. Chinese/Shanghai	–13.363	–13.363	–13.363	53.487	–13.363	–12.097	–12.097	–12.097	–12.097	–12.097
5. Chinese/other ports	–1.586	–1.586	–1.586	–1.586	6.677	–1.436	–1.436	–1.436	–1.436	–1.436
6. Non-Chinese/Dalian	–0.566	–0.566	–0.566	–0.566	–0.566	12.637	–0.566	–0.566	–0.566	–0.566
7. Non-Chinese/Tianjin	–0.805	–0.805	–0.805	–0.805	–0.805	–0.805	17.361	–0.805	–0.805	–0.805
8. Non-Chinese/Qingdao	–0.678	–0.678	–0.678	–0.678	–0.678	–0.678	–0.678	19.374	–0.678	–0.678
9. Non-Chinese/Shanghai	–3.807	–3.807	–3.807	–3.807	–3.807	–3.807	–3.807	–3.807	58.135	–3.807
10. Non-Chinese/other ports	–0.472	–0.472	–0.472	–0.472	–0.472	–0.472	–0.472	–0.472	–0.472	7.184

TABLE 7 Ship-Calls Elasticity of Port-Shipping Line Market Shares

Port-shipping line combinations	1	2	3	4	5	6	7	8	9	10
1. Chinese/Dalian	–33.716	5.458	5.458	5.458	5.458	4.941	4.941	4.941	4.941	4.941
2. Chinese/Tianjin	3.214	–19.997	3.214	3.214	3.214	2.908	2.908	2.908	2.908	2.908
3. Chinese/Qingdao	2.468	2.468	–21.217	2.468	2.468	2.233	2.233	2.233	2.233	2.233
4. Chinese/Shanghai	17.942	17.942	17.942	–71.816	17.942	16.243	16.243	16.243	16.243	16.243
5. Chinese/other ports	4.360	4.360	4.360	4.360	–18.358	3.947	3.947	3.947	3.947	3.947
6. Non-Chinese/Dalian	1.556	1.556	1.556	1.556	1.556	–34.742	1.556	1.556	1.556	1.556
7. Non-Chinese/Tianjin	0.953	0.953	0.953	0.953	0.953	0.953	–20.554	0.953	0.953	0.953
8. Non-Chinese/Qingdao	0.742	0.742	0.742	0.742	0.742	0.742	0.742	–21.204	0.742	0.742
9. Non-Chinese/Shanghai	5.112	5.112	5.112	5.112	5.112	5.112	5.112	5.112	–78.057	5.112
10. Non-Chinese/other ports	1.299	1.299	1.299	1.299	1.299	1.299	1.299	1.299	1.299	–19.752

The analysis of shippers' choice of shipping lines and ports is essential for policy formulation related to improving port infrastructure and services, in addition to market entry/exit decisions of shipping lines. This paper models the shipping line and port choice behavior of Chinese shippers. This is one of the few studies that models this behavior using an empirical model, and it may be the first study that models the joint choice of shipping lines and ports in China.

The data used in this paper are unique and from a survey of shippers conducted by the International Center for the Study of East Asian Development in Japan for 1998. Results indicate that Chinese shippers and forwarders are conservative and prefer Chinese shipping lines primarily because of long-established relationships and the availability of larger fleets. The shippers are indifferent to foreign shipping lines, but choose them based on the ports from which they would like to import or export their cargo. However, foreign shipping lines have only recently started operating in China and a change in the behavior of Chinese shippers can be expected in near future.

Of the variables we studied, high TEUs per berth indicate congestion and negatively affect shippers' decisions. High TEUs per crane indicate efficiency in moving cargo and show a positive coefficient. The distance of a port from a shipper's location and the number of ship calls at a port are important variables that determine the choice of a port. Distance and ship calls both have negative elasticity. Ports offering more routes have higher market share elasticity. Another variable that affects the choice of port is the handling value (thousand tons of cargo per meter of quay). This variable also captures congestion and has a negative coefficient.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to two anonymous referees whose comments helped in substantially improving the quality of this paper.

REFERENCES

Bayliss, B.T. and S.L. Edwards. 1970. *Industrial Demand for Transport*. London, England: Ministry of Transport.

Bird, J. 1988a. Freight Forwarders Speak: The Perception of Route Competition via Seaports in the European Communities Research Project (Part I). *Maritime Policy and Management* 15(1):35–55.

_____. 1988b. Freight Forwarders Speak: The Perception of Route Competition via Seaports in the European Communities Research Project (Part II). *Maritime Policy and Management* 15(2):107–125.

Boersch-Supan, A. and J. Pitkin. 1988. On Discrete Choice Models of Housing Demand. *Journal of Urban Economics* 24(2):153–172.

Brooks, M.R. 1984. An Alternative Theoretical Approach to the Evaluation of Liner Shipping: Part 1. Situational Factors. Maritime Policy and Management 11:35–43.

_____. 1985. An Alternative Theoretical Approach to the Evaluation of Liner Shipping: Part 2. Choice/Criteria. *Maritime Policy and Management* 12:145–155.

Chiu, R.H. 1996. Logistics Performance of Liner Shipping in Taiwan, Ph.D. dissertation. Department of Maritime Studies and International Transport, University of Wales, College of Cardiff, UK.

Gilmour, M.R. 1976. Some Policy Implications of Subjective Factors in the Modal Choice for Freight Movements. *The Logistics and Transportation Review* 12:39–57.

International Centre for the Study of East Asian Development. 1999. Survey of Cargo Flows in the Yellow Sea Rim (in Japanese). March.

Jamaluddin, T.T.M.S. 1995. Marketing of Freight Liner Shipping Services with Reference to the Far East-Europe Trade: A Malaysian Perspective", Ph.D. dissertation, Department of Maritime Studies and International Transport, University of Wales College of Cardiff, UK.

McFadden, D. 1981. Econometric Models of Probabilistic Choice. *Structure Analysis of Discrete Data with Econometric Applications*. Edited by C.F. Manski and D. McFadden. Cambridge, MA: MIT Press.

McGinnis, M.A. 1979. Shipper Attributes Towards Freight Transport Choice: A Factor Analytic Study. *International Journal of Physical Distribution and Materials Management* 10:25–34.

Measurement Publishing Co. 1994. *Surveying and Mapping Press* (in Chinese). Beijing, China.

Meyrick, S. and G. D'Este. 1989. More Than the Bottom Line: How Users Select a Shipping Service. *Proceeding of the 14th Australian Transportation Research Forum*, Perth, Australia, pp. 65–81.

Mongeluzzo, B. 2002. Spanning the Globe. *The Journal of Commerce,* November 18–24.

Ogden, K.W. and A.L. Rattray. 1982. Analysis of Freight Mode Choice. *Proceedings of the 7th Australian Transport Research Forum,* Hobart, Australia, pp. 249–276.

Pearson, R. 1980. Containerline Performance and Service Quality. Marine Transport Center, University of Liverpool, Liverpool, England.

People's Transportation Publishing House (PTPH). 2000. *China's Route Map* (in Chinese). Beijing, China.

Slack, B. 1985. Containerization, Inter-Port Competition and Port Selection. *Maritime Policy and Management* 12(4).

State Statistical Bureau (SSB). Annual. *China Statistical Yearbook* (in Chinese). Beijing, China.

Suthiwartnarueput, K. 1988. The Exploration of Sea Transport Efficiency: With a Concentration on the Case of Thailand, Ph.D. dissertation. Department of Maritime Studies and International Transport, University of Wales, College of Cardiff, UK.

Tiwari, P., H. Itoh, and M. Doi. 2003. Shippers' Port and Carrier Selection Behaviour in China: A Discrete Choice Model. *Maritime Economics and Logistics* 5(1).

Tongzon, J.L. 1995. Determinants of Port Performance and Efficiency. *Transportation Research: Part A* 29(3):245–252.

Water Transportation Ministry of the P.R.C. 1997. *The Rule of Port Charges (Foreign Trade)* (in Chinese). Beijing, China: People's Transportation Publishing House.

_____. 1999. *China Shipping Development Annual Report 1998* (in Chinese, a part in English). Beijing, China: People's Transportation Publishing House.

Wilson, F.R., B.J. Bisson, and K.B. Kobia. 1986. Factors that Determine Mode Choice in the Transportation of General Freight. *Transportation Research Record* 1061:26–31.

Year Book House of China Transportation & Communications. 1999. *Year Book of China Transportation & Communications* (in Chinese). Beijing, China.

Authors' addresses: Piyush Tiwari (corresponding author), Urban Specialist, Infrastructure Development Finance Co., Ltd., The Capital Court, 6th Floor, Olof Palme Marg, Munirka, New Delhi 110 067 India. Email: piyush@idfc.com.

Hidekazu Itoh, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba 305-8573 Japan. Email: hito@kwansei.ac.jp.

Masayuki Doi, Institute of Policy and Planning Sciences, University of Tsukuba, Tsukuba 305-8573 Japan. Email: doi@sk.tsukuba.ac.jp.

JOURNAL OF TRANSPORTATION AND STATISTICS

Guidelines for Manuscript Submission

Please note: Submission of a paper indicates the author's intention to publish in the *Journal of Transportation and Statistics* (JTS). Submission of a manuscript to other journals is unacceptable. Previously published manuscripts, whether in an exact or approximate form, cannot be accepted. Check with the Managing Editor if in doubt.

Scope of JTS: JTS has two primary areas of focus. We publish articles that look at transportation from a statistical perspective. We also publish articles that do not use or discuss statistical techniques but present information on areas of transportation research that are of interest to JTS readers. Papers that do not use or discuss statistical techniques may be considered for the Research Notes section.

Manuscripts must be double spaced, including quotations, abstract, reference section, and any notes. All figures and tables should appear at the end of the manuscript with each one on a separate page. Do not embed them in your manuscript.

Because the JTS audience works in diverse fields, **please define** terms that are specific to your area of expertise.

Electronic submissions via email to the Managing Editor are strongly encouraged. We can accept PDF, Word, Excel, and Adobe Illustrator files. If you cannot send your submission via email, you may send a disk or CD by overnight delivery service or send a hardcopy by the U.S. Postal Service (regular mail; see below). Do not send disks or CDs through regular mail.

Hardcopy submissions delivered to BTS by the U.S. Postal service are irradiated. Do not include a disk in your envelope; the high heat will damage it.

The cover page of your manuscript must include the title, author name(s) and affiliations, and the telephone number and surface and email addresses of all authors.

Put the **Abstract** on the second page. It should be about 100 words and briefly describe the contents of the paper including the mode or modes of transportation, the research method, and the key results and/or conclusions. Please include a list of keywords to describe your article.

Graphic elements (figures and tables) must be called out in the text. Graphic elements must be in black ink. We will accept graphics in color only in rare circumstances.

References follow the style outlined in the *Chicago Manual of Style*. All non-original material must be sourced.

International papers are encouraged, but please be sure to have your paper edited by someone whose first language is English and who knows your research area.

Accepted papers must be submitted electronically in addition to a hard copy (see above for information on electronic submissions). Make sure the hard copy corresponds to the electronic version.

Page proofs: As the publication date nears, authors will be required to return article page proofs to the Managing Editor within 48 hours of receipt.

Acceptable software for text and equations is limited to Word and LaTeX. Data behind all figures, maps, and charts must be provided in Excel, Word, or DeltaGraph. Data in American Standard Code for Information Interchange (ASCII) text will be accepted but is less desirable. Acceptable software for graphic elements is limited to Excel, DeltaGraph, or Adobe Illustrator. If other software is used, the file supplied must have an .eps or .pdf extension. We do not accept PowerPoint.

Maps are accepted in a variety of Geographic Information System (GIS) programs. Please save your GIS files using .eps or .pdf extensions. If this is not possible, maps may be provided in ArcView3.x project files (.apr), ArcGIS8.x map files (.mxd), or Arc/Info metafiles (.map). Send your files using the .apr or .mxd extensions on a CD-ROM via overnight delivery service.

Send all submission materials to:
Marsha Fenn
Managing Editor, *Journal of Transportation and Statistics*
Bureau of Transportation Statistics
U.S. Department of Transportation
400 7th Street, SW
Room 7412, K-15
Washington, DC 20590
Email: marsha.fenn@bts.gov

BUREAU OF TRANSPORTATION STATISTICS UNITED STATES DEPARTMENT OF TRANSPORTATION

JOURNAL OF TRANSPORTATION AND STATISTICS

Volume 6 Number 1, 2003
ISSN 1094-8848

CONTENTS