



*FINAL REPORT*

**Connecting People to Places:  
Spatiotemporal Analysis of Transit Supply  
Using Travel Time Cubes**

---

NITC-RR-662 ■ June 2016

*NITC is the U.S. Department of Transportation's national  
university transportation center for livable communities.*



**CONNECTING PEOPLE TO PLACES:  
SPATIOTEMPORAL ANALYSIS OF TRANSIT SUPPLY  
USING TRAVEL TIME CUBES**

**Final Report**

**NITC-RR-662**

by

Steven Farber  
University of Toronto Scarborough

for

National Institute for Transportation and Communities (NITC)  
P.O. Box 751  
Portland, OR 97207



**June 2016**



## Technical Report Documentation Page

1. Report No. NITC-RR-662	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Connecting People to Places: Spatiotemporal Analysis of Transit Supply Using Travel Time Cubes		5. Report Date June 2016	
		6. Performing Organization Code	
7. Author(s) Steven Farber		8. Performing Organization Report No.	
9. Performing Organization Name and Address		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No.	
12. Sponsoring Agency Name and Address  National Institute for Transportation and Communities (NITC) P.O. Box 751 Portland, Oregon 97207		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplementary Notes			
<p>16. Abstract</p> <p>Despite its importance, temporal measures of accessibility are rarely used in transit research or practice. This is primarily due to the inherent difficulty and complexity in computing time-based accessibility metrics. Estimating origin-to-destination travel times that include the “last mile” of travel between the transit network and actual start and endpoints of the trip is technically difficult. Not only do such estimations require multimodal network structures, they also require detailed knowledge of transit schedules and sophisticated algorithms for calculating shortest paths using such inputs. Recently, new standards for sharing transit schedules and geographic data, namely the General Transit Feed Specification (GTFS) have prompted innovations in the analysis of complex transit travel times using the Esri ArcGIS package with the Network Analyst extension. With continued development of the analytical capabilities of network analysis functionality, this project assesses spatiotemporal dynamics in transit supply through an investigation of scheduled travel time variability.</p> <p>This report consists of a collection of three stand-alone research papers. The first defines a new data object, the public transit travel time cube, and demonstrates its use in measuring changes in transit provision over time, changes in accessibility to jobs, and the impacts of bicycling on the last mile problem. The second paper consists of a study of temporal mismatch between observed travel patterns and the spatiotemporal patterns of transit supply in the Wasatch Front. The third paper expands the food desert literature by measuring temporal fluctuations in food accessibility over the course of a typical day, and examining the trends for disparities between socioeconomic subgroups. Our findings indicate that time-based measures of transit accessibility provide more nuanced abilities for us to understand how people are impacted by temporal variability in transit provision.</p>			
17. Key Words public transit; social equity; travel time cubes; transit mismatch; temporal mismatch; dynamic accessibility; food deserts.		18. Distribution Statement No restrictions. Copies available from NITC: <a href="http://www.nitc.us">www.nitc.us</a>	
19. Security Classification (of this report)  Unclassified	20. Security Classification (of this page)  Unclassified	21. No. of Pages	22. Price



## **ACKNOWLEDGEMENTS**

This project was funded by the National Institute for Transportation and Communities (NITC). The authors are thankful for generous support from University of Utah, Melinda Morang and Esri Ltd., Utah Transit Authority and TriMet.

## **DISCLAIMER**

The contents of this report reflect the views of the authors, who are solely responsible for the facts and the accuracy of the material and information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation University Transportation Centers Program [and other SPONSOR/PARTNER] in the interest of information exchange. The U.S. Government [and other SPONSOR/PARTNER] assumes no liability for the contents or use thereof. The contents do not necessarily reflect the official views of the U.S. Government [and other SPONSOR/PARTNER]. This report does not constitute a standard, specification, or regulation.



# TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY .....</b>	<b>1</b>
<b>1.0 DYNAMIC PUBLIC TRANSIT ACCESSIBILITY USING TRAVEL TIME CUBES: COMPARING THE EFFECTS OF INFRASTRUCTURE (DIS)INVESTMENTS OVER TIME .....</b>	<b>2</b>
1.1 ABSTRACT.....	3
1.2 INTRODUCTION .....	3
1.3 METHODS .....	4
1.4 CASE STUDY ONE: COMPARISON OF TRANSIT TRAVEL TIMES OVER TIME..	6
1.5 CASE STUDY TWO: ASSESSING ACCESS TO JOBS USING TRAVEL TIME CUBES .....	13
1.6 CASE STUDY THREE: THE EFFECTS OF BICYCLE USE ON THE LAST MILE PROBLEM.....	14
1.7 DISCUSSION AND CONCLUSIONS .....	17
<b>2.0 SPACE-TIME MISMATCH BETWEEN TRANSIT SERVICE AND OBSERVED TRAVEL PATTERNS IN THE WASATCH FRONT, UTAH: A SOCIAL EQUITY PERSPECTIVE.....</b>	<b>20</b>
2.1 ABSTRACT.....	20
2.2 INTRODUCTION .....	21
2.3 METHOD .....	22
2.3.1 The Public Transit Travel Time Cube .....	22
2.3.2 Spatiotemporal Measures of Transit Service .....	24
2.3.3 Comparing Travel Patterns to Transit Service.....	25
2.3.4 Data Used to Indicate Travel Patterns .....	26
2.4 RESULTS AND DISCUSSION .....	26
2.4.1 ANOVA .....	26
2.4.2 Multivariate Regression Results .....	29
2.4.3 Conclusions.....	32
<b>3.0 TEMPORAL VARIABILITY IN TRANSIT-BASED ACCESSIBILITY TO SUPERMARKETS .....</b>	<b>34</b>
3.1 ABSTRACT.....	34
3.2 INTRODUCTION .....	35
3.3 CASE STUDY – FOOD DESERTS – NEED FOR TIME-VARYING MEASUREMENTS .....	36
3.4 METHOD .....	37
3.4.1 Travel Time Estimation .....	40
3.4.2 Accessibility Metrification.....	40
3.4.3 Statistical Analysis.....	42
3.5 RESULTS .....	42
3.6 CONCLUSIONS.....	52



**APPENDICES**

APPENDIX A: DESCRIPTION OF OTHER FUNDED RESEARCH

APPENDIX B: TRAVEL TIME CUBE DOCUMENTATION

**LIST OF TABLES**

Table 1.1 Numerical Summary of Changes in Service Delivery ..... 7  
Table 2.1: Utah Household Travel Survey (UHTS) ANOVA Tests for Travel Time and Trip Speed..... 27  
Table 2.2: Utah Onboard Survey (UOBS) ANOVA Tests for Travel Time and Trip Speed ..... 28  
Table 2.3: Summary of Ordinary Least Squares Regression Analysis for Variables Predicting Travel Time Z-Scores using the UHTS data (N=34,409)..... 30  
Table 3.1 Interpretation of Average Travel Time and Dispersion Measures ..... 42

**LIST OF FIGURES**

Figure 1.1 Illustration of the Travel Time Cube and its Atomic Vectors..... 5  
Figure 1.2 Using Fourier Transforms to Discriminate between OD Pairs with Similar Travel Time Moments but Different Latent Frequencies of Connectivity..... 6  
Figure 1.3 Cartographic Depiction of Changes in Service Delivery between a) Time Period 1 and b) Time Period 2 ..... 8  
Figure 1.4 Depictions of Travel Time Changes, a) Numerically, b) By Time of Day for the Wasatch Front, c) By Time of Day for Portland, and d) By Change in Average Travel Times to and from All Zones in the City. Negative values represent decreases in travel time, while positive values are increases. .... 11  
Figure 1.5 Change in a) Average Travel Time Cycle and b) Jobs Accessibility..... 12  
Figure 1.6 A Comparison of Transit Travel Times Joint with Walking Versus Cycling. Average Travel Times are Depicted in A and C, while Standard Deviations of Travel Times are Depicted in B and D..... 15  
Figure 1.7 Transit Travel Time Improvements in the Center of Salt Lake City Associated with Bicycle Use. .... 16  
Figure 2.1: Wasatch Front Study Area Block Group Centroids ..... 23  
Figure 2.2: Illustration and Interpretation of Level of Service Z-scores ..... 25  
Figure 3.1 Maps of Cincinnati Depicting Supermarket Locations and a) Transit Frequency of Service, b) Distribution of the Black Population, c) Distribution of the Older Adult Population, and d) Distribution of Households Living in Poverty ..... 39  
Figure 3.2 Variation in Public Transit Travel Time Between Two Locations across a 24-Hour Period ..... 41  
Figure 3.3 Mean Travel Time to a) The Closest Supermarket, b) The Closest Three Supermarkets and Standard Deviation of Travel Time to c) The Closest Supermarket and d) The Closest Three Supermarkets ..... 44

Figure 3.4 Percent of Day with Access to One and Three Supermarkets within 10 Minutes (a and b) and 20 Minutes (c and d), Respectively .....	45
Figure 3.5 Scattergrams of Mean Travel Time to Closest Supermarket versus Standard Deviation of Travel Time .....	47
Figure 3.6 Percent of Day with 20-Minute Access to Supermarkets by Race.....	48
Figure 3.7 Percent of Day with 20-Minute Access to Supermarkets by Age .....	49
Figure 3.8 Percent of Day with 20-Minute Access to Supermarkets by Poverty Status .....	49
Figure 3.9: Zero-Vehicle Households and Transit-Based Access to Supermarkets .....	51



## **EXECUTIVE SUMMARY**

Transit planning traditionally emphasizes the spatial dimension of accessibility; networks are built to bridge locations in the city with the assumption that the provision of spatial connectivity is equivalent to providing people with access to their destinations. However, often underrepresented in transit analyses, is that travel time, not network proximity, is the fundamental unit of influence over people's travel behavior. It is the time lost in travel that drives whether or not people will make trips. Moreover, it is the time lost in travel that provides the clearest indicator of the burdens and benefits associated with spatial and temporal dimensions of transit provision to the public. Thus, understanding the supply of transit through an analysis of end-to-end travel times can help transit planners better explain patterns of ridership and implement changes to the network that will better serve its ridership. At the same time, by advancing the science of measuring transit accessibility, planners will be able to more accurately assess the equality with which their system provides opportunities to different locations and demographic sectors within their service area.

Despite its importance, temporal measures of accessibility are rarely used in transit research or practice. This is primarily due to the inherent difficulty and complexity in computing time-based accessibility metrics. Estimating origin-to-destination travel times that include the "last mile" of travel between the transit network and actual start and endpoints of the trip is technically difficult. Not only do such estimations require multimodal network structures, they also require detailed knowledge of transit schedules and sophisticated algorithms for calculating shortest paths using such inputs. Recently, new standards for sharing transit schedules and geographic data, namely the General Transit Feed Specification (GTFS), have prompted innovations in the analysis of complex transit travel times using the Esri ArcGIS package with the Network Analyst extension. With continued development of the analytical capabilities of network analysis functionality, this project aims to assess spatiotemporal dynamics in transit supply in the Wasatch Front and the Portland area through an investigation of scheduled travel time variability.

This report consists of a collection of three stand-alone research papers. The first defines a new data object, the public transit travel time cube, and demonstrates its use in measuring changes in transit provision over time, changes in accessibility to jobs, and the impacts of bicycling on the last mile problem. The second paper consists of a study of temporal mismatch between observed travel patterns and the spatiotemporal patterns of transit supply in the Wasatch Front. The third paper expands the food desert literature by measuring temporal fluctuations in food accessibility over the course of a typical day, and examining the trends for disparities between socioeconomic subgroups. Our findings indicate that time-based measures of transit accessibility provide more nuanced understandings of how people are impacted by temporal variability in transit provision.

# **1.0 DYNAMIC PUBLIC TRANSIT ACCESSIBILITY USING TRAVEL TIME CUBES: COMPARING THE EFFECTS OF INFRASTRUCTURE (DIS)INVESTMENTS OVER TIME**

**Steven Farber**

Department of Human Geography, University of Toronto Scarborough  
1265 Military Trail, Toronto ON, M1C 1A4, Canada  
Tel: (416) 208-5120 Fax: (416) 287-7283; Email: steven.farber@utoronto.ca

**Liwei Fu**

City of San José  
200 E. Santa Clara St., San José, California, 95113  
Email: folflw@gmail.com

*A version of this paper has been submitted and is currently under review for publication in an academic peer-reviewed journal.*

## 1.1 ABSTRACT

We put forward a new data object called the public transit travel time cube and demonstrate how the cube can be used in the analysis of transit travel time changes over space and time. The travel time cube contains the shortest path transit travel time between sets of origins and destinations in the city, at all times of day. Once computed, a wide range of investigations become readily available to the transit planner or transportation researcher. We conduct three demonstrative analysis using travel time cubes for the Wasatch Front, Utah and the Portland region in Oregon. Our studies investigate how travel times were impacted by service cuts and expansions in the two regions respectively, and the impact this had on jobs accessibility. We also use the travel time cube to study the last mile problem, and compute the travel time savings and the stability gained by solving the last mile problem with bicycling. The paper concludes with an expanded discussion on the merits of the travel time cube and outlines four avenues for continued research.

## 1.2 INTRODUCTION

Accessibility is largely understood as the ease with which people can reach destinations (Hansen, 1959). Cities in which inhabitants are more able to reach appropriately matched jobs, public services, leisure locations, and social contacts likewise produce opportunities for more rapid economic growth (Banister and Berechman, 2000), and a higher quality of life for inhabitants (Frank, 2000). In urban areas, public transportation plays a critical role in providing low-cost, energy-efficient and socially equitable means of accessibility (Pucher, 2004), yet techniques for measuring the spatiotemporal patterns of transit-based accessibility remain understudied. A large proportion of transit accessibility measures focus on the ease with which people can reach bus stops and train stations, rather than investigating how well the transit service provides access to actual destinations (Mavoia et al., 2012). Understanding access to destinations necessitates a travel time analysis through the transit network, which can be computationally cumbersome and requires access to digitized pedestrian and transport networks. When performing travel time analyses, it is important to realize that since travel by public transit is subject to the schedule-based fluctuations in the provision of service, incorporating temporal dynamics into our accessibility measures should increase their validity (Farber, Morang and Widener, 2014).

In this study, we explore the spatiotemporal patterns of public transit accessibility by focusing on schedule-based origin-to-destination (OD) travel times and their fluctuations over the course of the day. Our approach is based on the development and analysis of public transit travel time cubes, a data structure containing the estimated transit-based OD travel time between all locations in a region, at all times of day. In this article, we put forward our technique for creating the travel time cube and the technical procedures we developed to analyze the vast array of computed travel times, and demonstrate the use of travel time cubes in understanding changes in spatiotemporal patterns of accessibility through a series of case studies.

The rest of the paper is organized as follows: In the next section, we describe the public transit travel time cube and the techniques used to create and analyze it. Following this, we present three case studies that highlight the use of travel time cubes in the comparison of spatiotemporal patterns of accessibility. Finally, we conclude the paper with a summary of our contributions and ideas for future research. A formal review of the transit accessibility literature is purposefully excluded for the sake of brevity. Judicious citations to relevant literature have been placed throughout the text in lieu of a formal review, and interested readers may seek general reviews of the accessibility literature here (Geurs and Van Wee, 2004; Páez, Scott and

Morency, 2012; Levinson and Krizek, 2005), or more specific papers closely related to transit travel time cubes here (Farber, Morang and Widener, 2014; Ritter, 2014; Owen and Levinson, 2015; Lei and Church, 2010; Lei, Chen and Goulias, 2012; Mavoa et al., 2012; O'Sullivan, Morrison and Shearer, 2000).

### 1.3 METHODS

The public transit travel time cube is a three-dimensional array,  $T = [t_{i,j,m}]$ , of estimated transit travel times where  $i$  and  $j$  index locations and  $m$  indexes trip departure times. In the case where  $i$  and  $j$  are traffic analysis zones (TAZs) or census geography units, and  $m$  is computed for all minutes of the day, the dimension of the cube is  $N \times N \times M$ , where  $N$  is the number of zones in the region and  $M$  is the number of minutes in the day (i.e., 1,440). In this particular case, the travel time cube contains the public transit travel time from all origins to all destinations at all times of day, and thus represents the latent structure of public transit connectivity in a region.

In our implementation of the cube, travel times are computed using General Transit Feed Specification (GTFS) data, a pedestrian network file, and the *Add GTFS to Network Dataset* toolbox for *Esri ArcGIS Network Analyst*. Given a set of origins and destinations and a departure time, we use the *Esri OD Cost Matrix* tool to compute the shortest path travel time on the multimodal network (pedestrian plus transit) that includes ingress, egress, waiting, transfer and in-vehicle travel times. The algorithm returns a matrix of shortest path travel times which may include “walk-only” routes if walking is faster than using transit. Custom *Python* scripts are used to iterate matrix computations over the minutes of the day, and to assign concurrent *OD Cost Matrix* estimations to the available processors on a multiple core server running *Esri ArcGIS* in a *Windows* environment. A full suite of *Python* scripts and documentation are available from the authors upon request. Similar data objects built with tools by *Esri* and other developers have been used elsewhere in the literature to investigate issues of accessibility, social equity and mode choice (Lei, Chen and Goulias, 2012; Farber, Morang and Widener, 2014; Owen and Levinson, 2015)

Analysis of the travel time cube is achieved through summarizing and visualizing its various cross-sections. For example, holding an origin and a departure time constant, the vector  $T_{i,:m}$  contains the scheduled travel time from origin  $i$  to all destinations provided a trip start time,  $m$ . Similarly, the vector  $T_{:,j,m}$  contains travel times from all origins to a specific destination at a particular time of day. Each of these vectors is suitably visualized by mapping them either at zones of origin or destination. Alternatively, we can hold an OD pair constant and extract a third vector,  $T_{i,j,:}$  which can be plotted as the time series of travel times from  $i$  to  $j$  over the course of the day. Each of these queries can be thought of as a single column or row being extracted from the three-dimensional data cube (**Figure 1.1**), and the mapping and plotting of such vectors represents the atomic methods for visualizing the cube. In practice, multiple rows or columns can be summarized before visualization. For example, the mean travel time from an origin to all destinations over the course of the day is a meaningful measure of the accessibility of a place in a transit network, while the standard deviation of the vector  $T_{i,j,:}$  describes the amount of variation in travel time between two locations over the course of the day.

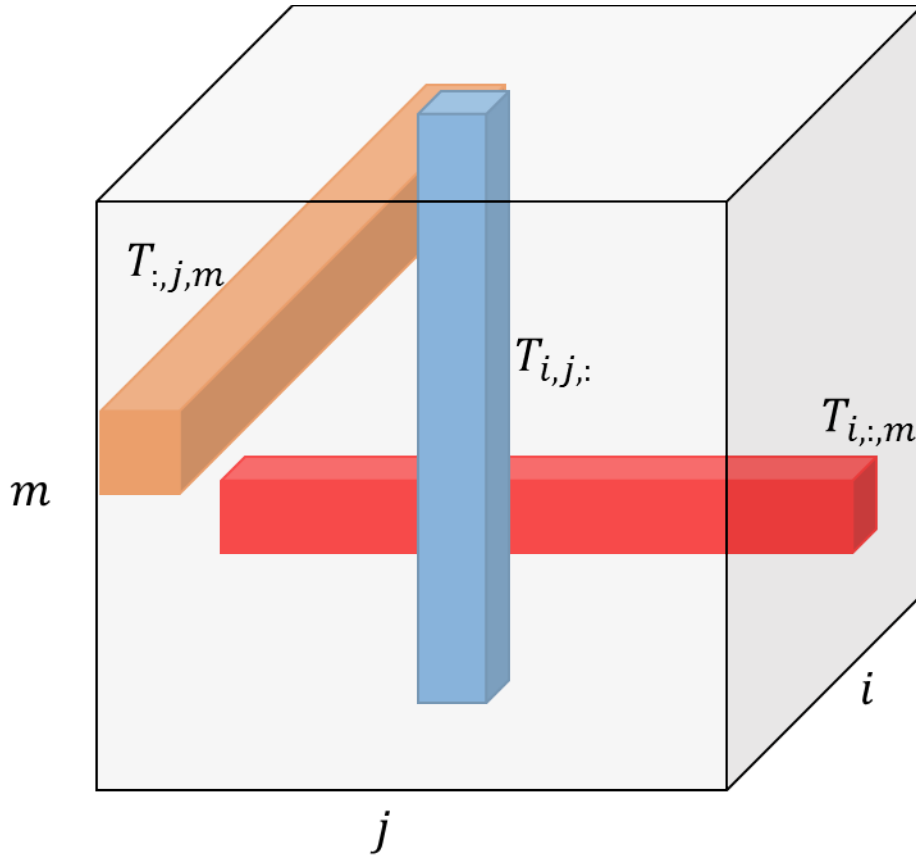


Figure 1.1 Illustration of the Travel Time Cube and its Atomic Vectors. The Axes Represent Origins ( $i$ ), Destinations ( $j$ ), and Minutes of the Day ( $m$ ). Each Location in this Three-Dimensional Matrix Contains a Shortest Path Travel Time from  $i$  to  $j$  at Time  $m$ .

While the mean and standard deviation of  $T_{i,j,:}$  are useful summary measures of average travel time and travel time variability, respectively, they are not sensitive to the periodicity in the time series. This means that OD pairs with vastly different levels of service caused by different vehicle headways may have similar means and standard deviations. **Figure 1.2** displays such a case. As one can observe, the time series for case A is characterized by hourly service throughout the daytime, while for B it is characterized by 30-minute service during peak periods and hourly service during the middle of the day. Despite these differences, the two time series have similar means and standard deviations of travel time. We introduce a measure based on Fourier Transforms (FT) in order to capture the dominant frequencies latent to the connectivity between each pair of locations in the region. Our method begins by decomposing a time series using the Fast Fourier Transform algorithm (Cooley and Tukey, 1965), which produces a distribution of frequencies that combine to reproduce the actual time series. We identify the mode of the frequency distribution and select all frequencies that appear with at least 75% of the frequency of the mode. This “short-list” of dominant frequencies is used to characterize the regularity of connectivity between locations. Importantly, we cannot use simple headway calculations from the schedule because our OD pairs consist of locations that are connected by multiple routes and routes that are composed of transfers. Using our FT-based measure, the differences in connectivity between the cases seen in **Figure 1.2** are numerically discernible.



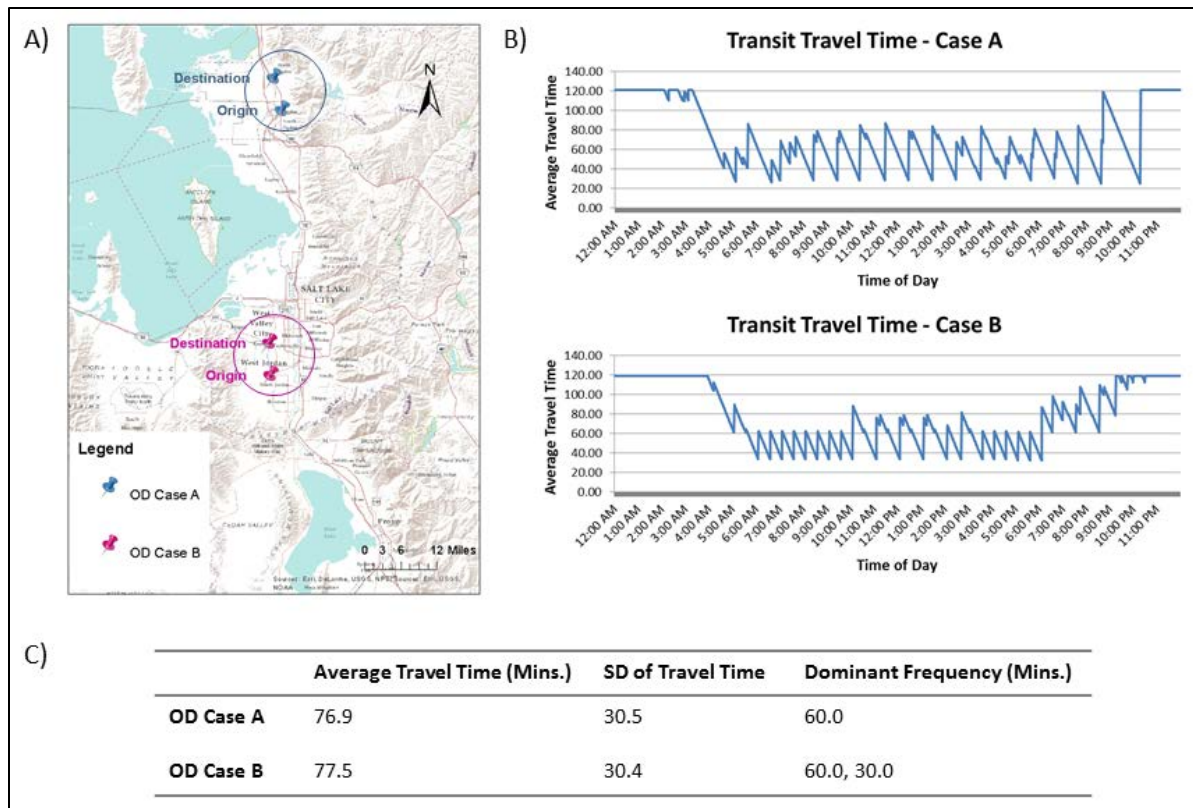


Figure 1.2 Using Fourier Transforms to Discriminate between OD Pairs with Similar Travel Time Moments but Different Latent Frequencies of Connectivity.

## 1.4 CASE STUDY ONE: COMPARISON OF TRANSIT TRAVEL TIMES OVER TIME

In this case study, we demonstrate how travel time cubes can be used to assess changes in travel times associated with changes to the provision of transit supply including network modifications (e.g., new routes, deletion of routes, and modification of routes) and level-of-service modifications (e.g., operating hours and headways). Such analyses are required to understand relationships between transit supply and mode share (Legrain, Buliung and El-Geneydy, 2015) or assessing whether social equity objectives are met (Foth, Manaugh and El-Geneydy, 2013). In this case, we conduct a retrospective analysis of travel time cubes for TriMet’s services in Portland, OR, and UTA’s services in the Wasatch Front, Utah. These two cases are selected because of the significant changes made in their networks, with TriMet cutting bus services by approximately 10% to accommodate for budget shortfalls following the Great Recession, and the UTA investing in both light rail (Trax) and commuter rail (FrontRunner) expansions. To conduct this case study, population-weighted census block group centroids form the spatial units in our travel time cubes, which are computed over a typical weekday’s 24-hour period in each time period for each city.

**Figure 1.3** presents route maps of the two systems at two different time points. These simple visualizations of the transit supply are neither useful in communicating overall changes in supply or the impact of these changes on travel times. As an alternative, the descriptive summary

of the GTFS data found in **TABLE 1.1** quickly informs on the scale and types of changes in supply. It contains a summary of transit supply in each region, at each point in time (2011 to 2014 for UTA and 2009 to 2013 for TriMet), as encapsulated in official GTFS packages shared with us by the transit agencies. Our summary focuses on typical weekday services, which do contain differences to the Saturday and Sunday levels of service. For UTA, observe that the period is marked by a 144% increase in route-kilometers travelled via light rail (+8,210 km) and a 103% increase in commuter rail (+3,717 km). This was met by an 11% decline (-10,575 km) in bus kilometers travelled. This type of modal trade-off between operational expenditures during times of rail expansion is well documented in the literature (Grengs, 2002), and is subject to criticism for unjustly servicing the needs of rail users (typically higher income and White) at the expense of bus users elsewhere in the city (typically lower income and racially diverse).

In the TriMet region, bus route-kilometers travelled was reduced by 10% (-11,374 km). At the same time, the increase in other modes is quite small, with no change in commuter rail delivery and only a 152-kilometer increase in light rail. So unlike Salt Lake City, the big cuts to bus services were not compensated by increased capacity on other modes.

**Table 1.1 Numerical Summary of Changes in Service Delivery**

	Wasatch Front			TriMet		
	Bus	Light Rail	Commuter Rail	Bus	Light Rail	Commuter Rail
<u>Number of Stops</u>						
<i>Time 1</i>	6,415	31	8	7,106	127	5
<i>Time 2</i>	6,118	56	16	6,732	159	5
<i>Change</i>	-297	25	8	-374	32	0
<i>Percentage Change</i>	-5%	81%	100%	-5%	25%	0%
<u>Number of Routes</u>						
<i>Time 1</i>	123	3	1	82	6	1
<i>Time 2</i>	114	4	1	79	6	1
<i>Change</i>	-9	1	0	-3	0	0
<i>Percentage Change</i>	-7%	33%	0%	-4%	0%	0%
<u>Route KMs</u>						
<i>Time 1</i>	4,959	68	71	2,582	184	25
<i>Time 2</i>	3,683	105	144	2,572	194	25
<i>Change</i>	-1,276	37	73	-10	10	0
<i>Percentage Change</i>	-26%	54%	103%	0%	5%	0%
<u>Route KMs travelled</u>						
<i>Time 1</i>	96,072	5,683	3,594	108,921	19,823	750
<i>Time 2</i>	85,497	13,893	7,311	97,547	19,975	750
<i>Change</i>	-10,575	8,210	3,717	-11,374	152	0
<i>Percentage Change</i>	-11%	144%	103%	-10%	1%	0%

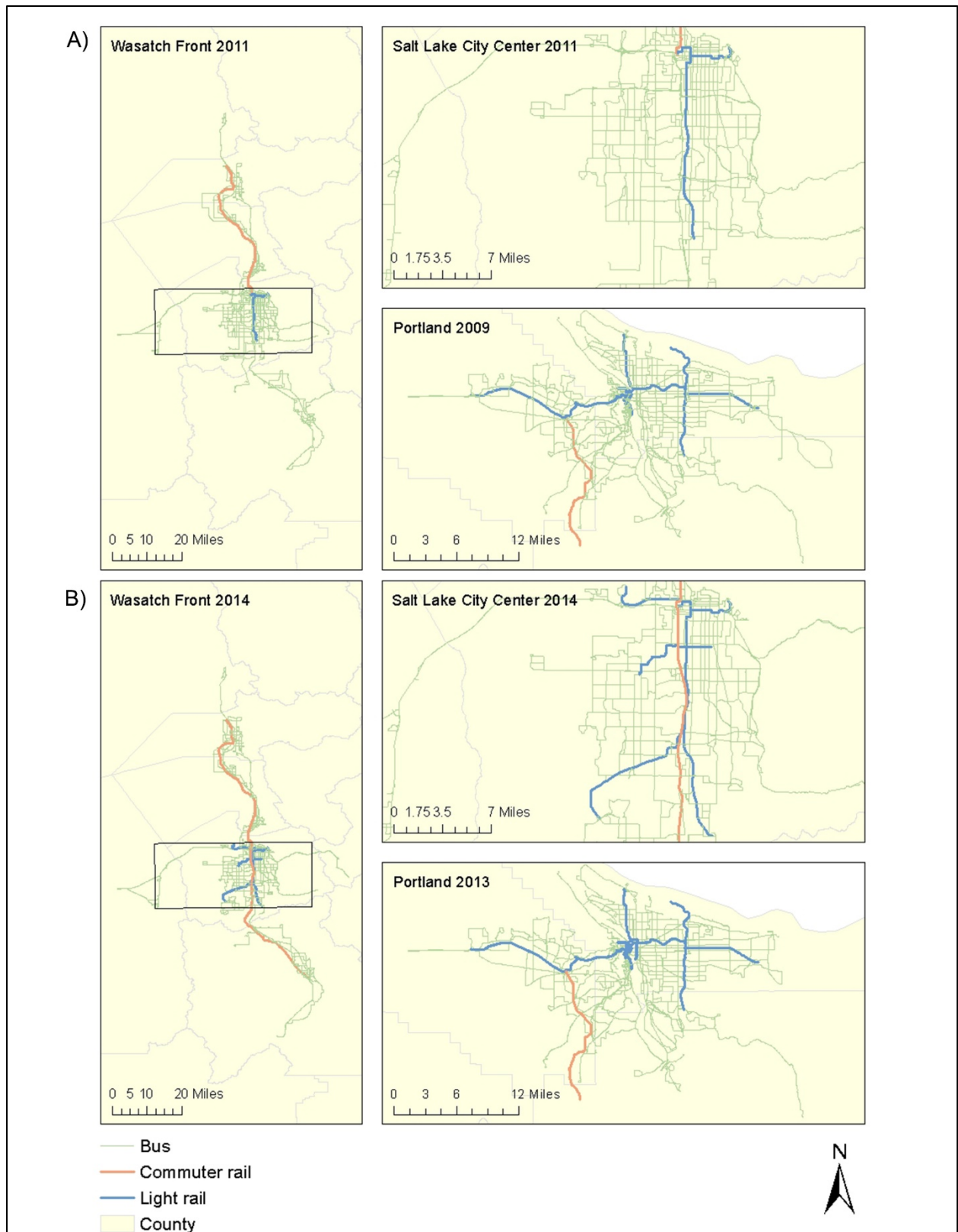


Figure 1.3 Cartographic Depiction of Changes in Service Delivery between a) Time Period 1 and b) Time Period 2

While numerical and (to a lesser extent) graphical summaries of transit networks are helpful in understanding the change in service from infrastructure- and operations-oriented points of view, they do very little to convey how these changes have affected travel times or accessibility in the region, arguably the most relevant ways to measure the distribution of benefits provided by the system. We therefore suggest the use of travel time cubes to summarize the impact of the changes on travel times in the region. **Figure 1.4A** displays the most aggregated type of summary measure of travel time cubes, the overall travel time average between all destinations in the region. We display this numerically and aggregated over all times of day:

$$\text{Overall Average Travel Time} = \left( \frac{\sum_{i,j,m} T_{i,j,m}}{N^2 M} \right) \quad (1-1)$$

and disaggregated over the course of the day as time series plots:

$$\text{Time Dependent Average Travel Time} = \left( \frac{\sum_{i,j} T_{i,j,m}}{N^2}, \forall m \right), \quad (1-2)$$

for the Wasatch Front (**Figure 1.4B**) and Portland (**Figure 1.4C**). These analyses yield some rather unique results. First, we can observe that the changes in service provision resulted in a 10-minute decline in average travel times in the Wasatch Front, and a 3.5-minute increase in Portland. Considering that these are averages over all OD pairs and at all times of day, the changes are actually quite large overall. The time series plots reveal that travel times in Utah were the most improved due to extended evening services and increased services during the morning rush. For Portland, the changes in travel times were more evenly distributed by time of day.

To delve deeper into the spatial patterns of travel time changes, we next explore zone-specific aggregations of the travel time cube and present these as maps. In this example, we hold an origin constant to produce the average travel time from that origin to all other destinations in the region, over all times of day:

$$\text{Average Travel Time from an Origin} = \left( \frac{\sum_{j,m} T_{i,j,m}}{NM}, \forall i \right), \quad (1-3)$$

and we keep that same zone constant as a destination to compute an average travel time from all origins to a specific zone:

$$\text{Average Travel Time to a Destination} = \left( \frac{\sum_{i,m} T_{i,j,m}}{NM}, \forall j \right). \quad (1-4)$$

These quantities can be mapped independently, but to determine how they have changed over the course of the years we map their percentage differences. So, for each location, we compute the percentage change in average travel time to and from that location and present this data in a bivariate choropleth map (**Figure 1.4D**). For the Wasatch Front, we see rather clear spatial patterns of travel time changes. Travel times to/from the southern areas were largely improved by the addition of commuter rail there. However, large swathes of the inner suburbs of Salt Lake City, along the eastern benches of the Wasatch Front and into the West Valley area, experienced increases in travel times, with these neighborhoods becoming harder to reach. Notably, the only places that became less accessible as an origin and a destination were affluent and touristic communities located within canyons heading east into the Wasatch Front. Transit services to

these locations are very limited, and the removal or addition of a single bus trip would have large ramifications on mean daily travel times to and from these locations.

The map of travel time changes in Portland (**Figure 1.4D**) tells a different story. Here, while changes overall are smaller, we do see that many places throughout the region experienced across-the-board travel time increases, or travel time increases from either the origin or destination perspective. The patterns are far less spatially clustered in comparison to the Wasatch Front, indicative of an overall reduction in service across the study area.

In addition to the changes in absolute travel times, we next turn our attention to the exploration of travel time fluctuations over the course of the day, and how changes in transit supply have impacted travel time variability from one time period to the next. To do this, we extract the most typical travel time cycle length for each OD pair using the Fourier Transform method described above. We then compute the mean cycle length for each origin in the region, and map the changes in average cycle length from one time period to the next. **Figure 1.5A** illustrates these changes cartographically. For the Wasatch Front, a clear pattern of improved (i.e., shortened) cycle length is observed along the north/south commuter rail corridors and toward the southwestern expanses of the Trax system. However, several corridors within the city have experienced decreases in service frequency, as seen in the inset map of the city center, in addition to some already inaccessible areas at the southern, western and eastern extremities of the region. These patterns loosely correspond to those found in **Figure 1.4D**, indicating that the changes in travel times are related to changes in service frequencies as well as the addition of new rail services in the region.

For Portland, the spatial pattern of increased travel time cycles are linear, with several axes corresponding to particular routes that experienced service cuts. On top of this, we do however observe large swathes of the city to the west and south of the center that have experienced decreases in cycle lengths, indicating a slight preference for suburban commuting over inner-city service provision.

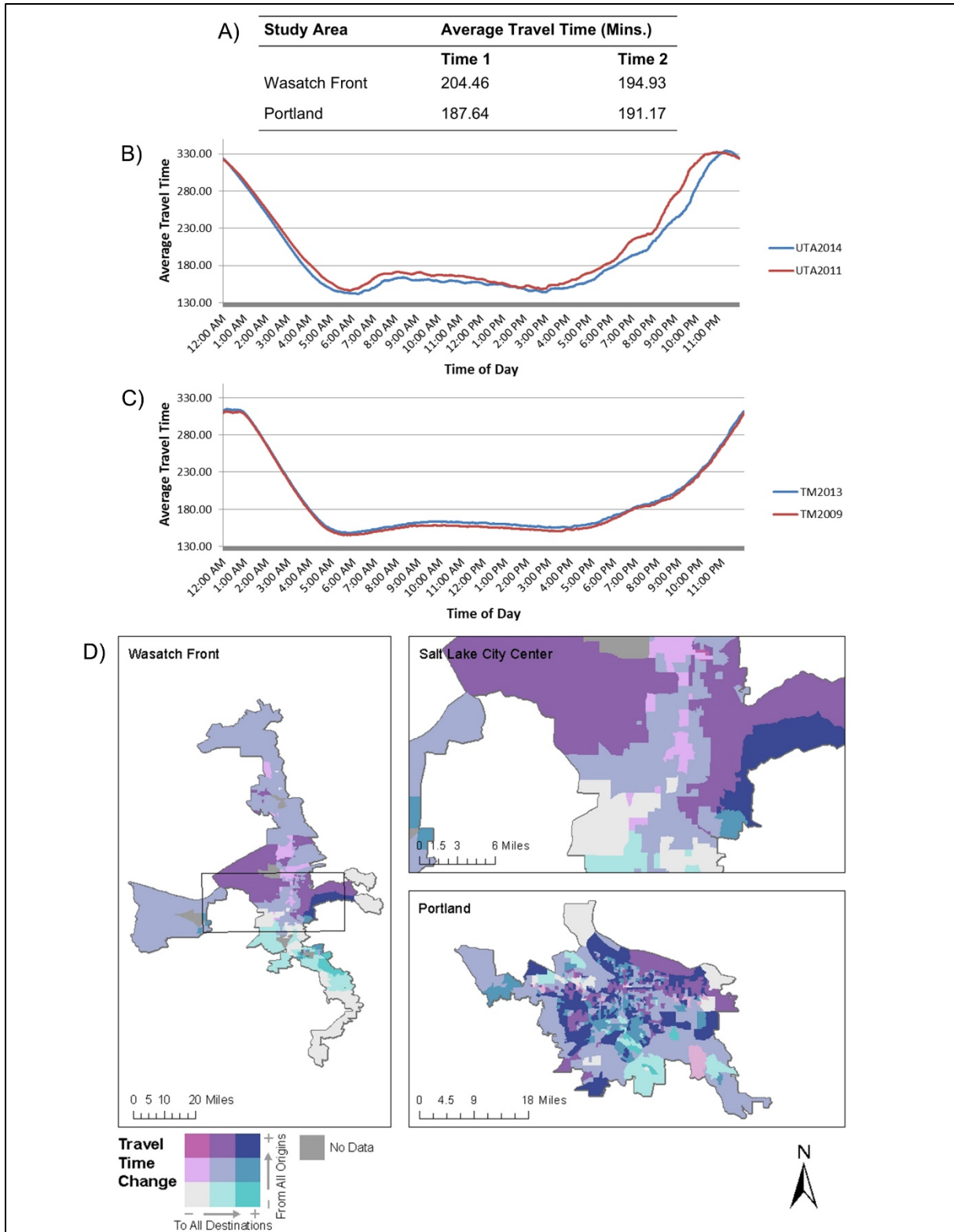


Figure 1.4 Depictions of Travel Time Changes, a) Numerically, b) By Time of Day for the Wasatch Front, c) By Time of Day for Portland, and d) By Change in Average Travel Times to and from All Zones in the City. Negative values represent decreases in travel time, while positive values are increases.

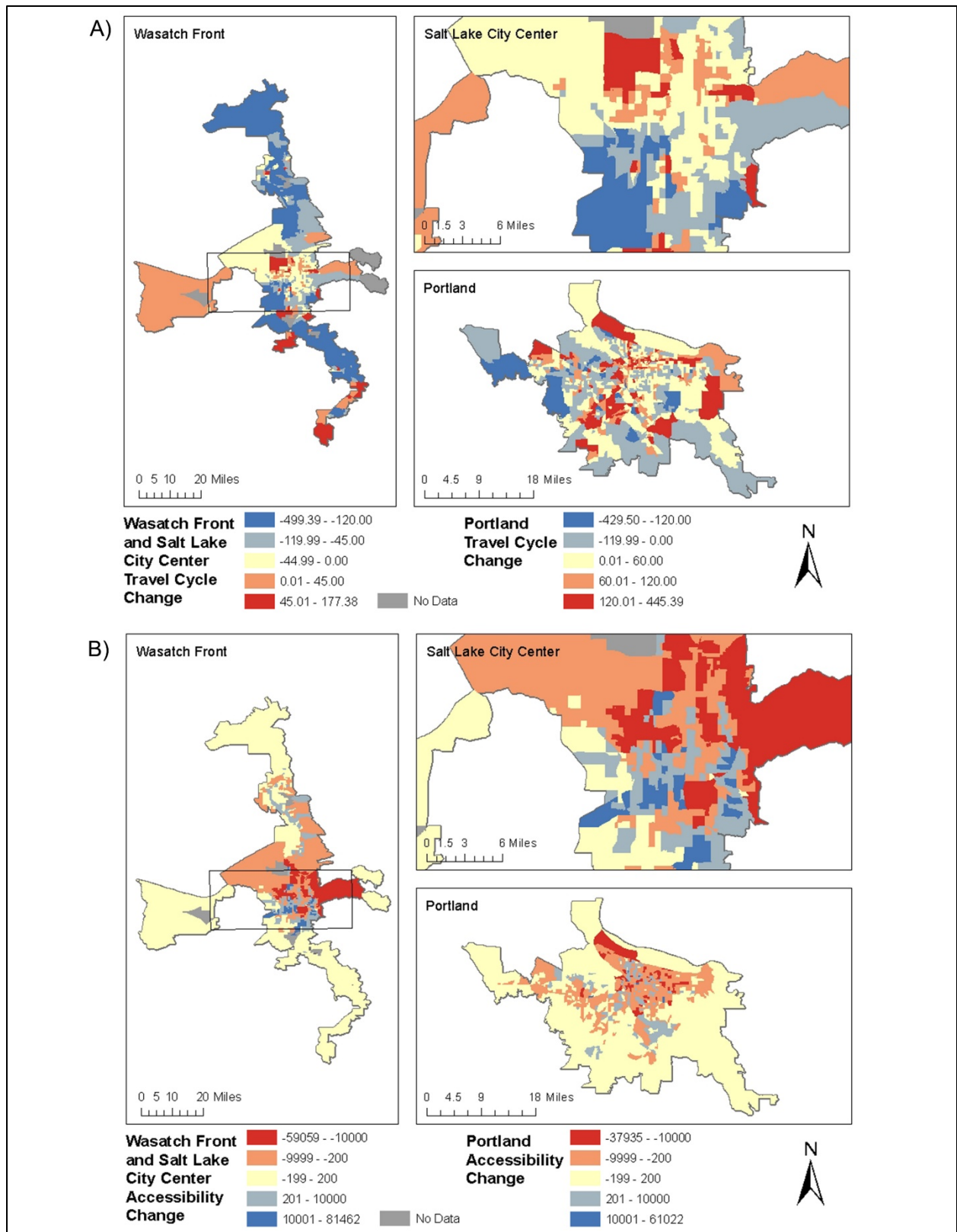


Figure 1.5 Change in a) Average Travel Time Cycle and b) Jobs Accessibility

## 1.5 CASE STUDY TWO: ASSESSING ACCESS TO JOBS USING TRAVEL TIME CUBES

In this case study, we use the travel time cube to assess how changes in travel times in each region impacted access to jobs. Cumulative opportunities measures of access to jobs are widely used in the literature to assess the distribution of transit benefits in a region (Foth, Manaugh and El-Geneidy, 2013; Páez and Farber, 2012; Sanchez, Shen and Peng, 2004) and to predict travel behavior (Owen and Levinson, 2015). Employment counts at the block-group level were acquired from the Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) data. A single jobs count file for 2011 was used for both time periods in order to control for accessibility differences caused by transit supply dynamics alone. At the time of analysis, this was the most recent vintage of the dataset available to the public. The LODES dataset has been used extensively over the past few years in accessibility research (Horner and Schleith, 2012; Owen and Levinson, 2015), and the paper by Spear (2011) provides a thoughtful overview of the pros and cons of its use in many aspects of transport planning in the United States, with the major issues including the absence of military jobs and potential miscoding of job locations for multiple-location employers.

Our accessibility score is computed as a cumulative opportunities measure within a 60-minute transit travel time buffer. This single threshold is selected for the sake of demonstration, while a more thorough analysis of accessibility using travel time cubes should always compare multiple thresholds depending on the normative values in the study region or specific population focus (Páez, Scott and Morency, 2012; Farber, Morang and Widener, 2014). For each block group centroid, we calculate the total number of jobs within a 60-minute journey at each minute of the rush-hour period (7-9 a.m.), and then compute an average accessibility score for each zone by averaging over the 120 minutes. This score is more sensitive to variabilities in accessibility that are introduced by fluctuations in travel times, something that is not captured in single-point-of-time estimates of accessibility, but which may be very large, depending on the level of transit service. To be precise, the accessibility measure for zone  $i$  is computed as:

$$A_i = \frac{\sum_{j,m} f(T_{i,j,m}) E_j}{M} \quad (1-5)$$

where  $E_j$  is the count of employment in zone  $j$ ,  $f(T_{i,j,m}) = \begin{cases} 1 & \text{if } T_{i,j,m} \leq 60 \\ 0 & \text{Otherwise} \end{cases}$ , and all other terms are defined as above. **Figure 1.5B** illustrates the changes in jobs accessibility between the two time periods. Accessibility was clearly impacted by the changes in transit service provision, with the population-weighted mean number of jobs accessible dropping from 35,694 to 33,243 in the Wasatch Front, and 23,410 to 21,666 in Portland. These figures are quite disturbing, considering the supposed increase in transit service provision by the UTA. The spatial patterns indicate that the expansion of commuter rail did little to improve the jobs accessibility in Utah County (towards the south) using the threshold of 60 minutes. At the same time, the removal of bus services in the center of the city had a clear negative impact on jobs accessibility throughout much of the city. The locations that experienced increases in jobs access are within very close proximity of the various light rail Trax expansions in the region. Combined, this indicates that the short-term effect of transit expansion on jobs accessibility puts light rail ahead of commuter rail.



## 1.6 CASE STUDY THREE: THE EFFECTS OF BICYCLE USE ON THE LAST MILE PROBLEM

The “last mile problem” in public transportation planning concerns the distance between the transit network and actual trip endpoints that are often too long to effectively accommodate by walking. Bicycle use is suggested as a potential solution to this issue due to its low cost and faster speed of travel (Shaheen, Guzman and Zhang, 2010; Liu, Jia and Cheng, 2012). Yet, there is little previous work that attempts to evaluate the impact of joint bicycle and transit use on reducing travel times and increasing travel time reliability. Travel time cubes are well positioned in the use of evaluating the latter, as they contain the full-day time-series of travel times from origins to destinations. In this case study, we compare travel time patterns using two travel time cubes. The first, as previously defined, assumes a 4.8 km/h walk speed for ingress, egress and transferring. The second assumes a 15 km/h bicycling speed for ingress, egress and transferring. In addition, due to increased boarding and alighting delays associated with travelling on transit with a bicycle, we double these delays from 15 to 30 seconds. Descriptive statistics of the travel time cubes are compared between scenarios in an attempt to quantify the time savings and increased stability associated with combining bicycle use with public transit.

Upon initial investigation of the results, we discovered that the major benefits associated with travel using the bicycle accrued to those with trip endpoints (i.e., origins or destinations) at great distances from transit stops. Finding this a somewhat banal example, we focused our analysis on the particular case of reaching destinations in downtown Salt Lake City by assuming a starting point at Salt Lake Central Station, where commuter rail, light rail and major bus routes have termini. In this way, we can evaluate how bicycle use helps arriving passengers in the central city reach their final destinations, which may require lengthy walks, and/or long waiting and in-vehicle travel times in the case of transfers.

Before delving into the downtown focus, we discuss the aggregate results pertaining to the two study areas found in **Figure 1.6**. First, we observe that cycling results in a 37% reduction in mean travel times throughout the 24-hour period in both study areas. Moreover, stability in travel times is increased via the introduction of the bicycle, with a 42% reduction in standard deviation found in Portland and a 39% reduction in Salt Lake City. The improvements in travel times tend to be consistent over the course of the day (as seen in **Figure 1.6A**), with slightly larger effects starting in the evening and lasting until the middle of the night. This is presumably due to the decreased transit service at night, and therefore a large increase in travel speed using a bicycle compared to trips composed of more walking and waiting. From a spatial perspective, the travel time savings are clearly most attributed to more peripheral origins, where long walking ingress times are most significantly reduced by faster bicycle speeds (**Figure 1.6C**). Interestingly, the impact of bicycling on travel time stability (**Figure 1.6D**) is found to have a donut shape, with higher impacts in the city center, followed by lower impacts in the inner suburbs, and then again higher in the outer periphery. The fact that stability is enhanced for trips in the center of the city indicates that the bicycle may truly be of benefit to the last mile problem there.

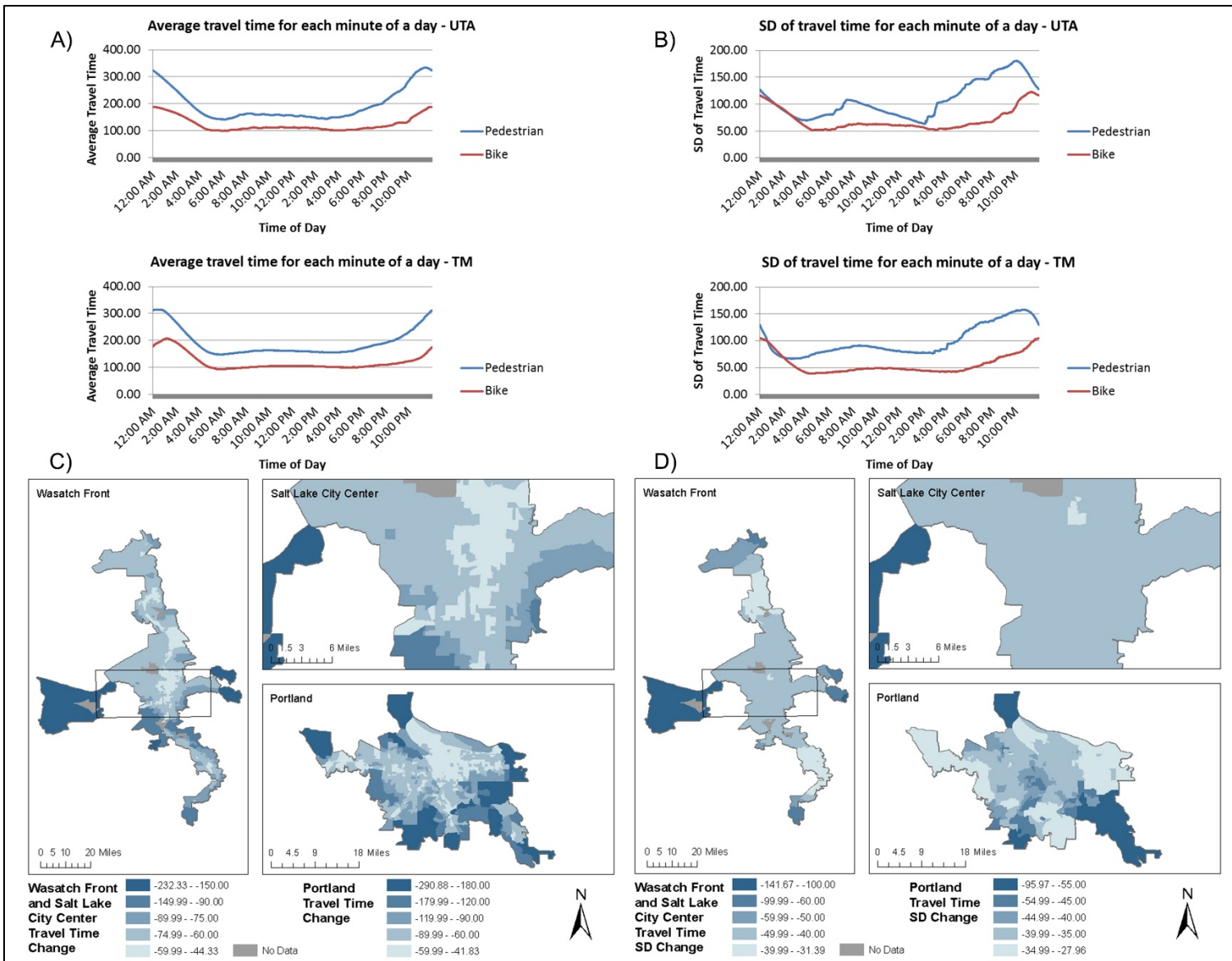


Figure 1.6 A Comparison of Transit Travel Times Joint with Walking Versus Cycling. Average Travel Times are Depicted in A and C, while Standard Deviations of Travel Times are Depicted in B and D

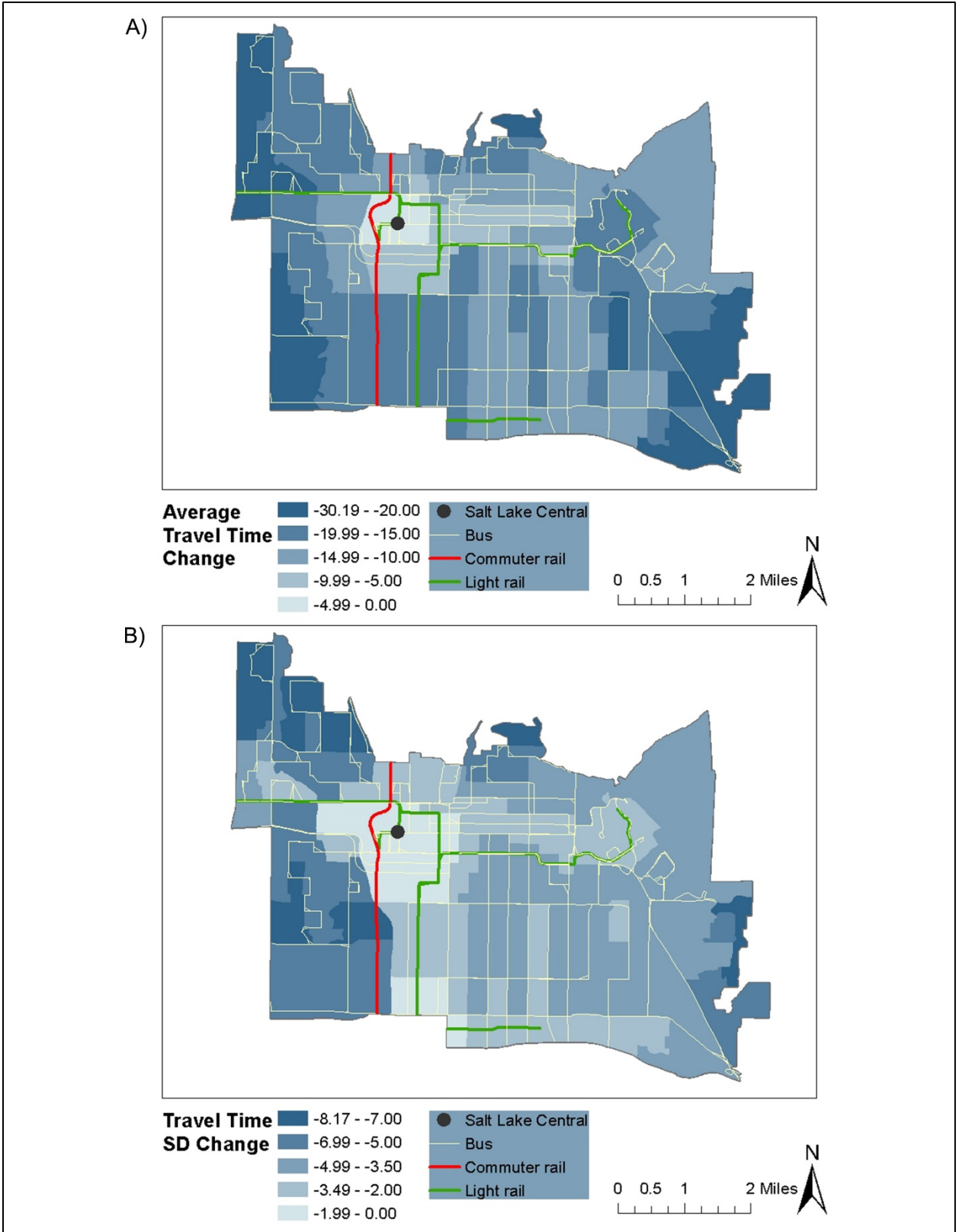


Figure 1.7 Transit Travel Time Improvements in the Center of Salt Lake City Associated with Bicycle Use.

Next we discuss the results pertaining to passengers arriving downtown at the intermodal hub, Salt Lake City Central Station. The set of destinations for these analyses are limited to those found in **Figure 1.7**, representing an area bounded by major municipal, land use and transportation barriers. This region comprises 9% of the region's population and 18% of its jobs. As seen in the figure, when replacing walking with cycling, travel from the intermodal hub to the rest of the study area is greatly improved. This is indicated by a 32% reduction in travel times within the morning rush hour (7-9 a.m.) and 47% reduction in standard deviation of travel time. The spatial patterns of travel time and variation improvements are quite similar, with increasing benefits accruing to farther away destinations. From the perspective of travel time reductions (**Figure 1.7A**), we see that the concentric pattern of improvements is modified by the transit network, where areas of sparser network coverage receive additional benefit from bicycle use. The same is true for travel time stability (**Figure 1.7B**). In both cases, travel to the west is more improved than the east, reflective of the relative share of transit routes travelling east versus west from the intermodal hub.

## 1.7 DISCUSSION AND CONCLUSIONS

In this paper we put forward a new data object called the public transit travel time cube and demonstrated its use in three case studies that evaluate how different types of changes to the public transit network impact a variety of travel time characteristics in a region. The novelty of this approach does not come from the travel time computations themselves, but rather the exhaustive precomputation of all possible travel times in the region, at all times of day. With advances in the abilities of commercial geographic information systems to compute transit-based travel times, coupled with the prevalence of cloud computing and off-the-shelf multiple core desktop computers, computing, storing and analyzing travel time cubes is only now a truly feasible endeavor. Through a series of case studies, this paper shows how travel time cubes can unlock new ways to describe spatiotemporal patterns of travel times, use these travel times in the development of more sophisticated measures of accessibility, and help quantify the impact of accommodating joint bicycle and transit trips.

Going forward, we foresee many directions for continued research on travel time cubes. First, from a technological perspective, we are in need of a custom-designed visualization tool for rapid and interactive exploration of the travel time cube. Collaboration is already underway with Miriah Meyer, a computer scientist and visualization researcher at the University of Utah, on the development of an open source visualization tool for travel time cubes.

Second, algorithmically, faster computation of travel time cubes would allow for more detailed comparisons of transit planning scenarios. We currently compute an OD cost matrix for every minute of the day, without any interaction between computations. It is likely that knowing the cost matrix for minute  $m$  will be useful in the computation of the cost matrix at minute  $m + 1$ , yet our algorithms do not yet exploit these structural dependencies. Interestingly, this dependence turns out to be very useful in the compression of data cubes for storage, as each OD time series can be losslessly stored as a series of coordinate plots with linear splines, but we are not exploiting this serial dependence in the cube's computation.

Third, the travel time cube only stores a single characteristic of public transit travel, travel time, yet other characteristics of each trip in the cube, such as the number of connections required, or a disaggregation of travel times by walking, waiting and in-vehicle times, may be very useful in future analyses of accessibility, mode choice modeling or social equity. This

would require a shift from the current GIS-based algorithms for computing travel times, to more specialized algorithms that can retrieve and store multiple characteristics of the computed shortest path in addition to travel time, essentially requiring an additional cube for each characteristic.

The aim of this paper was to introduce the reader to the vast potential of the development and analysis of public transit travel time cubes, but these demonstrations were not exhaustive. Elsewhere, travel time cubes have also been shown to be effective in broadening the investigation of food deserts into the temporal domain (Farber, Morang and Widener, 2014), understanding spatiotemporal mismatch between transit demand and supply from a social equity perspective (Ritter, 2014), and proving that temporal variation in transit-based jobs accessibility impacts one's mode choice decision (Owen and Levinson, 2015). The latter case is especially useful in pointing to future research in which travel time variability is more broadly used to identify markets of untapped potential transit riders, an objective shared by many transit agencies worldwide.



## **2.0 SPACE-TIME MISMATCH BETWEEN TRANSIT SERVICE AND OBSERVED TRAVEL PATTERNS IN THE WASATCH FRONT, UTAH: A SOCIAL EQUITY PERSPECTIVE**

### **Steven Farber**

Department of Human Geography, University of Toronto Scarborough  
1265 Military Trail, Toronto ON, M1C 1A4, Canada  
Tel: (416) 208-5120 Fax: (416) 287-7283; Email: steven.farber@utoronto.ca

### **Benjamin Ritter**

The Travelers Group  
Email: britterj@gmail.com

### **Liwei Fu**

City of San José  
200 E. Santa Clara St., San José, California, 95113  
Email: folflw@gmail.com

*A version of this paper has been published in the peer-reviewed Journal of Travel Behaviour and Society:*

Farber, S., Ritter, B., & Fu, L. (2016). Space–time mismatch between transit service and observed travel patterns in the Wasatch Front, Utah: A social equity perspective. *Travel Behaviour and Society*, 4, 40-48.

## **2.1 ABSTRACT**

In the absence of other alternatives, people who rely on public transportation to conduct their daily activities have travel patterns that differ from discretionary transit users, especially those who choose to use transit for work trips. At the same time, in many regions around the world, public transportation is primarily designed to accommodate peak-hour travel demands in order to reduce congestion and its impacts. It is theorized that this results in a mismatch between the demand and supply of public transportation amongst populations at risk of social exclusion. In this research, we characterize and compare the spatiotemporal patterns of travel demand and transit supply. Our analysis consists of a comparison between observed travel patterns and a new temporal measure of transit supply based on travel times. We measure travel demand with the observed trip-making characteristics (i.e., origin, destination, time-of-day) of the respondents to two transportation surveys conducted in Utah. Transit supply is characterized using a transit travel time cube, a three-dimensional array of origin-destination transit travel times computed for all origins, destinations and times of day. Mismatch is examined by descriptive and multivariate comparisons of observed trips and computed levels of transit provision. Our results confirm this theory: More marginalized groups demand travel between locations at times of the day that are

poorly served by transit. However, when controlling for all variables simultaneously in a multivariate regression, few socioeconomic factors remain significant, indicating the overall importance of employment status, making work trips and travelling during peak times, in explaining mismatch.

## 2.2 INTRODUCTION

In an urban context, a just society includes equitable access to public transportation (Martens, 2009; Martens, Golub and Robinson, 2012; Golub and Martens, 2014). One way to achieve equity, vertical equity in particular, is to provide transit service to those people who need it most, where need is most often assessed using considerations of socioeconomic status (Litman, 2002; Litman and Brenman, 2011; Bullard, Johnson and Torres, 2004). Of course, transportation planning incorporates a complex set of technical and political processes, and social equity is but one consideration in a multi-objective agenda that has often favoured more readily measurable, predictable and pecuniarily expressible outcomes of transport models like travel-time savings, congestion and throughput (Deka, 2004). As according to the adage, “we build what we measure,” transport planning has focused on achieving increased mobility rather than explicitly increasing accessibility or the equitable distribution of the accessibility benefit between modes (Martens, Golub and Robinson, 2012; Benenson et al., 2011; Golub and Martens, 2014; Kaplan et al., 2014), across space (Martens, Golub and Robinson, 2012; Welch and Mishra, 2013), and between population groups (Welch, 2013; Delbosc and Currie, 2011). As a result, some public transportation systems fall short of meeting the needs of those who depend on transit to participate in daily activities, putting people at risk of transport-related social exclusion (Church, Frost and Sullivan, 2000; Lucas, Grosvenor and Simpson, 2001; Hine, 2003; Kenyon, 2003; Raje, 2004; Kenyon, 2006; Preston and Rajé, 2007; Páez et al., 2009; Lucas, 2012).

Most studies of social equity and public transit accessibility entail an aggregate comparison between transit need and transit supply. In these studies, transit need is often established over space as a measure of socioeconomic status in neighborhood units while transit supply is typically measured at the neighborhood level as the ease of reaching transit facilities (Murray et al., 1998; O'Neill, Ramsey and Chou, 1992; Moniruzzaman and Páez, 2012), reaching transit facilities weighted by level of service (Rood and Sprowls, 1998; Henk and Hubbard, 1996; Ryus et al., 2000; Hunter-Zaworski, 2003; Drew and Rowe, 2010; Al Mamun and Lownes, 2011; Currie, 2010), or reaching actual destinations with transit (O'Sullivan, Morrison and Shearer, 2000; Lei and Church, 2010; Foth, Manaugh and El-Geneidy, 2013; Páez and Farber, 2012; Farber, Morang and Widener, 2014). One drawback of these studies is that transit need is poorly characterized by zonal population characteristics since different population groups demand travel to different types of destinations at different times of the day. Similarly, accessibility is poorly characterized by generalized measures of station access or destination access since travel times to destinations can be highly variable depending on time-of-day variations in schedules. Ignoring the temporal fluctuations in activity patterns and travel times makes it difficult to know whether the transit services that are “provided” are actually what is “needed” by different population groups at different times of day. In an effort to address temporal variations in transit access, Polzin, Pendyala and Navari (2002) conduct a time-of-day mismatch study of aggregate travel demand by computing the percentage of origin-destination flows in a region that could be feasibly met by the current provision of public transit. They consider feasibility as a threshold of acceptable wait time at the origin, and



frequency of service at the destination. Although their approach measures temporal mismatch in terms of total travel demand in a region, they make no attempt to further explore the distributional aspects of this mismatch between social groups. In fact, few equity studies have considered the unique spatiotemporal signatures of transit supply and travel demand for different population groups, yet doing so greatly increases the validity of the analysis and could lead to policies that more effectively increase equity in transit provision (Farber, Morang and Widener, 2014; Owen and Levinson, 2014; Ritter, 2014). The purpose of this study is to investigate how well public transit matches the spatiotemporal travel patterns for different population groups in the Wasatch Front, Utah. To accomplish this task, we: a) characterize travel patterns using observed trips from household travel and onboard passenger survey data; b) put forward a measure of spatiotemporal transit service based on origin-to-destination travel times; and c) determine whether socioeconomic status is associated with travel demands that are spatiotemporally mismatched with transit supply in the region.

The rest of the paper is organized as follows: In the next section we describe the travel time cube and how we use it to create a temporally dynamic measure of transit service. We also put forward our empirical analysis plan, and provide a description of our study area and datasets. In the third section we present and discuss the results of our descriptive and multivariate analysis of transit mismatch. We conclude the paper in Section 5 with a brief summary of results and a discussion about policy implications, and we propose several avenues for future research.

## 2.3 METHOD

### 2.3.1 The Public Transit Travel Time Cube

We propose a new data object, the public transit travel time cube, which can be used to establish spatiotemporal signatures of transit service in a region. The travel time cube is a three-dimensional array  $T = [t_{i,j,m}]$  where  $t_{i,j,m}$  is the shortest public transit travel time from location  $i$  to location  $j$  at time  $m$ <sup>1</sup>. In this case,  $i$  and  $j$  index population-weighted block group centroids in the region, and  $m$  is used to index the minutes in a day. So, for example,  $t_{4,10,480}$  is the travel time from block group 4 to block group 10 with a departure time of 8 a.m. (the 480<sup>th</sup> minute in the day).

In practice, the cube is computed in a GIS making use of a pedestrian network file (to model ingress and egress times), and a transit network and schedule stored as a General Transit Feed Specification (GTFS) package<sup>2</sup>. An *Esri ArcMap* plugin named *Add GTFS to Network Analyst* is used to create a routable multimodal *Network Dataset* and custom travel time evaluators which enable the use of many *Esri ArcMap Network Analyst* functions. We use the *Esri OD Cost Matrix* tool to compute shortest path travel times from centroid to centroid in the region, and custom Python scripts are employed to process the computational workflow of iterating cost matrix computations over every start-time minute of the day.

---

<sup>1</sup> The travel time recorded is the shortest path found in the multimodal network at a particular time of departure. It includes ingress, egress, waiting and transfer time associated with the fastest trip. If the shortest travel time is found by walking only, then the walking only trip time is recorded in the cube.

<sup>2</sup> Because of this, our transit travel time cube is based on scheduled travel times, and is not sensitive to real-world service disruptions or congestion. Future work investigating real-time or historical vehicle location data is one possible extension of the travel time cube research thread.

Similar data objects built with tools from Esri and other developers have been used elsewhere in the literature (Owen and Levinson, 2014; Lei, Chen and Goulias, 2012; Farber, Morang and Widener, 2014).

For our case study, we used the Utah Transit Authority (UTA) GTFS data to create travel time cubes for a typical weekday, Saturday and Sunday. The particular GTFS package used for this research consisted of service dates ranging from August 19<sup>th</sup> to December 7<sup>th</sup>, 2013 and included 122 transit routes, 6,202 transit stops, and 7,472 transit trips. Our study area contains 1,326 block groups, resulting in  $1326 \times 1326 \times 1440 \approx 2.5$  billion uniquely computed shortest path travel times per cube. Given the volume of computations and the ensuing data storage demands, we employed distributed processing in a windows ArcGIS environment to speed up the runtime of our computations. The study area, seen in Figure 2.1, was trimmed by excluding some peripheral block groups that either had no transit service, or only very specialized services for accessing ski resorts and distant urban settlements. Also, the use of population weighted centroids ameliorates the effects of varying block group sizes, especially at the periphery of the study area.

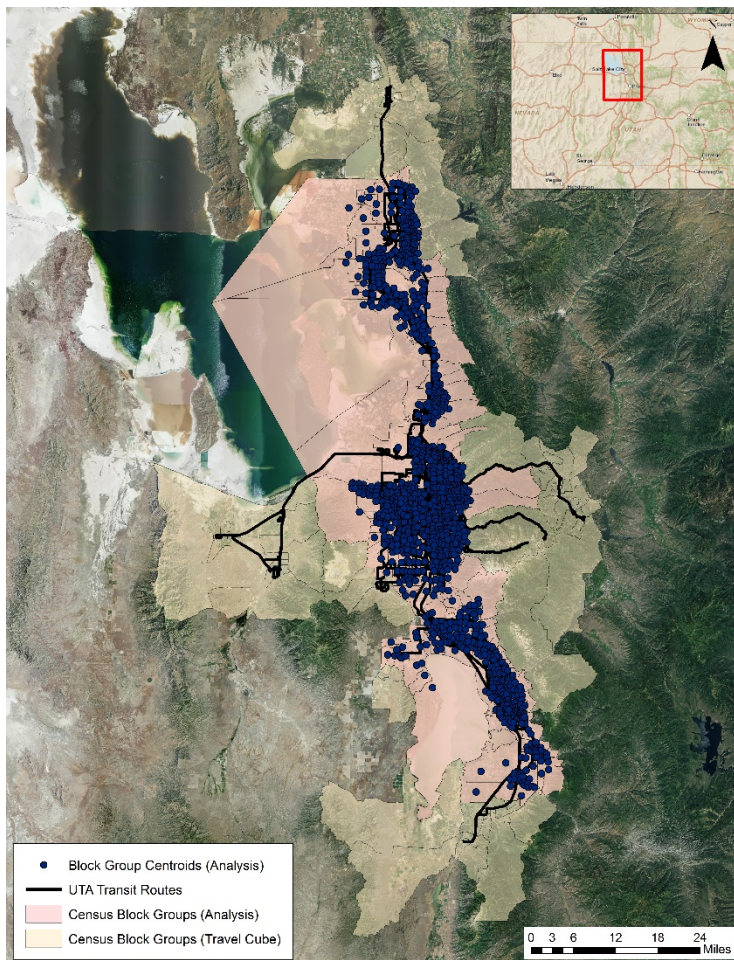


Figure 2.1: Wasatch Front Study Area Block Group Centroids

### 2.3.2 Spatiotemporal Measures of Transit Service

The individual trip records from the two surveys (see Section 2.4) were combined with the travel time cube in order to construct service measures for each observed trip. First, the origin and destination (OD) of each recorded trip was associated with an OD pair in the travel time cube based on a point-in-polygon assignment. Next, an average transit travel time for an hour-long period straddling the observed trip departure time was extracted from the travel time cube. The average travel time within a one-hour period is assumed to be representative of the transit service provided at the time of each recorded trip; this moving average is less sensitive to errors in trip time recording that may significantly impact the travel time extracted from the cube. The selection of a one-hour buffer (30 minutes before and after the observed trip) smoothes the travel time quite substantially (as seen in Figure 2.2). A sensitivity analysis found that a 15-minute buffer on each side obtained nearly identical travel time results (RMSE=1.13 minutes) and a 5-minute buffer on each side obtained a RMSE of 4.90 minutes. The degree of “smoothing” error for a given buffer is associated with the frequency of service and how rapidly frequencies of service change over the course of the day. While the size of the smoothing window we choose may impact the results slightly, we are more comfortable with a higher degree of smoothing than the potential for gross over- or underrepresentation of travel times associated with a buffer too small. A full sensitivity analysis of the use of different window thresholds is recommended for future research.

Next, we compare the local average travel time (i.e., within 1 hour) of the observed trip to the global average (i.e., across the entire day) to determine whether trips took place during periods of relatively high or low travel times. Since travel times for OD pairs are subject to vastly different magnitudes and scales of variation, the service measure associated with each observed trip is standardized into a Z-score as follows:

$$Z_q = \frac{X_q - \mu_q}{\sigma_q} \quad (2-1)$$

where  $X_q$  is the local average travel time for trip record  $q$ ,  $\mu_q$  is the global average trip time across the entire day for that OD pair, and  $\sigma_q$  is the standard deviation of trip times across the day for that OD pair. Thus,  $Z_q$  is an indicator of whether a trip is taken during a time period in which service frequencies are relatively higher or lower for that OD pair. Positive Z-scores indicate higher travel times at the time that an observed trip was made, while negative Z-scores indicate below average travel times. In addition to this measure, we similarly compute Z-scores based on travel speed by dividing the travel times in the travel time cube by the Euclidean distance between origins and destinations. Travel speed, calculated in this manner, is a measure of how well the transit network services the particular OD pair, regardless of distance traveled.

Figure 2.2 illustrates how the Z-scores are computed and interpreted. In this example, a pair of bus stops is connected by a single bus route. The in-vehicle journey time is fixed at 20 minutes. Further assume that the walking time between origin and destination is fixed at some duration greater than 49 minutes. During peak hours (7-9 a.m. and 4-6 p.m.) a bus departs every 10 minutes, and every 30 minutes at all other times. The blue time series depicts the estimated travel time (waiting time plus in-vehicle time) for the OD pair over the course of the day, from 6 a.m. to 10 p.m. The best case travel time is 20 minutes, when a journey begins at the exact moment of an arriving bus. The worst case occurs during the off-peak times when a journey begins immediately following the departure of a bus. In this case, the maximum estimated travel time is 49 minutes (29 minutes of waiting plus 20 minutes of in-vehicle time). The mean travel

time,  $\mu_q=32$  minutes, is depicted by the red horizontal bar, and the standard deviation of travel times,  $\sigma_q$ , not depicted graphically, is just less than 9 minutes. The yellow curve depicts the 60-minute moving average travel time,  $X_q$  from the equation above. The Z-scores for different times of the day are similarly depicted by the yellow curve, with their values appearing on the secondary vertical axis. As can easily be seen, when an individual makes a trip when service frequency is higher, the associated Z-score for travel time will be more negative. Similarly, during periods of relative poorer levels of service, Z-scores will be more positive. In reality, OD travel times, especially those connected via multiple routes or involving transfers, will not vary as regularly as the example provided. Using our methodology, we are able to discriminate between irregular fluctuations in travel times, not just those simply captured in a peak/off-peak dichotomy.

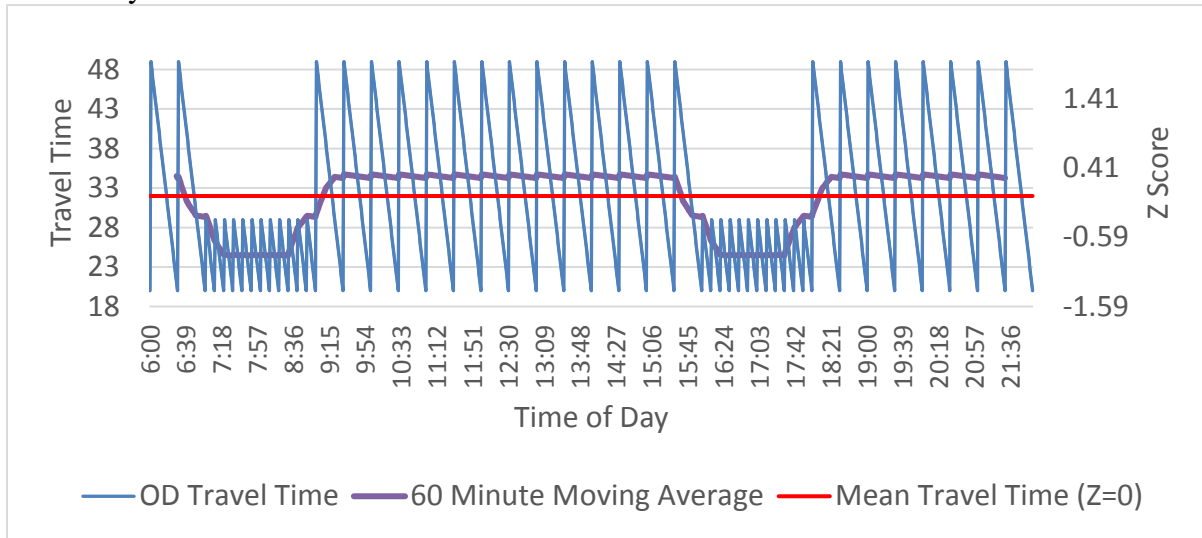


Figure 2.2: Illustration and Interpretation of Level of Service Z-scores

One additional note about the creation of Z-scores must be made. The travel time standard deviation,  $\sigma_q$ , equals 0 for a small number of observed trips that rely on walking only. These trips were excluded from the analysis since they have undefined Z-scores.

### 2.3.3 Comparing Travel Patterns to Transit Service

Two methods are used to determine whether there are systematic relationships between trip-maker characteristics and the supplied transit service at the time of their trips: ANOVA and multivariate ordinary least squares (OLS) regression. All available factors from the travel surveys were used in our analysis. These include person, household and trip-level characteristics (See Table 2.2 and Table 2.3 for a complete list of variables). In an effort to study service disparities by socioeconomic status, if socioeconomic status is found to have a significant and positive (negative) relationship with the standardized travel time (travel speed) Z-scores, our hypothesis that transit supply is less likely to meet the needs of those in need is supported. Conversely, if transit provision does not occur in synch with when less affluent people travel, then we could conclude that transit is less likely to meet their needs compared to meeting the needs of the affluent.

### **2.3.4 Data Used to Indicate Travel Patterns**

The Utah Household Travel Survey (UHTS) and the UTA Onboard Survey (UOBS) were used to indicate the spatiotemporal travel patterns for different social groups. These surveys are quite different in scope and sampling method. The UHTS is a state-wide household travel survey typical of those collected for U.S. metropolitan regions. The one-day trip survey recorded 101,404 trips taken by 27,046 individuals living in 9,155 randomly selected households across the state of Utah. The survey, being partially funded by Metropolitan Planning Organizations along the Wasatch Front, oversampled the urbanized regions of the state in which UTA operates. This provides a spatially dense sample of households considered to be within the operating district of the transit authority. The UHTS collected information about individual trips, the trip-makers and their households, and has been used as a data source in a number of peer reviewed articles appearing in the literature (Farber et al., 2014; Tian, Ewing and Greene, 2014; Burbidge, 2012; Liao, Farber and Ewing, 2015). Importantly, the observed trips include those made by all modes. This is in contrast to the trips observed in the UOBS, which only include those done by transit.

The UOBS was collected in 2011 with the goal of describing transit riders as well as calculating overall ridership levels on transit services offered by UTA. Respondents were intercepted randomly at transit stations and aboard transit vehicles. In total, 7,123 passengers provided usable responses which included some basic sociodemographic information, and detailed information on the particular trip being made when intercepted. Each respondent indicated whether or not a return trip was planned, and what the planned departure time of the return would likely be. Z-scores were created for these return trips and added to the analysis dataset despite not being observed directly.

The perspective taken in this paper is that the observed activity patterns, especially the timing of activities, represents an inherent desire to participate in activities at different times of the day, depending on the socioeconomic characteristics of the respondents. Amongst transit users, the above assumption is somewhat flawed, as a person will time the departure from their house with the scheduled arrival of a bus, but this type of timing adjustment occurs at a micro-scale, within the constraints of the route headway. The type of mismatch that we are investigating has to do with larger scale differences in the timing of activities (e.g., peak vs. off-peak, not 8 a.m. vs. 8:15 a.m.).

## **2.4 RESULTS AND DISCUSSION**

### **2.4.1 ANOVA**

The ANOVA results for the UHTS and UOBS datasets are presented in Table 2.2 and Table 2.3, respectively. The tables show the relationships between a host of independent factors and two measures of standardized transit service: transit travel time and transit travel speed. In both cases, the measures are standardized temporally against the mean travel times and speeds for each trip across the entire day. For example, in Table 2.2, observe that travelers from households earning over \$250,000 per year have a mean travel time Z-score of -0.124 while those earning less than \$10,000 have a mean travel time Z-score of 0.034. This indicates that those from high-income households typically have spatiotemporal travel demands during periods

of more frequent transit service than those in the low-income bracket. Similarly, for the same income categories, we see that travel speeds for high-income travelers have a Z-score of 0.156 compared to -0.056 for the low-income group. Thus, transit travel speeds for high-income travelers tend to be higher relative to the low-income riders. The significance of these group-wise differences are tested using one-way ANOVA F-tests, and are indicated with asterisks and plus signs for travel times and travel speeds, respectively.

The results indicate a large number of disparities in transit service by individual, household and trip-type characteristics. In terms of travel time, all variables are significant at least to the 0.05 level. For travel speed, all variables are significant except mobility status, Hispanic status and household income. We opt to further interpret the nature of these disparities based on the OLS regression results, where confounding factors can be suitably controlled. However, the reader can confirm that for most factors, socioeconomic disadvantage (e.g., lower income, Hispanic, only partial employment, higher rates of residential mobility, etc.) is associated with poorer levels of transit provision in terms of travel time and travel speed. Interestingly, the Z-scores for the on-board survey convey that transit trips are taking place during higher service periods as compared to the full set of trips observed in the UHTS. This is likely due to the rational decision by transit riders to travel during periods of better service. Be that as it may, there are still large transit service disparities between socioeconomic groups within the UOBS results. And, the fact that transit service provision is poorer for needier socioeconomic groups who are travelling by car (as observed in the UHTS data) may indicate that their temporal demands for activity participation are not being met by the temporally dynamic supply of transit services.

**Table 2.1: Utah Household Travel Survey (UHTS) ANOVA Tests for Travel Time and Trip Speed**

Factor	Level	Mean Time Z	Mean Speed Z	Factor	Level	Mean Time Z	Mean Speed Z
Res. Type	Single Family	-0.091	0.109	Mobility	Limited Mob.	-0.128	0.124
***	Townhouse	-0.106	0.131	**	No Limited Mob.	-0.079	0.096
++	Multi-Family	-0.053	0.015		N/A	0.181	-0.1
	< 3 Apts.	-0.209	0.204	Gender	Male	-0.099	0.122
	> 4 Apts.	0.03	-0.009	**/++	Female	-0.06	0.073
	Mobile	-0.082	0.187	Hispanic	Not-Hispanic	-0.083	0.1
	Dorm	-0.067	-0.018	*	Hispanic	0.002	-0.015
	Other	-0.666	0.573		N/A	0.013	0.077
Years at Res.	>20	-0.271	0.156	Income	Over 250K	-0.124	0.156
***	16-20	-0.168	0.157	**	200K – 250K	-0.122	0.232
+	11-15	-0.089	0.115		150K – 200K	-0.1	0.123
	6-10	-0.06	0.072		100K – 150K	-0.088	0.102
	1-5	-0.069	0.093		75K – 100K	-0.117	0.141
	< 1	0.028	-0.019		50K – 75K	-0.085	0.101
Place Type	City Downtown	0.028	-0.019		35K – 50K	-0.054	0.062
***	City Residential	-0.069	0.093		25K – 35K	-0.03	0.069
+	Suburbs (Mix)	-0.06	0.072		10K – 25K	-0.014	0.035
	Suburbs (House)	-0.089	0.115		Under 10K	0.034	-0.056
	Small Town	-0.168	0.157	Education	Grad. Degree	-0.124	0.173

	Other	-0.271	0.156	***	Bachelors	-0.098	0.114
Age	85 +	-0.158	0.28	+++	Technical	-0.09	0.116
***	75 – 84	-0.169	0.253		Some College	-0.045	0.032
+++	65 – 74	-0.133	0.19		Associates	-0.033	0.05
	55 – 64	-0.097	0.128		HS Diploma	-0.027	0.046
	45 – 54	-0.108	0.127		Less than HS	-0.021	0.051
	35 - 44	-0.089	0.094		Trip Purpose <sup>†</sup>		
	25 – 34	-0.068	0.069	***	HBW	-0.277	0.331
	18 – 24	0.092	-0.098	+++	HBSch	-0.144	0.222
					HBSHp	-0.016	-0.016
Employment Status	Full-Time	-0.115	0.12		HBPb	-0.173	0.212
***	Part-Time	-0.014	0.007		HBO	0.151	-0.17
+++	Self-employed	-0.03	0.052		NHBW	-0.254	0.289
	Student (>25)	0.086	-0.105		NHBNW	-0.044	0.067
	Student (<25)	0.042	-0.039		Travel Time <sup>††</sup>		
	Homemaker	-0.056	0.089	***/+	Peak	-0.505	0.639
	Not-employed	-0.046	0.09		Off-Peak	0.224	-0.289
	Retired	-0.137	0.211				

\*,\*\*,\*\*\* indicate significance of the ANOVA test for travel time at 0.05, 0.01 and 0.001, respectively.

+,++,+++ indicate significance of the ANOVA test for travel speed at 0.05, 0.01 and 0.001, respectively.† HBO: Home-based other; HBPb: Home-based personal business; HBSch: Home-based school; HBSHp: Home-based shopping; HBW: Home-based work; NHBW: Non-home-based work; NHBNW: Non-home-based non-work.

†† Peak: 6 a.m.-9 a.m. or 3p.m.-6 p.m.; Off-peak: all other times.

**Table 2.2: Utah Onboard Survey (UOBS) ANOVA Tests for Travel Time and Trip Speed**

Factor	Level	Mean Time Z	Mean Speed Z	Factor	Level	Mean Time Z	Mean Speed Z
Age	65 +	-0.439	0.986	Mode	FrontRunner	-0.537	1.295
***	45 – 64	-0.493	1.154	***	TRAX	-0.277	0.692
+++	25 – 44	-0.413	0.979	+++	Bus	-0.527	1.125
	18 – 24	-0.301	0.636		N/A	-0.524	0.979
Income	Over 75K	-0.549	1.239	Trip Type <sup>†</sup>	HBW	-0.534	1.272
***	50K–75K	-0.451	1.033	***	HBC	-0.341	0.746
+++	35K–50K	-0.391	0.959	+++	HBO	-0.177	0.420
	25K–35K	-0.339	0.738		NHB	-0.303	0.638
	15K–25K	-0.306	0.706	Fare type	Cash	-0.275	0.579
	< 15K	-0.276	0.660	***	One-Way	-0.326	0.880
No. of Vehicles	4 +	-0.485	1.082	+++	Reduced Fare	-0.257	0.404
***	3	-0.485	1.020		Day/Group	-0.187	0.406
+++	2	-0.468	1.074		Discounted	-0.333	0.741
	1	-0.351	0.854		Adult	-0.481	1.170
	None	-0.221	0.587		Ed/Eco/Annual	-0.532	1.214
Licensed	Licensed	-0.425	0.985		Senior	-0.534	1.046

***/+	Not Licensed	-0.271	0.592		Student	-0.33	0.761
Frequency	7 Days	-0.172	0.470		Medicaid	-0.27	0.709
***	6 Days	-0.295	0.726		Free Fare	-0.132	0.419
+++	5 Days	-0.464	1.096	Egress Mode	Drove	-0.496	1.082
	4 Days	-0.457	1.095	***	Walk	-0.394	0.926
	3 Days	-0.437	0.927	++	Bike	-0.265	0.635
	2 Days	-0.356	0.830	Ingress Mode	Drove	-0.462	1.062
	1 Day	-0.370	0.788	***	Walk	-0.389	0.900
	Less than 1 / week	-0.207	0.273	++	Bike	-0.273	0.655
	First Time	-0.315	0.879				

\*, \*\*, \*\*\* indicate significance of the ANOVA test for travel time at 0.05, 0.01 and 0.001, respectively.

+, ++, +++ indicate significance of the ANOVA test for travel speed at 0.05, 0.01 and 0.001, respectively.

† HBW: Home-based work; HBC: Home-based college/university/technical school as student; HBO: Home-based other; NHB: Non-home based.

## 2.4.2 Multivariate Regression Results

Next we present results for three models predicting the Z-score for travel time based on the Utah Household Travel Survey data (N=34,409). The first model is a full model that controls for all trip-level, individual, and household-level factors in the survey. In the second model, trip-level characteristics are removed. Finally, in the third model, a backwards stepwise regression using the Akaike Information Criterion is calibrated using model two as the starting point.

The purpose for estimating these regression models is to discover which of the correlated socioeconomic factors remain statistically significant while controlling for them simultaneously. This analysis is strictly exploratory, with the major emphasis on understanding the direction, magnitude and significance of the regression coefficients. Thus, the small but significant value of  $R^2$  for the models is not a large cause for concern (See Table 3). It should be noted that records with ( $|Z\text{-scores}| > 3$ ) were filtered before the regression so that undue influence from extreme values could be mitigated (2.2% of records were removed).

Since the variables are all categorically measured, the regression coefficients represent direct adjustments to the estimated Z-scores. We make several findings through a comparison of model results found in Table 3:

- The trip-level characteristics of Trip Purpose and Peak Travel are the most explanatory variables in the model. This makes sense, as we know that a) work trips are more likely to be made during peak hours, and b) service frequencies are likely to be higher during peak hours.
- Age. Compared to the middle-aged, youth are more likely to travel during lower periods of service, and the older one gets the more likely they are to travel during periods of higher service, even for those older than 65 years. These relationships hold even while controlling for trip-level characteristics, suggesting a strong mismatch between demand and supply for younger adults, and a better than average match for older adults.
- Gender. There is a slight gender effect with females being associated with modest levels of mismatch. This effect diminishes when controlling for trip characteristics.



- Employment status. This factor seems to be highly correlated with the trip-level characteristics. When trip characteristics are left uncontrolled, all employment statuses compared to full-time employed are associated with higher degrees of mismatch. When trip characteristics are controlled, being full-time employed is associated with more mismatch than the other statuses. This is most likely a result of the high degree of association between employment status and the timing, purposes and destinations of trips that are made (e.g., more work trips to highly served downtown during peak periods).
- Insignificant individual level factors. When controlling for other factors, there is no relationship between mismatch and education status, Hispanic status, race, having a driver's license, or having limited mobility.
- Household factors. People living alone are more likely associated with transit mismatch than those living in larger household arrangements. Also, self-reported housing type is consistently significant on several factors. Interestingly, those living in small apartment buildings (three or less units) have less mismatch than those in single detached housing, while those in larger apartment complexes (four or more units) are associated with more mismatch. This is true whether or not controls for trip characteristics are in place, and is perhaps an artifact of the spatial distributions of large and small apartment buildings with respect to the transit network. More spatial analysis is required for this relationship to be better understood.
- Insignificant household level factors. Once controlling for other factors, there is no consistent relationship between mismatch and household income or self-reported place of residence type. There are several factor levels that obtain significance in models one and two for these variables, but none that lend much interpretive value to the question at hand.

**Table 2.3: Summary of Ordinary Least Squares Regression Analysis for Variables Predicting Travel Time Z-Scores using the UHTS data (N=34,409)**

Variable	Model 1		Model 2		Model 3	
	Estimate	Std. Error	Estimate	Std. Error	Estimate	Std. Error
Intercept	-0.024	0.019	-0.179	0.019***	-0.166	0.016***
<i>Trip Factors</i>						
Trip Purpose (HBW) †						
HBSch	0.153	0.022***	-	-	-	-
HBSHp	0.251	0.014***	-	-	-	-
HBPb	0.105	0.016***	-	-	-	-
HBO	0.304	0.011***	-	-	-	-
NHBW	0.005	0.012	-	-	-	-
NHBNW	0.193	0.012***	-	-	-	-
Peak Travel (Off Peak) ††						
Peak	-0.482	0.007***	-	-	-	-
<i>Individual Factors</i>						
Age (35 - 44)						
18 - 24	0.066	0.016***	0.079	0.017***	0.081	0.017***
25 - 34	0.008	0.011	0.011	0.011	0.009	0.011
45 - 54	-0.017	0.012	-0.015	0.013	-0.014	0.013
55 - 64	-0.033	0.013**	-0.028	0.014*	-0.029	0.014*

65 - 74	-0.063	0.018***	-0.053	0.019**	-0.055	0.019**
75 - 84	-0.090	0.025***	-0.07	0.027**	-0.071	0.027**
85+	-0.098	0.053'	-0.113	0.058*	-0.113	0.057*
<b>Gender (Male)</b>						
Female	-0.001	0.008	0.016	0.008'	0.017	0.008*
<b>Employment Status (Full-time)</b>						
Part-time	-0.014	0.012	0.096	0.013***	0.097	0.013***
Self-employed	-0.034	0.015*	0.094	0.016***	0.095	0.016***
Student (>25)	0.015	0.024	0.084	0.026**	0.095	0.026***
Student (<25)	-0.079	0.019***	0.053	0.020**	0.060	0.020**
Homemaker	-0.078	0.013***	0.126	0.013***	0.126	0.013***
Not employed	-0.134	0.022***	0.063	0.023**	0.070	0.022**
Retired	-0.137	0.016***	0.066	0.017***	0.067	0.017***
<b>Education (Bachelors)</b>						
Grad. Degree	0.008	0.010	0.005	0.011	-	-
Technical	0.029	0.020	0.026	0.022	-	-
Some College	0.004	0.010	0.007	0.010	-	-
Associates	0.015	0.013	0.024	0.014'	-	-
HS Diploma	0.008	0.013	0.012	0.014	-	-
Less than HS	0.033	0.031	0.015	0.034	-	-
<b>Hispanic (No)</b>						
Yes	0.018	0.019	0.021	0.021	-	-
N/A	0.004	0.036	-0.002	0.039	-	-
<b>Race (White)</b>						
Other	-0.008	0.016	-0.029	0.018	-0.021	0.017
N/A	-0.042	0.030	-0.049	0.033	-0.044	0.023'
<b>Licensed (No)</b>						
Yes	-0.027	0.026	-0.048	0.028'	-	-
<b>Limited Mobility (No)</b>						
Yes	0.025	0.027	0.033	0.029	-	-
N/A	0.039	0.048	0.055	0.052	-	-
<b>Household Factors</b>						
Household Size	-0.012	0.003***	-0.013	0.003***	-0.014	0.003***
<b>Household Income (50k-75k)</b>						
Under 10k	0.034	0.029	0.037	0.031	-	-
10k-25k	0.006	0.017	0.022	0.018	-	-
25k-35k	0.013	0.015	0.009	0.017	-	-
35k-50k	-0.008	0.012	0.001	0.013	-	-
75k-100k	-0.025	0.011*	-0.010	0.012	-	-
100k-150k	0.005	0.011	0.007	0.012	-	-
150k-200k	-0.007	0.018	-0.002	0.020	-	-
200k-250k	0.042	0.029	0.082	0.031**	-	-
Over 250k	-0.024	0.028	-0.006	0.031	-	-
N/A	-0.004	0.013	0.001	0.014	-	-
<b>Place Type (Suburbs w/ Houses)</b>						

City Downtown	0.033	0.019'	0.048	0.021*	-	-
City Residential	-0.005	0.009	0.006	0.009	-	-
Suburbs (Mixed)	-0.008	0.009	0.005	0.010	-	-
Small Town	-0.027	0.017	-0.022	0.018	-	-
Other	-0.017	0.028	-0.013	0.030	-	-
Residential Type (Single Family)						
Townhouse	-0.013	0.015	-0.030	0.016'	-0.027	0.016'
Multi-Family	-0.021	0.023	-0.013	0.025	-0.004	0.024
<3 Apts.	-0.077	0.038*	-0.082	0.041*	-0.081	0.041*
>4 Apts.	0.034	0.013**	0.028	0.014'	0.044	0.013***
Mobile	-0.008	0.039	-0.035	0.042	-0.024	0.042
Dorm	0.008	0.057	-0.003	0.063	0.009	0.062
Other	-0.342	0.198'	-0.473	0.216*	-0.465	0.216*
Adjusted $R^2$	0.170		0.009		0.009	
$F$ for model fit	122.5 (58, 34350)***		7.0 (51, 34357)***		13.1 (25, 34383)***	

’, \*, \*\*, \*\*\* indicate significance at 0.1, 0.05, 0.01 and 0.001 respectively.

The reference category for each factor appears in parentheses.

† HBO: Home-based other; HBPb: Home-based personal business; HBSch: Home-based school; HBShp: Home-based shopping; HBW: Home-based work; NHBW: Non-home-based work; NHWNW: Non-home-based non-work.

‡‡ Peak: 6-9 a.m. or 3-6 p.m.; Off-peak: all other times.

## 2.4.3 Conclusions

A comparison of the descriptive statistics, ANOVA and multivariate results leads to some interesting interpretation and meaningful findings. On the one hand, the descriptive statistics of mean service Z-scores for different socioeconomic groups tell a very clear story of transit mismatch: Those who are most socioeconomically disadvantaged tend to have travel demand patterns that are out of synch with the provision of transit services in the region. The vast majority of the ANOVA F-tests confirm that the differences between socioeconomic groups are statistically significant. The regression results, however, challenge these descriptive findings in two ways. First, it appears that trip purpose and peak-period trip making are the most meaningful explainers of mismatch. Those who travel for work during peak periods are far more likely to travel during periods of higher service provision. Second, many of the factors found to have differences in the descriptive analysis lose importance when controlling for simultaneous effects. Those that remain significantly related to service level include: age, gender, employment status, household size and housing type. Importantly, from a civil rights and environmental justice policy perspective in the United States, race and income effects fall out of significance, indicating that the system isn't specifically discriminating along these lines. Nonetheless, due to the correlations between employment status, housing type and income, the fact remains that low-income individuals in the region experience more transit mismatch than those who earn more. It is just that the cause is more likely to be the trip-making patterns associated with un(der)employment, rather than income itself.

The importance of trip purpose and time-of-day in explaining transit mismatch hearkens to the long tradition of predicting these two outcomes using activity-based approaches to travel behavior modelling. Models relating socioeconomic factors of trip-makers to activity purposes and timing have been used to better understand behavior and to populate microsimulation

transportation models, but seldom have they been viewed through a social equity lens. The present papers suggest that this literature should be reviewed from the perspective of mismatch between travel demand and transit supply.

The methods developed in this research can be improved in a number of ways for future research. First, the creation of the Z-score, while necessary to control for heterogeneity in differences between local and global-average trip times across individuals and space, also somewhat obfuscates the interpretation of the regression coefficients. Marginal impacts on standardized Z-scores are hardly relatable to policy makers or indeed very meaningful in terms of valuing the absolute level of mismatch for an individual's trips or an entire social group's trips. In the future, more sophisticated modeling techniques that can account for size-effect heterogeneity across OD pairs could be employed so that the raw travel times or travel speeds can be used as dependent variables. This will allow for a more practical application of this research in terms of quantifying the experiences of individual trip-makers and socioeconomic groups.

Besides these technical refinements, there are clearly extensions of the research that should bear interest to the transportation research community. First, this research has identified a need for a spatial analysis of transit mismatch, with a particular emphasis on locating hotspots of mismatch at the origins or destinations of trips, and determining how these relate to current levels of transit provision. In particular, it will be important to determine whether mismatch can be better reduced through temporal policies (e.g., decreasing headways or shifting time-use behavior) or through spatial policies (e.g., expanding network coverage or decreasing the need for long-distance trips through land use policies). Second, it seems obvious that transit mismatch may be a root determinant of mode choice, and incorporating the degree of mismatch associated with a person's daily activity routine may yield interesting results in a mode choice model.

This paper has focused specifically on the equity considerations of a new type of transit mismatch, one that incorporates both spatial and temporal patterns of travel demand and public transit supply. It is important to recognize that the research was made possible through the development and use of the travel time cube, a new data object that embodies the spatiotemporal structure of transit connectivity in a region. Besides using the travel time cube to study transit mismatch, work is underway to investigate temporal variations in accessibility, new methods to quantify the effects of transit network modifications, and the impact of the bicycle in conquering the *last mile* problem in transit service provision. Indeed, the travel time cube is proving to be a useful tool that facilitates the study of the role of public transit in society.

### **3.0 TEMPORAL VARIABILITY IN TRANSIT-BASED ACCESSIBILITY TO SUPERMARKETS**

#### **Steven Farber**

Department of Human Geography, University of Toronto Scarborough  
1265 Military Trail, Toronto ON, M1C 1A4, Canada  
Tel: (416) 208-5120 Fax: (416) 287-7283; Email: steven.farber@utoronto.ca

#### **Melinda Z. Morang**

Environmental Systems Research Institute  
380 New York Street, Redlands, CA 92373  
Tel: (909) 793-2853 x3315; Email: mmorang@esri.com

Michael J. Widener (michael.widener@uc.edu)  
Department of Geography and Planning, University of Toronto  
100 St. George Street, Room 5047, Toronto ON, M5S 3G3  
Tel: (416) 946-0270; Email: michael.widener@utoronto.ca

*A version of this paper has been published in the peer-reviewed journal, Applied Geography:*

Farber, S., Morang, M. Z., & Widener, M. J. (2014). Temporal variability in transit-based accessibility to supermarkets. *Applied Geography*, 53, 149-159.

#### **3.1 ABSTRACT**

Food desert studies attempt to identify geographic areas in which people lack access to sources of healthy food. While academic and policy research often defines access to food stores using simple Euclidean distance or road network metrics, dense urban areas with large public transit systems call for more sophisticated methods of determining access. It is particularly important to understand the level of access the transit-dependent population has to healthy food vendors, as their mobility is largely constrained by the routes and scheduling of their local public transportation system. In this study, we analyze public transit access to supermarkets in Cincinnati, Ohio. Using General Transit Feed Specification (GTFS) data from the Southwest Ohio Regional Transit Authority (SORTA) and the Transit Authority of Northern Kentucky (TANK) and custom ArcGIS tools, we calculate the time it takes to travel from each Cincinnati census block to its nearest supermarkets at different times of the day. This transit-travel-time matrix allows us to investigate food deserts that change shape based on the time of day considered and to calculate the temporal variability in block-level access. Also, by combining this time-dependent analysis with census data on race, income and age, we search for disparities in average levels of accessibility. The results of this analysis identify the areas and subpopulations in Cincinnati with the greatest need for improved access to healthy food stores, and will demonstrate how schedule-dependent transportation can be factored into measures of accessibility. Ultimately, this study paints a more complete and realistic picture of food deserts in Cincinnati and helps policy makers better target interventions aimed at mitigating their effects.

## 3.2 INTRODUCTION

In transportation research, accessibility is understood as the ease with which interactions between people and places can take place (Hansen, 1959). While a plethora of measurement approaches exist (for reviews see: Geurs and Van Wee, 2004; Páez, Scott and Morency, 2012; Handy and Niemeier, 1997), accessibility is fundamentally at the intersection of mobility — the ease with which people travel — and the spatial configuration of destinations — the locations of stores, hospitals, workplaces, etc. to which one measures access (Páez et al., 2010). Holding the spatial distribution of destinations constant, it is clear that accessibility can be experienced disparately by different people based on varying levels of mobility (Farber and Páez, 2012). For example, in most North American cities, automobile drivers likely have higher levels of access to suburban employers than do transit riders because of differences in modal travel times. Taking this one step further, the accessibility provided by any given mode fluctuates throughout the course of the day. Both automobile and transit travel times fluctuate regularly due to congestion, but transit travel times are uniquely affected by service provision variables such as vehicle headways, scheduling, and the synchronization of transfers (Cooke and Halsey, 1966). This means that the accessibility provided by a public transport system is inherently dynamic, and that static measures of accessibility provide overly generalized indicators that may not suitably represent the actual levels of access for different population groups and activity purposes. Despite the growing adoption of accessibility as a transportation and land use planning goal (Levine et al., 2012), time-of-day variations in public transit accessibility have not been addressed extensively. This research addresses this gap through the computation of time-dependent origin-destination travel times and the development of metrics, graphs and maps to assess spatiotemporal variations in public transit accessibility.

Public transit accessibility has been extensively studied from the perspectives of station access (Murray et al., 1998; O'Neill, Ramsey and Chou, 1992; Moniruzzaman and Páez, 2012), average service frequency (Rood and Sprowls, 1998; Henk and Hubbard, 1996; Ryus et al., 2000; Hunter-Zaworski, 2003; Drew and Rowe, 2010; Al Mamun and Lownes, 2011) and origin-destination travel times (O'Sullivan, Morrison and Shearer, 2000; Lei and Church, 2010). These can all be considered static measures of access, since the score for a particular location does not vary temporally. Aside from a few exceptions (Lei, Chen and Goulias, 2012; Polzin, Pendyala and Navari, 2002; Ryus et al., 2000), very few dynamic measures of accessibility have been put forward by the research community.

Time-dependent travel times have been of interest to transportation scientists for two main reasons: traffic assignment (Ichoua, Gendreau and Potvin, 2003; Hill and Benton, 1992; Haghani and Jung, 2005) and the estimation of demand using flow matrices (Willumsen, 1984; Wong and Tong, 1998; Ashok and Ben-Akiva, 2000). Wong and Tong (1998) conducted a limited study of transit travel time variability that only considered travel between stations on the Hong Kong Mass Transit Rail network and therefore ignored ingress and egress times, and transfers between walking, bus and rail modes. Few studies of time-dependent travel times have been applied in an accessibility setting.

Works by Lei, Chen and Goulias (2012) and Owen and Levinson (2014) represent the closest predecessors of our research method. In Lei et al.'s paper, public transit travel times are used to measure accessibility from census block centroids to sets of activity locations within predefined travel-time buffers. The analysis compares public transit and automobile-based accessibility during morning and evening peak periods. Owen and Levinson calculate continuous

accessibility measures between census blocks in an effort to explain commuter mode choice. The premise behind their work is that fixed-time accessibility measures do not provide enough information about the accessibility between locations over time.

Our study departs from these by calculating and analyzing public transit travel times from all census block centroids at all times of the day in an attempt to investigate temporal variability in supermarket access. The next section briefly introduces the reader to the food desert literature. Following that, we describe the methods and datasets used to calculate and assess accessibility to supermarkets. This is followed by a discussion of the results and conclusions.

### **3.3 CASE STUDY – FOOD DESERTS – NEED FOR TIME-VARYING MEASUREMENTS**

One facet of urban accessibility that has garnered attention from a diverse set of stakeholders is that of spatial access to stores that sell healthy and affordable food options, like supermarkets or full-service grocery stores. Neighborhoods with low spatial and economic access to such stores are commonly referred to as “food deserts” (Larsen and Gilliland, 2008; Jiao et al., 2012; USDA, 2014), a term which relies heavily on the inherently spatial concept of a “desert” and implies the region is devoid of nutritious food options for its residents. A thorough review of research exploring the socioeconomic and spatial drivers and outcomes of food deserts can be found in Walker et al. (2010). Generally, food deserts are defined by identifying regions that have a large percentage of low-income residents and are spatially distant from affordable food stores with a wide selection of nutritious options (USDA, 2014). While food stores that carry nutritious options can include vendors as diverse as large format supermarkets, corner stores and mobile food trucks, large grocery stores and supermarkets are commonly used as proxies for healthy food stores, as they consistently stock reasonably priced and high-quality nutritious foods (Block and Kouba, 2006). Because of this, as well as the fact that spatial access is a prerequisite for purchasing and consuming healthy foods, this research focuses on spatial access to large grocery stores and supermarkets. Over the past decade, geographers (Wrigley et al., 2003; Smoyer-Tomic et al., 2006; Apparicio et al., 2007; Páez et al., 2010) and public health academics (Cummins and Macintyre, 2002), governments (Boyce and Treering, 2012; Chicago, 2013), and media institutions (Barclay, 2013; Ramsey, 2011; Kolata, 2012) have explored the phenomenon of food deserts and found their existence to be the result of complex social and structural processes with real public health consequences (Moore et al., 2008; McKinnon et al., 2009). While results have been mixed (Holsten, 2009; An and Sturm, 2012), some research points to residents residing in food deserts to be at risk of maintaining a less nutritious diet (Holsten, 2009), which can in turn lead to higher rates of common, but preventable, chronic illnesses like obesity, diabetes and some cancers (Bazzano et al., 2002; He et al., 2004; Higdon et al., 2007).

Recent research has documented the importance of creating more nuanced measures of the food environment by taking into account the spatiotemporal dynamics of urban life. For example, Widener et al. (2013) and Burgoine and Monsivais (2013) demonstrate how everyday-commuting patterns can result in different, and sometimes improved, accessibility outcomes than measures that only account for residents’ home locations. While these works demonstrate the importance of accounting for the dynamic lived spaces of residents when deriving accessibility scores, they do not adequately account for the differences incurred by the use of various modes of transportation.

Of particular interest is the level of spatial access transit-dependent riders have to healthy food options, though limited research is available on the topic. Burgoine and Monsivais (2013) do consider the potential of transit riders accessing food vendors during their walk to and from bus stops, but neglect to account for opportunities that might be available along the bus route. Fuller, Cummins and Matthews (2013) explored whether transportation mode to a resident's primary food store moderates the relationship between distance to the store and fruit and vegetable consumption. Their models found no relationship to exist. Lastly, Widener et al. (2014) demonstrate that many transit commuters have increased spatiotemporal access to healthy food vendors when accounting for their daily trip to and from work and a limited time budget.

While these papers establish a baseline from which to examine spatial access to healthy food for transit riders, they all neglect one of the most important limitations of traveling via transit: variability in spatial access dependent on the time of departure. Given this variability, it is possible that certain regions have acceptable access to healthy food stores via transit only during select periods. Highly variable transit travel times make shopping by public transit nearly impossible for a consumer who is trying to juggle work, child care, and other demands on his or her time — no matter what the average travel time may be. And low-frequency service could discourage residents from shopping at these healthier food vendors and result in grocery shopping taking place at closer, less healthy food stores like convenience markets (Hendrickson, Smith and Eikenberry, 2006).

This research directly addresses this issue by examining how often residents of Cincinnati, Ohio, have access to supermarkets within a predetermined time threshold via transit. By considering the variability of access over time, results show where and when spatial access to healthy food is an issue.

## **3.4 METHOD**

Our method is comprised of three stages: travel time estimation, accessibility metrification and statistical analysis. In the first stage, public transit travel times are estimated between all census block centroids in the city of Cincinnati and the 10 closest supermarkets in the Cincinnati metropolitan area. In the second stage, these travel times are converted into measurements of accessibility and temporal variability. Finally, in the third stage, the measurements are used to assess socioeconomic disparities in accessibility across different populations residing in Cincinnati.

### **3.4.1 Data**

Our analysis relies on transit and walking network data, demographics and food retailers. Transit stops, routes and schedules used in this analysis are obtained as General Transit Feed Specification (GTFS) datasets from the Southwest Ohio Regional Transit Authority (SORTA) and the Transit Authority of Northern Kentucky (TANK). The pedestrian network is characterized using street data from OpenStreetMap. Census blocks within the Cincinnati city limits from the 2010 Census are used as the demand locations, and we obtain 2010 Census attribute data for blocks and block groups from the National Historical Geographic Information System (Minnesota Population Center, 2011). For our network analysis, spatial representation of census blocks is based on centroids snapped to nearest street features.

Supermarket address data is acquired from the Orbis Business directory (Orbis) by selecting all businesses with the North American Industry Classification System code 445110



(Supermarkets and Other Grocery [except Convenience] Stores), more than 15 employees, and that are within one of the nine counties immediately surrounding Cincinnati (Ohio: Hamilton, Butler, Warren and Clermont counties; Kentucky: Boone, Kenton and Campbell counties; and Indiana: Dearborn and Franklin counties). This query results in a list of 175 stores. A number of stores are removed after checking their websites or calling their provided phone number, as many do not fit the assumed characteristics of a supermarket (e.g., having a wide variety of fresh produce and reasonable costs). These stores were generally gas stations or businesses related to the grocery industry, but not food vendors. Finally, the addresses are geocoded in ArcGIS 10.1 and their locations double-checked using Google Maps, resulting in a total of 117 supermarket locations. A conservative, manual procedure was used to remove 19 spatially outlying supermarkets, resulting in a set of 98 supermarkets for analytical use.

The four maps in Figure 3.1 depict the spatial extent of our study area, the locations of supermarkets, the intensity of transit level of service (in Figure 3.1a), and the spatial distributions of Black, older adults and impoverished populations (in Figure 3.1b-d). Cincinnati is home to a large, segregated community of African-Americans that also tends to reside in high poverty and low-income neighborhoods. The trips-per-hour map (Figure 1a) is created using the prototype BetterBusBuffers tool developed by Melinda Morang at Esri<sup>3</sup>. The map shows 10-minute walk distance buffers around transit stops, color-coded by the average number of unique transit trips per hour available at each location between 6 a.m. and 10 p.m. If multiple stops are reachable from a given area, trips visiting all reachable stops are considered. The map does not provide any information about the destinations of the available trips, simply their frequency. This level of service indicator is tabulated using the GTFS and walking network datasets. Figures 1b and 1d illustrate clustering and co-location of higher-percent Black and impoverished regions. Finally, hotspots of older adult population are identifiable in Figure 3.1c, but compared to the other demographic variables explored, the distribution of the older adult population is far more even.

---

<sup>3</sup> All of the tools used in this research to integrate GTFS data with ArcGIS Network Analyst are available for download at <http://www.transit.melindamorang.com>.

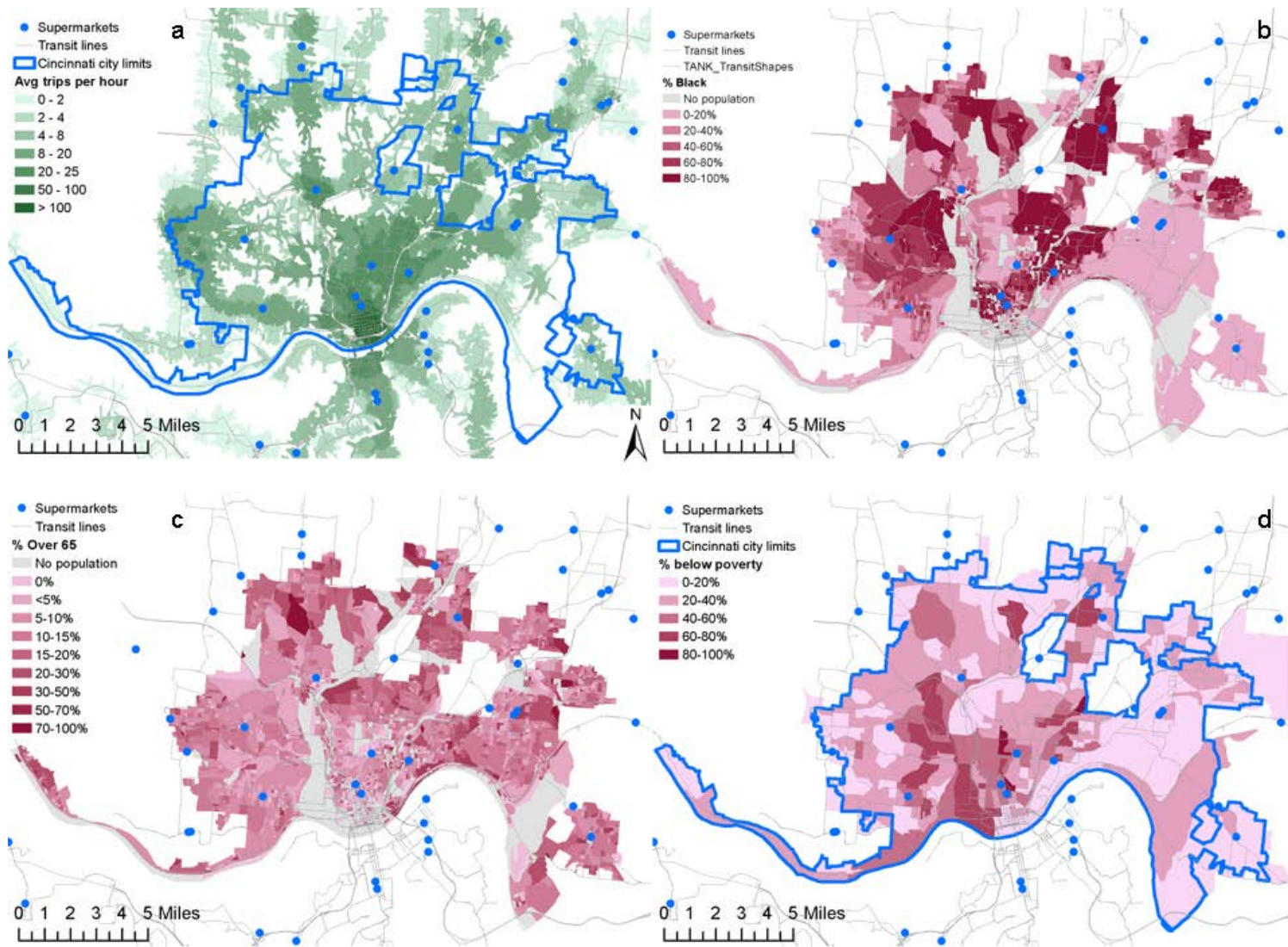


Figure 3.8 Maps of Cincinnati Depicting Supermarket Locations and a) Transit Frequency of Service, b) Distribution of the Black Population, c) Distribution of the Older Adult Population, and d) Distribution of Households Living in Poverty

### 3.4.2 Travel Time Estimation

We calculate the walking/transit travel time between each census block-supermarket pair using the Network Analyst OD Cost Matrix solver in ArcGIS 10.2, which uses Dijkstra’s algorithm to find the shortest path through the network. We use the prototype *Add GTFS to a Network Dataset* tools recently developed by Esri’s Network Analyst team to combine GTFS transit datasets with street data into an ArcGIS network dataset. We use the resulting multimodal network dataset to calculate transit trips through the network at specific times of day.

In setting up our network dataset, we assume an ingress, egress and transfer walking speed of 3 mph. We apply a delay of 0.3 minutes for transitions between streets and transit lines (to represent boarding a transit vehicle), and a delay of 0.1 minutes for transitions from transit lines to streets (to represent alighting). We also create a pedestrian restriction to prevent foot travel on limited-access highways and highway ramps.

To calculate the travel time along transit lines, our network dataset uses the TransitEvaluator from the *Add GTFS to a Network Dataset* tools. Each time a transit line in the network is queried by the OD Cost Matrix solver, the TransitEvaluator determines the travel time across that transit line by looking up the available transit trips in the GTFS schedules at the appropriate time of day and summing the wait time for the trip plus the ride time from the current stop to the next.

The total travel time between a census block and a supermarket is the time it would take to travel the optimal path through the network at a particular time of day. A travel time of 30 minutes could indicate 30 minutes of walking, 30 minutes of riding on transit, or any combination of walking, waiting and riding that adds up to 30 minutes. The optimization algorithm does not limit the pedestrian’s walking time, the number of transfers between transit lines, or the fare paid.

Because the travel time between origin-destination pairs changes throughout the day, we calculate the travel time matrix for every minute of the day. In order to reduce computation time, we only calculate the travel time for the 10 closest supermarkets to each block at each time of day. We use a python script to increment the start time and save the results, and we use Monday to represent a typical weekday. According to a national study, some 153 million grocery shopping trips take place every week from Monday to Friday across the United States (Goodman, 2008). This is compared to 74 million trips on the weekends. Thus, while investigating access using a Monday is not exhaustive, the analysis does represent the transit situation for approximately two thirds of all shopping trips.

### 3.4.3 Accessibility Metrification

Next, we describe how we convert travel time matrices into measurements of accessibility. Considering the eventual goal of describing variability in accessibility throughout the day, we make use of very simple measurements of accessibility for each time of day, and then analyze how these measurements fluctuate over the course of the day.

It is convenient to introduce some notation. Let  $T_{i,m,k}$  be the estimated transit travel time from block centroid  $i$ , at departure time  $m$  to the  $k^{\text{th}}$  closest supermarket (measured in minutes and where the ranking of supermarket proximity is specific to departure time). As previously stated, we compute  $T_{i,m,k}$  for all census blocks, at every minute of the day, to the closest 10 supermarkets. Despite storing the travel times in an indexed table, it is helpful to think of  $T$  as a

three-dimensional array, indexed by  $i$ ,  $m$ , and  $k$ . In this way, cross-sections of the array can be used to measure, graph and map spatiotemporal patterns in supermarket accessibility. For example, in Figure 3.2, we fix a specific origin and destination and chart  $T_{i,m,k}$  versus time-of-day,  $m$ . The resulting graph demonstrates how travel times from origin to destination can fluctuate over time depending on the frequency of bus service and level of synchronization between bus routes at transfer locations. The graph directly illustrates how “single-point” measures of transit-based accessibility are likely not representative of true levels of access. We see that travel times fluctuate between 23 minutes and 57 minutes, depending on the availability and timing of bus services. Even within a narrow window of 7-9 a.m., during peak service hours, we observe a range of nearly 30 minutes between the maximum and minimum required travel time. It is this variability in travel times that directly leads to variability in accessibility.

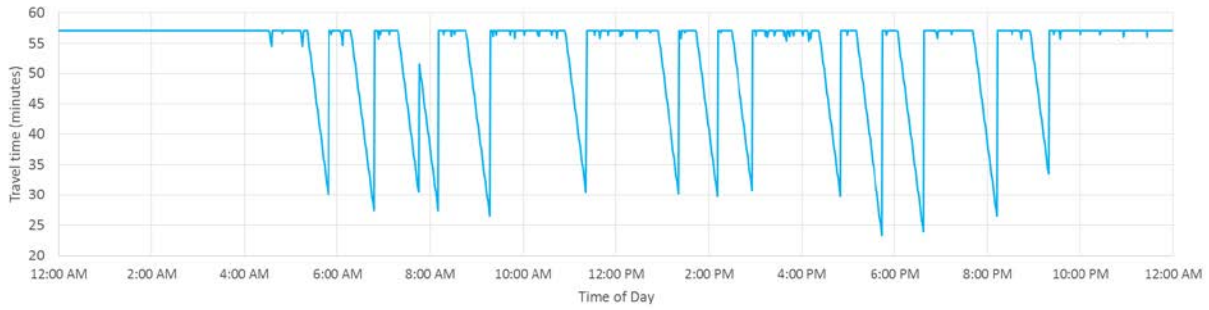


Figure 3.9 Variation in Public Transit Travel Time Between Two Locations Across a 24-hour Period

Moving from travel time estimates to measures of supermarket accessibility is quite straightforward. We use several easily interpretable summary statistics of travel times to measure access. First, for the closest  $K$  supermarkets, we define the average and standard deviation of travel time for each census block across the time period of 6 a.m.-10 p.m. We believe this window suitably captures the best case scenario of when supermarkets have the highest probability of being open to the shopper.

Let  $\mu_{i,m}^K = \frac{\sum_k T_{i,m,k}}{K}$  for  $k = 1..K$  be the average travel time to the closest  $K$  supermarkets at time  $m$  for census block  $i$ . Then  $\mu_i^K = \frac{\sum_m \mu_{i,m}^K}{M}$  for  $m = 1..M$  is the average travel time to the closest  $K$  supermarkets from block  $i$  across all the minutes  $m$  in the specified time window of size  $M$ . And

$\sigma_i^K = \sqrt{\frac{\sum_m (\mu_{i,m}^K - \mu_i^K)^2}{M-1}}$  is the standard deviation of minute-by-minute travel times around the average daily travel time. We make extensive use of  $\mu_i^K$  and  $\sigma_i^K$  in quantifying mean levels of accessibility to supermarkets and variations about the mean. We can then describe accessibility of a census block along these two dimensions. Table 3.1 illustrates how combinations of average travel time and dispersion about the mean lead to different accessibility outcomes. The best case scenario is for a block to have low mean travel time and low levels of dispersion. This indicates that throughout the day, people have short travel times to supermarkets with little variability. On the other hand, high average travel times with low levels of dispersion indicate that a block is very poorly connected to supermarkets. Low average with high dispersion indicates a generally strong accessibility with some aberration from the mean, while high average with high dispersion indicates poor accessibility with occasional peaks in higher access.

**Table 3.1 Interpretation of Average Travel Time and Dispersion Measures**

	<b>Low <math>\mu</math></b>	<b>High <math>\mu</math></b>
<b>Low <math>\sigma</math></b>	Best	Worst
<b>High <math>\sigma</math></b>	Mostly Good	Mostly Bad

In addition to standard summary statistics of central tendency and dispersion, we also use a simple indicator variable to determine whether a census block has access to supermarkets within a certain time threshold,  $h$ .

Let  $I_{i,m}^{K,h} = \begin{cases} 1 & \text{if } T_{i,m,K} \leq h \\ 0 & \text{otherwise} \end{cases}$ , be a variable indicating whether block  $i$  has access to  $K$  supermarkets within  $h$  minutes, then  $P_i^{K,20} = \frac{\sum_m I_{i,m}^{K,20}}{M}$  is the percentage of the time window that a block has 20-minute access to at least  $K$  supermarkets. A score of  $P_i^{K,20} = 1$  indicates that block  $i$  has 20-minute access to  $K$  supermarkets for the full duration of the time window, whereas a score of 0.5 means that the block can only reach  $K$  supermarkets within 20 minutes 50% of the time.

### 3.4.4 Statistical Analysis

Cincinnati is home to a large, segregated community of African-Americans that also tends to reside in high poverty and low-income neighborhoods. The aim of the statistical analysis is to explore the distribution of the accessibility metrics with particular reference to disparities in accessibility between different race, age and income groups. We make use of standard exploratory data analysis techniques to uncover any relationships between food accessibility, temporal variability, race, age and poverty status. Regression models (not shown) fail to find generalized relationships between socioeconomic status and accessibility. This is because both high and low socioeconomic status neighborhoods attain both high and low levels of accessibility.

In light of this, the analyses focus on quantifying the number of people in different socioeconomic groups attaining different levels of accessibility. This paints a more holistic picture of accessibility in general, while identifying pockets of mismatch that must be addressed from a policy perspective.

## 3.5 RESULTS

We present results of the accessibility analysis using maps, graphs and statistical analysis. First, we explore spatial patterns of daily average travel time to supermarkets, variation in daily access, and percent-of-day with 10- and 20-minute access. We believe that these “common sense” cutoffs suitably describe good and medium levels of access given the difficulty in carrying groceries for longer periods of time. After that, race, age and income disparities are explored using a combination of scatterplots and histograms.

Figure 3.3 depicts travel times to supermarkets in terms of the average and standard deviation of travel time to the closest supermarket (shown in 3.3a and 3.3c) and the closest three supermarkets (shown in 3.3b and 3.3d). The maps demonstrate that proximity to supermarkets results in faster travel times, and those areas with shorter duration travel times consistently also have lower standard deviations. We can observe locations that are not spatially proximate to

supermarkets but still acquire low travel times due to high levels of transit connectivity. Also, we find that shorter travel times extend over longer distances in the center of the city, where the street pattern is more rectilinear, in comparison to more suburban locales with less walkable and more convoluted street designs.

Interestingly, we see that those blocks with longer average travel times, while having higher standard deviations, vary to strikingly different degrees. For example, the western and eastern extensions of blocks along the Ohio River both have long average travel times to supermarkets, but those in the west have a high standard deviation while those in the east do not. This is a result of commuter-oriented bus routes that service the western blocks with faster but infrequent connections to supermarkets.

When comparing patterns of accessibility for one and three supermarkets, we see that accounting for the three closest markets results in much less heterogeneous map patterns. Simply, the averaging over three supermarkets smoothes the local accessibility gain to blocks within immediate vicinity of a supermarket. We hypothesize that access to three supermarkets, rather than just one, better guarantees access to affordable, healthy food due to competitive pricing and diversification of product (Pearson et al., 2014).

Figure 3.4 visualizes variability in travel times using the percent-of-day accessibility indicators described previously. In this case, Figure 3.4a and Figure 3.4b depict the percent of day that residents can reach one and three supermarkets within a 10-minute journey, while Figure 3.4c and Figure 3.4d depict the same measurements using a 20-minute threshold. In both cases, we only consider the time period between 6 a.m. and 10 p.m. when formulating the denominator of the percentages. There are several observations to be made. First, in Figure 3.4a, we see that only those blocks within walking distance to supermarkets can attain 100% access. Also, the level of accessibility drop-offs very abruptly; those blocks having partial transit-based, 10-minute access to a single supermarket only have that for 0-20% of the day. This, in combination with the sea of 0% access blocks across the study area, indicates a very low level of opportunity for transit-based grocery shopping beyond those places that are walkable. As before, the accessibility zones spreading out from supermarkets are larger in the downtown regions with more direct walkable routes.

While very few blocks attain any semblance of regular access to three supermarkets within 10 minutes (shown in Figure 3.4b), the accessibility landscape appears to be more promising when considering a 20-minute buffer shown in Figure 3.4c and Figure 3.4d. In contrast to the 10-minute buffer, percent-of-day figures drop off smoothly away from 100% zones, especially in areas along transit corridors that attain 40-80% access to one supermarket. Of course, this visual depiction of higher accessibility is dampened by the understanding that a 20-minute trip to a supermarket (plus a return trip) is not desirable, but perhaps feasible, for many individual shoppers.

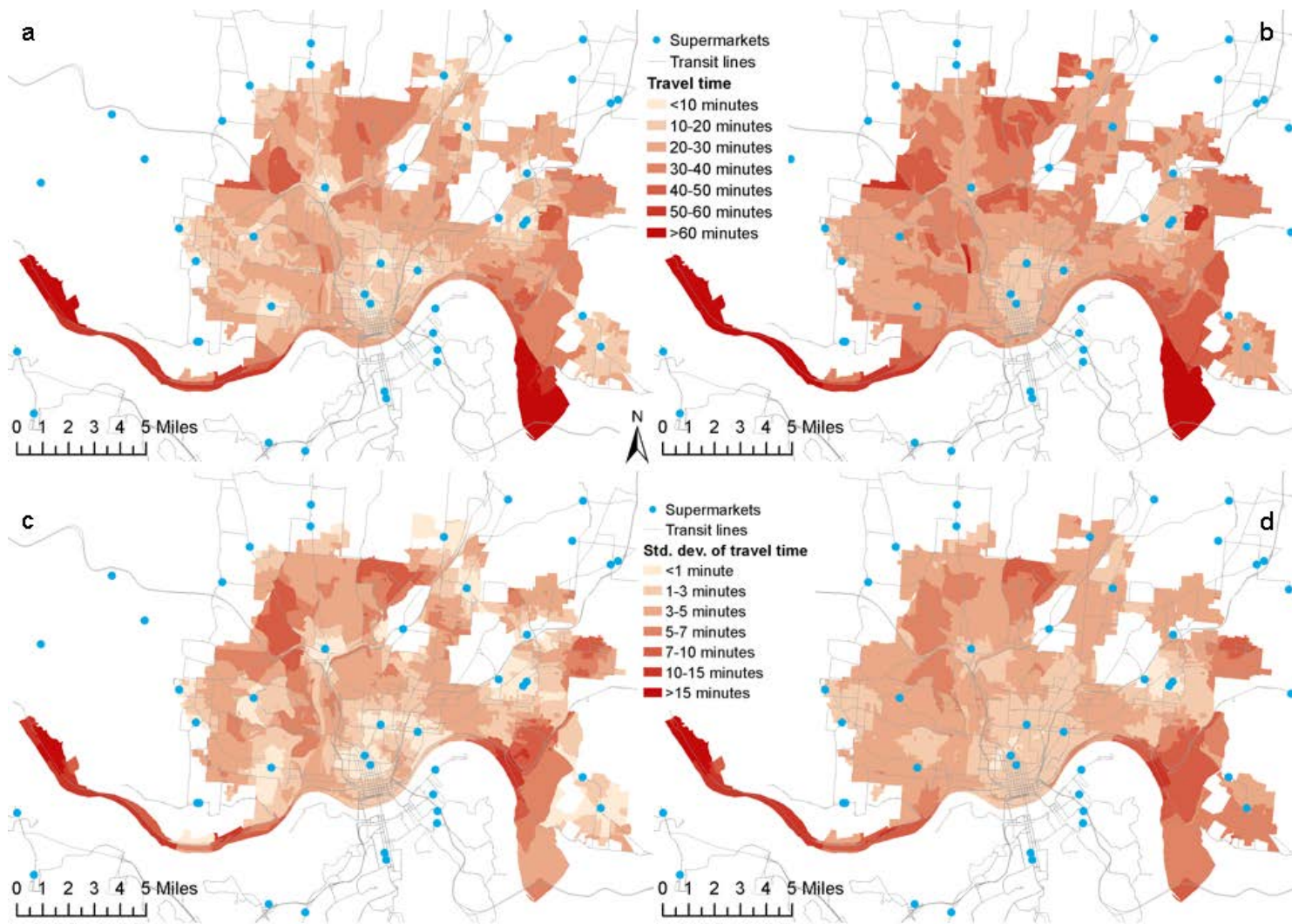


Figure 3.10 Mean Travel Time to a) The Closest Supermarket; b) The Closest Three Supermarkets and Standard Deviation of Travel Time to; c) The Closest Supermarket; and d) The Closest Three Supermarkets

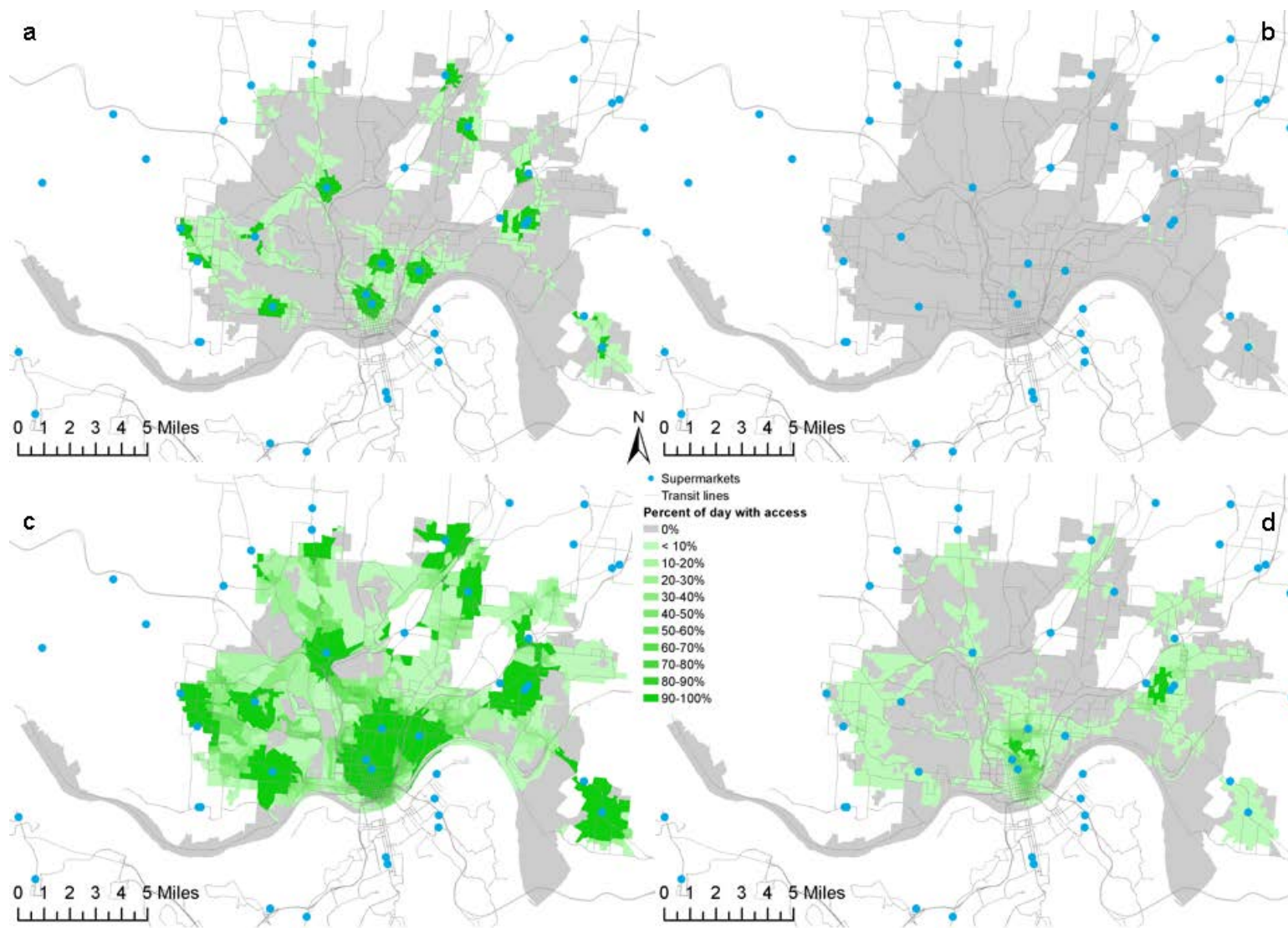


Figure 3.4 Percent of Day with Access to One and Three Supermarkets within 10 Minutes (a and b) and 20 Minutes (c and d), Respectively





Figure 3.5 depicts disparities in access to supermarkets between White and Black people. Each subplot contains a scatterplot of average versus standard deviation of travel time to the nearest supermarket. The subplots contain census blocks with increasing concentration of the African-American population. The red iso-lines mark different values of the coefficient of variation ( $CV = \frac{\sigma}{\mu}$ ) as labeled in the bottom-left plot.

In general, most of the blocks across all four plots are concentrated to the left of 30 minutes and below the CV=0.3 line. This indicates quite good access overall when considering average travel times over the course of the day (6 a.m.-10 p.m. in this case). Focusing attention on the top-left plot, we observe the existence of several dozen non-Black blocks with mean travel times in excess of 60 minutes. Contrastingly, there is only one block amongst the most concentrated Black neighborhoods with an average travel time greater than 42 minutes (FIPS 390610104001013 with total population of 13 people). The longer travel time blocks identified in the top-left plot can be readily classified into high and low levels of dispersion. The blocks with long travel times and low dispersion are much worse off than those with high levels of dispersion since shorter duration transit trips to the supermarket in the latter group *may* be possible.

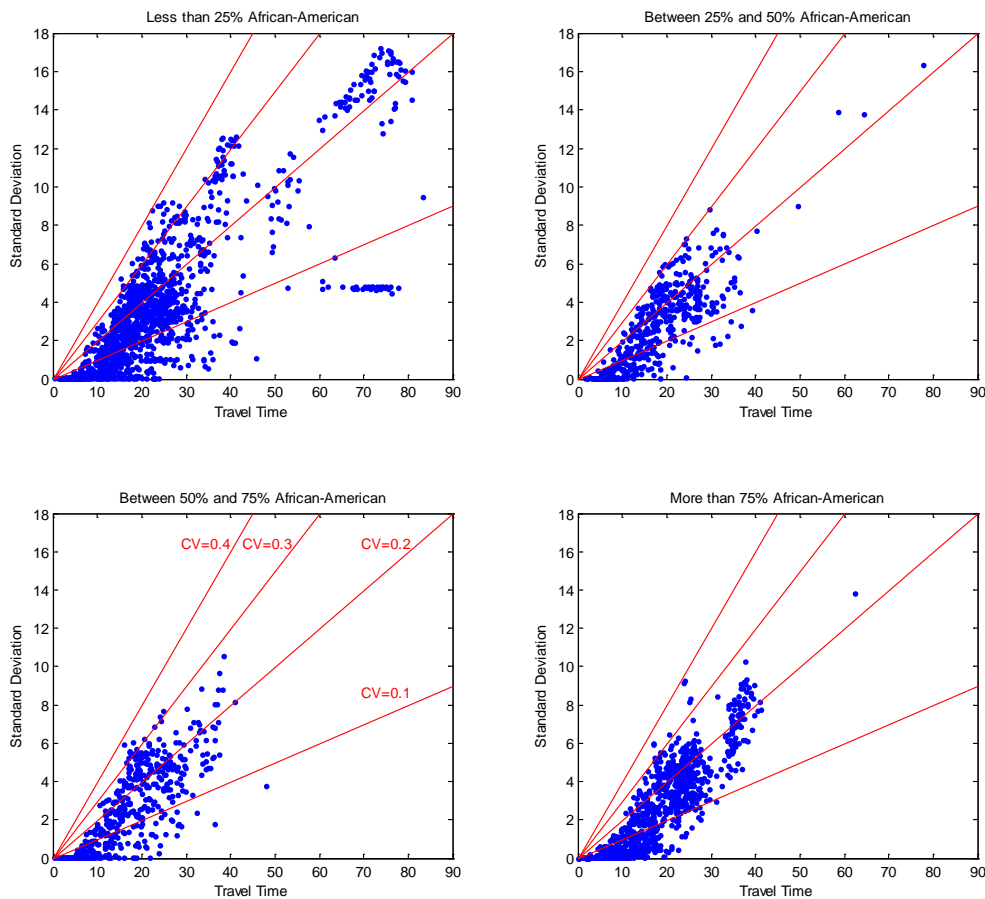


Figure 3.5 Scattergrams of Mean Travel Time to Closest Supermarket versus Standard Deviation of Travel Time (6 a.m.-10 p.m.)

Figures 3.6-3.8 depict the social disparities in supermarket access by race, age and income. Each histogram plots the percentage of the population in each social grouping (vertical axes) that attains 20-minute access to its nearest supermarket for different percentages of the day (horizontal axes; using 6 a.m.-10 p.m. as the denominator). Accordingly, Figure 3.6 indicates that roughly 41% of the White population and 36% of the Black population can reach a supermarket within a 20-minute transit journey at all times of the day. Interestingly, Black people have lower representation in the 0% access category, 7% versus 10% for White people, but are more likely to reside in blocks with partial access to supermarkets. This indicates that the Black population is subjected to more uncertainty with respect to when they might have access to supermarkets when compared to the White population.

Figure 3.7 indicates that older adults are more likely to reside in neighborhoods with less than 20% access, and less likely to reside in blocks with 100% access. In reality, if we were to relax the constant walking speed assumption embedded in our travel time calculations, we would likely find even lower levels of access for the older adults than those observed.

Finally, we repeat the analysis for those living below the poverty line in Figure 3.8. We use five-year 2012 American Community Survey data at the block-group level of aggregation to report on poverty. That is the reason for the general shape of the poverty histogram to be so different to the rest, which are generated using block-level decennial census counts. For this analysis, block-level accessibility scores are aggregated to the block group using a population-weighted average. The histogram indicates that those living below the poverty line are actually more likely to be within higher-accessibility block groups. Despite this, more than 40% of below-poverty individuals can only reach a supermarket within 20 minutes during 40% or less of the day. So, despite the absence of gross disparities between rich and poor, overall levels of transit-based accessibility for the poor are quite low except for those living within walking distance of markets (approximately 25% of those in poverty).

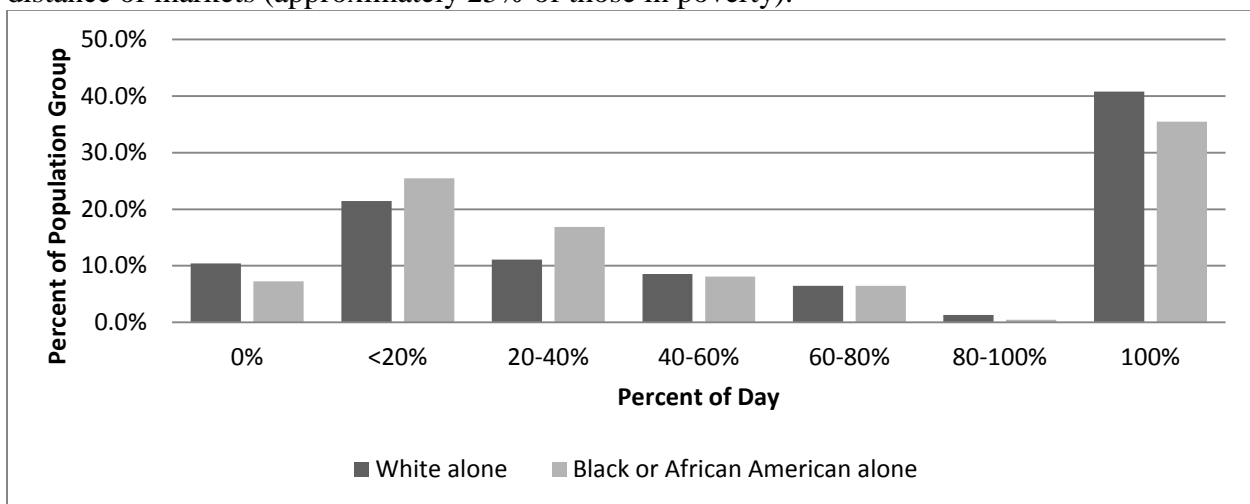


Figure 3.11 Percent of Day with 20-Minute Access to Supermarkets by Race

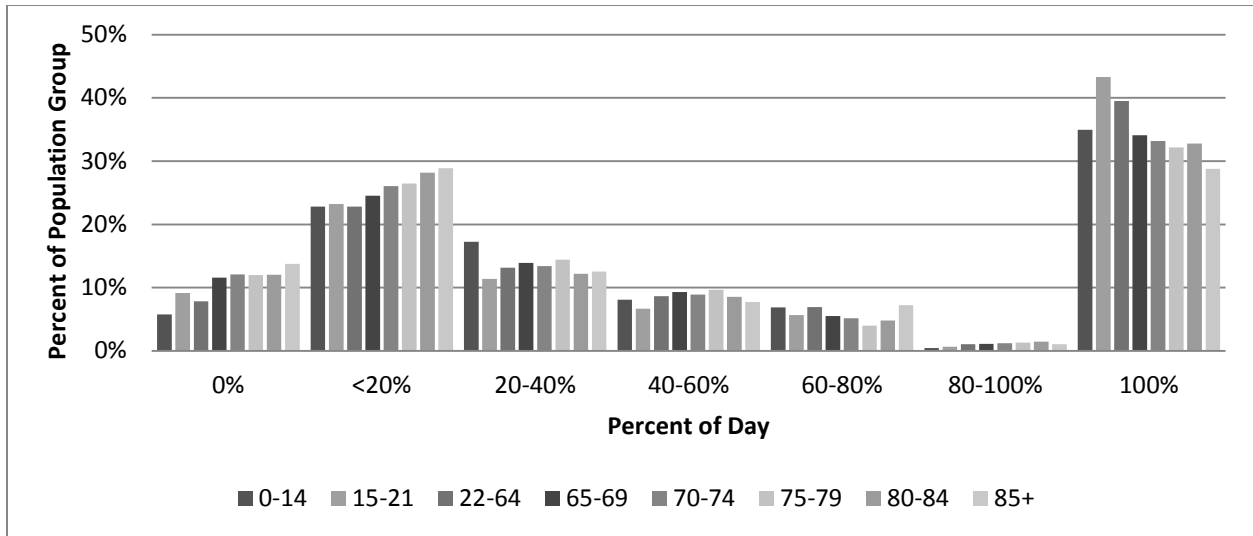


Figure 3.12 Percent of Day with 20-Minute Access to Supermarkets by Age

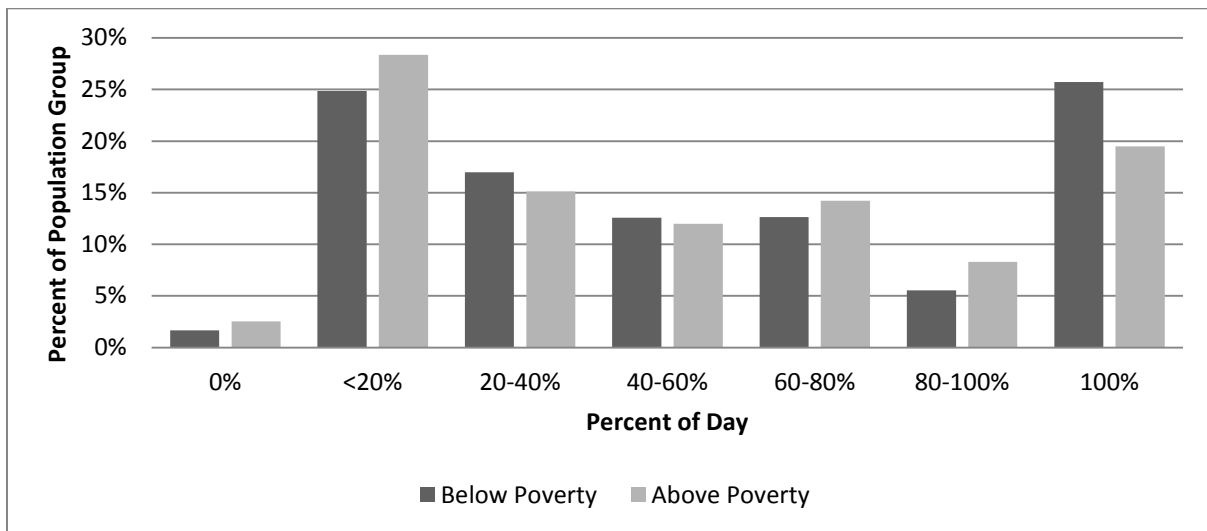


Figure 3.13 Percent of Day with 20-Minute Access to Supermarkets by Poverty Status

Automobile ownership is another policy dimension that has not yet been explored in this paper. While the demographic variables (race, age and poverty status) allow us to look for evidence of discrimination, a comparison of automobile ownership with transit-based access to supermarkets can more precisely identify the neighborhoods in the city where poor transit-based access is most likely to cause poor accessibility to healthy foods.

American Community Survey 2008-2012 five-year data provide estimates for the percentage of households with zero vehicles at the census-tract level of aggregation. We focus on zero-vehicle households because we assume that households with at least one vehicle will tend to use that vehicle for grocery trips. A selection of census tracts that approximately covers Cincinnati is made for this analysis (n=119). According to this data, there are 30,321 (20.8%) zero-vehicle households in Cincinnati. The percentage of zero-vehicle households per census tract is spatially concentrated (Moran's I of 0.47 using row-standardized queen contiguity), with a map pattern that closely resembled that for poverty seen in Figure 1d.

Figure 3.9 depicts the spatial distribution of zero-vehicle households (each dot represents 40 households) overlaid on the block-level supermarket accessibility scores. In Figures 3.9a and 3.9b, the blocks are colored according to the percentage of the day (6 a.m.-10 p.m.) that at least one supermarket is reachable within a 10- or 20-minute trip, respectively. Evidently, very few zero-vehicle households have consistent 10-minute transit access to a single supermarket. The case for 20-minute access is slightly more optimistic, with large numbers of downtown, zero-vehicle households attaining regular access. Still, even for the 20-minute threshold, several low-access suburban neighborhoods with large numbers of zero-vehicle households can be identified.

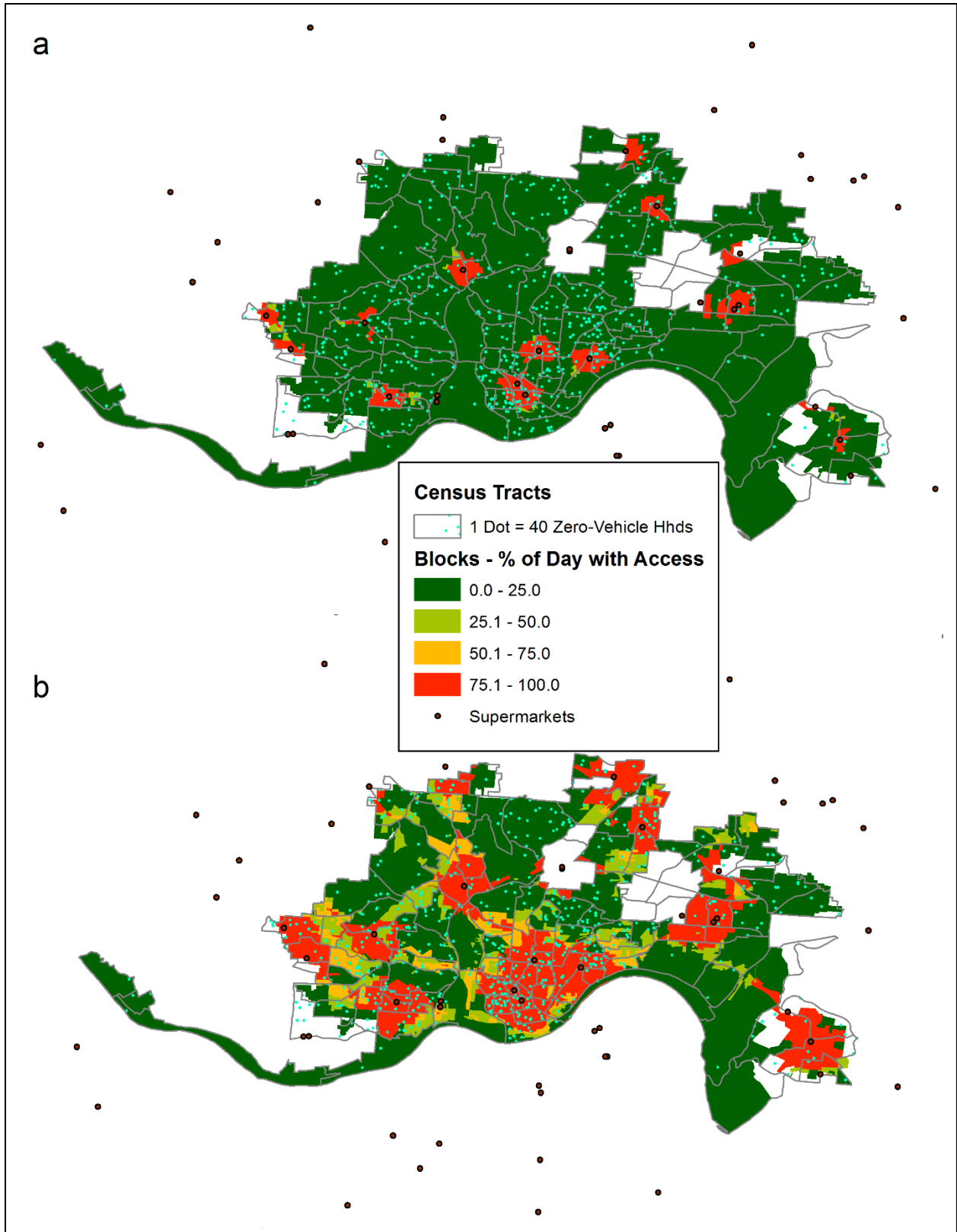


Figure 3.14: Zero-Vehicle Households and Transit-Based Access to Supermarkets

### 3.6 CONCLUSIONS

In this study we have put forward new techniques for measuring transit-based accessibility to destinations. In particular, we demonstrate the degree to which schedule-based, fixed-route transit systems provide time-varying levels of accessibility to supermarkets in Cincinnati, Ohio. We also explore disparities in accessibility and accessibility variations among race, age and income groups. Several conclusions regarding the science of accessibility, and how it pertains to food shopping, can be drawn from the analysis.

First, the research clearly identifies that access to destinations via public transit is characterized by a high degree of variability over time. Moreover, this variability depends on the distance from origin to destination, the existence of transit options, and the headways along transit routes. Not only does this research suggest the inadequacy of using a single time of day for transit-based accessibility research, the methods put forward also paint a more complete picture of the degree to which individuals can reach their destinations via transit. It is therefore a more nuanced and valid measure of accessibility.

Second, in terms of the specific case study findings, the continuous accessibility measures identify spatial, temporal and demographic trends in accessing supermarkets via transit. The analysis clearly shows that outside of walkable buffers surrounding supermarkets, very few people attain consistent access to supermarkets throughout the day. In other words, once transit is actually involved in the trip, many Cincinnatians can only reach a supermarket within 20 minutes during 20% of the day or less. Thus, despite the potential ability for people to align shopping trips with the transit schedule, the degree of flexibility for transit riders to choose when they would like to go shopping appears to be severely limited. In a time-crunched modern world, scheduling barriers such as these are likely to result in fewer shopping trips being made.

There are several caveats concerning this methodology that should be addressed in future work. First, the method only uses one-way travel times in measuring accessibility. In reality, shoppers consider the total travel time when conducting their trip. Second, the measure of travel time does not differentiate between different components of a multimodal trip (walk time, wait time, transfer time and in-vehicle time). Trips of equal travel time requiring different levels of in-vehicle vs. walking time are experienced differently by different types of people. For example, elderly travelers may find longer in-vehicle travel times suitable in exchange for shorter walking times. Third, we do not consider weekend shopping trips, on days when U.S. transit agencies typically run at reduced levels of service. An interesting follow-up study would be to compare accessibility between weekdays and weekends.

The results of the demographic analysis are mixed. People identifying as Black or African-American and older adults tend to have lower levels of consistent access to supermarkets throughout the day. Furthermore, there exist several predominantly White neighborhoods with terribly low levels of access to supermarkets throughout the entire day, while no such predominantly Black neighborhoods were found. In terms of income, those below the poverty line actually have slightly higher levels of transit-based access compared to those not living in poverty. However, it should be cautioned that 40% of the population living in poverty has less than 40% access to a supermarket within 20 minutes. So, despite there being little inequality between income groups, levels of access are still very poor among those least likely to own automobiles. This conclusion is clearly supported by the vehicle ownership overlays, which identify widespread lack of accessibility at the 10-minute threshold, and more concentrated pockets of inaccessibility for zero-vehicle households at the 20-minute threshold. If a 10-minute

grocery trip is deemed desirable for a policy goal, the abundance of zero-vehicle households not attaining this level is quite worrisome; at this level of widespread inaccessibility, policies aimed at increasing mobility (such as automobile ownership or taxi vouchers) may provide the most benefit given that alternative policies, such as providing more transit routes or increasing the density of supermarkets, could not be concentrated in targeted areas. Comparatively, if a 20-minute travel time is considered sufficient, place-based policies such as transit expansion or supermarket development could be targeted in several low-vehicle tracts with poor access. It is quite fascinating that the most suitable policy direction clearly hinges on whether 10- or 20-minute access is deemed sufficient. Thus, more research into the impact of accessibility on food shopping behaviors is needed to inform this debate.





## REFERENCES

- Al Mamun, Md, and Nicholas E Lownes. 2011. "A composite index of Public transit accessibility." *Journal of Public Transportation* 14 (2):69-87.
- An, R., and R. Sturm. 2012. "School and Residential Neighborhood Food Environment and Diet Among California Youth." *American Journal of Preventive Medicine* 42 (2):129-135.
- Ashok, Kalidas, and Moshe E Ben-Akiva. 2000. "Alternative approaches for real-time estimation and prediction of time-dependent origin–destination flows." *Transportation Science* 34 (1):21-36.
- Banister, D., and J. Berechman. 2000. *Transport Investment and Economic Development*. London: Routledge.
- Barclay, Eliza. 2013. "Can Star Power Make New Orleans' Food Deserts Bloom." Accessed 06/03. <http://www.npr.org/blogs/thesalt/2013/05/15/183992818/can-star-power-make-new-orleans-food-deserts-bloom>.
- Bazzano, LA, J He, LG Ogden, CM Loria, S Vupputuri, L Myers, and PK Whelton. 2002. "Fruit and vegetable intake and risk of cardiovascular disease in US adults: the first National Health and Nutrition Examination Survey Epidemiologic Follow-up Study." *American Journal of Clinical Nutrition* 76 (1):93.
- Benenson, Itzhak, Karel Martens, Yodan Rofé, and Ariela Kwartler. 2011. "Public transport versus private car GIS-based estimation of accessibility applied to the Tel Aviv metropolitan area." *The Annals of Regional Science* 47 (3):499-515. doi: 10.1007/s00168-010-0392-6.
- Boyce, Katina, and David Treering. 2012. Cincinnati Fresh Food Retail Financing Fund: Proposed Elements. Cincinnati, OH: Cincinnati City Council.
- Bullard, R.D., G.S. Johnson, and A.O. Torres, eds. 2004. *Highway Robbery: Transportation Racism & New Routes to Equity*. Cambridge, MA: South End Press.
- Burbidge, S.K. 2012. Identifying a Profile for Non-Traditional Cycle Commuters. Salt Lake City: Utah Department of Transportation Research Division.
- Burgoine, Thomas, and Pablo Monsivais. 2013. "Characterising food environment exposure at home, at work, and along commuting journeys using data on adults in the UK." *International Journal of Behavioral Nutrition and Physical Activity* 10 (1):85.
- Chicago, City of. 2013. "Mayor Emanuel Announces Release of Food Desert Data and New Interactive Efforts to Combat Food Deserts in Chicago." Accessed 11/19/13. [http://www.cityofchicago.org/city/en/depts/fss/provdrs/emerg/news/2013/aug/mayor\\_emanuel\\_announcesreleaseoffooddesertdataandnewinteractive.html](http://www.cityofchicago.org/city/en/depts/fss/provdrs/emerg/news/2013/aug/mayor_emanuel_announcesreleaseoffooddesertdataandnewinteractive.html).
- Church, A., M. Frost, and K. Sullivan. 2000. "Transport and social exclusion in London." *Transport Policy* 7 (3):195-205.
- Cooke, Kenneth L, and Eric Halsey. 1966. "The shortest route through a network with time-dependent internodal transit times." *Journal of mathematical analysis and applications* 14 (3):493-498.
- Cooley, James W, and John W Tukey. 1965. "An algorithm for the machine calculation of complex Fourier series." *Mathematics of computation* 19 (90):297-301.
- Cummins, S., and S. Macintyre. 2002. "'Food deserts'—evidence and assumption in health policy making." *BMJ: British Medical Journal* 325 (7361):436-438.
- Currie, Graham. 2010. "Quantifying spatial gaps in public transport supply based on social needs." *Journal of Transport Geography* 18 (1):31-41.

- Deka, D. 2004. Social and Environmental Justice Issues in Urban Transportation. In *The Geography of Urban Transportation*, edited by S. Hanson and G. Giuliano. New York: Guilford Press.
- Delbosc, Alexa, and Graham Currie. 2011. "Using Lorenz curves to assess public transport equity." *Journal of Transport Geography* 19 (6):1252-1259.
- Drew, Keith, and Martin Rowe. 2010. "Applying accessibility measures to assess a transport intervention strategy: A Case Study of Bromsgrove." *Journal of Maps* 6 (1):181-191.
- Farber, S., M.Z. Morang, and M.J. Widener. 2014. "Temporal variability in transit-based accessibility to supermarkets." *Applied Geography* 53:149-159. doi: <http://dx.doi.org/10.1016/j.apgeog.2014.06.012>.
- Farber, S., and A. Páez. 2012. "Activity Spaces and the Measurement of Clustering: A Case Study of Cross-linguistic Exposure in Montreal." *Environment and Planning A* 44 (2):315-332.
- Farber, Steven, Keith Bartholomew, Xiao Li, Antonio Páez, and Khandker M Nurul Habib. 2014. "Assessing social equity in distance based transit fares using a model of travel behavior." *Transportation Research Part A: Policy and Practice* 67:291-303.
- Foth, Nicole, Kevin Manaugh, and Ahmed M El-Geneidy. 2013. "Towards equitable transit: examining transit accessibility and social need in Toronto, Canada, 1996–2006." *Journal of Transport Geography* 29:1-10.
- Frank, L. D. 2000. "Land Use and Transportation Interaction: Implications on Public Health and Quality of Life." *Journal of Planning Education and Research* 20 (1):6.
- Fuller, Daniel, Steven Cummins, and Stephen A Matthews. 2013. "Does transportation mode modify associations between distance to food store, fruit and vegetable consumption, and BMI in low-income neighborhoods?" *The American Journal of Clinical Nutrition* 97 (1):167-172. doi: 10.3945/ajcn.112.036392.
- Geurs, Karst T, and Bert Van Wee. 2004. "Accessibility evaluation of land-use and transport strategies: review and research directions." *Journal of Transport geography* 12 (2):127-140.
- Golub, Aaron, and Karel Martens. 2014. "Using principles of justice to assess the modal equity of regional transportation plans." *Journal of Transport Geography* 41 (0):10-20. doi: <http://dx.doi.org/10.1016/j.jtrangeo.2014.07.014>.
- Grengs, Joe. 2002. "Community-based planning as a source of political change: The transit equity movement of Los Angeles' Bus Riders Union." *Journal of the American Planning Association* 68 (2):165-178.
- Haghani, Ali, and Soojung Jung. 2005. "A dynamic vehicle routing problem with time-dependent travel times." *Computers & operations research* 32 (11):2959-2986.
- Handy, Susan L, and Debbie A Niemeier. 1997. "Measuring accessibility: an exploration of issues and alternatives." *Environment and planning A* 29 (7):1175-1194.
- Hansen, Walter G. 1959. "How Accessibility Shapes Land Use." *Journal of the American Institute of Planners* 25 (2):73-76. doi: 10.1080/01944365908978307.
- He, K, FB Hu, GA Colditz, JE Manson, WC Willett, and S Liu. 2004. "Changes in intake of fruits and vegetables in relation to risk of obesity and weight gain among middle-aged women." *International Journal of Obesity* 28 (12):1569-1574.
- Hendrickson, D, C Smith, and N Eikenberry. 2006. "Fruit and vegetable access in four low-income food deserts communities in Minnesota." *Agriculture and Human Values* 23 (3):371-383.

- Henk, Russell H, and Sarah M Hubbard. 1996. "Developing an index of transit service availability." *Transportation Research Record: Journal of the Transportation Research Board* 1521 (1):12-19.
- Higdon, JV, B Delage, DE Williams, and RH Dashwood. 2007. "Cruciferous vegetables and human cancer risk: epidemiologic evidence and mechanistic basis." *Pharmacological Research* 55 (3):224-236.
- Hill, Arthur V, and WC Benton. 1992. "Modelling intra-city time-dependent travel speeds for vehicle scheduling problems." *Journal of the Operational Research Society*:343-351.
- Hine, J. 2003. "Social exclusion and transport systems." *Transport Policy* 10 (4):263-263.
- Holsten, Joanna E. 2009. "Obesity and the community food environment: a systematic review." *Public Health Nutrition* 12 (03):397-405. doi: doi:10.1017/S1368980008002267.
- Horner, Mark W., and Daniel Schleith. 2012. "Analyzing temporal changes in land-use–transportation relationships: A LEHD-based approach." *Applied Geography* 35 (1–2):491-498. doi: <http://dx.doi.org/10.1016/j.apgeog.2012.09.006>.
- Hunter-Zaworski, Katherine. 2003. Transit capacity and quality of service manual. In *Transit Cooperative Research Program*, edited by Transportation Research Board.
- Ichoua, Soumia, Michel Gendreau, and Jean-Yves Potvin. 2003. "Vehicle dispatching with time-dependent travel times." *European journal of operational research* 144 (2):379-396.
- Kaplan, Sigal, Dmitrijs Popoks, Carlo Giacomo Prato, and Avishai Avi Ceder. 2014. "Using connectivity for measuring equity in transit provision." *Journal of Transport Geography* 37:82-92.
- Kenyon, S. 2003. "Understanding social exclusion and social inclusion." *Proceedings of the Institution of Civil Engineers-Municipal Engineer* 156 (2):97-104.
- Kenyon, S. 2006. "Running on empty: Transport, social exclusion and environmental justice." *European Planning Studies* 14 (2):287-289.
- Kolata, Gina. 2012. "Studies Question the Pairing of Food Deserts and Obesity." Accessed 11/19/13. [http://www.nytimes.com/2012/04/18/health/research/pairing-of-food-deserts-and-obesity-challenged-in-studies.html?\\_r=1&](http://www.nytimes.com/2012/04/18/health/research/pairing-of-food-deserts-and-obesity-challenged-in-studies.html?_r=1&).
- Legrain, Alexander, Ron Buliung, and Ahmed El-Geneidy. 2015. "Who, what, when, and where: Revisiting the influences of transit mode share." *Transportation Research Record: Journal of the Transportation Research Board*.
- Lei, Ting L, Yali Chen, and Konstadinos G Goulias. 2012. "Opportunity-Based Dynamic Transit Accessibility in Southern California." *Transportation Research Record: Journal of the Transportation Research Board* 2276 (1):26-37.
- Lei, TL, and RL Church. 2010. "Mapping transit-based access: integrating GIS, routes and schedules." *International Journal of Geographical Information Science* 24 (2):283-304.
- Levine, Jonathan, Joe Grengs, Qingyun Shen, and Qing Shen. 2012. "Does Accessibility Require Density or Speed? A Comparison of Fast Versus Close in Getting Where You Want to Go in US Metropolitan Regions." *Journal of the American Planning Association* 78 (2):157-172.
- Levinson, D.M., and K. J. Krizek. 2005. *Access to Destinations*. Amsterdam: Elsevier.
- Liao, Felix Haifeng, Steven Farber, and Reid Ewing. 2015. "Compact development and preference heterogeneity in residential location choice behaviour: A latent class analysis." *Urban Studies* 52 (2):314-337.
- Litman, T. 2002. "Evaluating Transportation Equity." *World Transport Policy & Practice* 8 (2):50-65.

- Litman, T., and M. Brenman. 2011. *A New Social Equity Agenda For Sustainable Transportation*. Victoria: Victoria Transport Policy Institute.
- Liu, Zhili, Xudong Jia, and Wen Cheng. 2012. "Solving the last mile problem: Ensure the success of public bicycle system in Beijing." *Procedia-Social and Behavioral Sciences* 43:73-78.
- Lucas, K., T. Grosvenor, and R. Simpson. 2001. *Transport, the environment, and social exclusion*. York: Joseph Roundtree Foundation.
- Lucas, Karen. 2012. "Transport and social exclusion: Where are we now?" *Transport policy* 20:105-113.
- Martens, K. 2009. "Justice in transport as justice to access: applying Walzer's "Spheres of Justice" to the transport sector." 88th Annual Meeting of the Transportation Research Board, Washington, DC, USA, January 11-15.
- Martens, Karel, Aaron Golub, and Glenn Robinson. 2012. "A justice-theoretic approach to the distribution of transportation benefits: Implications for transportation planning practice in the United States." *Transportation Research Part A: Policy and Practice* 46 (4):684-695. doi: <http://dx.doi.org/10.1016/j.tra.2012.01.004>.
- Mavoia, Suzanne, Karen Witten, Tim McCreanor, and David O'Sullivan. 2012. "GIS based destination accessibility via public transit and walking in Auckland, New Zealand." *Journal of Transport Geography* 20 (1):15-22.
- McKinnon, R.A., J. Reedy, M.A. Morrissette, L.A. Lytle, and A.L. Yaroch. 2009. "Measures of the Food Environment: A Compilation of the Literature, 1990-2007." *American Journal of Preventive Medicine* 36 (4):124-133.
- Minnesota Population Center. 2011. *National Historical Geographic Information System: Version 2.0*. edited by University of Minnesota. Minneapolis, MN.
- Moniruzzaman, Md, and Antonio Páez. 2012. "Accessibility to transit, by transit, and mode share: application of a logistic model with spatial filters." *Journal of Transport Geography* 24:198-205.
- Moore, Latetia V., Ana V. Diez Roux, Jennifer A. Nettleton, and David R. Jacobs. 2008. "Associations of the Local Food Environment with Diet Quality—A Comparison of Assessments based on Surveys and Geographic Information Systems: The Multi-Ethnic Study of Atherosclerosis." *American Journal of Epidemiology* 167 (8):917-924. doi: 10.1093/aje/kwm394.
- Murray, Alan T, Rex Davis, Robert J Stimson, and Luis Ferreira. 1998. "Public transportation access." *Transportation Research Part D: Transport and Environment* 3 (5):319-328.
- O'Neill, Wende A, R Douglas Ramsey, and JaChing Chou. 1992. "Analysis of transit service areas using geographic information systems." *Transportation Research Record* (1364).
- O'Sullivan, David, Alastair Morrison, and John Shearer. 2000. "Using desktop GIS for the investigation of accessibility by public transport: an isochrone approach." *International Journal of Geographical Information Science* 14 (1):85-104.
- Orbis. "Company Information." Bureau Van Dijk Accessed June, 2013. <http://www.bvinfo.com/products/company-information/international/orbis>.
- Owen, A., and D.M. Levinson. 2014. "Modeling the Commute Mode Share of Transit Using Continuous Accessibility to Jobs." 93rd Annual Meeting of the Transportation Research Board, Washington, D.C., 01/15/2014.

- Owen, Andrew, and David M. Levinson. 2015. "Modeling the commute mode share of transit using continuous accessibility to jobs." *Transportation Research Part A: Policy and Practice* 74 (0):110-122. doi: <http://dx.doi.org/10.1016/j.tra.2015.02.002>.
- Páez, A., and S. Farber. 2012. "Jobs and the Single Parent: An Analysis of Accessibility to Employment in Toronto." *Urban Geography* Accepted.
- Páez, A., R. G. Mercado, S. Farber, C. Morency, and M. Roorda. 2009. Mobility and Social Exclusion in Canadian Communities: An Empirical Investigation of Opportunity Access and Deprivation. Report to Policy Research Directorate, Strategic Policy and Research, Human Resources and Social Development Canada.
- Páez, A., R. G. Mercado, S. Farber, C. Morency, and M. J. Roorda. 2010. "Relative Accessibility Deprivation Indicators for Urban Settings: Definitions and Applications to Food Deserts in Montreal." *Urban Studies*.
- Páez, Antonio, Darren M. Scott, and Catherine Morency. 2012. "Measuring accessibility: positive and normative implementations of various accessibility indicators." *Journal of Transport Geography* 25 (0):141-153. doi: <http://dx.doi.org/10.1016/j.jtrangeo.2012.03.016>.
- Polzin, Steven E, Ram M Pendyala, and Sachin Navari. 2002. "Development of time-of-day-based transit accessibility analysis tool." *Transportation Research Record: Journal of the Transportation Research Board* 1799 (1):35-41.
- Preston, John, and Fiona Rajé. 2007. "Accessibility, mobility and transport-related social exclusion." *Journal of Transport Geography* 15 (3):151-160. doi: <http://dx.doi.org/10.1016/j.jtrangeo.2006.05.002>.
- Pucher, J. 2004. "Public Transportation." In *The Geography of Urban Transportation*, edited by S. Hanson and G. Giuliano, 199-236. New York: Guilford Press.
- Raje, F. 2004. "Engineering social exclusion? Poor transport links and severance." *Proceedings of the Institution of Civil Engineers-Municipal Engineer* 157 (4):267-273.
- Ramsey, Krista. 2011. "Lost in the Food Desert." Gannett Accessed 11/19/13. [http://news.cincinnati.com/article/20111112/COL01/111130305/?nlick\\_check=1](http://news.cincinnati.com/article/20111112/COL01/111130305/?nlick_check=1).
- Ritter, B. 2014. "When & Where: Temporal Analysis of the Wasatch Front's Public Transit Network." Master of Science, Department of Geography, University of Utah.
- Rood, Timothy, and Sharon Sprowls. 1998. *The local index of transit availability: An implementation manual*: Local Government Commission.
- Ryus, Paul, Jon Ausman, Daniel Teaf, Marc Cooper, and Mark Knoblauch. 2000. "Development of Florida's transit level-of-service indicator." *Transportation Research Record: Journal of the Transportation Research Board* 1731 (1):123-129.
- Sanchez, Thomas W, Qing Shen, and Zhong-Ren Peng. 2004. "Transit mobility, jobs access and low-income labour participation in US metropolitan areas." *Urban Studies* 41 (7):1313-1331.
- Shaheen, Susan, Stacey Guzman, and Hua Zhang. 2010. "Bikesharing in Europe, the Americas, and Asia: past, present, and future." *Transportation Research Record: Journal of the Transportation Research Board* (2143):159-167.
- Spear, BD. 2011. Improving employment data for transportation planning. In *American Association of State Highway and Transportation Officials*.
- Tian, Guang, Reid Ewing, and William Greene. 2014. "Desire for Smart Growth: A Survey of Residential Preferences in the Salt Lake Region of Utah." *Housing Policy Debate* (ahead-of-print):1-17.

- Welch, Timothy F. 2013. "Equity in transport: The distribution of transit access and connectivity among affordable housing units." *Transport Policy* 30:283-293.
- Welch, Timothy F, and Sabyasachee Mishra. 2013. "A measure of equity for public transit connectivity." *Journal of Transport Geography* 33:29-41.
- Widener, Michael J, Steven Farber, Tijds Neutens, and Mark W Horner. 2013. "Using urban commuting data to calculate a spatiotemporal accessibility measure for food environment studies." *Health & place* 21:1-9.
- Widener, Michael John, Steven Farber, Tijds Neutens, and Mark W. Horner. 2014. "Spatiotemporal Accessibility to Supermarkets Using Public Transit: Interaction Potential Approach in Cincinnati, Ohio." TRB 93rd Annual Meeting, Washington, DC.
- Willumsen, LG. 1984. "Estimating time-dependent trip matrices from traffic counts." Ninth International Symposium on Transportation and Traffic Theory, VNU Science Press.
- Wong, SC, and CO Tong. 1998. "Estimation of time-dependent origin–destination matrices for transit networks." *Transportation Research Part B: Methodological* 32 (1):35-48.

## **APPENDIX A**



In addition to the research in this project report, this NITC project funded several collaborations that have resulted in academic publications not directly contained herein. Below is a list of outputs and their abstracts.

1. Lamers, J., Farber, S., & Habib, K. N. (2016). Accessibility to Community Services in the City of Toronto: A Perspective on Social Disadvantage and Homelessness. In *Transportation Research Board 95th Annual Meeting* (No. 16-2863).

Local governments and non-profit organizations fund and operate many services and programs throughout the city including those which are specifically for the benefit of homeless, low-income or otherwise socially disadvantaged people. Accessing these services can sometimes be difficult, especially for those who do not have access to private automobiles. For such transit dependent people, transit accessibility to these services are crucial. This research measures the travel time by transit to various community services in the City of Toronto from each census tract and homeless shelter as a measure of transit accessibility. Each census tract was scored with an index of social disadvantage using four characteristics from census data. It was found that access was highest in the old City of Toronto due the dense concentration of services and more frequent transit supply. Places in the outer parts of the city face concerning levels of inaccessibility to community services. However, transit access to community services do not vary across the levels of social disadvantage. The travel time matrices generated in this research can serve as a guide for policy analysis for future service expansion.

2. Fransen, K., Neutens, T., Farber, S., De Maeyer, P., Deruyter, G., & Witlox, F. (2015). Identifying public transport gaps using time-dependent accessibility levels. *Journal of Transport Geography*, 48, 176-187.

One of the concerns that has aroused much scholarly attention in transport geography lately is the extent to which public transport provision enables the less privileged population segments, especially those without privately owned motorized vehicles, to participate in activities that are deemed normal within the society they live in. This study contributes to this line of inquiry by proposing a methodology for identifying public transit gaps, a mismatch between the socially driven demand for transit and the supply provided by transit agencies. The methodology draws on the latest accomplishments in the field of modeling time-continuous, schedule-based public transport accessibility. Accessibility levels to key destinations are calculated at regular time intervals, and synoptic metrics of these levels over various peak and off-peak time windows are computed for weekdays and weekends. As a result, a temporally reliable picture of accessibility by public transport is constructed. The obtained index of public transport provision is compared to a public transport needs index based on the spatial distribution of various socio-demographics, in order to highlight spatial mismatches between these two indices. The study area consists of Flanders, which is the northern, Dutch-speaking region of Belgium. The results indicate that mainly suburban areas are characterized by high public transport gaps. Due to the time-variability of public transport frequencies, these gaps differ over time.

3. Widener, M. J., Farber, S., Neutens, T., & Horner, M. (2015). Spatiotemporal accessibility to supermarkets using public transit: an interaction potential approach in Cincinnati, Ohio. *Journal of Transport Geography*, 42, 72-83.

Improving nutrition in urban regions involves understanding which neighborhoods and populations lack access to stores that sell healthy foods, such as fruits and vegetables. To this end, recent work has focused on mapping regions without access to places like supermarkets, often terming them 'food deserts'. Until recently, this work has not considered residents' mobility as facilitated by transportation systems, and even among those that do, few have considered alternative forms of transportation, like public transit, opting for automobile-oriented travel assumptions. This paper analyzes people's spatio-temporal constraints to accessing supermarkets, and focuses on the transit commuting population. Analysis of commute data from Cincinnati, Ohio shows there are a significant number of residents that have improved access to supermarkets when a grocery shopping trip is made on the way home from work, than if they were to depart from their home location. These results extend previous work showing relatively few automobile commuting residents have better access to supermarkets given their work locations.



## **APPENDIX B**

## **TRAVEL TIME CUBE DOCUMENTATION**

Instructions to create and analyze travel time cubes using GTFS data and ArcGIS

**Contact:**

Dr. Steve Farber

University of Toronto Scarborough

1265 Military Trail, Toronto, ON M1C 1A4, Canada

(416) 208-5120

[steven.farber@utoronto.ca](mailto:steven.farber@utoronto.ca)

**Prepared by:**

Liwei Fu

City of San José

200 E. Santa Clara St., San José, California, 95113

[folflw@gmail.com](mailto:folflw@gmail.com)

## BRIEF OVERVIEW

Calculating travel time cubes in ArcGIS requires the following steps which will be explained in further detail in this instruction manual:

- 1) Remote desktop connection
- 2) Create network dataset
- 3) Generate origin-destination matrix
- 4) Analyze travel time cube

If you wish to create OD matrix from locations that are not directly connected to your street files (e.g. census block centroids), start from Part 1, connecting to the remote desktop and follow Parts 1-3. It will help you to create connectors between locations and street line shapefiles so that travel times from locations to nearest streets can be accurately estimated.

Part 2, creating the network dataset, is heavily based on the instructions for the ArcGIS add-in toolbox [Add GTFS to a Network Dataset](#). Any user should consult the [users' guide](#) when creating their network dataset. This is also a helpful source of information on how the shortest path algorithms function, and how to troubleshoot using GTFS data with the network analyst.

If you've already created your network dataset, you should connect to the geoprocessing server using Step 1, and then skip to Step 3 to generate travel cube.

## REMOTE DESKTOP CONNECTION

In order to use the servers for travel time cube calculation, you need to use the secure remote desktop gateway to access farber[1-3].geog.utah.edu servers.

To access the servers from off campus on a Window PC:

1. Launch Remote Desktop Connection. It should be located in Start->All Programs->Accessories. Enter the computer's hostname in the Computer field (for example, farber1.geog.utah.edu). Enter your username as ADCSBS\u0xxxxxx.

2. Click options below the address window. Go to Advanced. Click on Settings... under Connect from anywhere and change the following settings:

- a. Select the radio button titled "Use these RD Gateway server settings"
- b. Enter server name: csbs-rdg1.csbs.utah.edu
- c. Select logon Method: Ask for password (NTLM)
- d. Check Bypass RD Gateway Server for local addresses
- e. Check Use my RD Gateway credentials for the remote computer. Hit OK to save settings.

3. Click Connect.

To access the servers from off campus on a MAC:

1. Download Microsoft Remote Desktop from the App Store. Note that Remote Desktop Connection, commonly installed with Office:Mac 2011, will not work because it doesn't support the appropriate Remote Gateway protocol.

2. Launch Microsoft Remote Desktop. Click Preferences. Change the following settings:

- a. Gateway name: csbs-rdg1.csbs.utah.edu
- b. Server: csbs-rdg1.csbs.utah.edu
- c. User name: ADCSBS\U0xxxxxx

- d. Password: Ask user to enter, or leave blank. Click the red X in the upper left of the window. Settings should save automatically.
3. In the main Microsoft Remote Desktop Window, click the New button in the upper-left corner. In the new window that pops up, change the following settings:
  - a. Connection name: hostname of remote computer.
  - b. PC name: FQDN of remote computer (for example: farber1.geog.utah.edu)
  - c. User name: ADCSBS\U6xxxxxxx
  - d. Password: Ask user to enter, or leave blank. Again, click the red X to close the window.
4. Click Start to initiate a connection. The user will probably see a certificate warning. Instruct them to simply click Continue if this appears.



## CREATE NETWORK DATASET

### 1.0 DOWNLOAD AND INSTALL ADD GTFS TO A NETWORK DATASET TOOL

This tool requires to register a GTFS transit evaluator for ArcGIS Network Analyst extension. In order to install the evaluator, download *Add GTFS to a Network Dataset* from [ArcGIS Online](#) and install it following step 1) in the [user's guide](#).

### 1.1 FUNCTIONS

This toolbox creates a network dataset based on a pedestrian network and General Transit Feed Specification (GTFS) package. The major difference between these guidelines, and the ones included with the Add GTFS to Network Dataset toolbox is that link connectors between locations and the pedestrian network are explicitly created in these steps. This means that travel times from locations to street networks are estimated. This is important when locations are not geocoded within a quick walking distance of the street network. If this is not a concern, the documentation supplied with the Add GTFS to Network Dataset toolbox is recommended.

### 1.2 COMPONENTS

The travel cube calculation folder contains the following files and folders:

**Scripts:** GenerateConnectors.py, GenerateTransitSchedule.py

**Templates Folder:** UtahTransit.gdb, GTFS data

### 1.3 PREPROCESS DATABASE

Before creating network dataset, area of interest needs first to be determined. The sample data uses Salt Lake Valley area. Preprocess database by the following steps:

1. Gather required data (GTFS data, Census Population and Boundary Data, Road Data (update roads near transit stops & Origins/Destinations ));
2. Select only BGs within 5 km of transit stops.
3. Select only blocks within these selected BG's.

4. Generate census block points with population from block polygons.
5. Generate census block group population-weighted mean centroids (Mean Center) to use as origins/destinations.
6. From these, select only centroids within 5 km of transit stops.
7. Edit road data for analysis purposes (added missing roads near transit stops and block group centroids).
8. Move centroids closer to roads by placing them on populated areas closest to original point.
9. Created topology and verified that no “real” dangles were in the study area

## 1.4 GENERATE CONNECTORS

This script creates connectors from census block group centroids to road network as well as stop-street connectors.

**Scripts:** GenerateConnectors.py

The script uses following input variables:

**outFD:** Location of file geodatabase where network dataset will be created. It is recommended to specify a feature dataset for the purpose of creating network dataset in the next steps.

**inGTFSdir:** Location of GTFS directories. The template uses Utah transit network from April 13<sup>th</sup>, 2014 to August 16<sup>th</sup>, 2014. GTFS files can be acquired from <http://www.gtfs-data-exchange.com/>

**OrgDest:** Location of census block group centroids feature class as origins and destinations.

**origRoad:** Transit road network including roads near transit stops and BG centroids.

The script creates following outputs under the select geodatabase directory (outFD):

**TransitLines:** Line feature class of straight lines between each pair of stops that is directly connected by a transit trip. Note that these are not the actual transit routes but merely straight lines between transit stops.

**Stops\_Snapped2Streets:** Point feature class of GTFS transit stops snapped to the closest street lines.

**Connectors\_Stops2Streets:** Lines feature class of connectors between transit stops and the corresponding snapped-to-street points.

**Streets\_UseThisOne:** A copy of the street lines feature that selected as input with vertices at the locations of snapped-to-street points. This is the streets feature class you should use in your network dataset instead of your original streets feature class.

**GTFS.sql:** SQL database containing processed GTFS data. No further process need to be made to this file, but don't delete it because it will be used by the network dataset during analysis.

## 1.5 CREATE NETWORK DATASET

After creation of street lines and connector, you need to create network dataset from the previous output feature dataset (outFD). See [Create your network dataset using correct connectivity groups](#) for further instructions.

## 1.6 GENERATE TRANSIT SCHEDULE TABLES

This script generates file geodatabase tables of GTFS schedule information in the correct format for use in the GTFS evaluator in the network dataset. It also retrieves the network dataset's EIDs for the transit lines features and adds the EIDs to the transit schedule table used in the GTFS evaluator. The network dataset must be built prior to running this tool, and the tool must be re-run every time the network dataset is rebuilt in order to update the EID values.

**Scripts:** GenerateTransitSchedule.py

The script uses following input variables:

**inNetworkDataset:** Network dataset created in the last step containing TransitLines

**scriptsDir:** Location of the scripts

The script creates following outputs under the geodatabase directory containing selected network dataset . Note that it should be the same FGDB containing "GTFS.sql":

**TransitScheduleTable:** Start time and end time for transit lines.

**TransitSchedule\_Calendar:** Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.

**TransitSchedule\_CalendarExceptions:** Exceptions for the service IDs defined in the calendar.txt file.

## **GENERATE ORIGIN-DESTINATION MATRIX**

This toolbox creates travel time cubes using network dataset created in the last step. It estimates shortest path travel time from location to location (census block group centroid) making use of a public transit schedule. And it calculates OD cost matrix for every minute of a day.

Before using this tool, make sure the GTFS transit evaluator file is registered on the machine. If not, follow step 1.0 in *Create network dataset* to install it.

This script automates the travel cube calculation process by looping OD matrix calculation over every minute. It further stores the results in HDF5 files. Every dataset in a file represents an OD matrix of a single minute. The X axis represents origins, while Y axis represents destinations. The sequence of origins and destinations in the matrix will match the uid (ie. Object ID) of input points.

The multiprocessing Python module is used to distribute workload to multiple processes on a server. Each CPU core processes network dataset individually and output a HDF5 file. The start time of each cube is defined by data given to worker function, and the size of travel cube is defined within the worker function:

```

import multiprocessing
from multiprocessing import Process
import sys

# Define worker function to calculate travel time cube at given time
def CalculateTravelCube((Hour, Minute)):
    ...
    #Calculate 30 minute at a time
    LoopTime = 30
    ...

if __name__ == '__main__':
    #Use multiprocessing module to 30 minute travel cube on each processor
    #Create a pool class to run the job
    pool = multiprocessing.Pool()
    print "Start to run CalculateTravelCube"
    #Define the number of jobs. We use 16 here, each maps to a processor, calculating total 8 hours
of travel cube. Job is defined as [hour, minute]
    data = ([0, 0], [0, 30], [1, 0], [1, 30], [2, 0], [2, 30], [3, 0], [3, 30], [4, 0], [4, 30], [5, 0], [5, 30], [6, 0],
[6, 30], [7, 0], [7, 30])
    #Synchronize the main process with the job processes
    pool.map(CalculateTravelCube, data)
    pool.close()
    print "Done with CalculateTravelCube"
    #End main

```

If you wish to calculate for a different time period, change *data* and *LoopTime* accordingly. If the calculation amount is relatively small, you can choose not to use the multiprocessing module and run *CalculateTravelCube* function directly.

Scripts: GenerateODMatrix.py

The script uses following input variables:

**InNetworkDataset:** Network dataset created after running last steps.

**OrgDest:** Point layer to be analyzed as origins and destinations. If you wish to calculate OD matrix from different origins and destinations, define and add then separately after creating OD matrix (in *arcpy.AddLocations\_na* after *arcpy.MakeODCostMatrixLayer\_na*).

**Date:** Assign travel date to be analyzed. Note that data must be selected within the selected GTFS package time period

**matrix\_size:** number of census block group centroids

**OutFile:** Directory of output travel time cubes.

Be sure to match census block group centroid uid name in `arcpy.AddLocations_na` function, here we use 'BGGEOID'

The script creates a series of HDF5 travel time cubes. Each half hour is calculated simultaneously and stores as a HDF5 file, containing OD matrix of 30 minutes. For example, the travel time cube for the first half hour of a day will be stored as "TravelcubeUT14\_000AM\_30.h5py".

The sample script makes use of multiprocessing function (`multiprocessing.py`) in order to speed up calculation process. In the main function, we mapped a total of 16 cores, with each CPU calculate travel cube of half an hour. As a result, it creates travel cube of 8 hours at a time. With all three servers full functioning, the 24 hour cube can be calculated at the same time by assigning different time periods to different servers. This function cannot print outputs interactively in python shell, so a log file is created in the directory of the scripts to print out the progresses of each processor. For example, the output of the processor that calculates the first half hour is stored in "\Log\out-00.log".

# TRAVEL TIME CUBE ANALYSIS TOOL

## 1.1 FUNCTIONS

This toolbox reads and analyzes the travel time cube, creates summary statistic information, Fourier transform results of travel frequencies, and accessibility analysis.

## 1.2 COMPONENTS

The travel cube analysis folder contains the following files and folders:

**Scripts:** TravelCubeIO.py, CalStatistics.py, Fourier.py

**Templates Folder:** TravelCube, LEHD

## 1.3 READ AND WRITE TRAVEL CUBE

This script reads the whole travel time cube into memory. A travel cube with 1197 origins/destinations takes 16GB memory and will function quickly with the capacity of the server (32GB memory). As the cubes are stored in a total of 48 files, it takes a great amount of time for HDF5 file I/O. This script also saves the travel time cube into one single file and provides I/O for that file.

**Scripts:** TravelCubeIO.py, CalStatistics.py, Fourier.py

The script uses following input variables:

**Filepath:** Location of HDF5 files. It is defined to match the names of each travel time cube.

**msize:** Size of input matrix length to store travel cube, which is equal to the number of O/D locations.

**outfile:** Location of a single hdf5 travel cube file

The script creates following outputs under the select directory (outfile):

**TravelCube.hdf5:** A single file containing OD matrix of every minute of a day, the size of this file is  $1440*n*n$ , in which  $n$  represent number of locations in the OD matrix. The first axis (axis = 0) indexes minute of a day, the second axis (axis = 1) indexes destinations, and the third axis (axis = 2) indexes origins.



## 1.4 GENERATE SUMMARY STATISTICS

This script calculates summary statistics of the travel time cube, including mean, standard deviation, and coefficient of variance.

**Scripts:** CalStatistics.py

The script makes use of the 3D travel cube matrix that is already read into memory in the last step (TravelCube). It uses the following input variable:

**msize:** Size of input matrix length to store travel cube, which is equal to the number of O/D locations.

The script outputs the mean, std. and CV of the entire cube (ie. mean = 996.72). It also creates the following outputs under the script directory:

**MinuteMean.txt:** A list of average travel time for every minute of a day. The length should equal to number of minutes (1440).

**Org\_mean.txt:** A list of average travel time from every origin to all destinations. The length should equal to number of locations in file.

**Dest\_mean.txt:** A list of average travel time from all origins to every destination. The length should equal to number of locations in file.

**Org\_std.txt:** A list of travel time std. from every origin to all destinations. The length should equal to number of locations in file.

**Dest\_std.txt:** A list of travel time std. from all origins to every destination. The length should equal to number of locations in file.

**MinCV.txt:** A list of travel time CV for every minute of a day. The length should equal to number of minutes (1440).

**CVOrg.txt:** A list of travel time CV from every origin to all destinations. The length should equal to number of locations in file.

**CVDest.txt:** A list of travel time CV from all origins to every destination. The length should equal to number of locations in file.

## 1.5 FOURIER TRANSFORM

This script calculates average travel cycle as well as std. of travel cycle using Fourier Transform. As the travel frequency at night time with no or little transit services will be approaching to 0, with cycling approaching to infinite, this script uses travel time at daytime from 6am to 7pm ( $t = [360:1140]$ ).

**Scripts:** Fourier.py

The script makes use of the 3D travel cube matrix that is already read into memory in the last step (TravelCube). It uses the following input variables:

**msize:** Size of input matrix length to store travel cube, which is equal to the number of O/D locations.

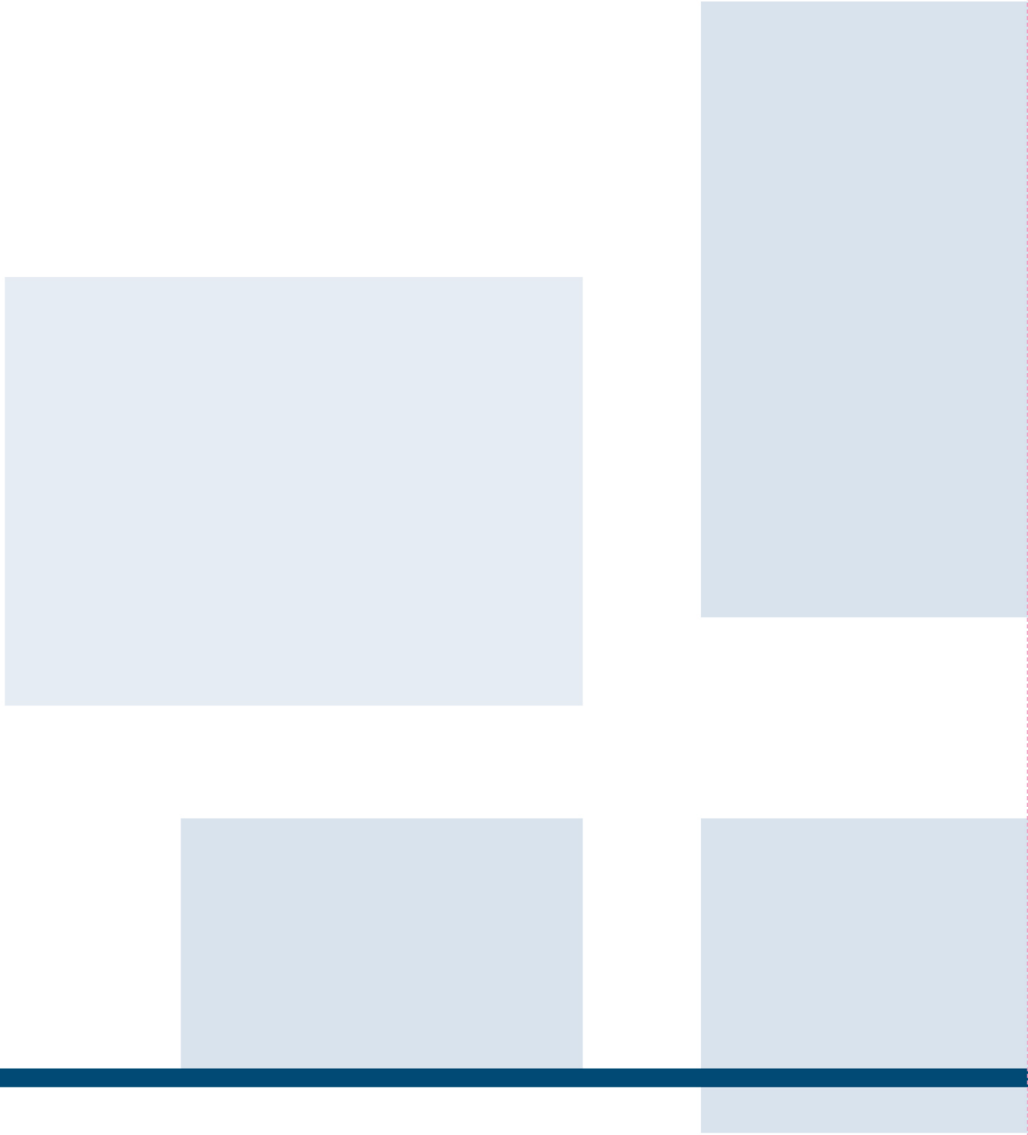
**threshold:** The threshold number to filter peak frequencies. The default is set as 0.75, suggesting that the amount of frequencies that is larger than 75% the maximum frequency amount is considered as a peak frequency and will be stored in the result dictionary.

The script outputs the mean, std. and CV of the entire cube (ie. mean = 996.72). It also creates the following outputs under the script directory:

**FourierDict:** A dictionary containing frequencies for every OD pair. The key is the OD pair, while the values are the detected frequencies of the OD pair. For example, the travel cycle from point 0 to point 1 is:  $\text{FourierDict}(1, 0) = 30, 60, 180, \dots$

**FreMeanUT14:** A matrix of average travel cycle for every OD pair.

**FreStdUT14:** A matrix of travel cycle std. for every OD pair.



Transportation Research and Education Center  
Portland State University  
1900 S.W. Fourth Ave., Suite 175  
Portland, OR 97201