INVESTIGATING TRAVEL BEHAVIOR IN TRANSIT-ORIENTED DEVELOPMENT: TOWARD SUSTAINABLE AND MULTIMODAL MOBILITY

A Dissertation
Presented to
The Academic Faculty

by

Yunkyung Choi

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of City and Regional Planning

Georgia Institute of Technology
December 2021

INVESTIGATING TRAVEL BEHAVIOR IN TRANSIT-ORIENTED DEVELOPMENT: TOWARD SUSTAINABLE AND MULTIMODAL MOBILITY

Approved by:

Dr. Subhrajit Guhathakurta, Advisor
School of City and Regional Planning
Georgia Institute of Technology

Dr. Catherine Ross
School of City and Regional Planning
Georgia Institute of Technology

Dr. Patricia Mokhtarian
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Clio Andris
School of City and Regional Planning &
School of Interactive Computing
Georgia Institute of Technology

Dr. Giovanni Circella
School of Civil and Environmental Engineering
Georgia Institute of Technology

Date Approved: July 20, 2021

ACKNOWLEDGEMENTS

I am grateful to everyone who has helped me through my academic journey at the Georgia Institute of Technology. First of all, I would like to express my sincere gratitude to my advisor, Dr. Subhrajit Guhathakurta, for his insightful guidance, invaluable support, and constant encouragement. He has opened many opportunities for me to work on various research projects that have strengthened my capacity to complete this dissertation. I would also like to thank my committee members: Dr. Catherine Ross, Dr. Patricia Mokhtarian, Dr. Clio Andris, and Dr. Giovanni Circella, for their valuable guidance. This dissertation would not be possible without their unique insights and perspectives on the substantive topics of this dissertation and critical advice on improving this dissertation. I am also thankful to all the faculty, staff, and students at the School of City and Regional Planning for their support during this graduate journey.

Last and most importantly, I would love to thank my family for their unconditional love and support. I cannot thank my husband, Yeongjun, and my daughter, Leah, enough for accompanying this journey. I would also like to give special thanks to my mother and father, who taught me to love learning and do my best in all my endeavors.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

ACS	American Community Survey
API	Application Programming Interface
ARC	Atlanta Regional Commission
CTA	Chicago Transit Authority
CTOD	Center for Transit-Oriented Development
FTA	Federal Transit Administration
HHI	Herfindahl-Hirschman Index
LEHD	Longitudinal Employer Household Dynamics
LODES	Longitudinal Origin-Destination Employment Statistics
MARTA	Metropolitan Atlanta Rapid Transit Authority
MSA	Metropolitan Statistical Area
NHTS	National Household Travel Survey
OR	Odds Ratio
OSM	Open Street Map
PSM	Propensity Score Matching
SA	Station Area
TAZ	Traffic Analysis Zone
TNC	Transportation Network Company
TOD	Transit-Oriented Development
VMT	Vehicle Miles Traveled

SUMMARY

An extensive literature has shown that transit-oriented development residents (TOD) have lower automobile use and diverse travel modes due to easy access to transit, better walkability, and proximity to various amenities. While such benefits of TOD are generally expected, the degree to which TODs influence travel behavior is still debatable. Besides, TOD implementation differs by context, and not all transit areas are developed along TOD principles. This variation in transit areas leads to different impacts on transportation outcomes. Although past studies have developed different TOD typologies, they are limited to a particular city or region.

The other ongoing debate in land use and travel behavior field is the emergence of new mobility services that enable users to utilize a mode of transport on an as-needed basis. Recent advances in information technologies have facilitated new mobility services that meet travelers' diverse needs, such as transportation network companies (TNCs), ridesharing, car sharing, bike sharing, microtransit, and shared autonomous vehicles. While new mobility services are expected to play an important role—either positive or negative—in planning how TODs can be implemented, the impacts and consequences of such services on traditional modes of transport such as public transit are still not well understood.

In doing so, this dissertation investigates different modes of transportation in TOD areas by posing the following research questions: 1) do people walk more in transit-oriented developments? 2) are residents more multimodal in transit-oriented developments? and 3) what is the potential impact of new mobility services on public transit demand?

For the first question, this dissertation addresses the effect of rail transit access on walking behavior in TOD areas. TODs are compared to other similar areas without rail transit access to determine whether people tend to walk in TOD areas for purposes other than transit use in Atlanta. The second question is addressed by identifying different TOD types on their impacts on residents' multimodal behavior to capture various conditions of existing TODs and their heterogeneous outcomes. This research identifies different types of 4,400 transit areas—a half-mile buffer area from rail station—in the U.S. and develops several analytic models to explain the multimodal traveler behavior in the 2017 NHTS. The third question examines the potential impacts of TNCs on transit demand in Chicago, with a particular focus on understanding heterogeneity in the effects by employing fixed effects panel regression models. By investigating various travel behavior around transit station areas, this dissertation provides insights on how TODs can be better implemented to promote sustainable and multimodal travel behavior.

CHAPTER 1. INTRODUCTION

Transit-oriented development (TOD) is a planning approach for areas around transit stations that promotes higher densities and a mix of land uses together with walkable streets to incentivize non-motorized travel and the use of transit. Over the last decade, the proponents have claimed that TODs can relieve various urban problems, such as air pollution, traffic congestion, affordable housing shortages, and sprawl (Cervero et al., 2002). Therefore, local governments have established and implemented plans for TODs as an effective solution in promoting social, economic, and environmental sustainability within neighborhoods.

Among all the benefits of TOD, an extensive literature on land use and transportation interaction largely agreed that various land use characteristics, such as density, diversity, design, distance to transit, and destination accessibility, generate synergistic effects on travel behavior. TODs are a perfect example of integrating various land use and transportation characteristics that encourage residents to use public transit and non-motorized transportation instead of relying on private automobiles (Greenwald & Boarnet, 2001; Ewing & Cervero, 2010; Renne, Hamidi, & Ewing, 2016). Most studies regarding TOD examined its impacts on travel behavior, with a specific emphasis on how effectively it reduces car usage, promote non-motorized travel, and encourage transit ridership (Cervero, 1993; Boarnet & Crane, 2001; Chatman, 2006; Arrington & Cervero, 2008; Hale, 2014; Nasri & Zhang, 2014; Ewing & Hamidi, 2014; Langlois et al., 2015; Laham & Noland, 2017; Park et al., 2018). In general, scholars have found a high correlation between sustainable travel modes and TODs. However, the question that remains insufficiently addressed is whether the increased sustainable mode of transport is associated with urban form characteristics in the TODs independent of transit access. In other words, I am echoing Chatman's (2013) question: "Does TOD need the T?"

Given that Chatman and I both examine the impact of transit access in TODs, Chatman's paper is highly relevant to this discussion. Still, the objective of his study was somewhat different from the purpose of this research. While his research focuses on automobile miles driven in TOD areas, I examine walking activities that originate from such areas in this dissertation. Because researchers have tested primarily on the effects of TODs on automobile and transit use, we know far less about walking behavior in TODs. The limited number of studies that have analyzed the impact of TODs on walking behavior predominantly focused on ingress and egress modes of walking trips to transit, mainly for commuting purposes (Loutzenheiser, 1997; Alshalalfah & Shalaby, 2007). Since commuting trips account for a relatively small proportion of all travel, the effect of TODs on non-work-related walking behavior remains largely unknown. The share of walking trips for commuting purposes in 2017 was 7.8% of total walking trips, while non-work/school-related walking trips accounted for 81.6% (U.S. Department of Transportation, 2017). In Atlanta, the percentage of non-work/school-related walking trips was approximately 61% in 2011, lower than the national average but significant nonetheless. Thus, an analysis of non-work-related walking behavior in addition to the existing studies on walking trips for commuting can offer valuable insights on sustainability and community health in TOD areas.

While a large volume of existing literature examined the identification of TOD and evaluated the performance of TODs, more recent literature demonstrates a growing interest in typologies of station areas for the evaluation of TODs. Previous studies have shown that living in TOD increases the probability of diversifying travel modes due to easy access to transit, better walkability, and proximity to diverse amenities. While such benefits of TOD are generally expected, its implementation differs by context and the application of TOD principles. This variation in transit areas leads to different impacts on transportation outcomes. In empirical studies, many researchers have not shown consistent effects of different transit areas on transportation outcomes such as car ownership, VMT, and travel mode share. Although several studies have shown that residents near TODs use more non-motorized travel modes, these studies often fail to control for residential self-selection and the variation in TOD types (Higgins & Kanaroglou, 2016; Renne, Hamidi, & Ewing, 2017). Also, a limited number of studies covered transit areas of the entire U.S. and presented different effects of TOD types in a rigorous way. Thus, an analysis that captures various conditions of existing TODs and their heterogeneous outcomes on travel behavior can support planners in developing a context-based TOD planning approach.

The other ongoing debate in the land use and travel behavior field is the emergence of new mobility services that enable users to utilize a mode of transport on an as-needed basis. Particularly, the rapid growth of new mobility services has attracted significant attention from researchers and planners due to its transformative impacts on the transport sector. Recent advances in information technologies have facilitated new mobility services that meet travelers' diverse needs, such as ridehailing, ridesharing, car sharing, bike sharing, microtransit, and shared autonomous vehicles. While new mobility services are expected to play an important role, the impacts and consequences of such services on traditional modes of transport are still not well understood.

Theoretically, how new mobility services affect other modes of transport could be explained through two mechanisms (Hall, Palsson, & Price, 2018). On the one hand, new mobility services could complement public transit by expanding travel options for people without private automobiles. For example, ridesourcing services, such as Uber and Lyft, could fill the temporal and spatial gap in public transit's fixed route and fixed schedule. They could further reduce parking needs in TOD areas by reducing the need for car ownership and its associated costs. On the other hand, ridesourcing services could be an alternative mode of travel by encouraging riders to shift from public transit. Its flexible and convenient service based on real-time information could make public transit less competitive. Also, ridesourcing would generate new trips—induced travel—that did not exist in the absence of ridesourcing service. To date, however, the literature on the relationship between ridesourcing and public transit has offered differing results that largely depend on specific contexts and assumptions due to a lack of publicly available data on ridesourcing services.

This dissertation will fill in these gaps by examining travel behavior in transit-oriented development through three main research questions:

- Do people walk more in transit-oriented developments?
- Are residents more multimodal in transit-oriented developments?
- What is the potential impact of new mobility services on public transit demand?

This dissertation first examines the association between various attributes of TODs and the prevalence of walking trips for purposes other than transit use by testing the hypothesis that rail transit access by itself would generate more walking trips regardless of built environment characteristics. I first divided traffic analysis zones (TAZs) into two groups to achieve said objective. The first group consists of TAZs within the catchment area of the Metropolitan Atlanta Rapid Transit Authority (MARTA) rail stations, and the other group consists of TAZs outside the catchment area. I then identified pairs of TAZs from the two groups with similar built environment characteristics, making rail transit access the key differentiator. A propensity score matching method was employed to compare the built environment features of various TAZs. Finally, I excluded ingress and egress trips to and from transit stations and examined the remaining walking trips that originate from these two TAZ groups while controlling for sociodemographic and travel characteristics. This modeling approach allows me to determine whether the presence of rail transit, independent of built environment characteristics, significantly influences walking behavior. In conclusion, I found a strong positive association between the presence of rail transit access and the level of walking activity for commuting and non-commuting trips, ceteris paribus, suggesting that the presence of transit also encourages walking for all purposes.

Second, this dissertation estimates the effects of land use attributes around transit facilities on their residents' multimodal travel behavior. The multimodal travel pattern, or multimodality, refers to individuals using multiple travel modes for certain periods (Buehler & Hamre, 2015). Researchers are especially interested in multimodality because incentives and regulations for popularizing a particular mode of travel tend to be most effective for multimodal travelers. Thus, identifying travelers' multimodality and the drivers of their multimodality is a critical first step (Ralph et al., 2016). Given that living in TOD increases the probability of diversifying travel modes due to easy access to transit, better walkability, and proximity to shops, restaurants, and offices, it is likely that residents in TODs would be multimodal. To estimate the impact of different TOD types on residents' multimodal behavior, this work follows three analytical steps: 1) the classification of areas within a half-mile radius from the 4,400 fixed-guideway transit facilities in the U.S. using factor- and cluster-analysis techniques, 2) the measurement of multimodal travel behavior for all individuals in the 2017 NHTS, and 3) the development of a series of regression models to explain non-automobile mode share and multimodality level of residents in each type of TOD with the help of data about socio-demographics of individuals and the type of TOD where these individuals lived. The models also include the probability of individuals living in the other types of TOD except one's own as a control for residential self-selection. The results suggest that not all land use attributes promote sustainable and multimodal travel behavior in various contexts. The effects of land use variables on residents' multimodality for work and non-work tours show heterogeneity across station area types. In sum, this study sheds light on how much difference in multimodal travel behavior planners can expect by supporting to convert a transit area from one TOD type to another.

The third question is addressed by examining how ridesourcing services offered by transportation network companies (TNCs) are associated with public transit demand in Chicago, with a particular focus on understanding heterogeneity in the effects of the type of ridesourcing services. This study started from the fact that new mobility services have grown rapidly in recent years, while transit ridership has declined or stagnated over the same period. There is a growing body of literature on the impacts of TNCs on existing modes of travel, particularly focusing on the relationship of whether they are complementary or substitutionary. However, the literature has provided mixed results primarily due to the relative novelty of TNC services and the lack of publicly available trip-level data on those services. Thus, this study assesses the relationship between TNCs and public transit using a panel data set that combines information on rail transit ridership with TNC trip data published by Chicago Transit Agency (CTA). It employs panel regression models at the census tract level due to a data structure that combines cross-sectional and time-series data. Other sociodemographic factors contributing to transit demand, such as population density, employment density, household income level, car ownership, are considered in the models. The results suggest that exclusive TNC service complements rail transit ridership, while shared TNC service substitutes them. The findings are meaningful for local and regional agencies to craft policies that encourage the complementary effect.

This dissertation consists of three studies, each one employing a different data and analytical method to investigate various travel behavior in TOD. Chapter 2 compares sustainable modes of travel in TODs and non-TODs to examine the role of transit access on the prevalence of walking behavior around station areas for purposes other than transit use by employing propensity score matching and multi-level logistic regression models. Chapter 3 identifies different types of TOD using factor- and cluster-analysis techniques and measures the multimodality level of residents in each station area based on the Herfindahl-Hirschman Index (HHI) and Shannon's entropy. The study estimates different effects of land use attributes around transit facilities on their residents' multimodal travel behavior for work and non-work tour purposes through several regression models. Chapter 4 reviews the literature on disruptive technologies and their impacts on traditional modes of travel and implements the panel regression models by the type of TNC services to examine the relationship between TNCs and rail transit demand. Each chapter summarizes the literature review, presents data and methods used for the study, and discusses the results of the models. Chapter 5 concludes all the findings, contributions, limitations, and directions for future research.

CHAPTER 2. DO PEOPLE WALK MORE IN TRANSIT-ORIENTED DEVELOPMENTS?

2.1 Introduction

The vast and growing literature on the relationship between built environment and travel behavior has generally indicated that particular urban forms, such as transit-oriented development (TOD), encourage the use of public transit and non-motorized transportation (Greenwald & Boarnet, 2001; Ewing & Cervero, 2010). TOD refers to the design of residential and commercial areas around transit stations that maximizes transit access and minimizes automobile use. These areas tend to be high-density areas with mixed-land uses that are pedestrian-friendly. Several studies have shown that proper design of TODs encourages people to own fewer vehicles, drive less, and use more non-motorized modes of travel (Pushkarev & Zupan, 1977; Cervero et al., 2004; Evans et al., 2007; Haas et al., 2010; Suzuki, Cervero, & Iuchi, 2013; Gallivan et al., 2015). These studies also identified high-density, mixed-use, pedestrian-friendly environments, and quality public transit facilities and services as key characteristics of TODs that encourage people to drive less and walk more. In general, researchers have found a high correlation between TODs and walking activity. However, the question that remains inadequately addressed is whether increased walking is related to urban form characteristics in the TODs independent of transit access. In other words, this study echoes Chatman's (2013) question: "Does TOD need the T?"

Given that we both examine the impact of transit access in TODs, Chatman's paper is highly relevant to this discussion, but the objective of his study was somewhat different from the purpose of this paper. While his research focuses on automobile miles driven in TOD areas, this study examines walking activities that originate from such areas in this paper. Because researchers have focused primarily on testing the effects of TODs on automobile and transit use, we are relatively in the dark regarding the specifics of walking behavior in TODs. The limited number of studies that have tried to examine the impact of TODs on walking behavior focused predominantly on ingress and egress modes of walking to transit, primarily for commuting (Loutzenheiser, 1997; Alshalalfah & Shalaby, 2007). Since commute travel accounts for a relatively small proportion of all travel, the impact of TODs on non-work-related walking behavior remains largely unknown. In fact, the share of commuting trips in 2017 was 7.8% of total walking trips, whereas non-work/school-related walking trips accounted for 81.6% (U.S. Department of Transportation, 2017). In Atlanta, the percentage of non-work/school-related walking trips was approximately 61% in 2011, which was lower than the national average at the time but significant nonetheless. Therefore, an analysis of non-work-related walking behavior in addition to the existing studies on commuting trips can offer valuable insights on sustainability and community health in TOD areas.

This paper aims to investigate the relationship between various characteristics of TODs and the prevalence of walking trips for purposes other than transit use by testing the hypothesis that rail transit access by itself would generate more walking trips regardless of built environment characteristics. This study first divided traffic analysis zones (TAZs) into two groups to achieve said objective. The first group consists of TAZs within the catchment area of the Metropolitan Atlanta Rapid Transit Authority (MARTA) rail stations, and the other group consists of TAZs outside the catchment area. This study then identified pairs of TAZs from the two groups that have similar built environment characteristics, making rail transit access the key differentiator. A propensity score matching method was employed to compare the built environment features of various TAZs. Finally, this study excluded ingress and egress trips to and from transit stations. It examined the remaining walking trips that originate from these two TAZ groups while controlling for sociodemographic and travel characteristics. This modeling approach allows us to determine whether the presence of rail transit, independent of other built environment characteristics, significantly influences walking behavior. In conclusion, this study found a strong positive association between the presence of rail transit access and the level of walking activity for both commuting (the odds ratio of 2.504) and non-commuting trips (the odds ratio of 1.655), ceteris paribus, which may have significant planning implications.

2.2 Literature Review

2.2.1 Transit-Oriented Developments and Their Impact on Travel Behavior

In the early 1990s, Peter Calthorpe introduced transit-oriented development (TOD), a community development model that promotes efficient and environmentally sensitive development patterns in newly developing areas (Calthorpe, 1993). Many different definitions of TODs have been suggested since its introduction, but most agree that TODs refer to compact, mixed-use developments with walkable environments within a specified geographical area near transit services (Calthorpe, 1993; Bernick & Cervero, 1997; Boarnet & Crane, 1998; Parker, 2002; Cervero, Ferrell, & Murphy, 2002; Cervero et al., 2004).

As discussed extensively in the literature, the identification of TODs depends on the assessment of a variety of land use characteristics, which are often referred to as the "D" variables (Austin et al., 2010; Kamruzzaman et al., 2014; Nasri & Zhang, 2014; Higgins & Kanaroglou, 2016; Ralph et al., 2016). Cervero and Kockelman (1997) coined the term "three Ds," which stands for development density, land use diversity, and pedestrian-friendly design. Studies from later periods built on this idea and introduced a four "D" variables system: destination accessibility, distance to transit, demand management, and demographics (Ewing & Cervero, 2010).

The proponents of TOD believe that TODs can contribute to relieving various urban problems such as traffic congestion, air pollution, affordable housing shortages, and sprawl (Cervero et al., 2002). Therefore, TODs could be considered an effective solution in promoting social, economic, and environmental sustainability within communities. Most studies about TODs analyzed their impacts on travel behavior, with a specific emphasis on how effectively they reduce car usage, encourage transit ridership, and promote non-motorized travel (Cervero, 1993; Boarnet & Crane, 2001; Chatman, 2006; Arrington & Cervero, 2008; Hale, 2014; Nasri & Zhang, 2014; Ewing & Hamidi, 2014; Langlois et al., 2015; Laham & Noland, 2017; Park et al., 2018). For instance, Nasri and Zhang (2014) found that residents living in TOD areas were more likely to have between 21% and 38% lower vehicle miles traveled (VMT) than those living in non-TOD areas. Arrington and Cervero (2008) analyzed 17 TOD projects in urbanized areas and similarly observed that TOD commuters typically took transit about two to five times more than other commuters in the region. Also, Langlois et al. (2015) found that newcomers in TOD areas were more likely to use sustainable travel modes for amenities and leisure trips.

While TODs are generally believed to reduce automobile use and promote the use of sustainable travel modes, there is some debate about how components of TODs influence the travel behavior of residents living in such areas. Several studies have empirically examined various determinants that influence mode choice for informing planners and policymakers to develop appropriate policies that facilitate more sustainable travel behavior. Cervero (1993) claimed that proximity to a transit station effectively promotes more transit ridership than what can be expected from just a mixed-land use and walkable environment. Laham and Noland (2017) also found that proximity to transit stations leads to more walking for restaurants–coffee trips and grocery–food shopping trips. In addition, Arrington and Cervero (2008) found that TOD's mixed land use attribute is a key factor in facilitating transit use for various trip purposes and that the combination of high population/employment density and small-sized street blocks encourages more transit use. Similarly, Vale and Pereira (2016) claimed that the built environment of a workplace and its accessibility significantly affect walking behavior. Elsewhere, Park et al. (2018) found that transit accessibility, land use diversity, and street network design of a station area are strongly associated with transit use and walking, but density not so much.

On the other hand, some studies claimed that a predisposition of self-selecting a residential location is a major determinant of mode choice (Cervero et al., 2002; Bhat and Guo, 2007; Cao et al., 2009; Salon, 2015; Scheiner et al., 2016). The underlying notion of residential location choice is that people choose to reside in neighborhoods that fulfill their travel needs and preferences. Cervero et al. (2002) observed that TODs experience demographic changes over time, such as increasing numbers of childless couples, growing shares of people who want to downsize their living space, and an increasing influx of foreign immigrants who may come from countries with a preference for transit-oriented living. In other words, TODs attract particular types of households that seek higher levels of transit accessibility. This group of researchers addressed that empirical results may be biased without controlling for residential self-selection when evaluating the relationship between built environments and travel behavior. However, most travel surveys often limit scholars to test and control self-selection due to a lack of information about travel attitudes or previous residential locations.

Various methodological approaches have been employed in previous studies to respond to the challenge. In their review of 38 empirical studies, Cao et al. (2009) summarized nine strategies to capture the effect of residential self-selection as follows: direct questioning, statistical control, instrumental variables, sample selection models, propensity score, structural equations models, joint discrete choice models, mutually dependent discrete choice models, and longitudinal designs. However, most of these approaches still require particular data sets, including travel attitude/preference or previous residential location information. In the more recent study, Salon (2015) applied an econometric model, which is similar to the previous model by Dubin and McFadden (1984) and Bento et al. (2005), to control the endogenous bias between residential location choice and travel choice. In particular, her approach helps deal with limited data through two separate models. First, her model estimates a multinomial logit model of residential neighborhood type choice and predicts the probabilities of choosing to live in each type of neighborhood. In the second step, her model analyzes determinants of travel behavior—

VMT in this study—by including the selection variables. This approach that estimates self-selection variables will be more suitable for studies covering a large geographical scope where attitudinal information does not exist or a longitudinal data set is unavailable.

While admitting the presence of self-selection, another group of researchers claimed that the effect of self-selection is limited compared to other more dominant factors related to TODs (Chatman, 2009; Nasri et al., 2018). Nasri et al. (2018) found that self-selection accounted for roughly 40% of the effect of TODs in lowering auto trips in both Washington, D.C. and Baltimore. Despite the considerable effect of self-selection, they found that TOD still plays an essential role in influencing the mode choice of residents. Chatman (2009) also found that residential self-selection tends to enhance built environmental influences rather than diminish those impacts, which suggests that self-selection may actually downplay the effects of the built environment.

The key takeaway from the above literature review is that most studies have primarily focused on the relationship between the various characteristics of TODs and the reduction in automobile usage or the increase in transit use. Others have also analyzed the use of non-motorized travel in TOD areas (Greenwald & Boarnet, 2001; Rodríguez & Joo, 2004; Schwanen & Mokhtarian, 2005; Durand et al., 2016), but their studies were limited to walking access to and from transit only. Thus, whether the presence of rail transit stations in TOD areas leads to increased walking activities after excluding ingress and egress trips to and from the station remains unclear up to this point.

2.2.2 Prevalence of Walking Trips and Behavioral Theories

There are two theoretical propositions—behavioral spillover effects and social interaction effects — that can guide to addressing 'Why the higher level of non-motorized travel modes for purposes other than transit use is expected in TOD areas than in non-TOD areas ever after controlling for built environmental and demographic characteristics.' Unfortunately, these propositions have received very little attention in empirical studies despite their theoretical relevance.

In the field of economics and psychology, the theory of behavioral spillover effects has gained a substantial foothold in recent years. According to the theory of behavioral spillover effects, each behavior influences the next in that one behavior engenders a similar or complementary behavior that follows (Dolan & Galizzi, 2015). Often, the sequence of behaviors is pre-planned to ensure that they can be executed with high efficiency and minimum obstructions. In the case of a trip chain, people tend to decide the travel mode by considering the entire tour that includes the first and last trips as well as intermediate stops (Frank et al., 2008). In essence, if a person walked to a transit station from home, she is more likely to walk back home. Based on the behavior spillover theory, walking trips to and from transit stops, common in neighborhoods with transit stations, may lead to other walking trips, such as picking up a child or groceries on the way home, simply because they are on the same trip chain. In fact, the theory of behavioral spillover has already been applied to travel behavior related studies in examining the relationship between an individual's climate-relevant behavior and travel mode choice (Lanzini & Thøgersen, 2014; Lanzini & Khan, 2017).

The theory of social interaction effects, on the other hand, captures the propensity of individuals to behave similarly to others in their vicinity, and this sociological concept has been widely used in the fields of economics and psychology as well. In his study, Manski (2000) identified two types of social interactions – endogenous and exogenous (contextual) – to explain why people in the same group tend to behave similarly. He argued that endogenous interactions lead to a similarity in behavior because of the presence of a dominant behavior within the group and that exogenous interactions result in similar behavior due to the social characteristics of the group. A popular example of endogenous interactions is Schelling's residential segregation, which describes individuals' propensity to live in neighborhoods where the share of residents of their race is above a certain threshold (Manski, 1993). Low graduation rates in more impoverished communities and high graduation rates in more affluent neighborhoods are both examples of exogenous interactions. When I apply this concept to walking mode choice decisions, endogenous interactions exist if a person's propensity to walk increases with the number of neighbors who walk, and exogenous interactions are present if a person's propensity to walk relies on the socioeconomic attributes of those neighbors (Goetzke & Andrade, 2010).

Of the two kinds of social interactions, endogenous interactions have been more widely applied to explain various travel-related observations. For instance, Young (1996) showed that endogenous interactions heavily influence the formation of driving conventions in the absence of road laws. In his study, Young found that the choice of each driver between driving on the left or the right side of the road depends on the other drivers' decisions on the same road. A similar example is provided by Sidharthan et al. (2011), who found that parents tend to allow their children to walk to school if many other children in the same neighborhood walk to school. Notwithstanding many other factors, such as safety and supportive infrastructure, that influence a parent's decision to allow their children to walk to school, Sidharthan et al. (2011) claimed that the prevalence of walking to school created a favorable environment for walking, which in turn had a significant positive effect on a parent's mode choice. Based on this premise that social interactions affect travel behavior, endogenous interactions across individuals may also arise in clusters around transit stations because these areas generally have larger pedestrian traffic. In such an environment, people who are not necessarily transit riders may be influenced by the high frequency and volume of walking activities around transit stations.

The other possible explanation for the prevalence of walking in transit-accessible places is car shedding. When alternative travel modes become more attractive in terms of convenience, cost, and time efficiency, people may shed all or some of their vehicles (Carroll, Caulfield, & Ahern, 2017). TOD is a policy incentive to encourage people to drive less by providing more sustainable modes of transport. If those using rail transit to get to and from work eventually shed their vehicles, they would naturally walk more for other purposes or activities.

2.3 Data and Methods

2.3.1 Propensity Score Matching

This study employs propensity score matching (PSM), which is a method used in comparative studies to construct control groups that are matched with treated groups with respect to the observed characteristics. PSM is widely used in various fields, including social sciences and economics, in which a randomized experiment is often limited. Unlike controlled experiments, observational studies do not allow for random assignment of treatments to the population, which introduces a bias in estimating the treatment effect. PSM provides an opportunity to mitigate such bias by balancing the distribution of observed characteristics of control groups corresponding to treated groups using propensity scores, thus providing more precise estimates of the true treatment effects (Rosenbaum & Rubin, 1983). The propensity score is a single scalar that is estimated from a probit regression, where such scores measure the conditional probability of selecting the treatment (Thoemmes & Kim, 2011). The major advantage of using PSM, it finds matched groups based on the propensity scores that integrate all the covariate information regardless of the number of covariates in the model (D'Agostino Jr., 1998). In conventional matching techniques, it is difficult to find close matches between treated and control groups when many covariates are included in the model, increasing the dimensionality of matches.

The objective of PSM in this study is to find two TAZ groups that have similar built environment characteristics but are distinguished by the presence or absence of a transit station. Before applying the PSM analysis, I followed several steps for data preparation. First, I identified which TAZs are located within a rail catchment area. The catchment area is defined as a one-mile walking distance along the street network from the nearest rail station to the centroid of each TAZ. To identify catchment areas, I used the OSMnx street network that is based on OpenStreetMap. Among different OSMnx network types, I employed the walk network that includes all streets and paths for pedestrian use. I then applied PSM to match TAZs within rail catchment areas (treated group) to TAZs without access to rail stations (control group) based on its built-environmental attributes. Since PSM only accounts for observed covariates, any missing data or latent variable may lead to biased estimates (Garrido et al., 2014). To reduce bias, I included built environment attributes that were commonly used in previous studies. As a result, a binary probit model to estimate propensity scores contains the following "D" variables: activity density, a balance between population and all jobs (or retail/service jobs), land use diversity, intersection density, the proportion of four-way intersections, average block length, sidewalk density, open space access, and transit access to bus stops. A binary treatment variable in the probit model takes 1 when the TAZ is located within a rail catchment area and 0 otherwise. In estimating the propensity scores, I used 1:1 matching for the nearest neighbor with a replacement option and a caliper of the 0.25 standard deviation of the propensity scores of treated TAZs.

To check for the robustness of PSM, I evaluated the balance between treated and control groups. The results of PSM often exhibit a substantial overlap between treatment and control groups when the sample size is limited (Stuart, 2010). Although the minimum requirement of sample size for PSM has not yet been determined, existing literature suggests evaluating the balance between the covariates in two groups with a standardized difference, which is the mean difference. The standardized difference is calculated as follows: $(\bar{x}_t - \bar{x}_c)/\sqrt{\{(s_t^2 + s_c^2)/2\}}$, where $\bar{x}_t$ refers to the mean of the treated cases, $\bar{x}_c$ the mean of the control cases, and s_t and s_c, the corresponding standard deviations (d'Agostino, 1998). According to Rubin (2001), the absolute standardized difference in means should be less than 0.25. To satisfy this recommendation, I reduced the caliper of PSM from 0.25 to 0.10 standard deviation of the propensity scores of treated TAZs.

2.3.2 *Multi-Level Logistic Regression Model*

Many studies employing PSM have examined whether treated and control groups are systematically different in travel behavior by comparing the distribution of travel behavior between the two groups. Thus, I conducted a Chi-square test of homogeneity to examine whether the treated and control TAZ groups have the same distribution of a single categorical variable, which is the observed travel pattern (i.e., walking and non-walking trips). The null hypothesis of a Chi-square test is that there is no difference in the distribution between two TAZ groups in terms of walking trips. To determine the rejection of the null hypothesis, I compared the P-value to the significance level of 0.05.

However, researchers including Ho et al. (2007) and Stuart (2010) suggest that a more meaningful result could be derived by employing regression analysis, which controls for covariates that affect the outcome of interest on matched samples. Since the combination of PSM and regression analysis provides double-robustness in removing estimation bias of treatment effect due to confounding variables, I employed a multi-level logistic regression model to compare walking behavior in the treated and control TAZs (i.e., TOD and non-TOD areas).

Among various types of logistic regressions, multi-level logistic regression analysis is a suitable approach for this study due to data structure. A multi-level model is widely used to evaluate a clustered structure where elementary units are nested within a hierarchical structure (Bhat, 2000). In this study, people in a given TAZ are likely to be influenced by the walking behavior of other people in the same TAZ. Since this dependency among the observational units violates the independence assumption, standard errors of regression coefficients may be underestimated in standard logistic regression models. On the other hand, multi-level logistic regression models estimate unbiased standard errors of the regression coefficients by including cluster-level characteristics in the model to account for the dependence in a nested data structure (Raudenbush & Bryk 2002).

Also, a multi-level model disentangles the within-cluster effects from the between-cluster effects. It distinguishes those two sources of variations by formulating a model at the macro-level of clusters in addition to the micro-level of individuals (Bhat, 2000). In this study, a multi-level model estimates two variances: 1) within-TAZ effects, the extent to which individual-level characteristics are associated with the odds of choosing to walk, and 2) between-TAZ effects, the extent to which TAZ-level attributes are related to the odds of choosing to walk. The variance of within-TAZ effects is also known as fixed effects, and the estimates of the effects are reported as odds ratios (OR). The variance of between-TAZ effects represents unobserved TAZ attributes affecting individual behaviors after controlling for the explanatory variables, called random effects.

This study developed multi-level logistic regression models incrementally to test different model specifications based on three sets of explanatory factors: 1) sociodemographic characteristics, 2) travel-related attributes, and 3) rail transit access. The first and second models add individual-level variables, including sociodemographic characteristics and travel attributes, respectively. The final model adds a TAZ-level factor, which is the rail transit access variable, in addition to the second model. That is, the final model estimates the odds of walking as a function of both individual and TAZ characteristics.

The unit of analysis is individual trips, and the dependent variable is mode choice, which takes the value of 1 for walking and 0 otherwise. I analyze trip-based travel instead of tour-based travel, and I focus on walking trips that are not involved with other modes of travel in a tour. Based on this premise, walking trips to and from transit are excluded from the analysis since those trips are linked to transit trips in its tour. Thus, walking trips includes all purpose of activities except ingress and egress to stations. I develop models to examine walking trips for commuting and non-commuting purposes that are not relevant to transit use. Non-commuting activities include shopping, eating out, household errands, health care, social, religious, and recreational purpose of activities.

2.3.3 Data Structure

Data for PSM analysis were extracted from various sources, and Table 1 contains a list of built environment variables and corresponding descriptions of measurement at the TAZ level. For sociodemographic and employment information, I used the 2007-2011 American Community Survey (ACS) 5-year estimates (https://www.census.gov/programs-surveys/acs) and the 2011 Longitudinal Employer Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES) (https://lehd.ces.census.gov/data). To measure land use characteristics, I used the 2014 Tax Parcel Records for Fulton and DeKalb Counties. From the Atlanta Regional Commission (ARC) (https://opendata.atlantaregional.com), I used the 2006 green space data for open space access, and the 2016 ARC transit stop information for bus and rail transit.

The 2016 OpenStreetMap (OSM) (http://download.geofabrik.de/north-america/us.html) was used for the road network and sidewalk information. The OSM data has been frequently used to calculate street network characteristics because it is open-source and up-to-date. However, the quality of OSM data has been a concern because the data are collected by volunteers who are not trained in data collection procedures. The limited ability of the volunteers may result in incomplete and inconsistent data, and missing street or sidewalk information in the OSM data may cause lower intersection or sidewalk density than the actual level. Despite its limitation, the OSM data is still a powerful source of information due to its high coverage that includes areas where official data is not available. Also, previous studies on assessing the completeness of OSM data found that street and sidewalk information has been increased in the OSM platform, making the data more complete. In fact, the OSM data imported TIGER/Line as a foundational data source in 2008, and numerous improvements have been made by including additional features such as sidewalks and bike lanes (Craun & Chih-Hung, 2017; Zielstra, Hochmair, & Neis, 2013).

Table 1 – Measures of built environmental characteristics

Variable	Description
Activity density	This measures sum of population and employment per acre in TAZ.
Jobs (or retail/service jobs) to population balance	These variables measure all jobs (or retail/service jobs) to population ratio in a TAZ as compared to the same ratio in the county as a whole. It ranges from 0 for a TAZ with residents but no jobs (or only jobs, no residents) to 1 for a TAZ with the same ratio of all jobs (or retail/service jobs) to the population as that of the county as a whole. It is calculated from the following equation: $$B_{TAZ} = 1 - \left\lvert \frac{(S_{TAZ} - aP_{TAZ})}{(S_{TAZ} + aP_{TAZ})} \right\rvert$$ where: B_{TAZ} = Jobs (or retail/service jobs) to population balance in TAZ; S_{TAZ} = Jobs (or retail/service jobs) in TAZ; P_{TAZ} = Population in TAZ; and a = the ratio of jobs (or retail/service jobs) to population in the county.
Land use diversity	Inverse Simpson's index of diversity is computed to derive land use diversity based on six land use categories. These categories include residential, commercial, office, institutional, recreational/open space, and utilities. If land use is homogeneous, it takes a diversity score of 1, and a higher score indicates diverse land use. The index is calculated as follows: $$D_{TAZ} = 1/\sum_{i=1}^{6}(n_i/N)^2$$ where: D_{TAZ} = Diversity of land use in TAZ; n_i = Total land area of land use type i in the TAZ; and N = Total land area in TAZ.
Intersection density	This variable is a measure of the number of intersections per acre in TAZ.
Four-way intersection proportion	This variable is a percentage of four-way intersection in TAZ.
Average block length	This measures the average length of blocks in TAZ.
Sidewalk density	This measures the length of sidewalks per square mile in TAZ. Footway and pedestrian classes in the OSM data were used to calculate sidewalk density.
Open space access	This variable is the percent of the total area in TAZ that within 1 mile of recreational/open space.
Bus stop density	This variable is a measure of the number of bus stops per acre in TAZ.
Distance to bus stops	This measures the average distance to the nearest three bus stops from each residence in TAZ.
Rail transit access	This is a measure that indicates whether or not the TAZ is located within the rail catchment area. I employed network analysis to identify TAZs which centroids are located within a one-mile walking distance from the nearest rail station.

The primary source of data used for the multi-level logistic regression models is the 2011 household travel survey obtained from ARC, which was conducted as an activity-based survey following a 24-hour travel diary between February 2011 and October 2011. This data contains information on sociodemographic and travel behavior characteristics of 10,278 households in the 20 counties of the Atlanta metropolitan region. Among the 20 counties, I used data for Fulton and DeKalb, which contain all MARTA rail stations except for the airport station. The study area has 636 TAZs, and each walking trip is coded according to its origin TAZ. I selected the following individual-level variables based on existing literature and the availability of information in the ARC household travel survey. Sociodemographic characteristics include age (over 15 to 95), gender (male/female), ethnicity (non-Hispanic others/Hispanic), driving license ownership (yes/no), household income groups (from 1 to 10), and the number of vehicles per household size (from 0 to 5). Travel-related attributes are represented by the trip length.

2.4 Results

2.4.1 Identifying Treatment and Control Areas

As noted earlier, this study examines the effect of transit access on the prevalence of walking in the Atlanta metropolitan area. To investigate this effect, I divided TAZs in the Atlanta metro region into either the treated group (TOD areas) or the control group (non-TOD areas) using PSM so that the differences on each of the covariates across the two groups are reduced to the minimum. In this step, I employed a binary probit model to estimate the probability of each TAZ being located within the rail catchment area, which is the propensity score. From the total TAZs (n=636), this study finds 73 treated TAZs and

73 control TAZs, forming pairs of comparable built environment characteristics distinguished by the presence or absence of a rail station. Specifically, 62% of treated TAZs (73 out of 118) are matched with similar control TAZs. The unmatched TAZs in the treated group are mainly located in the business district. Due to their unique characteristics in terms of the built environment, they are not matched with TAZs from the control group. The first map in Figure 1 presents the locations of 118 TAZs which are located within a rail catchment area with the coverage of MARTA rail stations in the study area. The second map in Figure 1 displays the paired 73 treated TAZs and 73 control TAZs. While some TAZs are located close to rail stations, the walking distances along the street network between their centroids to the nearest rail station exceed one mile. Thus, I classify these TAZs into the control group.

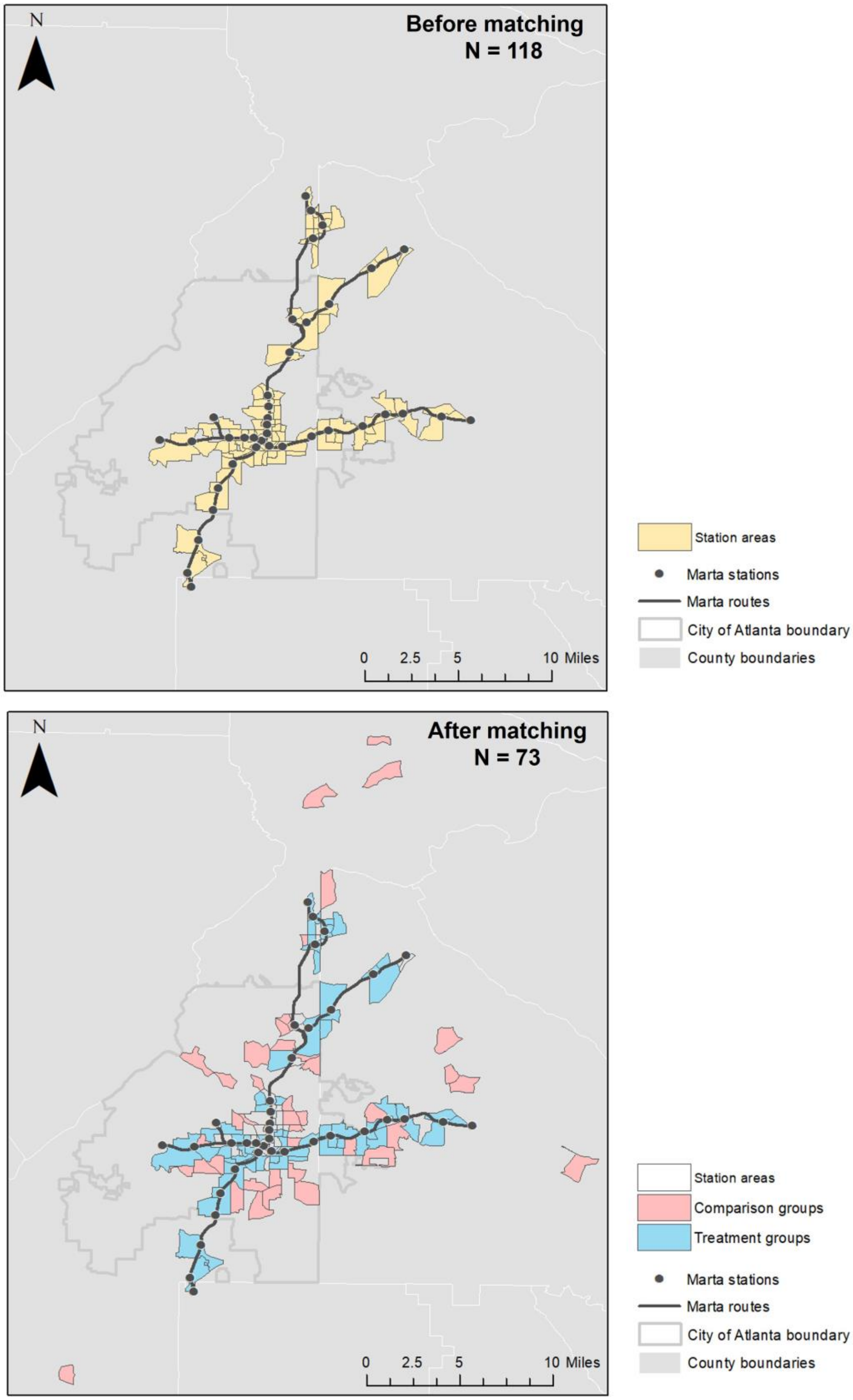

Figure 1 – The locations of treated and control TAZs in the study area

Figure 2 shows propensity scores before and after matching, and it reveals that PSM reduces the imbalance between treated and control TAZ groups after matching. Table 2 presents observed built environment characteristics of the treated and control TAZs before and after matching, and the balance of the covariates is checked with the standardized difference in mean. The treated and control TAZ groups show substantial initial differences in all built environmental characteristics with large standardized differences in mean. As expected, the difference in the observed built environment characteristics between the two groups was reduced after matching by having the absolute values of the standardized differences in mean below 0.25.

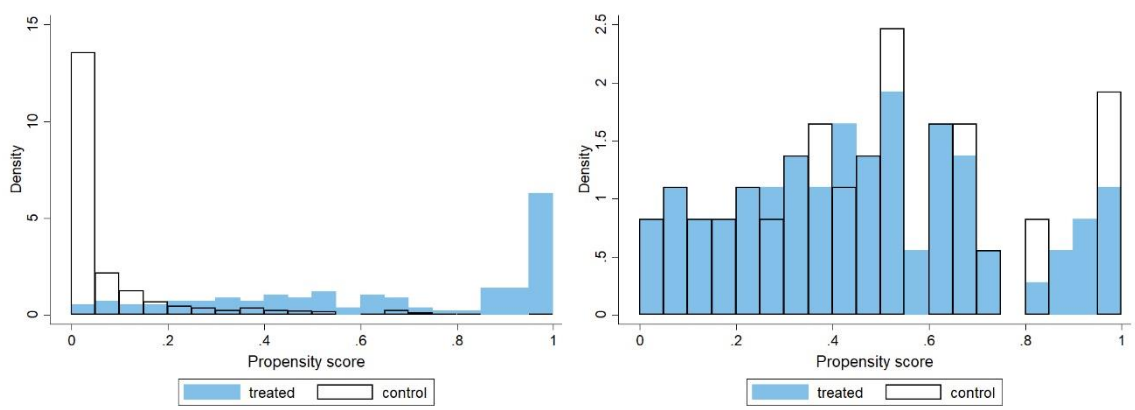

Figure 2 – Propensity scores of treated and control TAZs: before (left) and after matching (right)

Table 2 – Summary statistics of treated and control TAZs

Variable	Treated TAZs		Control TAZs		Std. Dif.[1]
	Mean	Std. Dev.	Mean	Std. Dev.	
Before matching: the summary statistics for treated TAZs (n=118) and control TAZs (n=518)					
Activity density ((population+jobs)/acre)	62.54	114.84	6.32	6.18	0.69
Land use diversity (Inverse Simpson's index)	2.32	0.93	1.77	0.66	0.68
Balance between population and all jobs	0.70	0.72	0.48	0.41	0.38
Balance between population and retail/service jobs	0.78	0.88	0.50	0.49	0.39
Intersection density (intersections/acre)	0.90	0.60	0.26	0.27	1.38
Four-way intersection proportion (%)	0.18	0.12	0.09	0.06	0.95
Average block length (mile)	0.11	0.04	0.15	0.11	-0.48
Sidewalk density (mile/square mile)	0.18	0.17	1.08	3.13	1.05
Bus stop density (bus stops/acre)	0.14	0.17	0.02	0.03	0.98
Average distance to bus stops from each residence (mile)	0.16	0.14	1.19	1.57	-0.92
Open space access (%)	0.87	0.27	0.36	0.34	1.66
After matching: the summary statistics for treated TAZs (n=73) and matched control TAZs (n=73)					
Activity density ((population+jobs)/acre)	15.21	12.71	13.81	10.48	0.12
Land use diversity (Inverse Simpson's index)	2.14	0.77	2.00	0.76	0.18
Balance between population and all jobs	0.49	0.55	0.48	0.34	0.01
Balance between population and retail/service jobs	0.54	0.71	0.54	0.43	0.00
Intersection density (intersections/acre)	0.68	0.43	0.64	0.37	0.10
Four-way intersection proportion (%)	0.14	0.08	0.14	0.07	-0.04
Average block length (mile)	0.12	0.04	0.11	0.04	0.19
Sidewalk density (mile/square mile)	6.04	9.13	5.33	7.70	0.08
Bus stop density (bus stops/acre)	0.07	0.06	0.06	0.04	0.24
Average distance to bus stops from each residence (mile)	0.19	0.15	0.20	0.14	-0.10
Open space access (%)	0.82	0.32	0.87	0.17	-0.18

[1] The standardized difference is the mean difference as the average standard deviation: $(\bar{x}_t - \bar{x}_c)/\sqrt{\{(s_t^2 + s_c^2)/2\}}$, in which $\bar{x}_t$ refers to the mean of the treated cases, $\bar{x}_c$ the mean of the control cases, and s_t and s_c, the corresponding standard deviations. Boldface numbers indicate absolute values > 0.25.

Figure 3 shows satellite images of two matched pairs in the study area. Due to similar built environment attributes, TAZ image "a" is matched with TAZ image "b" near the Ashby station. For instance, both TAZs have a high density and balanced land use mix between housing and employment locations. These areas also consist of mid- to high-rise buildings of various uses and have well-connected networks to support a high volume of the active mode of transport. Similarly, TAZ image "c" is matched with TAZ image "d" where the Hamilton E. Holmes and West Lake stations are located. These TAZs are located primarily in residential districts with lower densities, and there are small-scale mixed-use developments around the station areas and few areas of pedestrian connectivity.

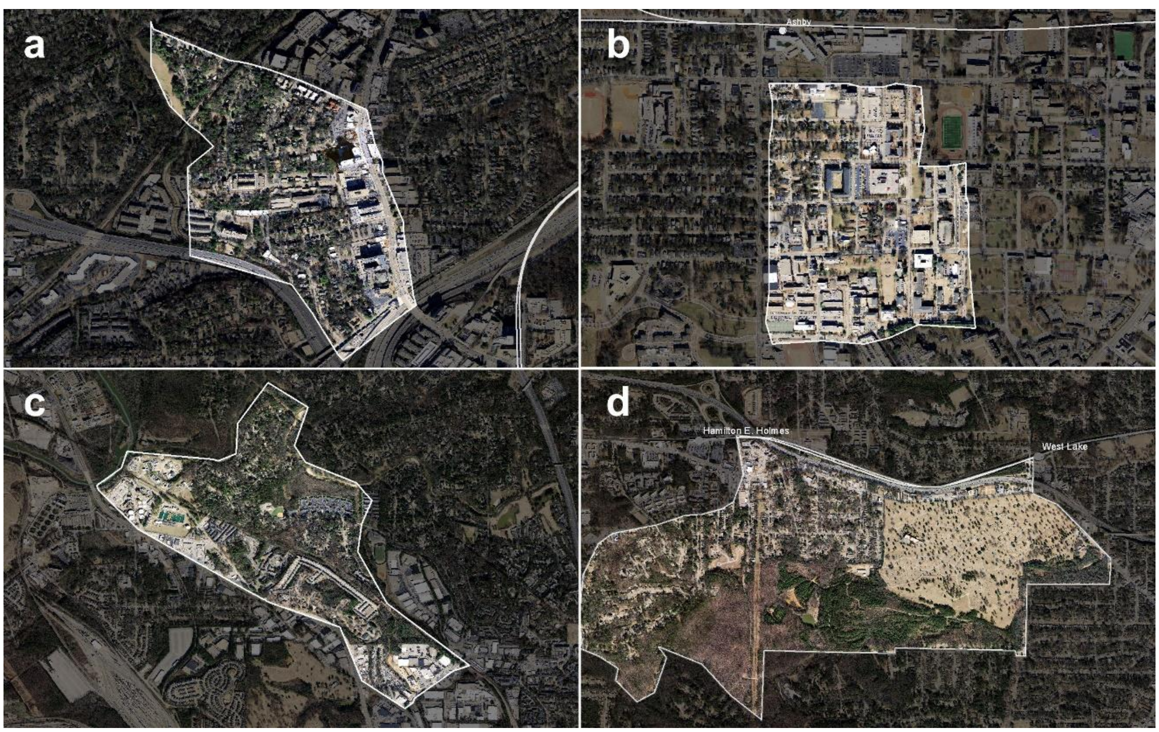

Figure 3 – Examples of matched control TAZs (left) and treated TAZs (right)

2.4.2 *Walking Behavior in Transit-Oriented Developments*

The travel survey data for the matched sample shows a higher percentage of walking trips in the treated TAZs with rail transit access than those in the control TAZs without rail transit access, as shown in Table 3. Walking trips that originate from the treated TAZs are 9.8%, while those from the control TAZs are only 7.8%. The significant result of the Chi-square test ($\chi 2 = 8.95$, $p = 0.002$) indicates that two groups—control TAZs and treated TAZs—have a statistically significant difference in the share of walking trips.

Table 3 – Number of trips in the study area

	Non-walking trips		**Walking trips**		**Total trips**	
	n	**%**	**n**	**%**	**n**	**%**
Treated TAZs	4,113	90.2	449	9.8	4,562	100.0
Control TAZs	2,486	92.2	209	7.8	2,695	100.0
Total trips	6,599	90.9	658	9.1	7,257	100.0

This information, however, does not tell us the relative importance of various factors that impact walking trips for both commuting and non-commuting activities. Thus, I examine the relationship between walking behavior and the presence of rail transit access by employing multi-level logistic regression models. Table 4 presents summary statistics for 6,436 observations, and these are broken down by two types of activities: 1) commuting trips and 2) non-commuting related trips. Descriptive statistics between the treated TAZs and control TAZs show that walking trips' mean values for commuting and non-commuting purposes are higher in the treated TAZs. While the mean value of trip distance for commuting purposes is lower in the treated TAZs, that of trip distance for non-commuting purposes is higher in the treated TAZs. Other variables have relatively similar mean values between treated TAZs and control TAZs.

Table 4 – Summary statistics of all trips except transit access & egress trips

Variable	All trips in treated TAZs		All trips in control TAZs	
	Mean	**Std. Dev.**	**Mean**	**Std. Dev.**
Commuting purpose: number of trips in treated TAZs (n=644) and control TAZs (n=696)				
Walking trip (1=walk, 0=other modes)	0.15	0.35	0.08	0.28
Age	44.07	12.20	45.26	11.66
Gender (1=female, 0=male)	0.44	0.50	0.48	0.50
Hispanic (1=Hispanic, 0=non-Hispanic)	0.04	0.20	0.08	0.28
License ownership (1=yes, 0=no)	0.97	0.18	0.97	0.16
Number of vehicles per household member	0.88	0.43	0.84	0.40
Income less than $10,000	0.01	0.10	0.02	0.15
$10,000 to $19,999	0.02	0.12	0.04	0.20
$20,000 to $29,999	0.03	0.17	0.04	0.20
$30,000 to $39,999	0.05	0.22	0.03	0.18
$40,000 to $49,999	0.05	0.21	0.05	0.22
$50,000 to $59,999	0.09	0.28	0.05	0.22
$60,000 to $74,999	0.11	0.31	0.06	0.25
$75,000 to $99,999	0.22	0.41	0.28	0.45
$100,000 to $149,999	0.22	0.41	0.19	0.39
$150,000 or more	0.21	0.41	0.22	0.41
Log transformed trip distance	0.60	1.63	0.78	1.56
Rail transit access	1.00	0.00	0.00	0.00
Non-commuting purpose: number of trips in treated TAZs (n=2,474) and control TAZs (n=2,622)				
Walking trip (1=walk, 0=other modes)	0.11	0.31	0.09	0.29
Age	46.60	13.58	47.27	14.14
Gender (1=female, 0=male)	0.55	0.50	0.60	0.49
Hispanic (1=Hispanic, 0=non-Hispanic)	0.04	0.20	0.05	0.22
License ownership (1=yes, 0=no)	0.95	0.22	0.95	0.23
Number of vehicles per household member	0.83	0.48	0.82	0.41
Income less than $10,000	0.05	0.21	0.04	0.19
$10,000 to $19,999	0.06	0.23	0.04	0.21
$20,000 to $29,999	0.06	0.23	0.05	0.22
$30,000 to $39,999	0.06	0.24	0.06	0.24
$40,000 to $49,999	0.06	0.24	0.07	0.26
$50,000 to $59,999	0.06	0.24	0.04	0.21
$60,000 to $74,999	0.07	0.26	0.08	0.28
$75,000 to $99,999	0.19	0.39	0.17	0.38
$100,000 to $149,999	0.19	0.39	0.25	0.43
$150,000 or more	0.21	0.41	0.18	0.39
Log transformed trip distance	0.84	1.68	0.50	1.65
Rail transit access	1.00	0.00	0.00	0.00

Results from the multi-level logistic regression models indicate a strong association between sociodemographic, travel, rail transit access, and walking trips for commuting. Table 5 presents fixed and random effects from three multi-level logistic regression models for commuting walking trips. Model 1, which includes only individual-level variables, shows that some sociodemographic characteristics of individual travelers are associated with walking trips for commuting purposes. Among various sociodemographic factors, having a driver's license and the number of vehicles per household member are statistically significant. The odds ratios of having a driver's license (0.295) and the number of vehicles per household member (0.478) indicate that they tend to lower the probability of commuting walking trips within TAZs.

Model 2 adds travel attributes as an explanatory variable in addition to the individual sociodemographic variables in model 1. The results of model 2 present the effect of travel characteristics on walking behavior. As expected, trip distance is negatively associated with walking trips for commuting by having the odds ratio of 0.184. In other words, longer trip distance reduces the likelihood of walking for commuting since the distance is strongly associated with the disutility of traveling. Therefore, trip-makers are likely to choose a faster mode than walking as the distance to the destination increases. After adjusting for the travel attribute in the model, I find some changes in the influences of sociodemographic characteristics on walking trips for commuting purposes. In model 2, the association between the number of vehicles per household member and commuting walking trips disappears while the influence of having a driver's license on commuting trips persists. Individuals with a driver's license are 0.275 times less likely to walk for commuting purposes than those without a driver's license. The households with incomes between \$10,000 to \$19,999 (income group 2) show a statistically significant odds ratio of 12.471. However, this association does not appear across all income categories.

Model 3 is the final model, which includes the presence of rail transit access as a key explanatory variable at the TAZ level in addition to sociodemographic and travel characteristics. The result of the log-likelihood test between the unrestricted model with rail transit access variable and the restricted model without the variable indicates that the unconstrained model, which is the final model, is better at the 99 % confidence level. The results of model 3 show that the rail transit access variable is significant at a 95% confidence level after accounting for individual-level variables. The 2.504 odds ratio of rail transit access implies that the odds of choosing walking mode is 2.504 times larger in the treated TAZs than in the control TAZs. That is, people are more likely to choose walking in areas with rail transit access than those in areas without rail transit access. When I translate the impact of rail transit access on the prevalence of walking into probability, it is easier to understand the trend. The probability of choosing to walk in the treated TAZs is 7.4%, whereas that in the control TAZs is only 3.1%. Similar to the results from the previous two models, some sociodemographic characteristics exhibit distinct influences on walking trips for commuting in model 3 as well. Having a driver's license tends to lower the probability of walking trips for commuting purposes. The lower-income group is also significantly associated with commuting walking trips. However, these influences are less profound than those in the previous two models. In terms of travel attributes, I find that trip distance is the most critical determinant of walking behavior for commuting, similar to the results of the second model.

Table 5 – Multi-level logistic regression models for commuting walking trips

Variable	Model 1 Sociodemographic characteristics			Model 2 Sociodemographic and travel characteristics			Model 3 Sociodemographic, travel, and rail access		
	OR		SE	OR		SE	OR		SE
Constant	0.624		0.688	1.147		1.488	0.796		1.053
Rail transit access							2.504	***	0.766
Age	1.035		0.052	0.968		0.061	0.963		0.062
Age squared	0.999		0.001	1.000		0.001	1.000		0.001
Female	0.988		0.194	0.881		0.230	0.928		0.246
Hispanic	0.415		0.216	0.539		0.427	0.502		0.394
Driving license	0.295	***	0.135	0.275	**	0.174	0.329	*	0.204
Vehicles per H.H. size	0.478	***	0.128	0.705		0.242	0.680		0.235
Income group 1 (Reference)									
Income group 2	4.078	*	3.029	12.471	***	11.754	12.501	***	12.022
Income group 3	0.585		0.468	0.347		0.334	0.286		0.281
Income group 4	1.151		0.878	1.849		1.737	1.386		1.337
Income group 5	0.979		0.730	3.580		3.117	2.728		2.430
Income group 6	0.812		0.602	2.659		2.272	1.728		1.523
Income group 7	0.650		0.480	1.129		0.989	0.729		0.657
Income group 8	0.805		0.543	1.992		1.505	1.746		1.350
Income group 9	0.425		0.299	1.425		1.144	1.014		0.838
Income group 10	0.377		0.265	0.909		0.713	0.642		0.520
Trip distance (log)				0.184	***	0.025	0.186	***	0.025
N			1,340			1,340			1,340
Log(L)			-432.019			-221.571			-216.286
ρ^2 (market share model as base)			0.081			0.529			0.540
Adj. ρ^2 (market share model as base)			0.049			0.495			0.504

*Significant at 90% **Significant at 95% ***Significant at 99%

Table 6 presents the multi-level logistic regression models of factors associated with non-commuting walking trips. Similar to model 1, the independent variables in model 4 are sociodemographic characteristics of the individual traveler. Model 4 for non-commuting trips presents more sociodemographic variables that are statistically significant than the models for commuting trips. The results reveal that attributes such as gender, ethnicity, having a driver's license, number of vehicles per household member, and income level show distinct influences on walking for non-commuting trips. The odds ratios of female (0.694), having a driver's license (0.230), and the number of vehicles per household member (0.217) indicate that they are likely to lower the probability of walking trips for non-commuting related purposes. The odds ratio of the percent of Hispanic persons in the population suggests that Hispanic individuals are 2.605 times as likely to choose walking for non-commuting-related trips than non-Hispanic individuals. While not all income groups are statistically significant, the income level is negatively related to walking for non-commuting trips, indicating that a household with higher income except for income groups 2 and 4 is less likely to walk.

Similar to model 2, model 5 adds travel attribute variables to sociodemographic variables. Model 5 presents the odds ratio of 0.184 for trip distance, indicating that the odds of choosing walking for non-commuting trips decreases by 0.184 for one unit increase in trip distance. After controlling for the travel attribute, the associations between sociodemographic characteristics and walking for non-commuting trips still exist. Age and percent of Hispanic persons in the population variables significantly predict walking trips for non-commuting purposes. In addition to these attributes, females, having a driver's license, and the number of vehicles per household member are less likely to make walking trips for non-commuting purposes by presenting the odds ratios of 0.780, 0.163, and 0.303, respectively. Even though there is a statistically significant negative association between income and non-commuting walking trips, this trend does not appear across all income groups.

Table 6 – Multi-level logistic regression models for non-commuting walking trips

Variable	Model 4 Sociodemographic characteristics			Model 5 Sociodemographic and travel characteristics			Model 6 Sociodemographic, travel, and rail access		
	OR		**SE**	**OR**		**SE**	**OR**		**SE**
Constant	1.519		0.679	0.356	*	0.201	0.287	**	0.162
Rail transit access							1.655	***	0.309
Age	1.021		0.020	1.071	***	0.027	1.067	**	0.027
Age squared	1.000		0.000	0.999	***	0.000	0.999	**	0.000
Female	0.694	***	0.074	0.780	*	0.101	0.782	*	0.101
Hispanic	2.605	***	0.518	2.822	***	0.719	2.897	***	0.733
Driving license	0.230	***	0.044	0.163	***	0.041	0.168	***	0.042
Vehicles per H.H. size	0.217	***	0.036	0.303	***	0.061	0.303	***	0.060
Income group 1 (Reference)									
Income group 2	0.652		0.173	0.648		0.231	0.659		0.233
Income group 3	0.437	***	0.119	0.878		0.314	0.900		0.319
Income group 4	0.646		0.177	1.795		0.648	1.800		0.648
Income group 5	0.244	***	0.075	0.440	*	0.185	0.452	*	0.190
Income group 6	0.334	***	0.112	0.587		0.255	0.576		0.249
Income group 7	0.194	***	0.065	0.401	**	0.163	0.414	**	0.168
Income group 8	0.283	***	0.070	0.507	**	0.163	0.521	**	0.167
Income group 9	0.357	***	0.088	0.677		0.221	0.709		0.231
Income group 10	0.504	***	0.122	1.059		0.340	1.075		0.343
Trip distance (log)				0.295	***	0.016	0.295	***	0.016
N			5,096			5,096			5,096
Log(L)			-1399.63			-944.03			-940.45
ρ^2 (market share model as base)			0.133			0.415			0.418
Adj. ρ^2 (market share model as base)			0.124			0.406			0.407

*Significant at 90% **Significant at 95% ***Significant at 99%

Model 6 is the final model with rail transit access variable for non-commuting related trips. The results of model 6 present that all sociodemographic and travel characteristics are associated with walking for non-commuting trips. Model 6 also presents the statistical significance of rail transit access on the probability of choosing to walk for non-commuting trips after controlling for all other individual-level variables. The odds of choosing to walk for non-commuting related trips is 1.655 times higher in the treated TAZs compared to the control TAZs. It means that the treated TAZs show a high probability of choosing to walk (44.7%) compared to the control TAZs (32.8%). Since this trend persists in both commuting and non-commuting trips, I conclude that the "T" is an essential element in increasing walking trips for all purposes in TOD areas.

2.5 Discussion and Conclusion

This study revisits Chatman's (2013) question: "Does TOD need the T?" by addressing the role of transit access in influencing walking behavior in TOD areas. In the existing literature, high density, mixed land use, pedestrian-friendly environments, and quality public transit facilities and service are major components of TODs in promoting active modes of transport. Among these various attributes of TOD, I particularly evaluated the role of rail transit access on walking trips that are generated from TOD areas. To estimate the true effect of transit access on walking trips, I first identified the treated TAZs and control TAZs that have similar built environment characteristics using the PSM technique. PSM used the presence of rail transit access as a key differentiator between the treated and control TAZ groups. I then compared walking trips for commuting and non-commuting purposes between the two TAZ groups by employing multi-level logistic regression models. Since TOD areas typically generate more walking trips to transit stations than non-TOD areas, I excluded any walking trips related to transit use. This unique research design provided an opportunity to reduce bias in samples and examine walking behavior in the Atlanta metropolitan area.

The major finding from this study is that the presence of rail transit access does have a measurable association with walking trips for all purposes that do not involve transit use after controlling for sociodemographic and travel characteristics. In other words, "T" is a critical element in TOD. Two theoretical propositions—behavioral spillover effects and social interaction effects—can explain the prevalence of walking that is not relevant to transit use in TOD areas. Based on the behavioral spillover theory, adopting one behavior leads to the additional adoption of related behaviors. Since a relatively large number of people walk to and from transit stations in TOD areas, their behavior may lead to more walking trips to other destinations than transit stations. These additional walking trips can be linked to a trip chain to and from transit stops or an individual trip. According to the social interaction theory, people within the same group are likely to behave similarly. This implies that people's propensity to walk would increase when there is a high volume of pedestrians in TOD areas.

The research finding also supports current policies that target compact and dense urban forms around transit facilities to promote sustainable transportation to destinations other than the transit stops. Currently, the U.S. Department of Transportation's Federal Transit Administration (FTA) is offering supportive programs and technical assistance to localities to advance sustainable modes of travel. FTA has funded about 20 transit organizations across the country to support their TOD projects to improve public transit access, and the amount of funding has increased from $14.7 million in 2016 to $19.2

million in 2019 (FTA, 2019). Our finding of the positive association between TODs and walking behavior for commuting and non-commuting purposes supports the soundness of such investments. Well-planned TODs have successfully served neighborhoods by connecting transit to surrounding places with diverse amenities such as jobs, housing, retail, restaurants, open spaces, and pedestrian-friendly environments. This study indicates that TODs may have also helped improve overall walkability, which benefits the environment and supports a healthy lifestyle.

Our findings indicate an important relationship between transit access and walking behavior; however, this study has some limitations. First, the main threat to this study is self-selection bias occurring when individuals who like to walk choose to live in TOD areas. Because our analytical models did not control for residential self-selection due to the data structure, the estimated treatment effect might be overestimated if there are strong residential and travel-related preferences in the study area. A longitudinal research design or a model that controls for individuals' travel-related preferences may be helpful to deal with the self-selection issue.

Second, previous studies noted that people are less likely to own a vehicle and have a driver's license when they live in transit-accessible areas (Ewing & Hamidi, 2014). The limited access to an automobile because of higher transit access may have direct and indirect effects on walking trips. However, this study is limited to estimating only the direct effect on walking trips. This fact may lead to underestimating rail transit's contribution to this study. Future studies can employ path analysis, structural equation modeling, or other adequate models to address this limitation. Also, the number of vehicles per household may be less reliable for some ethnic and income groups. Instead of the number of vehicles per household, access to vehicles, such as owning cars, leasing cars, and carpool, can better capture vehicle availability of households. This study did not include access to vehicles per household in the model due to a lack of data.

Third, this study did not differentiate TOD types in the model specification. Considering that transit agencies have developed TODs for different goals based on where they are located, analyzing the effects of TODs using separate models for urban and suburban areas may provide useful insights in improving TOD plans and guidelines, particularly TODs which aim to support broader transit networks that cover both urban and suburban areas. Unfortunately, I found that most TAZs in our study area are categorized as urban TAZs. Thus, a future study may need to expand the geographical scope, covering multiple metropolitan regions.

Fourth, this study only evaluated the influence of origins on walking trips. Recent studies have noted that destination and route attributes are also associated with walking behavior (Moran, Rodríguez, & Corburn, 2018; Vale & Pereira, 2016). The model specifications considering built environmental attributes of both origins and destinations may provide more concrete results. In addition, this study may need additional variables in terms of omitted variables such as parking prices and availability and crime that may also be associated with walking trips. However, given prior studies, the omission of crime and parking variables might suggest that our estimates for walking in TOD areas are conservative or overly generous.

Fifth, this study did not investigate the impact of individual built environment variables on walking behavior. Since the main objective of this study was to examine whether increased walking is related to the presence of rail transit independent of built environment characteristics, PSM used the presence of rail as a key differentiator between treated and control groups. Even if PSM finds two comparable TAZ groups that share similar built environment characteristics except for rail transit, there may still be differences in built environment variables between the two groups. Thus, future study is required to test all built environment variables one by one by including them in the regression models and examine whether any of them appears to have a significant effect on walking.

Finally, PSM method requires discarding about one-third of treated TAZs, leaving a particular sample of treated units to be compared to a similarly unrepresentative matched control sample. In fact, the treated TAZs' means on built environment variables are largely different for the matched group (n=73) compared to the entire group (n=118). The treated TAZs' means of the variables decreased after matching due to the unmatched treated TAZs' unique built environment characteristics, such as high activity density, diverse land use, high balance between population and jobs, and high intersection density. Since Atlanta is one of the most sprawling cities in the U.S., PSM may not find appropriate matches for the unmatched treated TAZs, which are mainly located in the central business district within the region. Thus, future research may test PSM method in other cities, including both sprawling and non-sprawling cities.

CHAPTER 3. ARE RESIDENTS MORE MULTIMODAL IN TRANSIT-ORIENTED DEVELOPMENTS?

3.1 Introduction

Transit-Oriented Development (TOD) is a planning approach for areas around transit stations that promotes higher densities and a mix of land uses together with walkable streets to incentivize non-motorized travel and the use of transit. As discussed extensively in the literature on land use and transportation interaction, various land use characteristics, such as density, diversity, design, distance to transit, and destination accessibility, generate synergistic effects on travel behavior. TODs are a perfect example of integrating various land use and transportation characteristics that encourage residents to be less reliant on private automobiles (Renne, Hamidi, & Ewing, 2017). Unfortunately, not all transit areas are developed along TOD principles. Therefore, previous studies have not shown consistent effects of different transit areas on transportation outcomes such as car ownership, VMT, and travel mode share. Although several studies have shown that residents near TODs use more non-motorized travel modes, these studies often fail to control for residential self-selection and the variation in TOD types (Higgins & Kanaroglou, 2016; Renne, Hamidi, & Ewing, 2017). Also, no study to my knowledge covers transit areas of the entire US and presents different effects of TOD types in a rigorous way.

This study attempts to estimate the effects of land use attributes around transit facilities on the multimodal travel behavior of their residents. The multimodal travel pattern, or multimodality, refers to individuals using multiple travel modes for certain periods (Buehler & Hamre, 2015). Researchers are especially interested in multimodality because incentives and regulations for popularizing a particular mode of travel tend to be most effective for multimodal travelers. Thus, identifying multimodal travelers and the drivers of their multimodality is a critical first step (Ralph et al., 2016). Given that living in TOD increases the probability of diversifying travel modes due to easy access to transit, better walkability, and proximity to shops, restaurants, and offices, it is likely that residents in TODs would be multimodal.

This study follows three analytical steps, as shown in Figure 4. First, it uses factor- and cluster-analysis techniques to classify areas within a half-mile radius from the 4,400 fixed-guideway transit stops in the U.S. This classification scheme generated four distinctive TOD types: business district, town center, neighborhoods, and suburban station areas. Second, it measures the level of multimodality for all individuals in the 2017 NHTS. In calculating the multimodality indicators, I extract trip data for the following time periods; travel mode in the past week, survey day, and work and non-work tours on the survey day. Finally, this study employs a series of regression models to explain the multimodality index with the help of data about socio-demographics of individuals, the type of TOD where these individuals lived, and other selection variables. The selection variables are the probability of individuals living in the other types of TOD except one's own, and this study uses them as a control for residential self-selection. In sum, this paper will shed light on how much difference in multimodal travel behavior planners can expect by a transit area from one TOD type to another.

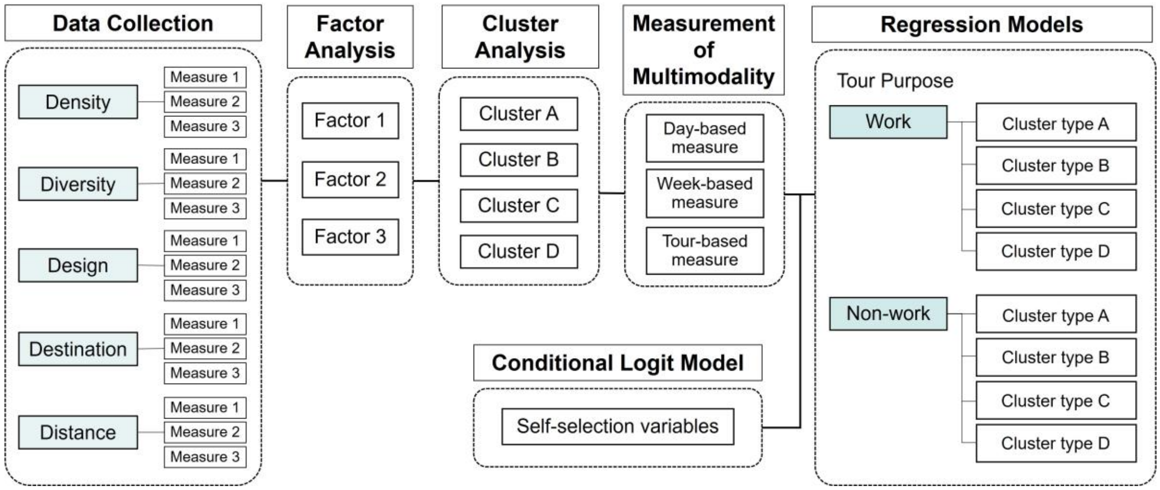

Figure 4 – Research process flowchart

The remainder of this study consists of four parts. The second section reviews and summarizes the results of existing research. The third section describes the methodology and the data source for the analysis. The fourth section discusses the findings and concludes policy implications, research limitations, and directions for future studies.

3.2 Literature Review

3.2.1 Transit-Oriented Development Typology

While a large volume of existing literature examined the identification of TOD and evaluated the performance of TODs, more recent literature demonstrates a growing interest in typologies of station areas for the evaluation of TODs. Since the implementations of TOD and its expected outcomes vary by context, it is necessary to identify a range of TOD types as a critical tool for deriving context-based TOD planning. That is, a typology of TOD with desired features supports planners and policymakers to develop an optimal strategy for a TOD project under specific conditions at a given site (Belzer & Autler, 2002). Based on this notion, many studies have classified station areas into several types based on their existing characteristics or performance measures (Bertolini, 1999; Reusser et al., 2008; Austin et al., 2010; Atkinson-Palombo & Kuby, 2011; Zemp et al., 2011; Kamruzzaman et al., 2014; Higgins & Kanaroglou, 2016).

Bertolini (1999) developed a conceptual framework on a node-place typology of TOD in classifying train stations in the Netherlands. He identified four TOD types using the node index for each station using the connectivity, frequency, and diversity of transport services and the place index based on walkable distance from the stations. Reusser et al. (2008) and Zemp et al. (2011) extended the work of Bertolini (1999) by using additional measures to derive the node and place index to classify rail stations in Switzerland. Both studies employed a hierarchical cluster analysis, which resulted in five TOD types (smallest, small, mid-sized in populated areas, mid-sized but unstaffed, and large-to very-large stations) and seven TOD types (central stations, large connectors, medium commuter feeders, small commuter feeders, tiny tourist stations, isolated tourism nodes, and remote destinations), respectively.

Atkinson-Palombo and Kuby (2011) also used a hierarchical cluster analysis in classifying transit-oriented overlay zoning around 27 LRT stations in the Metropolitan Phoenix. They conducted a factor analysis to reduce the multicollinearity in 13 indicators regarding the characteristics of a node, people, and places in the station areas. The derived factors were used in the hierarchical cluster analysis, which resulted in five-station types: transportation nodes, high population rental neighborhoods, areas of urban poverty, employment and amenity centers, and middle-income mixed-use. They found an uneven distribution of overlay zoning across station area types.

Kamruzzamam et al. (2014) classified 1,734 census collection districts in Brisbane using six TOD indicators: employment density, residential density, land use diversity, intersection density, cul-de-sac density, and public density transit accessibility. By employing a two-step cluster analysis, they identified four TOD types: neighborhood residential TODs, activity center TODs, potential TODs, and non-TODs (not suitable for TOD). The validation of the typologies against a travel survey found that people living in TOD clusters are more likely to use public transit and active modes than those in non-TOD groups.

Higgins and Kanaroglou (2016) classified transit stations into a set of more homogeneous station types by distilling measures of station areas, including distance to transit, density, diversity, design, and destination accessibility. They used latent class clustering to accommodate unscaled or unstandardized variables. As a result, they identified ten different TOD types from inner urban to outer suburban in the Toronto region. The features of urban station type in their study exhibit TOD's ideal characteristics, while suburban stations show room for improvement toward the TOD concept.

While the existing literature on the classification of TODs provides some empirical results on TOD types' heterogeneity, most studies have focused on a single city or region. Only a few studies covered the whole country to ensure external validity (Austin et al., 2010; Chatman et al., 2014; Renne et al., 2016).

Recent studies on TOD classification use indicators that show a dichotomy between TOD inputs and TOD outcomes. In general, TOD inputs refer to ‘D’ variables such as density, diversity, design, distance to transit, and destination accessibility. At the same time, TOD outcomes indicate the performance measurements for mobility, such as transit ridership and vehicle miles traveled (VMT). Unlike the majority of studies that use either inputs or outcomes of TOD as an indicator, Reusser et al. (2008) and Austin et al. (2010) simultaneously applied both input-based and outcome-based measures of TOD to classify station types. For example, Austin et al. (2010) identified fifteen different station types in the U.S. with the use of both a place indicator (e.g., use-mix measure) and a performance indicator (e.g., household VMT). Zemp et al. (2011) criticized the use of measures for TOD outputs in classifying station types in that TOD outputs—such as passenger frequency—are not context-based measures. Instead, it results from the interaction between various contextual inputs within the station area.

3.2.2 Multimodal Travel Behavior

One of the expected benefits of TOD is a modal shift from private automobiles to sustainable transport modes, such as public transit, walking, and bicycle. In this context, extensive previous literature on travel behavior and TODs have concluded that TODs tend to reduce car usage and increase transit ridership (Cervero, 2004; Chatman, 2013; Nasri and Zhang, 2014; Langlois et al., 2015; Kwoka et al., 2015). However, these studies have primarily focused on household VMT or discrete mode-choice outcomes, representing a single or partial aspect of travel behavior.

Recently, modal variability or flexibility has gained increasing attention to encourage people to use sustainable modes under the circumstances where they are sufficiently competitive while using private automobiles for only some trips (Heinen & Mattioli, 2019). Based on this notion, there is a growing body of scholarship on the multimodal travel behavior (referred to as multimodality), which captures the use of various travel modes within a specific time period (Nobis, 2007; Kuhnimhof et al., 2012; Buehler & Hamre, 2015; Clauss & Doppe, 2016; Molin et al., 2016; Ralph et al., 2016; Scheiner et al., 2016). Previous literature agreed upon the general definition of multimodality, but measuring it varies between studies. Specifically, determining the duration of time and a range of travel modes in measuring multimodality largely depended on data availability and the purpose of the study.

In studies of multimodality, three types of observation periods have been used: one-day (Blumenberg & Pierce, 2014), weeklong (Buehler & Hamre, 2015; Heinen and Chatterjee, 2015; Scheiner et al., 2016), and multi-week (Vij et al., 2013). Blumenberg and Pierce (2014) used a one-day travel survey from the 2009 National Household Travel (NHTS) to classify travelers into three groups: multimodal (individuals who used more than one mode, regardless of the number of trips or tours they took on the travel day), unimodal (individuals who used only one mode on the travel day), and non-travelers (individuals who did not make trips on the travel day). They found that income is positively associated with the likelihood of multimodal travel. Specifically, low-income adults are less multimodal than those with higher incomes.

Buehler and Hamre (2015), Heinen and Chatterjee (2015), and Scheiner et al. (2016) employed a one-week travel survey. With the 2001 and 2009 NHTS data, Buehler and Hamre (2015) investigated the intensity of multimodality during a typical week and distinguished three traveler groups: monomodal car users, multimodal car users, and walk, bicycle, public transportation only users. They found that majority of Americans were multimodal car users who drive and make at least one weekly trip by walk, bicycle, or public transportation, while only 28% of the American population were monomodal car users.

Contrary to this approach, Heinen and Chatterjee (2015) attempted to measure a multidimensional aspect of multimodality using continuous indicators over weekly travel. They used the 2010 National Travel Survey of Great Britain to derive continuous indicators of individual modal variability. Herfindahl-Hirschman Index (HHI) represents the equality of distribution of mode choices across the options with a value of 1/N, where N is the number of mode options, representing equality and a value of one representing a concentration of mode choices on one mode option. They generated two HHI values from eight mode categories (walk, bicycle, car driver, car passenger, bus, rail, taxi, and other) and for three-mode categories (private transport, public transport, and active transport). This study showed that about 69% of adults are multimodal over their weekly travel. It also revealed that the reduced modal variability is associated with mobility difficulties, age over 60, non-white, full-time working, lower household income, smaller settlement, access to a car, no public transport pass/ticket, and no access to a bicycle.

Similarly, Scheiner et al. (2016) used the German Mobility Panel for the period 1994 to 2912 to measure continuous four indicators—the share of trips, HHI, Shannon's entropy, and the number of modes used—of six mode use—walking, cycling, public transport, car driver, car passenger, and others—in a week. Their finding showed that some life course events are significantly related to changes in multimodality. While only a few socio-demographic life events showed significant associations, they found that multimodality was positively associated with an urban environment, such as the public transit system's quality and reduction in parking spaces.

While some studies claimed that short period of travel surveys (e.g., one week) tend to capture typical variability in travel behavior (Nobis, 2007), other studies used a longer period of travel surveys to capture both everyday habitual and occasional travel behavior (Molin et al., 2016; Ralph et al., 2016; Vij et al., 2013). Molin et al. (2016) estimated five multimodal travel groups—car multimodal, bike multimodal, bike and car, car mostly, and public transit—in the Netherlands based on the frequency of use of various modes including car, bicycle, train, and bus for a weekly, monthly, and yearly basis in their latent class choice model. By including socio-demographic and work-related variables in their model to predict the probability of belonging to each of the five groups, they found different attitudes towards travel mode among the five groups. Specifically, solo car drivers have more negative attitudes towards public transit and bicycle, while frequent car drivers who also use public transit have less negative public transit attitudes.

From the 2009 NHTS, Ralph et al. (2016) identified four types of travelers in the U.S.—drivers, long-distance trekkers, multimodals, and carless—by employing a latent class analysis with seven travel variables for various time extent over the short- and longer-term from daily to annual travel pattern. Using a multimodal logistic regression, they found that neighborhood type (e.g., urban residential and mixed-use) was an important predictor of being multimodal. Similarly, Vij et al. (2013) employed a six-week travel survey in Germany and estimated three groups with distinct modality styles over multiple work and non-work tours—habitual drivers (16.5% of the sample population), time-sensitive multimodals (44.2%), and time-insensitive multimodals (39.3%)—from the latent class choice models.

The review of all the above studies indicates that there is little guidance in measuring the degree of multimodality regarding a range of travel modes. Determining the number of transport modes in measuring multimodality largely depends on data availability and the purpose of the study. Some studies also argued the exclusion of walking modes from the analysis due to the difficulty of measuring walk trips accurately (Nobis, 2007). This is because survey respondents often drop short walk trips in reporting their travel diary, as every trip begins with a walk.

Also, our understanding of the multimodal behavior of households living in TOD areas is limited. This research gap comes from the differences in the measurement of multimodal travel, which often limits the comparability of the results among existing studies. In addition, not all transit areas are developed along TOD principles. Thus, previous studies have not shown consistent effects of different transit areas on travel behavior. While the existing literature on TOD typology provides some empirical results on the heterogeneity in transit stations, few studies covered the whole country to ensure external validity.

3.3 Data and Methods

3.3.1 Factor and Cluster Analysis

This study employs factor- and cluster-analysis techniques to classify various development patterns around transit facilities across the U.S. The common factor analysis technique, which assumes that the unique parts of the variables are uncorrelated with each other and with their common parts, is applied for dimensionality reduction of various land use characteristics of a transit station area. Next, the k-means cluster analysis is employed on the extracted factors from the previous step to classify all transit station areas into different clusters. To find the best solution when employing these two techniques, the literature advises applying several criteria (e.g., selecting the number of factors based on the eigenvalue measures, scree test, and size of the loadings; identifying the final solutions to be interpretable and matched the knowledge of the local context).

3.3.2 Measurement of Multimodal Travel Behavior

This study employs several measures to capture the degree of multimodal travel behavior by residents around transit facilities in the U.S. As for measures on a continuous scale, previous studies have used the share of trips made by certain modes, the Herfindahl Hirschman Index (HHI) and Shannon’s entropy index, each of which has its own merits and shortcomings. The share of travel mode has an intuitive meaning of how much individuals employ a set of desirable modes for their entire travel demand (e.g., walking, biking, and public transit). However, it does not capture how their use of various modes is distributed. With this approach, various travel patterns are simplified to the same share value, making it difficult for planners to develop differentiated demand management strategies for those with the same value.

The HHI and Shannon’s entropy share a similar strength. They capture how individuals’ mode use patterns are distributed. The HHI and Shannon’s entropy generate different values for the two cases above. Since HHI calculates and sums the square of each share of trips by various modes, as shown in equation (1), if individuals use fewer modes, their profiles get higher values. Since HHI ranges from 1/N to one, where N is the number of travel modes, I normalized HHI using the equation (2). While large HHI values reflect a strong concentration of mode choices on one mode option, Shannon's entropy presents higher values for those who use more travel modes (see equation (3)). However, their weakness is also apparent: these measures do not differentiate the use of personal vehicles from the use of less-polluting or more sustainable modes. In brief, this study tests all of these measures and searches for their consistent relationships to land use attributes of transit-facility areas.

$$HHI = \sum_{i=1}^{N} S_i^2 \tag{4}$$

$$Normalized\ HHI = \frac{HHI - \frac{1}{N}}{1 - \frac{1}{N}} \tag{5}$$

$$Shannon's\ entropy = -\sum_{i=1}^{N} (p_i * \log(p_i)) \tag{6}$$

3.3.3 Controlling for Residential Self-Selection

Several studies have employed a set of methods to control for residential self-selection when analyzing the effect of land use attributes on the travel behavior of individuals (Cervero et al., 2002; Bhat and Guo, 2007; Cao et al., 2009; Salon, 2015; Scheiner et al., 2016). In this study, I employ a method that does not require a survey dataset on qualitative variables, such as attitudes towards or preferences for certain neighborhood types. Instead, it accounts for individuals residing in a similar built environment but differs in the characteristics. In this method, I adopt a two-stage approach. The first stage estimates the probabilities of individuals choosing various types of neighborhoods from traditional residential neighborhoods in the central city to suburban subdivisions with large parcels. The probabilities are estimated with the multinomial logit model of chosen neighborhood types based on available individual attributes such as demographic and socioeconomic attributes. Then, I include the new terms from these probabilities to the second stage, where several multimodality measurements are tested to detect any effect by land use attributes. This stage controls for residential self-selection, or the effect by individuals choosing residence by their preferences while separating the "true" effect of land use attributes on multimodal travel behavior. One limitation of this approach is that it requires separate modeling of individuals by the type of neighborhoods, which is the type of transit-facility area in this study, leading to a smaller sample size of each model than that of the pooled model.

3.3.4 Data Structure

This study employs three sets of data: the first dataset provides the location of all fixed-guideway transit facilities across the U.S., the second includes land use attributes for the areas around these transit facilities, and the third contains travel behavior of those residents who resided within or outside of these transit-facility areas. TOD Database is a data collection and maintenance project that the Center for Transit-Oriented Development (CTOD) conducted in 2011 with support from the Housing and Urban Development and the Federal Transit Administration. It contains various transit and land use information of all existing, planned, or proposed fixed-guideway transit facilities across the U.S. From the TOD Database, this study extracts the location of the 4,400 transit facilities.

To measure land use attributes for the areas around these transit facilities, I employed various data sources, including ACS 5-year estimates, LEHD, OpenStreetMap (OSM), and Tiger shapefile. I employ two additional sources to complement land use attributes: Google Places Application Programming Interface (API) and walkscore.com. Google Places API allows users to extract the location of various types of businesses from their server, which is updated by Google or users' requests. While public data sources such as the U.S. Census County Business Patterns or LODES share the count or the location of businesses either at aggregated geographic and industrial levels or as an estimation, Google Places API provides detailed information for each business under more than 50 categories. For the factor analysis of this study, I test the density of cafes, bars, and nightclubs in proximity to the fixed-guideway transit facilities. In addition, to capture the conduciveness of the built environment for walking, biking, and transit trips, I include in the factor analysis three composite measures from walkscore.com: walk score, bike score, and transit score. These indices obtained from the walkscore website have been used in several studies to estimate the effect of the built environment on walking behavior in reliable ways (Manaugh & El-Geneidy, 2011; Carr, Dunsiger, & Marcus, 2011).

The 2017 National Household Travel Survey (NHTS) offers the trip diary of individual members of households—923,572 trips made by 264,234 persons living in 129,696 households—for their assigned travel day, which this study employs to calculate several multimodality indexes for these individual travelers. Also, their residential and work or school locations are used to estimate the probability of individual households choosing one type of neighborhood over the others in the first step. Note that the public version of the 2017 NHTS only contains the residence of metropolitan statistical areas and no information for the location of work or school because of the confidentiality agreement with the survey participants. Thus, I requested and gained access to the unrestricted version of the 2017 NHTS, including the residence at the census block group level and the work or school location at the Census Tract level. Estimating the probability employs household- and individual-specific attributes from the household and person tables of the 2017 NHTS. These attributes are included both in the first and second modeling steps. Table 7 presents the variable description and its data source.

Table 7 – List of data source

Variable	Description	Source
Transit facilities	The location information of 4,400 fixed-guideway transit facilities across the U.S.	CTOD database
Population density	Population density	ACS
Activity density	Activity density ((Population + Employment)/area)	ACS, LEHD
Housing density	Housing density	ACS
Multifamily housing	Percentage of multifamily housing units	ACS
Renters	Percentage of renters	ACS
Median block size	Median block circumference length	TIGER
Job population balance	Job to population balance	ACS, LEHD
Job to household	Jobs per household	ACS, LEHD
Retail/service job to household ratio	Retail/service job per household	ACS, LEHD
Food/drink service density	The location of various types of businesses (café, bars, and nightclubs)	Google Places API
Land use entropy	$E = -\sum_{i=1}^{N}(p_i * \ln(p_i))$	LEHD
Street density	Street miles per square mile	OSM
Intersection density	Intersection counts per square mile	OSM
Four-way intersection density	Four-way intersection counts per square mile	OSM
Edge to node ratio	Edges per node in an area	OSM
Betweenness centrality	The percentage of shortest paths that pass through the most important node/edge	OSM
Degree centrality	The number of links incident upon a node	OSM
Regional access	Access to jobs within 45 min driving time	LEHD
A composite index of sustainable travel mode	walk score, bike score, and transit score	Walkscore.com
Travel behavior	Trip diary of individual members of households for their assigned travel day	2017 NHTS Add-on PERSON data

3.4 Results

3.4.1 Typology of TODs

To classify the areas around the 4,400 fixed-guideway transit stops in the U.S., I first identified the half-mile buffer from the transit stops and defined the area inside the buffer as the station area. The station areas may overlap one another in some metropolitan areas where the closest transit stops to a transit stop are often within a half mile. Since land use attributes are not available at the station area level but at the Census Block Group level, I clipped the block group polygons within the half-mile buffer. I collected only the segments of block groups that fell inside the buffer and calculated the area-weighted averages of land use attributes for each station area. These weighted averages capture the development patterns that people experience at a randomly selected location inside the half-mile buffer from a transit stop.

Table 8 presents the final factor solution of 21 land use attributes of a station area. With the Oblimin rotation and Kaiser Normalization, the factor analysis provides three factors: development density, job diversity, and street connectivity. The loadings below 0.3 are not shown in Table 8. The factor correlation matrix in Table 9 indicates that development density and job diversity factors have a low correlation with a magnitude of 0.155, while the street connectivity factor is moderately correlated with development density (0.482) and job diversity factors (0.328). Since the factors show some levels of correlations with each other, individual land use attributes instead of the three factors are tested in the final regression models.

Table 8 – Factor analysis of land use attributes

Variables	Development Density	Job Diversity	Street Connectivity
ln(population density)	0.984		
ln(housing density)	0.964		
ln(block group size)	-0.958		
ln(activity density)	0.835		
transit score	0.761		
ln(regional job accessibility)	0.735		
walk score	0.702		
% multi-family housing units	0.693		
% renters	0.616		
ln(food/drink service density)	0.520		
bike score	0.509		
ln(job to household ratio)		0.960	
ln(retail/service job to household ratio)		0.912	
land use entropy		0.685	
job to population balance		-0.454	
ln(intersection density)			0.841
ln(betweenness centrality)			-0.759
ln(four-way density)			0.735
ln(degree centrality)			-0.734
ln(street density)			0.731
ln(edge to node ratio)			0.700

Table 9 – Factor correlation matrix

Factors	Development Density	Job Diversity	Street Connectivity
Development Density	1.000		
Job Diversity	0.155	1.000	
Street Connectivity	0.482	0.328	1.000

I employed the K-means cluster analysis to classify individual station areas with the extracted factors capturing three dimensions of development patterns around the fixed-guideway transit facilities. Each cluster contains relatively homogeneous station areas with respect to land use. After testing three through six clusters, I chose four based on interpretability and knowledge of the local context. Figure 5 displays a few metropolitan areas with various types of station areas in them.

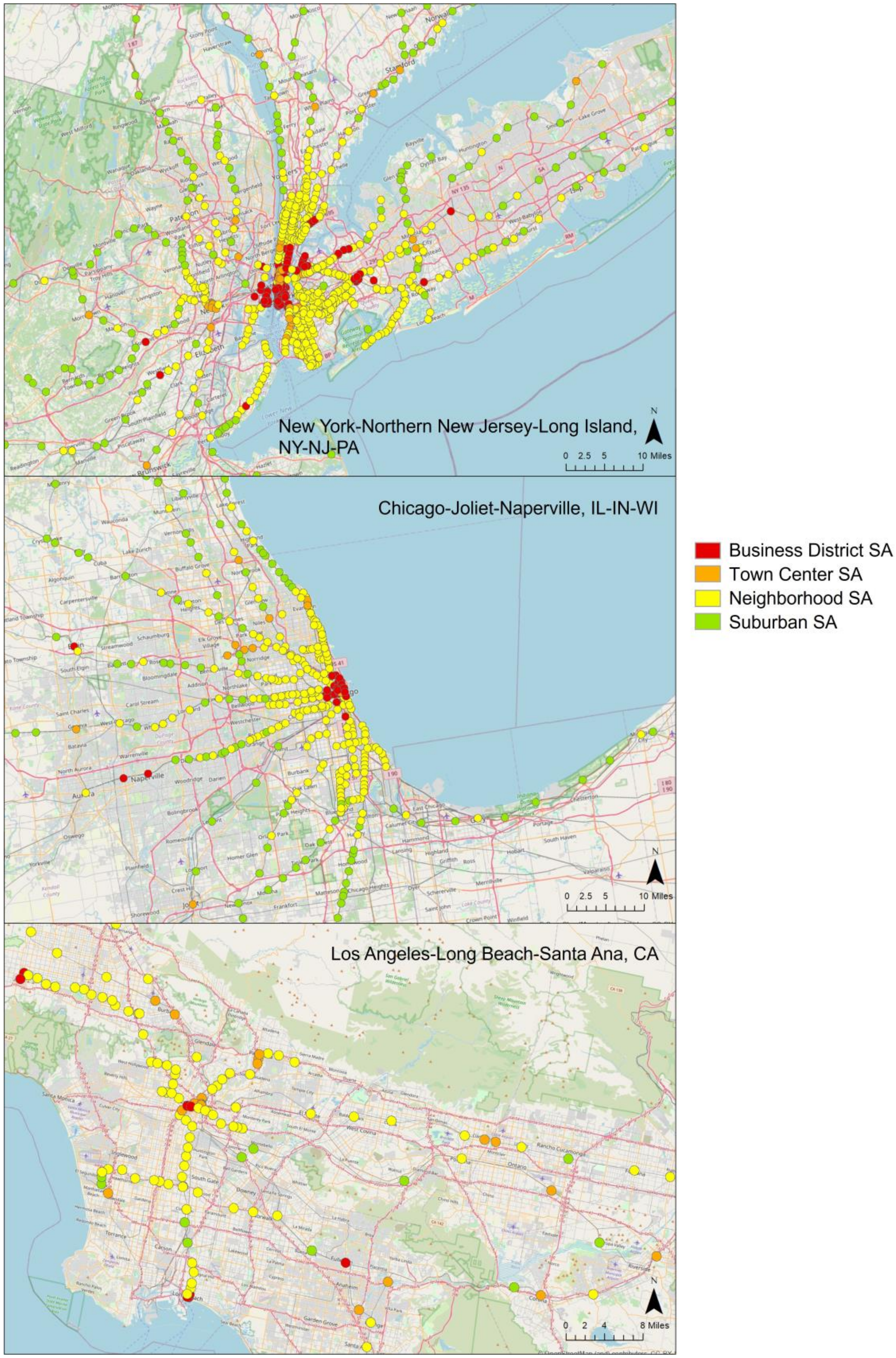

Figure 5 – The types of stations in major metropolitan areas

The cluster analysis resulted in four unique clusters, and Table 10 presents means of factor scores and land use attributes by the types of station areas. Based on the factor scores, the four clusters can be interpreted as business district (type 1), town center (type 2), residential (type 3), and suburban station areas (type 4).

Type 1 is classified as station areas located in a business district. About 20% of station areas (892) are identified in this type. The station areas of this type demonstrate higher densities due to their central location and higher street connectivity due to grid-street network design in the business district. Since these areas are often more oriented to employment and commercial land uses, they have a moderate level of job diversity (or job to housing balance). Type 2 has about 14% of station areas (634) featuring the highest job diversity. The station areas in this type differ from those in the business district in that development is comparatively less dense with low to mid-rise buildings rather than high-rise. Also, they tend to have a more balanced mix of uses than those in the business district. Type 3 has the largest group of station areas (45%), and a total of 1,958 station areas fall into this type. This type presents a moderate development density, but it has a lower job diversity and street connectivity level. The station areas of this type are located primarily in residential districts such as lower-density single- or multi-family housing. Type 4 includes a suburban station area that lacks land use density, job diversity, and street connectivity by having the lowest factor scores compared to other station areas. Around 21% of station areas (916) in the U.S. fall in this type, and they are mainly located at the edge of the city or in a less dense area of the region. These station areas seem to require both land use and transport investments to qualify as a TOD. Surprisingly, station areas in type 3 have a lower job diversity level than those in type 4. While the station areas in type

3 have higher densities than those in type 4, they show the lowest level of job to household ratio, retail/service job to household ratio, and land use entropy among all the station types.

Table 10 – Means of factor scores and land use variables by the types of stations

Factors/variables	Type 1 Business district SA	Type 2 Town center SA	Type 3 Neighbor-hood SA	Type 4 Suburban SA
Development density	**0.826**	**0.068**	**0.187**	**-1.253**
Population density	59.038	23.844	42.011	5.991
Housing density	33.200	13.122	17.062	2.441
Block group size	0.258	1.291	0.284	4.113
Activity density	125.796	77.740	24.710	4.774
Transit score	80.348	58.697	55.977	15.191
Regional job accessibility	436,445.700	316,146.300	320,585.500	147,281.900
Walk score	91.739	74.612	77.204	39.615
Share of multi-family housing units	0.843	0.753	0.540	0.261
Share of renters	0.733	0.683	0.568	0.326
Food/drink service density	0.131	0.085	0.032	0.007
Bike score	79.102	65.915	60.878	32.567
Job diversity	**0.632**	**1.402**	**-0.606**	**-0.291**
Job to household ratio	10.235	257.517	1.204	5.883
Retail/service job to household ratio	7.165	137.015	0.895	2.957
Land use entropy	0.592	0.630	0.401	0.528
Job to population balance	0.385	0.205	0.671	0.626
Street connectivity	**1.422**	**0.085**	**-0.126**	**-1.174**
Intersection density	2.818	1.269	0.892	0.411
Betweenness centrality	0.018	0.029	0.035	0.065
Four-way density	1.357	0.431	0.297	0.075
Degree centrality	0.004	0.008	0.012	0.036
Street density	170.572	103.058	87.566	46.251
Edge to node ratio	2.814	2.570	2.599	2.507
N = 4,400	892 (20%)	634 (14%)	1,958 (45%)	916 (21%)

3.4.2 Measuring Multimodality in Travel Behavior

Multiple indicators were employed to measure multimodal travel behavior in the station areas in this study. Before calculating indicators, I measured the shares of modes in trips by day, week, and trip chain levels. Specifically, I included the following modes of transport: car, walking, cycling, public transit, and ridesourcing service. Figure 6 presents trends in the mode share for a day, week, and chained trip levels by the types of station areas. During a travel day, most residents in all types of station areas relied solely on automobiles to get around, ranging from 40.9% to 83.8%. Similarly, the most chained trips were made exclusively by automobile in all types of station areas ranging from 41% to 83.6%. The share of automobile trips during the past week, by contrast, are smaller than those on the travel day for all types except type 1. The walking, cycling, public transportation, and ridesharing percentage are highest in type 1 at all levels. On the contrary, type 4 has the lowest share of non-motorized modes of travel at all levels.

As a measure of multimodality, I calculated the HHI for the five-mode categories. HHI emphasizes modes with large shares, and it ranges from 1/N to one, with N being the number of modes. Another measure of multimodality is Shannon's entropy which ranges from zero to ln(N). While zero entropy indicates that a person uses only one mode, the maximum value (ln(N)) means that all modes considered are used with equal shares. Similarly, I calculated the entropy for the five-mode categories, and Figure 7 reports mean values for the two multimodality indicators by the types of station areas.

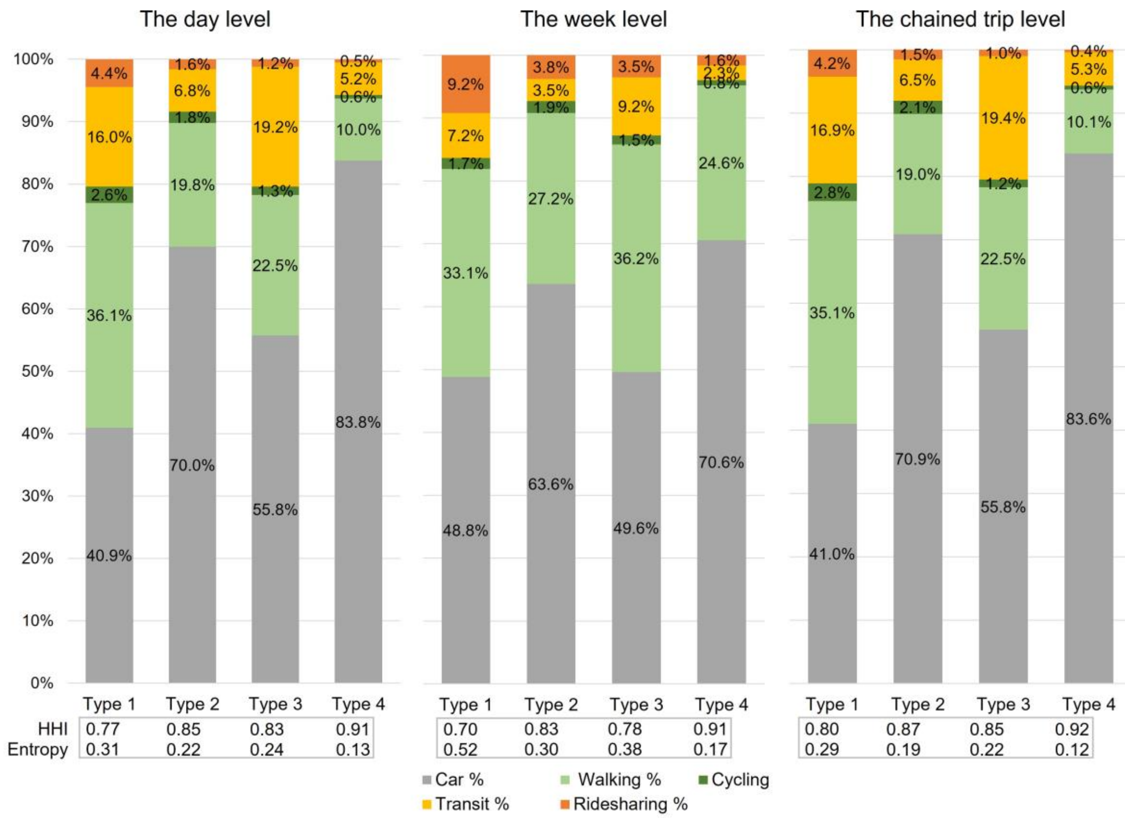

Figure 6 – Mode share at the day, week, and chained trip levels by types of station areas

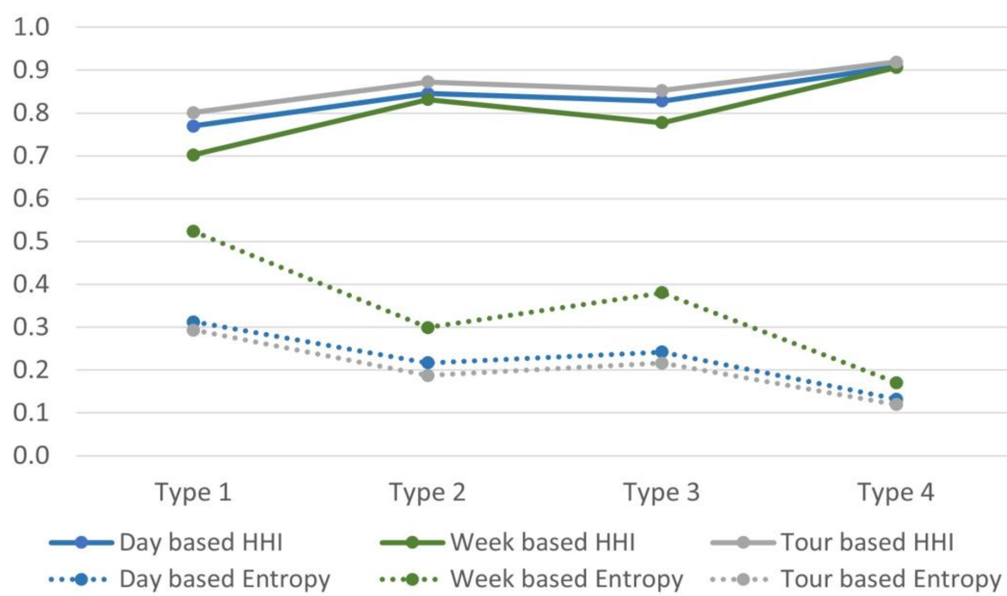

Figure 7 – Multimodality indicators by the types of station areas

All indicators show that residents in type 1 are more multimodal than those in other types of station areas. Specifically, HHIs at all levels (i.e., day, week, and tour-based) tend to increase from type 1 to another. Since a higher value of HHI indicates a strong concentration of mode, residents in type 4 with the highest HHI values (ranging from 0.91 to 0.92) are less likely to be multimodal than those in type 1 with the lowest HHI values (ranging from 0.77 to 0.80). Similarly, entropy shows the same trend by having the highest values (ranging from 0.31 to 0.52) for residents in type 1 and the lowest values (ranging from 0.13 to 0.17) for those in type 4. Based on two multimodality indicators, types 2 and 3 are roughly comparable, but type 3 seems slightly more multimodal than type 2. Even though type 2 has higher factor scores of job diversity and street connectivity than those of type 3, I identified a different spatial distribution (or location) pattern between the two types. Most of the station areas in type 3 are closer to inner cities than those in type 2. Although not always, station areas in type 2 are mainly located on the outer edge of a city and act as connectors between stations in types 1 and 3.

Before running the regression models, I examined the cross-sectional correlations between the multimodality indicators to determine which indicators represent different dimensions. I also looked at correlations between the indicators and the underlying mode shares to illustrate the links between multimodality and modes used. As shown in Table 11, the multimodality indicator that reflects modal concentration (HHI) levels is negatively correlated to the indicators that reflect variability (entropy). Although it is negative, the correlation between HHI and entropy measures is strong at all levels. The multimodality indicators are from moderately to weakly correlated to mode shares. The strongest correlations appear between multimodality and the share of trips made by car (correlations ranging between absolute values of 0.42 and 0.48), which is positively correlated with HHI's modal concentration and negatively correlated with mode entropy. Correlations with all other modes than car are in the opposite direction, and they are weaker than those between multimodal indicators and car use (correlations ranging between absolute values of 0.07 and 0.39). The strongest correlations besides driving can be seen with walking. These results suggest that more multimodality is typically associated with lower levels of driving and vice versa. All directions of correlations are consistent with Scheiner et al. (2016).

Table 11 – Correlations between multimodality indicator and mode shares

	Multimodal indicators					
	HHI			Entropy		
	Day	Week	Tour	Day	Week	Tour
Multimodality						
Week based HHI	0.318					
Tour based HHI	0.911	0.282				
Day based entropy	-0.982	-0.298	-0.902			
Week based entropy	-0.343	-0.987	-0.306	0.322		
Tour based entropy	-0.908	-0.281	-0.996	0.904	0.305	
Mode share						
Car	0.458	0.484	0.420	-0.435	-0.423	-0.415
Walking	-0.370	-0.218	-0.390	0.370	0.168	0.384
Cycling	-0.144	-0.252	-0.126	0.156	0.247	0.128
Public transit	-0.172	-0.514	-0.102	0.172	0.454	0.099
Ridesharing	-0.183	-0.376	-0.139	0.068	0.372	0.146

The correlation results show a high correlation exists around 0.91 between day-based and tour-based multimodality indicators for both HHI and entropy. Thus, I combined week-based and tour-based measures to generate a composite multimodality index instead of combining all three day, week, and tour-based indicators. Also, I normalized the HHI in the range of zero to one since the range of HHI is between 1/N to one, with N being the number of modes. As shown in Figure 8, which displays the distribution of normalized HHI and entropy by the types of station areas, residents show different multimodality levels based on where they live. Specifically, residents in the business district station area (type 1) tend to have relatively equal distribution over the range of both multimodality indicators. On the other hand, residents in the suburban station area (type 4) show a large percentage of high normalized HHI level closest to a maximum value of one and low entropy level closest to a minimum value of zero.

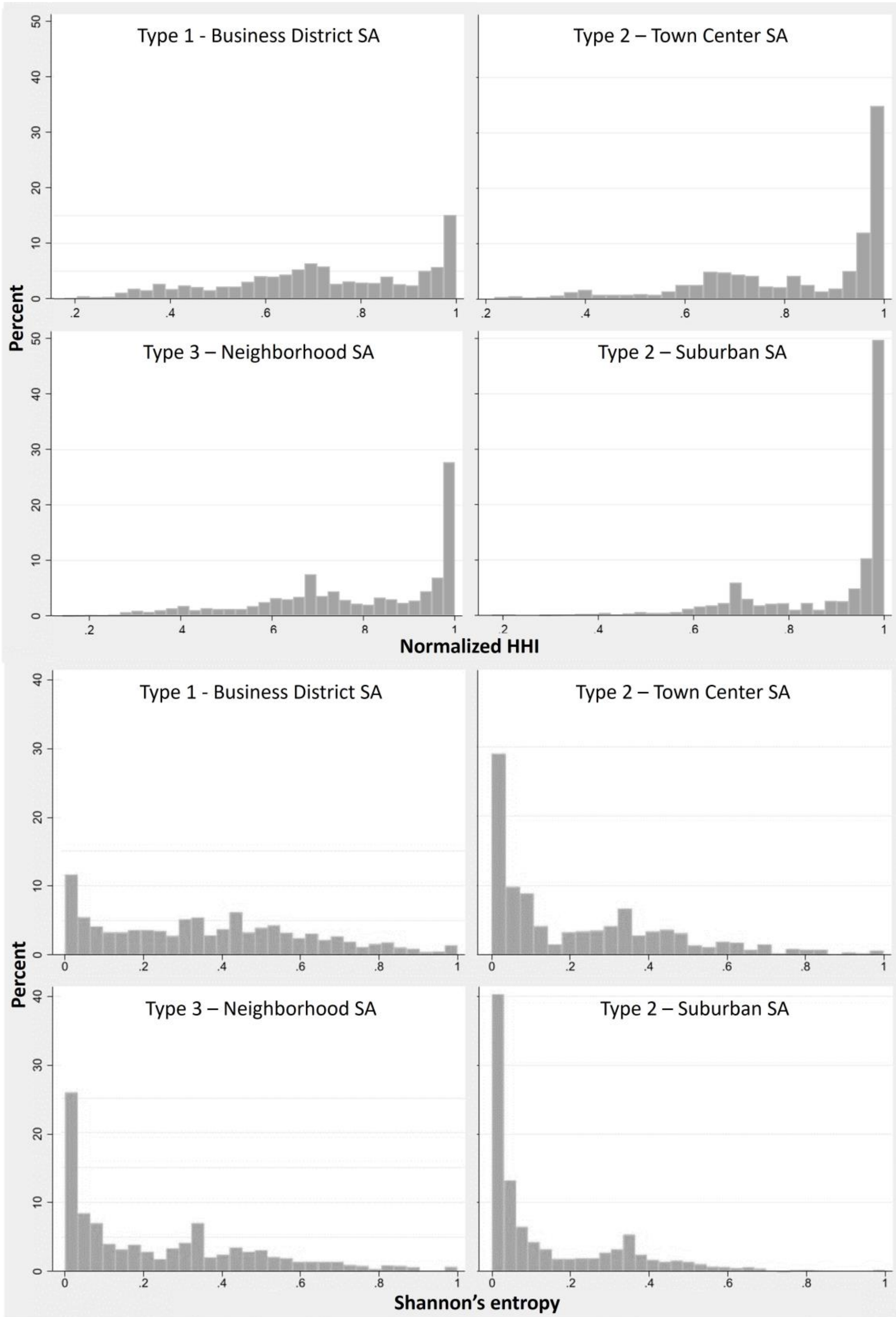

Figure 8 – Normalized HHI and entropy by the types of station areas

3.4.3 *Controlling for Residential Self-Selection*

To control for residential self-selection, I first estimated the probabilities of households choosing various types of station or non-station areas as their residence. In doing so, I limited my data to those households whose member(s) commuted to a workplace or school that was located in the Urban Area in the U.S. Census. That is, these households may have considered locating in station areas because they often traveled to the dense part of metropolitan areas. Next, I processed the dependent variables, a categorical variable containing the choice set of households, in a way in which I take into account the unequal choice set for individual households. After all, not all households were given a full choice set of the four station types to choose from. For example, the households in Buffalo-Cheektowaga-Niagara Fall metropolitan area could not choose the business district station area because their area did not have any. Thus, the final data has the maximum size of choice set as five, including business district station area, town center station area, neighborhood station area, suburban station area, and non-station area. With the final sample size of 35,521 households, Table 12 presents the conditional logit model outcome of residential location choice with the non-station area as the base category. The model includes race, life cycle, highest educational attainment among household members, household income, car per driver, home ownership, and transit score of the workplace. If there is more than one worker in a household, I calculated an average value of transit scores of the workplaces of every household member.

The life cycle of households affects their probabilities of choosing three types of station areas over the non-station area: compared to retiree households without children, multiple-adult households without a child or households with a child are less likely to choose business district station area, town center station area, or suburban station area. Households with children may prefer less congested or safer residential areas with good public schools. The highest educational attainment among the household members accounts for their likelihood of location in one type of neighborhood or another. Compared to graduate or professional degrees, households with lower levels in educational attainment are less likely to reside in business district or town center station areas but more likely to locate in the non-station area. Interestingly, the size of the significant coefficients for varying educational attainment levels decreases as the level increases.

The coefficients of household income levels show that less wealthy households are less likely to choose business district or town center station areas over non-SA. As for household vehicle ownership, if the vehicle(s) of a household is more available for use, households tend to choose non-SA over business district, town center, and neighborhood station areas, except the households without vehicles. These non-vehicle households are more likely to locate in business district or neighborhood station areas, at which living without cars is not as inconvenient as doing so at non-station areas. However, note that the vehicle ownership and residential location of a household are inherently interconnected, or endogenous: Households may choose business district station area because they do not own a car, or those who do not think that a car is necessary may locate in business district station area so that they can live without cars. The coefficients of homeownership show that homeowners are less likely to choose any of the station areas except suburban station areas over the non-station area. We include the average transit score of household members' commute destinations to control the possibility of households commuting by public transit.

The positive coefficients of transit score support that households may be more likely to live in station areas instead of the non-station area if transit commuting is feasible.

Table 12 – The conditional logit model with the unequal choice set of station area type choice

Reference: non-SA	Business district SA		Town center SA		Neighbor-hood SA		Suburban SA	
Race (ref: others)								
White					-0.348	**		
Black					-0.222			
Asian only					-0.491	***		
Life cycle (ref: retiree(s), no child)								
Single, no child	0.951	***	1.484	***			-0.471	*
2+ adult, no child	0.438	**	0.910	**			-0.687	***
Youngest child 0-5	-0.393		-0.584	*			-0.583	**
Youngest child 6-15	-0.816	**	0.748	*			-0.315	
Youngest child 16-21	-1.141	***	0.282				-0.234	
Highest educational attainment (ref: graduate degree)								
No high school diplomat	-23.524	***	-1.361	*				
High school graduate	-1.576	***	-1.275	***				
Some college or associate	-1.251	***	-0.960	**				
Bachelor's degree	-0.522	***	-0.338					
Household income (ref: greater than 100k)								
Under 30k	0.071		-0.553	*				
30k - 60k	-0.511	***	-0.415					
60k - 100k	-0.171		-0.538	**				
Car per driver (ref: more than one)								
No car	0.974	***	0.845		0.747	***		
Less than one car per driver	-2.574	***	-1.692	***	-1.349	***		
One car per driver	-1.226	***	-0.745	**	-0.688	***		
Home ownership	-0.685	***	-0.677	***	-0.532	**		
Transit score of workplace	0.044	***	0.019	*	0.061	***	0.038	***
Constant	-4.033	***	-4.221	***	-4.820	***	-4.948	***
N			35,521					
Log pseudolikelihood (full model)			-20,888,105					
Log pseudolikelihood (market share)			-26,042,575					
Pseudo ρ^2 (market share model as base)			0.20					

*Significant at 90% **Significant at 95% ***Significant at 99%

3.4.4 Effects of Land Use Attributes on Non-Auto Mode Share

Before examining the effect of desirable TOD characteristics on multimodal travel behaviors of residents, I investigated the relationships between land use attributes of TOD and residents' non-auto mode share, including walking, cycling, and public transit trips. I first divided the sample into five groups of residents by the type of station areas in which these residents resided: the business district, the town center, the residential, suburban, and non-station area. Then, I employed a separate regression model for each of the first four groups of individuals who lived within a half mile from any station. Although station areas in one type are distinctive from those in the other types, these station areas present some levels of variation in land use attributes. Thus, by running separate regression models for individuals residing at the same type of station area, I estimated the effect of various land use attributes on non-auto mode share for those individuals who chose to live at the same type of station area. This approach allows me to identify possibly differential strategies that promote residents' sustainable and multimodal travel behavior by their station type. Also, by including self-selection terms in these separate regressions, I controlled for effect by individual attributes and preferences (e.g., urban lifestyles) on the use of various travel modes.

Table 13 presents the outcomes from eight regression models by station area types and tour purposes. For each model, I tested the share of non-automobile mode as a dependent variable, and I tested the share for two tour purposes: work and non-work tour. In this study, I defined a tour as a sequence of trip links that start from and end at one's residence, regardless of the length of the time individuals spent at intermediate stops. If a tour contains at least one trip link whose purpose was related to work, I classified it as a work-related tour, and if not, then I classified it as a non-work tour. Among 21 land use attributes of TODs, I selected five key variables, including activity density, regional job accessibility, service job density, job to population balance, and intersection density, that are frequently employed in the previous studies. In all models, the dependent variable and land use variables were transformed by taking natural logarithms for two reasons: 1) dealing with the non-normal distribution of data and 2) interpreting parameter estimates as elasticities.

As shown in Table 13, the estimated effects of land use variables on the non-automobile mode share exhibit heterogeneity across the station area types and the tour purposes. Specifically, all land use attributes of a station area except regional job accessibility account for more use of non-automobile mode for work-related tours of those who live in the business district SA (type 1). The elasticities of non-automobile mode share with respect to activity density (0.632), food/drink service density (0.419), job to population balance (0.704), and intersection density (0.250) suggest that every 1% increase in those variables are associated with 0.63, 0.42, 0.70, and 0.25% increase in the use of non-automobile mode, respectively. For residents in the town center SA (type 2), the share of non-automobile mode for work-related tours increases with a higher job to population balance of a station area: an increase of job to population balance by 1% is associated with the increase in the non-automobile mode by 0.51%. Activity density (0.551), regional job accessibility (0.701), and job to population balance (0.849) are positively associated with the non-automobile mode share for work-related tours of residents in the neighborhood SA (type 3). The non-automobile mode share for work-related tours increases with higher regional job accessibility and food/drink service density of a station area in the suburban

SA (type 4): 1% increase in regional job accessibility and food/drink service density are associated with 0.29 and 0.33% increase in the share of non-automobile mode.

Similarly, the regression models for non-work tours in Table 13 show that the effects of land use variables on the share of non-automobile mode are also variable across the types of station areas. To be specific, the elasticities of non-automobile share for non-work related tours with respect to activity density (0.521), food/drink service density (0.612), and job to population balance (0.583) for the residents in the business district SA (type 1) indicate that an increase by 1% of those variables are associated with the increase in the share of non-automobile mode by 0.52, 0.61, and 0.58%, respectively. The food/drink service density of a station area accounts for more use of non-automobile for non-work related tours of those who live in the town center SA (type 2). The elasticity of non-automobile mode share for the food/drink service density (0.463) suggests that an increase in the food/drink service density by 1% is associated with a 0.46% increase in the share of non-automobile mode for non-work related tours. For residents in the neighborhood SA (type 3), the share of non-automobile mode for non-work related tours increases with higher activity density, regional job accessibility, and job to population balance of a station area. The elasticities of non-automobile mode share with respect to activity density (0.663), regional job accessibility (0.449), and job to population balance (0.378) indicate that every 1% increase in those variables are related to 0.66, 0.45, and 0.38% increase in the use of non-automobile mode for non-work tours in the neighborhood SA, respectively. For those who live in the suburban SA (type 4), the share of non-automobile mode for non-work related tours is positively related to the food/drink service density: an increase of the food/drink service density by 1% is associated with the increase in the non-automobile mode by 0.20%. However, intersection density does not account for the share of non-automobile mode for non-work-related tours in any station area.

In addition to the land use variables, Table 13 presents the associations between socio-demographic attributes and the non-automobile mode share for work and non-work related tours. The age group is statistically significant in all station areas for work-related tours, while the association becomes insignificant in town center SA and suburban SA type for non-work tours. In general, old age groups tend to decrease their non-automobile mode compared to the young adult group aged below 25. The racial group is statistically significant in particular types of station areas for both work and non-work related tours. Interestingly, the black tends to use less non-automobile mode for work and non-work-related tours than the white in town center SA. In neighborhood SA, another racial group, not including the black and Asian, tends to decrease the use of non-automobile mode for work-related tours compared to the white, while the Asian group tends to increase the use of non-automobile mode for non-work related tours compared to the white. Although not always, education level is negatively associated with the share of non-automobile mode for both work- and non-work tours in all SA except town center SA. Interestingly, the association becomes opposite in business district SA for work-related tours. The life cycle variable shows inconsistent results by the type of station area. For example, a retiree with no child group tends to increase their use of non-automobile mode for both work- and non-work tours in town center SA compared to single with no child group. However, this association becomes the opposite in neighborhood SA for work-related tours. Household income levels are negatively associated with the share of non-automobile mode for both work- and non-work tours in all SA except suburban SA. Compared to households with an income level below 35k, more wealthy households are more likely to reduce the use of non-automobile mode for both work- and non-work tours. The ratio of a car per driver is also negatively associated with the share of non-automobile mode for both work- and non-work tours in the town center and neighborhood SAs. The association also appears in business district SA for non-work-related tours: an increase in car availability is related to a decrease in the share of non-automobile mode.

Table 13 – Regression models of non-auto mode share (log-transformed)

Variables	Work tour				Non-work tour			
	Business district SA	Town center SA	Neighbor-hood SA	Suburban SA	Business district SA	Town center SA	Neighbor-hood SA	Suburban SA
Land use attributes								
Activity density	0.632***	0.349	0.551***	-0.180	0.521**	-0.276	0.663***	0.052
Regional job accessibility	-0.309	0.050	0.701***	0.294**	-0.452	0.347	0.449***	0.109
Food/drink service density	0.419***	0.422	-0.134	0.330**	0.612**	0.463**	-0.147	0.196*
Job to population balance	0.704***	0.509**	0.849**	-0.407	0.583***	0.159	0.378*	-0.112
Intersection density	0.250***	0.099	0.074	0.203	0.110	-0.142	0.229	0.020
Cohort (ref: age<25)								
25<=age<=34	-0.452**	-1.234	-0.193	0.265	-0.332		-0.373*	
35<=age<=44	-0.710***	-1.153	-0.774**	0.522	-0.318		-0.480**	
45<=age<=54	-0.864***	-1.310	-0.614	-0.284*	-0.825**		-0.540	
55<=age<=64	-0.868**	-0.965	-0.574	-0.135	-0.776		-0.439	
age>=65	-0.385	-2.031*	-0.186	-0.463	-0.540*		-0.249	
Female								
Race (ref: White)								
Black		-1.350***	0.240			-1.474***	0.060	
Asian Only		0.150	0.196			-0.386	0.649***	
Others		0.619	-0.462**			0.466	0.113	
Education (ref: no HS)								
HS graduate	0.475			-0.762*	-0.485*		-0.741**	-1.137**
Some college	0.859***			-0.607	-0.379		-0.881***	-1.422***
Bachelor's degree	0.926***			0.052	-0.070		-0.706***	-1.361***
Graduate/professional	0.950***			-0.128	0.074		-0.404	-1.092**
Life cycle (ref: single, no child)								
2+ adults, no child	0.210	0.680	0.005			-0.343		
Youngest child 0-5	0.602	1.161	-0.012			-0.257		
Youngest child 6-15	0.260	0.635	0.205			0.073		
Youngest child 16-21	0.849**	0.084	0.487			0.623		
Retiree(s) no child	0.554	1.305*	-0.390*			1.140***		

Table 13 continued

Income (ref: 0-34,999)								
35,000 – 74,999	-0.188	-0.998*	-0.415*			-0.491	-0.463**	
75,000 – 124,999	-0.480	-0.554	-0.833***			-0.609	-0.665***	
Above 125,000	-0.564**	-0.649	-0.441**			-0.666*	-0.584*	
Car per driver (ref: no car)								
Less than one		-2.337*	-0.974***		-0.807**	-2.157***	-1.105***	
One		-1.509**	-2.010***		-0.992	-1.203**	-2.191***	
More than one		-0.784	-2.309***		-1.123	-2.162***	-2.202***	
Self-selection term								
Business district SA		0.205	-0.148	0.425		-0.201**	-0.323***	1.266
Town center SA	-0.796***		-0.054	-0.884	-0.842***		-0.106	-1.632
Neighborhood SA	0.468**	-0.014		-1.344	0.461***	0.093		-1.128
Suburb SA	-0.070	0.322**	0.207		-0.214*	0.259***	-0.071	
No SA	0.745***	0.193	-0.294	1.176	0.577***	0.511**	0.042	1.123
Constant	1.976	2.706	-11.641***	-6.969**	5.250	1.453	-9.491***	-3.199**
N	436	253	932	563	644	330	1,341	891
Adjusted R-squared	0.36	0.54	0.37	0.13	0.24	0.54	0.36	0.08

*Significant at 90% **Significant at 95% ***Significant at 99%

3.4.5 *Effects of Land Use Attributes on Multimodality*

The results from the regressions in Table 14 and Table 15 present the effects of various TOD characteristics on the multimodal travel behavior of residents. I tested two multimodality indicators (log-transformed normalized HHI and Shannon's entropy) as a dependent variable for two tour purposes. Similar to the previous models on the non-automobile mode share, Table 14 and Table 15 show that the estimated effects of land use variables on normalized HHI and Shannon's entropy show heterogeneity by the type of station areas and tour purposes.

The models for the work tours in Table 14 show that food/drink service density and intersection density account for more multimodality for work-related tours of residents in business district SA (type1). As a large value of HHI reflects a strong concentration, the negative associations indicate that the two land use variables tend to reduce the level of mode concentration for work-related tours of residents in business district SA. Specifically, the elasticity of normalized HHI with respect to food/drink service (-0.092) and intersection density (-0.024) suggest that a 1% increase in those variables is associated with 0.09 and 0.02% decrease, respectively, in the normalized HHI. For residents in the town center SA (type 2), regional job accessibility (-0.116), food/drink service density (-0.082), and job to population balance (-0.110) are negatively associated with the normalized HHI for work-related tours. The normalized HHI for work-related tours of those who live in neighborhood SA (type 3) decreases with higher activity density, regional job accessibility, and job to population balance of a station area. The elasticities of normalized HHI with respect to activity density (-0.060), regional job accessibility (-0.066), and job to population balance (-0.076) indicate that every 1% increase in those variables are associated with 0.06, 0.07, and 0.08% decrease in the normalized HHI, respectively. For residents in suburban SA (type 4), the normalized HHI for work-related tours decreases with higher food/drink service density of a station area: an increase of food/drink service density by 1% is associated with the decrease in the normalized HHI by 0.03%.

For the non-work tours, Table 14 shows that food/drink service density and intersection density tend to reduce the mode concentration of residents in business district SA (type 1). The elasticities of normalized HHI with respect to food/drink service density (-0.043) and intersection density (-0.026) suggest that an increase in those variables by 1% is associated with a 0.04 and 0.03% decrease in the normalized HHI, respectively. For residents in town center SA (type 2), the elasticities of normalized HHI with respect to food/drink service density (-0.052) and job to population balance (-0.062) indicate that every 1% increase in those variables is associated with the decrease in the normalized HHI for non-work related tours by 0.05 and 0.06%, respectively. The regional job accessibility of a station area is negatively related to the normalized HHI for non-work related tours of those who live in neighborhood SA (type 3): an increase of regional job accessibility by 1% is associated with the decrease in the normalized HHI by 0.06%. For residents in suburban SA (type 4), higher food/drink service density accounts for less multimodality for non-work related tours: a 1% increase in the food/drink service density is associated with a 0.03% decrease in the normalized HHI for non-work-related tours. Interestingly, activity density does not account for the normalized HHI for non-work-related tours in any station area.

Table 14 also presents the relationship between residents' sociodemographic characteristics and their multimodality for work and non-work-related tours. In general, age groups have a positive relationship with the normalized HHI for both work- and non-work tours in most station types, suggesting that the elderly tend to increase their mode concentration that fits their needs. This is in line with expectations, as young adults and teenagers are associated with multiple ranges of activity spaces and associated mode options, while older people may limit their travel to fewer modes. However, the age group between 25 and 34 is negatively related to the normalized HHI for work tours in neighborhood SA (type 3) and non-work tours in town center SA (type 2), suggesting an increase in multimodality. This particular age group may have more access to various travel modes than the youngest under 25.

Gender is statistically significant in particular types of station areas for both work and non-work tours. Specifically, female is positively related to an increase in mode concentration for work tour in town center SA (type 2). However, the direction of the relationship becomes opposite in neighborhood SA (type 3) for both work- and non-work tours. Therefore, the gender effect is not fully clear.

In terms of racial groups, the black group tends to decrease their multimodality for work tours in neighborhood SA (type 3), while the group increases their multimodality for non-work tours in town center SA (type 2) compared to the white group. In general, Asian only group is positively associated with the mode concentration for both work- and non-work tours in most station types compared to the white group. However, the group decreases their mode concentration for work tours in town center SA (type 2), suggesting an increase in multimodality.

Education level shows decreases in mode concentration in business district SA (type 1) for both works- and non-work tours. The association tends to be stronger for those with a university entrance qualification or higher education level than those with no high school diploma. However, residents with high school diploma increase their mode concentration for work tours in suburban SA (type 4) compared to those with no high school diploma, suggesting a decrease in multimodality.

The life cycle variable shows inconsistent results regarding multimodality. For example, all life cycle groups compared to single with no child tend to decrease their mode concentration for work tours in business district SA (type 1). Some of these associations are still statistically significant for work tours in suburban SA (type 4) and non-work tours in business district SA (type 1). Conversely, life cycle groups are positively associated with mode concentration for work and non-work tours in neighborhood SA (type 3) compared to a single group with no child. Also, having children is associated with an increase in mode concentration for work tours in town center SA (type 2).

Household income levels are negatively associated with the normalized HHI for both work- and non-work tours in most station types except town center SA (type 2). Compared to households with an income level below 35k, more wealthy households tend to reduce their mode concentration. The ratio of a car per driver is positively associated with the normalized HHI for both work and non-work tours across the models except business district SA (type 1). Car owners tend to limit their travel to using private automobiles, while those who have no car available use a mix of various modes other than car.

Table 14 – Regression models of normalized HHI (log-transformed)

Variables	Work tour				Non-work tour			
	Business district SA	Town center SA	Neighbor-hood SA	Suburban SA	Business district SA	Town center SA	Neighbor-hood SA	Suburban SA
Land use attributes								
Activity density	0.035	-0.064	-0.060**	-0.018	-0.040	-0.050	-0.031	-0.010
Regional job accessibility	-0.050	-0.116*	-0.066**	0.021	0.074	-0.015	-0.058**	0.004
Food/drink service density	-0.092***	-0.082**	0.016	-0.027**	-0.043**	-0.052*	-0.002	-0.026*
Job to population balance	-0.024	-0.110***	-0.076*	-0.000	-0.025	-0.062***	-0.032	-0.027
Intersection density	-0.024*	0.035	-0.009	-0.015	-0.026*	0.015	-0.003	0.024
Cohort (ref: age<25)								
25<=age<=34	0.025		-0.130**			-0.105*	-0.056	0.106
35<=age<=44	0.079		0.036			-0.069	0.015	0.110*
45<=age<=54	0.165**		0.005			-0.003	0.032	0.108*
55<=age<=64	0.078*		0.034**			0.050	0.056**	0.148**
age>=65	-0.019		-0.007			-0.034	0.054*	0.104**
Female		0.076**	-0.045**				-0.023**	
Race (ref: White)								
Black	0.014	-0.109	-0.061*		0.024	0.110*		-0.014
Asian Only	0.149*	-0.117**	0.028		0.060**	0.214***		0.072**
Others	0.003	-0.093	0.032		0.038	-0.022		-0.102
Education (ref: no HS)								
HS graduate	-0.051			0.135*	0.007			
Some college	-0.213**			0.089	-0.041			
Bachelor's degree	-0.220***			0.018	-0.084*			
Graduate/professional	-0.208***			0.052	-0.132***			
Life cycle (ref: single, no child)								
2+ adults, no child	-0.089**	-0.029	0.128**	-0.042	-0.085***		0.038	
Youngest child 0-5	-0.074*	0.109	0.138**	-0.235**	0.003		0.045	
Youngest child 6-15	-0.076*	0.025	0.161***	-0.042	-0.059		0.058**	
Youngest child 16-21	-0.399***	0.213***	0.150***	0.022	-0.125		0.091***	
Retiree(s) no child	-0.087*	-0.006	0.153***	-0.026	-0.027		0.024	

Table 14 continued

Income (ref: 0-34,999)								
35,000 – 74,999	-0.097**		-0.056***		-0.106**		-0.037	-0.004
75,000 – 124,999	-0.061		-0.015		-0.050		-0.019	0.002
Above 125,000	-0.125***		-0.116***		-0.068		-0.076***	-0.085**
Car per driver (ref: no car)								
Less than one		0.312	0.078***	0.207		0.139	0.084**	0.459***
One		0.309*	0.256***	0.298		0.101	0.194***	0.407***
More than one		0.341*	0.233***	0.372*		0.197*	0.221***	0.503***
Self-selection term								
Business district SA		-0.022	0.030**	0.072		0.034	0.016	0.167
Town center SA	0.011		0.004	0.453	0.071***		0.027	-0.438
Neighborhood SA	0.004	0.016		-0.447	-0.037*	0.010		-0.032
Suburb SA	0.017	-0.030**	-0.025		0.031**	-0.026**	-0.037***	
No SA	-0.113***	0.054	0.017	-0.131	-0.101***	-0.044	0.016	0.307*
Constant	-0.095	0.630	0.636*	-1.119***	-1.144	-0.479	0.478	-0.636***
N	435	263	932	564	633	340	1,460	929
Adjusted R-squared	0.20	0.59	0.28	0.18	0.11	0.35	0.18	0.15

*Significant at 90% **Significant at 95% ***Significant at 99%

The regression results of Shannon's entropy in Table 15 also present the heterogeneity in the effects of land use variables on multimodality by type of station areas and tour purposes. As a maximum value of entropy reflects a uniform distribution of travel mode use, the elasticities of entropy for food/drink service density (0.440), job to population balance (0.251), and intersection density (0.236) suggest more multimodality for work-related tours of residents in business district SA (type1). For non-work tours, all land use attributes except regional job accessibility account for more multimodality of residents in business district SA (type 1). Specifically, activity density (0.287), food/drink service density (0.247), job to population balance (0.191), and intersection density (0.108) are positively associated with multimodality.

For residents in town center SA (type 2), job to population balance (0.747) and intersection density (0.272) are positively associated with multimodality for work tours. However, the effect of intersection density disappears in the non-work tour model. The elasticities of entropy with respect to food/drink service density (0.543) and job to population balance (0.354) suggest that every 1% increase in those variables is associated with the increase in the entropy for non-work tours by 0.54 and 0.35%, respectively.

A high activity density and regional job accessibility are associated with an increase in multimodality of residents for both work- and non-work tours in neighborhood SA (type 3). Specifically, a 1% increase in activity density tends to increase the entropy of residents by 0.32 – 0.34%. The increase in regional job accessibility by 1% is associated with the increase in entropy by 0.28 – 0.22%.

For residents in suburban SA (type 4), food/drink service density accounts for more multimodality for both work- and non-work tours. A 1% increase in food/drink service density is associated with the increase in the entropy of residents by 0.17 – 0.13%.

In terms of socio-demographic attributes, the estimated results of the normalized HHI models and entropy models show similar patterns. Age groups are negatively associated with the entropy for work and non-work tours in most station types. In general, the elderly above 45 are likely to decrease their multimodality compared to the youngest under 25. However, the age group between 25 and 34 increases their multimodality for work and non-work tours in neighborhood SA (type 3).

The effect of gender is not fully clear since the effect is only statistically significant in particular types of station areas. Females tend to decrease their multimodality for work tours in business district SA (type 1), while the group increases for non-work tours in neighborhood SA (type 3).

Particular racial groups show significant effects in specific types of station areas. For example, Asian only group decreases their multimodality for work tours in neighborhood SA (type 3) and non-work tours in town center SA (type 2) and suburban SA (type 4). Other racial groups, not including the black and Asian only, are also negatively associated with multimodality for work tours in neighborhood SA (type 3) and non-work tours in business district SA (type 1). However, the black group is not statistically significant in any model.

In business district SA (type 1), residents with higher education tend to increase their multimodality for work and non-work tours compared to those with no high school diploma. Also, having a graduate or professional degree is positively associated with multimodality for non-work tours in suburban SA (type 4). However, residents with high school diploma in suburban SA (type 4) decrease their multimodality for work tours compared to those without a high school diploma.

The life cycle variable shows mixed results regarding multimodality. Compared to single with no child, a household with two adults with no child increases their multimodality for work and non-work tours in business district SA (type 1). A household with the youngest child aged between 16 – 21 is also positively associated with multimodality for work tours of the same type. However, the direction of this effect becomes opposite in town center SA (type 2) and neighborhood SA (type 3).

Although not always, higher household income levels account for more multimodality for both work and non-work tours. For example, more wealthy households are positively associated with the entropy for work tours in business district SA (type 1) and neighborhood SA (type 3) compared to households with an income level below 35k. This association also appears in the model for non-work tours in neighborhood SA (type 3) and suburban SA (type 4). The ratio of a car per driver is negatively associated with multimodality for work and non-work tours in most station types. Thus, an increase in car availability tends to decrease the entropy for both work- and non-work tours across the models. However, the association disappears in business district SA (type 1) and suburban SA (type 4).

Table 15 – Regression models of Shannon's entropy (log-transformed)

Variables	Work tour				Non-work tour			
	Business district SA	Town center SA	Neighbor-hood SA	Suburban SA	Business district SA	Town center SA	Neighbor-hood SA	Suburban SA
Land use attributes								
Activity density	0.065	0.627	0.321**	0.011	0.287*	-0.112	0.343**	0.096
Regional job accessibility	0.097	-0.133	0.278*	0.052	-0.465	0.281	0.216*	-0.046
Food/drink service density	0.440***	0.037	-0.021	0.167*	0.247**	0.543**	-0.062	0.126*
Job to population balance	0.251***	0.747***	0.389	-0.062	0.191**	0.354*	0.264	-0.023
Intersection density	0.236***	0.272*	-0.135	0.053	0.108**	0.109	0.063	-0.033
Cohort (ref: age<25)								
25<=age<=34	-0.191	-0.331	0.476**		0.002		0.323*	-0.373
35<=age<=44	-0.261	0.216	-0.018		0.063		-0.011	-0.596*
45<=age<=54	-0.542**	-0.515	0.093		-0.426***		-0.037	-0.533*
55<=age<=64	-0.485***	-0.396	-0.212		-0.540***		-0.241	-0.549*
age>=65	0.012	-1.894**	0.104		0.080		-0.206	-0.434
Female	-0.126*						0.161*	
Race (ref: White)								
Black			-0.043		-0.008	-0.062		0.228
Asian Only			-0.384**		-0.036	-0.551**		-0.865***
Others			-0.518*		-0.370**	0.399		0.966
Education (ref: no HS)								
HS graduate	-0.001			-0.957***	0.354			0.032
Some college	0.671**			-0.209	0.667**			0.189
Bachelor's degree	0.688***			0.368	0.663			0.340
Graduate/professional	0.809***			0.366	0.789			0.738**
Life cycle (ref: single, no child)								
2+ adults, no child	0.270***	-0.029	-0.487**	0.104	0.344***		-0.267**	
Youngest child 0-5	0.032	-0.922	-0.791***	1.254***	0.112		-0.433**	
Youngest child 6-15	0.118	-0.505	-0.705***	0.000	0.515		-0.395**	
Youngest child 16-21	0.477***	-2.350***	-0.622***	0.147	0.599		-0.451***	
Retiree(s) no child	-0.062	0.379	-0.475***	0.059	0.014		-0.112	

Table 15 continued

Income (ref: 0-34,999)								
35,000 – 74,999	0.060		0.313**				0.327**	0.131
75,000 – 124,999	-0.109		0.084				0.205**	0.197
Above 125,000	0.202*		0.546***				0.547***	0.742**
Car per driver (ref: no car)								
Less than one		-0.469	-0.116			-0.321	-0.461***	-2.820***
One		-0.897**	-0.793***			-0.549	-1.128***	-2.648***
More than one		-1.181*	-0.687**			-0.992**	-1.317***	-3.254***
Self-selection term								
Business district SA		0.091	-0.003	1.900		0.002	0.050	-0.177
Town center SA	-0.313**		-0.343**	-5.170***	-0.340***		-0.493***	1.792
Neighborhood SA	0.119	-0.056		0.351	0.140**	-0.017		0.072
Suburb SA	-0.071	0.177	0.321***		-0.060	0.153**	0.279***	
No SA	0.356***	-0.424	-0.035	2.734**	0.429***	0.071	-0.032	-1.601
Constant	-2.503**	-2.467	-5.087**	-1.218	3.744	-1.886	-5.993***	-0.182
N	445	263	932	564	640	341	1,460	866
Adjusted R-squared	0.29	0.45	0.26	0.19	0.26	0.34	0.22	0.16

*Significant at 90% **Significant at 95% ***Significant at 99%

3.5 Discussion and Conclusion

This study provides a comprehensive assessment of residents' sustainable and multimodal travel behavior in different types of station areas. First, this study identified various factors representing land use attributes in categorizing types of station areas. With three factors—development density, job diversity, and street connectivity—extracted from factor analysis, this study classified four types of station areas in the U.S.: business district, town center, neighborhood, and suburban SA. Residents in each type of station area show different non-automobile mode share and multimodal travel behavior. For example, residents in the business district SA tend to have relatively equal distribution in mode shares between driving, walking, cycling, public transit, and ridesharing at the day, week, and tour levels. On the contrary, residents in suburban SA show a large share of driving ranging from 71 to 84% at all levels of the trip. Similarly, the multimodality indicators—the normalized HHI and entropy—show a different pattern by the types of station areas. All multimodality measures indicate that residents in business district station areas are more multimodal than those in other types of station areas.

While each type of station area presents different travel behavior, it is clear that business district SA (type 1) reflects TOD as a concept that induces residents to be more multimodal. To separate out the effect of desirable TOD characteristics on multimodal travel behavior of residents, several regression models were employed by the types of station areas and tour purposes. Although the five land use attributes—activity density, regional job accessibility, food/drink service density, job to population balance, and intersection density—are desirable elements of the TOD, regression analysis results suggest that not all of these attributes promote sustainable and multimodal travel behavior at various contexts. For example, the regression results in Table 13 show that activity density is statistically significant in particular types, including business district and neighborhood SAs (types 1 and 3). Compared to the elasticities of walking (0.04 - 0.07) and transit mode share (0.01 - 0.07) from Ewing and Cervero (2010), activity density in this study shows quite large elasticities ranging from 0.52 to 0.66. Regional job accessibility is statistically significant in the neighborhood and suburban SAs (types 3 and 4), having elasticities ranging from 0.29 to 0.70. Food/drink service density and job to population balance are statistically significant in the five out of eight models for non-automobile mode share. The elasticities of non-automobile mode share with respect to food/drink service density range from 0.20 to 0.61 in most of the types except neighborhood SA (type 3). Job to population balance has the greatest elasticities ranging from 0.38 to 0.85 in all types except suburban SA (type 4). Interestingly, intersection density is strongly associated with non-automobile mode share only for work tours in business district SA (type 1). The magnitude of elasticity of intersection density (0.25) in this study is similar to those in Ewing and Cervero (2010).

Table 14 and Table 15 also present that the effects of land use variables on residents' multimodality for both work and non-work tours show heterogeneity across station area types. Specifically, activity density around station areas accounts for more multimodality of residents in neighborhood SA (type 3). Regional job accessibility is statistically significant in neighborhood SA (type 3) for work and non-work tours. Higher food/drink service density is associated with more multimodality of those in all station types except neighborhood SA (type 3). Job to population balance leads to less mode concentration of residents in the town center and neighborhood SAs (types 2 and 3), and it also accounts for higher entropy of those in business district and town center SAs (types 1 and 3). Intersection density is statistically significant in business district SA (type 1).

While the effects of land use variables on the share of trips made by non-automobile mode and multimodality indicators are statistically significant, the relationships are inelastic. For example, the elasticity of non-automobile mode share for work tours with respect to food/drink service density (0.419) in business district SA (type 1) suggests that a 1% increase in food/drink service density tends to increase the non-automobile mode share by 0.42%. Similarly, an increase in food/drink service density in type 1 by 1% is associated with a decrease in normalized HHI by 0.09% and increased entropy by 0.44% for work tours. In real life, a reduction in normalized HHI of 0.09% can be explained in the following examples. The value of normalized HHI of a resident in type 1 would be 0.153 (HHI value of 0.322 with five mode categories) if the share of non-automobile mode is 55%, which is a combination of walking (35%), cycling (3%), and transit (17%). In this case, a decrease of 0.09% in the normalized HHI (from 0.153 to 0.152) could correspond to an increase in the non-automobile mode share by 1 – 9% (e.g., an increase in walking mode by 7% together with a 2% increase in cycling). With the same non-automobile mode share of 55%, the value of Shannon’s entropy of the resident in type 1 would be 0.55. An increase in the entropy by 0.44 with respect to the rise in food/drink service density by 1% could correspond to an increase in the non-automobile mode share by 1 – 10% (e.g., an increase in walking by 10%). Although individual elasticities are small, the combined effect of several land use variables on multimodality for certain station areas and tour purposes could be quite large.

In addition to land use attributes, the regression analysis confirms that younger age, higher education, higher household income, and higher car availability level are strongly associated with sustainable and multimodal travel behavior. These relationships are consistent with the findings from previous studies (Buehler & Hamre, 2015; Scheiner et al., 2016). Notably, the elasticities of non-automobile mode share, normalized HHI, and entropy with respect to the car per driver variable present the greatest magnitude of 2.34, 0.50, and 3.25, respectively. This finding suggests that policies that reduce the need to own a car may effectively promote sustainable and multimodal travel behavior. Such policies may include intensifying housing supply near stations and offering accommodation for households with fewer cars that tend to reside in proximity to transit.

The findings from several regression analyses suggest that planners and policymakers should take different strategies to encourage sustainable and multimodal travel behavior by the types of station areas. First, business district SA (type 1) presents land use attributes around stations that reflect TOD characteristics, including high density, mixed-use, and high intersection density. Since they are mainly located in the central urban area, which tends to be the most intensely developed area, the improvement of existing station areas and further development of TOD construction may be limited. Thus, specific policy incentives, such as offering density bonuses, optimizing land use patterns and street networks, and alleviating land costs around station areas, may be needed. Second, town center SA (type 2) has a high potential for effective TOD, promoting multimodality. These station areas are mainly located between stations in type 1 and those in type 3, interconnecting two station areas. Since the regression results reveal that food/drink service density and job to population balance tend to increase sustainable and multimodal travel behavior in these areas, the investment in these areas, such as offering service areas that meets residents' needs and attracting workplaces and workers, may be required. As the largest group of station areas, neighborhood SA (type 3) has a good basis for TOD implementation. Since these areas are primarily located inner cities closer to the city center, increasing regional job accessibility and obtaining a higher level of job to population balance would be adequate to promote multimodality in these areas. Compared to the types mentioned earlier, suburban SA (type 4) has limited potential for TOD implementation because of its location at the end of a transit network. The regression models confirmed that the effects of land use attributes around station areas in type 4 on sustainable and multimodal travel behavior are relatively small than those in other types. Also, only one of the five land use variables, food/drink service density, is statistically significant in the models. Thus, investment in station areas of this type to develop as a TOD would be less effective in the short term.

This study has several limitations that future research could address. First, it is important to note that more multimodality does not necessarily indicate more sustainable travel behavior. A higher level of multimodality may also have resulted from a frequent user of public transit who starts driving to work. To better understand how sustainable travel can be attained by increasing multimodality, it would be useful to enhance the measurement of multimodality with more information, such as detailed categories of trip purpose or mode. Also, there are multiple ways to measure multimodal travel behavior in addition to HHI and entropy. With the same information about trips, HHI and entropy show a slightly different level of multimodality because each measure captures different dimensions of travel behavior. This indicates that the regression analysis results may depend on the indicator chosen. Thus, future research can develop more complex measures that take into account as much information as possible on multimodal travel behavior. Also, future research can test group membership which categorizes individuals into several traveler types (e.g., multimodal vs. monomodal travelers) in addition to multimodality indicators. The classification of multiple groups of multimodal travelers may allow direct consideration of the specific mode used. By employing latent class cluster analysis, which generates the probabilities, the group membership can have continuous variables and be used as the dependent variable for the regression models.

Second, the list of variables needs to be extended. The regression models in this study did not include some variables that were employed in previous studies, such as parking supplies and prices, quality of transit service, environmental awareness, or public safety due to limitations in data availability. This study covers station areas in the entire U.S., and these variables are unavailable for such a large geographic area.

Finally, this study relies on cross-sectional data—the 2017 NHTS—in estimating the effects of TOD characteristics on multimodal travel behavior. Thus, the findings are reported based on correlation, not causal inferences. Future research can investigate changes in multimodality over time at the individual level by the types of station areas. Although both the 2009 and 2017 NHTS are cross-sectional data that do not trace changes in the travel behavior of the same person over time, the comparison of the two data may provide a better understanding of the change in multimodal travel behavior changes around station areas.

CHAPTER 4. THE POTENTIAL IMPACT OF NEW MOBILITY SERVICES ON TRANSIT DEMAND

4.1 Introduction

With the emergence of advanced technology, new mobility services have grown rapidly and have disrupted urban transportation systems. In the sharing economy, new mobility service has become more prevalent as an innovative transportation strategy that enables users to access a mode of transport on an as-needed basis (Cohen & Shaheen, 2018). New mobility services generally include various service models and transportation modes, such as ridehailing, ridesharing, car sharing, bike sharing, microtransit, and shared autonomous vehicles, that meet the diverse needs of travelers.

Among others, ridesourcing services continued to expand across urban areas as innovative mobility-on-demand services. Ridesourcing refers to services that connect passengers with drivers who provide rides in their private vehicles through smartphone apps. Transportation network companies (TNCs), such as Uber and Lyft, operate these services in over 900 large metropolitan areas and 10,000 cities worldwide (Uber, 2020). Since these services are more flexible and convenient based on real-time information, TNCs have had a profound impact on reshaping urban transportation systems and have rendered the traditional modes such as car, taxi, transit, walk, and bike less competitive (Brazil & Kirk, 2016). In fact, major U.S. cities have experienced a growing population of ridesourcing services and a stagnant or declined transit ridership at the same time (Graehler, Mucci, & Erhardt, 2019).

The explosive growth of ridesourcing services has stimulated a debate on whether they predominantly complement or substitute for public transit. How ridesourcing affects public transit could be explained through two mechanisms (Hall et al., 2018). Theoretically, ridesourcing could be an alternative mode of travel by encouraging riders to shift from public transit. Its convenient service and reduced cost from a shared ride option could make ridesourcing more attractive. On the other hand, ridesourcing could complement public transit by filling the temporal and spatial gap in public transit's fixed route and fixed schedule. They could further reduce parking needs in TOD areas by reducing the need for car ownership and associated cost. In addition, ridesourcing could complete trips for transit passengers by providing first- and last-mile services. While new mobility services are expected to play an important role in planning how TODs can be implemented, the impacts and consequences of such services on traditional modes of transport are still not well understood.

This study examines the potential impact of transportation network companies (TNCs) on transit demand by examining whether ridesourcing services affect ridership of rail transit systems in Chicago, focusing on understanding heterogeneity in the effects. Specifically, this study assesses the impact of different types of ridesourcing services, including exclusive and shared services on rail transit, using a panel data set that combines information on ridesourcing trips with rail transit ridership data published by Chicago Transit Agency (CTA). It employs panel regression models at the census tract level due to a data structure that combines cross-sectional and time-series data. Other built environment and socio-demographic factors contributing to rail transit ridership, such as population density, employment density, household income level, and car ownership, are considered in the models. The findings of this study will offer insights into the role of the built environment and sociodemographic attributes in rail transit demand, in addition to the relationship between the ridesourcing service and rail transit demand. These insights will equip the local and regional agencies to craft policies that encourage a complementary relationship between the two services.

4.2 Literature Review

Overall, the understanding of the impact of TNCs on existing modes and travel behavior is growing but is still limited due to the relative novelty of these services and the lack of publicly available data on ridesourcing services (Loa, Hossain, & Habib, 2021). Multiple studies have often found a detrimental impact of ridesourcing on taxi services (Rayle et al., 2016; Alemi et al., 2018; Brodeur & Nield, 2018; Oviedo, Granada, & Perez-Jaramillo, 2020), but the literature regarding the relationship between ridesourcing and public transit has offered mixed results that largely depend on specific contexts and analytical methods.

Some studies developed a survey design to examine the use of ridesourcing and its impact on the use of public transit. Using survey results for San Francisco, Rayle et al. (2016) found that ridesourcing competes with public transit for some trips because of travel time savings. Specifically, 33% of the ridesourcing users reported that they would have taken public transit if ridesourcing had not been available. Also, they found an induced travel effect by ridesourcing. Specifically, 8% of the respondents said they had made their trip because of the availability of ridesourcing services. The survey results conducted by Clewlow and Mishra (2017) provided both complementary and substitutive effects of ridesourcing on transit use by the type of transit service. While ridesourcing services tend to attract people away from light rail service (a 3% reduction in light rail transit use), they complement commuter rail services (a 3% increase in commuter rail transit use). Using an online survey of millennials in California, Alemi et al. (2018) analyzed the self-reported behavioral changes in response to the use of ridesourcing services. The majority reported they had reduced the amount of driving, active modes, and public transit due to the use of ridesourcing. Also, most respondents answered that their last trip with ridesourcing would have been made by a taxi if ridesourcing service had not been available. These findings support the substitution effect of ridesourcing on other modes of travel. Young and Farber (2019) found that the rise of ridesourcing corresponds to a significant decrease in ridership of taxis and a rise in public transit and active modes of transit using household travel survey data in Southern Ontario.

While previous studies conducted a descriptive analysis of survey data, Circella & Alemi (2018) employed a latent class analysis to classify the impact of ridesourcing use on the use of other travel modes. Among three latent classes, they found that first class members tend to live in urban neighborhoods with high transit accessibility and are more multimodal than other classes. Also, they reported a reduction in the amount of driving, public transit, and active modes due to their last ridesourcing trip. Conversely, the last class includes individuals who mainly live in the suburbs with low transit accessibility. Surprisingly, the members of this class tend to be multimodal and increase the use of public transit as a result of the use of ridesourcing. By employing logistic regression models on the survey of ridesourcing passengers in the Greater Boston region, Gehrke et al. (2019) found that respondents with mobility tools, such as monthly transit pass and private vehicles, are more likely to have substituted ridesourcing services for the use of public transit. Mostofi et al. (2020) employed binary logistic regression with survey data to compare the probability of frequent use of public transit between regular ridesourcing users (who regularly use ridesourcing service as their primary mode for at least one of their trip purposes) and non-regular users (who do not use ridesourcing service as their primary mode for their trip purposes). They found that the association between the regular use of ridesourcing and the use of public transit depends on the contexts of cities. For example, regular ridesourcing users are more likely to use public transit than non-regular users in Cairo, while this association is negative in Tehran. By employing structure equation models on the web-based survey in Toronto, Loa et al. (2021) revealed that individuals who are students, younger than the age of 25, earning less than $50,000 annually, not owning a private vehicle tend to substitute ridesourcing trips for public transit, while respondents with old age and high income over $50,000 annually are likely to replace their ridesourcing trips for a taxi.

Other studies employed empirical data to examine the effects of ridesourcing on public transit and the factors that impact the effect. Since ridesourcing data at a trip level is not publicly available, some studies used a binary dummy variable to represent the presence of ridesourcing services in the study area. For example, Hall et al. (2018) and Babar and Burtch (2020) evaluated the effects of ridesourcing service entry on public transit use by employing a difference-in-differences approach. Hall et al. (2018) estimated the timing of Uber entry and the intensity of uber penetration using the share of google searches for “Uber” in each metropolitan statistical area (MSA). Their finding indicates that Uber complements the use of public transit, increasing ridership by 5%. Also, they found the heterogeneity effect of Uber on transit ridership. Specifically, Uber's entry reduces transit ridership in smaller MSAs by 5.9%, while it increases ridership in larger cities by 0.8%. Using monthly observations of transit use in urbanized areas together with a binary indicator of ridesourcing service availability, Babar and Burtch (2020) found that ridesourcing complements the use of commuter rails over the 12 months following the entry of ridesourcing.

Zhang and Zhang (2018) and Deka and Fei (2019) used individual-level trip frequency data from the 2017 NHTS in their zero-inflated negative binomial models. Both studies examined the relationship between the frequency of ridesourcing use and the frequency of public transit use. Zhang and Zhang (2018) found that public transit use is positively associated with the frequency and the probability of ridesourcing use. Their findings also indicate that the positive relationship between the two modes was more evident for people living in neighborhoods with higher density or households with fewer vehicles. While Deka and Fei (2019) did not show whether two modes are complements or substitutes, they found that ridesourcing frequency is higher for people living near transit stations. That suggests that people may use ridesourcing instead of transit even when the accessibility to transit services is prevalent in a neighborhood. Kong et al. (2020) collected 181,172 ridesourcing trip data, which took place in November 2016 in Chengdu, China. They developed a three-level structure to differentiate the complementary and substitution effect between ridesourcing and public transit: identifying transit coverage, estimating travel time difference, and quantifying service quality. They assumed that ridesourcing substitutes public transit when a ridesourcing trip meets the following criteria: 1) ridesourcing trip picked up/dropped off within the transit coverage, 2) the travel time difference between ridesourcing and public transit was shorter than a pre-defined threshold, and 3) ridesourcing does not improve service quality much. Their results corroborate that ridesourcing both complements and substitutes for public transit. The substitution effect was more apparent in the city center around transit lines, while the complementary effect was prevalent in the peripheral areas. By combining detailed data of ridesourcing trips and subway disruption data in Toronto, Hawkins and Habib (2020) found that many subway users tend to use ridesourcing services rather than other public transit modes with parallel service routes in response to disruption in the subway systems. Specifically, subway users are likely to switch to TNCs from the subway after a service delay of 7 minutes with a confidence bound between 3 and 12 minutes. Loa et al. (2021) employed the recursive regression model and the bivariate ordered probit model to examine the role of socio-demographic and land use attributes in the generation of public transit and ridesourcing demand. Using trip-level TNC data and a regional household travel survey in Toronto, they revealed that the number of ridesourcing trips generated in a dissemination area with a population of 400-700 persons is positively associated with the number of transit trips generated in the area. Also, they found that higher densities of commercial and recreational establishments and increases in the coverage of transit routes tend to affect the generation of both ridesourcing and public transit trips. These findings suggest a complementary relationship between ridesourcing services and public transit.

While the existing literature has attempted to examine new mobility services in station areas, the lack of publicly available data regarding ridesourcing prevents researchers from examining how new mobility services impact other modes of transportation. Also, previous studies have presented the heterogeneity in the effects of ridesourcing on public transit that largely depends on the context and the specific area. Whether ridesourcing services complement or substitute for public transit is still unclear. Such inconsistent results in the literature review indicate the need for further investigation of the impacts of ridesourcing on public transit demand. Also, existing studies are limited to examining different types of ridesourcing services (i.e., exclusive vs. shared rides) on transit demand. Thus, this study will fill in these gaps by developing a longitudinal analysis to capture different impacts of each type of ridesourcing on transit demand, rather than treating ridesourcing services as a single mode.

4.3 Data and Methods

4.3.1 Conceptual Framework

This study examines the potential effects of ridesourcing services on demand for rail transit, focusing on understanding heterogeneity in the effects of the type of ridesourcing services. To achieve said objective, I first categorize TNC trips into two groups: ridehailing (exclusive service trips) and ridesharing (shared service trips), then identify each trip's pickup location based on whether it is originated within transit coverage. While two terms have been used interchangeably, ridehailing and ridesharing services are different from each other. The main idea of ridehailing is that a rider hires or hails a personal driver to reach the exact location where he needs to go. By contrast, ridesharing services allow a rider to share a ride with other passengers in his journey to the destination. While major TNCs provide both services, each service may impact transit demand differently. By differentiating between exclusive and shared TNC trips, I expect to capture the heterogeneity in two types of ridesourcing services.

Figure 9 describes possible relations between ridesourcing and public transit. On the positive side, ridesourcing could complement public transit by expanding travel options for people without private automobiles. It could fill the temporal and spatial gap in public transit's fixed route and fixed schedule. The service could further reduce parking needs in station areas by reducing the need for car ownership and its associated costs. On the other hand, ridesourcing services could be an alternative mode of travel by encouraging riders to shift from public transit. Its flexible and convenient service based on real-time information could make public transit less competitive. Also, ridesourcing would generate new trips—induced travel—that did not exist in the absence of ridesourcing services. The endogenous variables—ridesourcing and public transit services—are influenced by several exogenous variables, such as socio-demographic and land use attributes. For example, income growth could increase car ownership and decrease transit ridership. Car ownership is another logical determinant of transit ridership, with zero-car households primarily dependent upon transit.

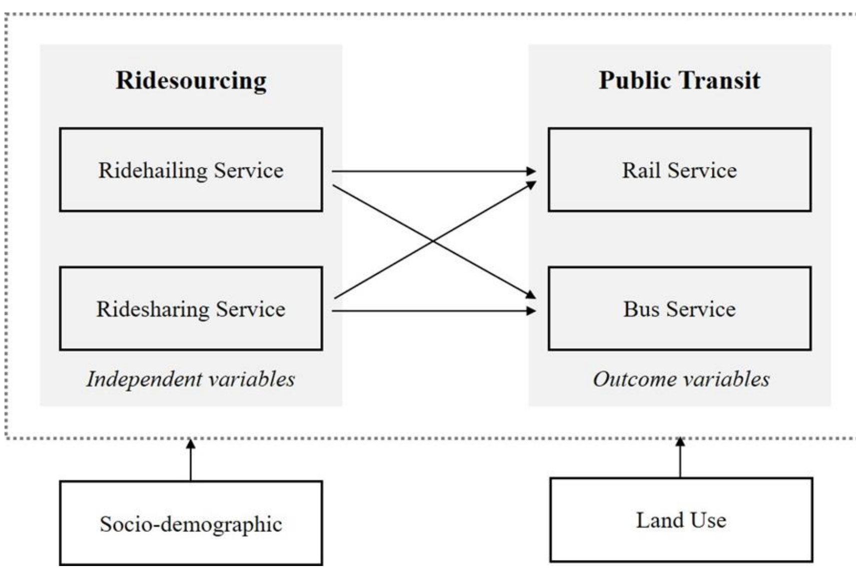

Figure 9 – Conceptual framework

4.3.2 Panel Regression Models

This study employs panel regression models to assess the effects of TNCs on public transit demand in Chicago. Since CTA publishes public transit ridership in a format of time-series data that observed the ridership from each station across time, a longitudinal analysis is a more appropriate estimation approach than a cross-sectional analysis. Using monthly transit ridership data, coupled with data obtained from various sources for the City of Chicago at the census tract level, this study develops panel regression models: fixed effect (FE) and random effect (RE) models, as shown in equation (7). Each model quantifies each factor's—exclusive and shared TNC trips--effects on transit ridership.

$$\mathcal{Y}_{it} = \alpha + \beta x_{it} + \mathcal{u}_i + e_{it} \tag{8}$$

where $\mathcal{Y}_{it}$ represents a dependent variable, rail transit ridership, at census tract *i* for time *t*, x_{it} represents a column vector of attributes at census tract *i* for time *t*, β represents the corresponding coefficient column vector of parameters, $\mathcal{u}_i$ represents a spatial specific effect for all the census tract-specific time-invariant unobserved attributes, and e_{it} represents the idiosyncratic error or time-varying unobserved attributes for each census tract in each time. $\mathcal{u}_i$ is added for each census tract *i* as an intercept in the FE model, while it is treated as a random term that is independent and identically distributed in the RE model.

FE model has shown great value when the number of factors is high, and a correlation exists within factors in the panel data. FE model adds individual-specific intercepts $\mathcal{u}_i$ (i = 1,…, n) to each census tract that captures heterogeneities across census tracts and eliminates spatial correlation. On the other hand, RE models assume that factors within panel data are independent. The independent assumption can be an over-simplification for plenty of datasets, giving sub-optimal results compared to FE models. This study conducts the Hausman test to examine the correlation among factors. The null hypothesis is that the FE model and RE models are statistically similar at a p-value of 0.05. RE models should be chosen because of their computational efficiency when FE and RE models are not significantly different.

4.3.3 Data Structure

The outcome variable in this study is rail transit ridership, which is a primary indicator of transit demand. The CTA publishes the ridership data on a monthly basis. Since rail system ridership is primarily counted as boardings at each station, this study aggregated the rail station's ridership at the census tract level based on the location of stations. The key independent variable of interest is TNC trips published by the City of Chicago. The TNC trip data is publicly accessible through Chicago Data Portal (https://data.cityofchicago.org/Transportation/Transportation-Network-Providers-Trips/m6dm-c72p). Since November 2018, the City of Chicago ordinance has required transportation network providers to report all TNC trips within the city boundary on a quarterly basis. Thus, this study collected data from November 1, 2018 to February 28, 2020, which includes a total of about 105 million trips (average of 216,793 trips per day). I selected this time frame because transportation network providers suspended their shared-ride options in March 2020 in the U.S. due to the COVID-19.

Each trip record contains various attributes, including trip duration, distance, pickup location, drop-off location, fare, and ride option, whether an exclusive ride or a shared ride. Due to privacy protection, the data does not provide latitude and longitude points of a trip. Instead, it gives the approximate location of a trip, the geographic coordinates of the pickup and drop-off census tracts' centroid. Also, trips that departed from or arrived outside the city boundary do not have location information. Thus, this study excludes those trips from the final dataset. I divided TNC trips into two groups in terms of ride options: ride-hailing trips and ride-sharing trips. By doing so, I expect to estimate the different impacts of TNCs on transit demand.

In addition to transit ridership data and TNC trip data, this study uses a list of socioeconomic and demographic variables obtained from the American Community Survey (ACS) 5-year estimates data, including population density, median household income, and percentage of zero-vehicle households. Some employment characteristics are obtained from the Longitudinal Employer-Household Dynamics (LEHD) data. Furthermore, I collected the quality score of transit service of a census tract's centroid from Walkscore. While TNC trips and transit ridership data are available until 2020, no publicly available data for other determinants exists after 2019 (the year 2018 for the LEHD data). Thus, I employed a linear forecasting method to extrapolate data to 2020 based on observed patterns between 2010 and 2019 (between 2010 and 2018 for the LEHD data) at the census tract level. I collected nine consecutive years of ACS 5-year estimates, starting from 2010 (2006-2010) to 2019 (2015-2019) data. Table 16 presents the selected variables with their descriptions, data source, and descriptive statistics.

Table 16 – List of data source

Variable	Description	Source
Dependent variable		
Monthly rail ridership	Natural log of weekday rail ridership Natural log of weekend rail ridership	CTA Nov. 2018 – Feb. 2020
Independent variables		
Monthly TNC trips	Natural log of weekday single trips Natural log of weekend single trips Natural log of weekday shared trips Natural log of weekend shared trips Percentage of weekday trips occurring out of rail transit service hour Percentage of weekend trips occurring out of rail transit service hour	City of Chicago Nov. 2018 – Feb. 2020
Population density	Total number of populations per square mile	2010 (2006-2010) – 2019 (2015-2019) ACS 5-year estimates
Employment density	Total number of the primary jobs per square mile	2010 – 2018 LEHD
Household income	Percentage of households with income below \$25k Percentage of households with income between \$25-50k Percentage of households with income between \$50-75k Percentage of households with income between \$75-100k Percentage of households with income above \$100k	2010 (2006-2010) – 2019 (2015-2019) ACS 5-year estimates
Zero-vehicle households	Percentage of zero-vehicle households	2010 (2006-2010) – 2019 (2015-2019) ACS 5-year estimates
Transit services	Transit score (out of 100) capturing the quality of all public transit offerings	Walkscore.com

4.4 Results

4.4.1 Exploring Temporal and Spatial Distribution of TNC and Rail Transit

Before running the regression models, I examined ridership trends of 143 rail stations that exist in the study area over time. Figure 10 presents the monthly totals for ridership of different travel modes in Chicago. Although rail transit ridership is roughly two times bigger than TNC trips, overall, monthly rail transit ridership declined by 11%, while monthly TNC trips increased by 3% between November 2018 and February 2020. Interestingly, TNC ridership shows two divergent trends over the same period. Specifically, monthly exclusive TNC trips show an increase of more than 18%, while monthly shared TNC trips decreased by 47%. The share of exclusive TNC trips has steadily increased from 77% to 89%. The decline in ridership for rail transit and shared TNC trips indicates that people may prefer exclusive TNC trips, which provide convenient service, assuming it is affordable.

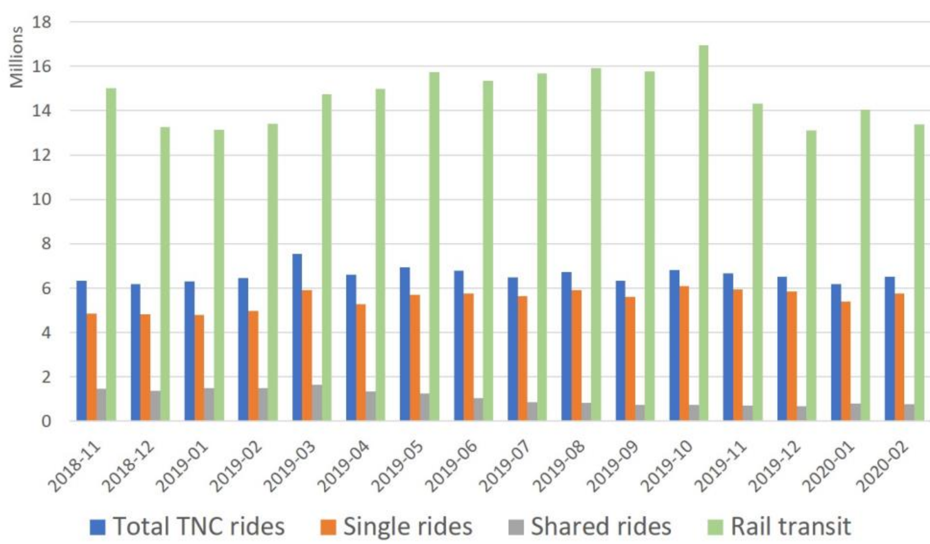

Figure 10 – Monthly ridership by mode of travel

The two maps shown in Figure 11 depict the spatial distributions of daily ridership for rail transit and overall TNC trips that originated from each census tract in Chicago over the 16 months. Due to the different ranges of the number of trips made by two modes, natural breaks (Jenks) classification has been employed in displaying the maps. In both cases, the maps reveal the spatial heterogeneity because most of the trips took place in the central areas and the northwest part of Chicago. After examining the flow between origins and destinations of TNC trips, I found that TNC trips predominantly began around the central business district and two airports—Midway International Airport and O'Hare International Airport. Since these airports are the largest generators of TNC trips within the region and the sociodemographic information for these areas is not available, this study considered TNC trips from the two airports as outliers and excluded them from the analysis.

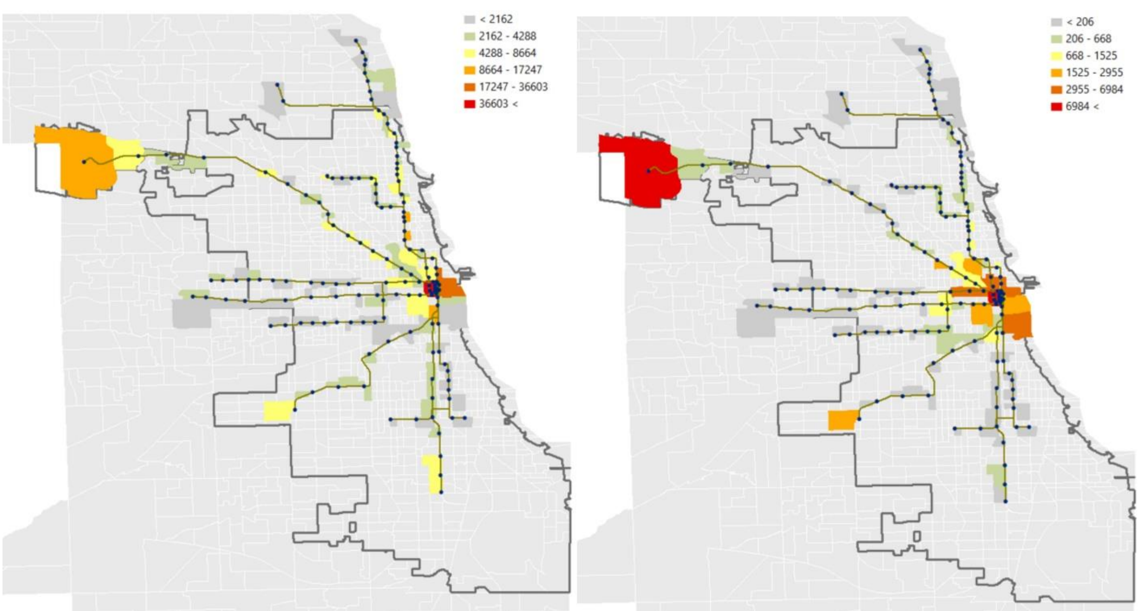

Figure 11 – Daily Rail Transit Ridership (left) and Daily TNC Trips (right) by Census Tract in the City of Chicago

Descriptive statistics in Table 17 – Descriptive statistics show that the average monthly ridership of rail transit is larger during weekday than on weekend. This trend consistently appears in the average ridership of two ridesourcing services. Interestingly, the percentage of monthly TNC trips occurring out of rail transit service hours is higher during the weekend than weekday. Although the information on trip purposes is not available in this study, the finding suggests that travelers tend to use ridesourcing services for their social and recreational activities on the weekend night when rail transit service is unavailable.

Table 17 – Descriptive statistics of variables in original metrics

Variable	**Mean**	**Standard deviation**
Dependent variable		
Monthly rail transit during weekday	106,692	184,262
Monthly rail transit during weekend	22,006	31,686
Independent variables		
Monthly single TNC trips during weekday	13,468	33,182
Monthly single TNC trips during weekend	6,220	13,243
Monthly shared TNC trips during weekday	2,452	5,499
Monthly shared TNC trips during weekend	900	1,615
% of monthly weekday TNC trips occurring out of rail transit service hour	0.006	0.030
% of monthly weekend TNC trips occurring out of rail transit service hour	0.021	0.061
Population density	18,276	13,982
Employment density	21,306	82,270
% of households with income below $25k	0.270	0.170
% of households with income between $25-50k	0.184	0.086
% of households with income between $50-75k	0.142	0.053
% of households with income between $75-100k	0.105	0.050
% of households with income above $100k	0.300	0.194
% of zero-vehicle households	0.309	0.148
Transit score (out of 100)	70.6	16.28

4.4.2 *Effects of Ridesourcing Service on Rail Transit Ridership*

To measure the effect of ridesourcing services on rail transit ridership, I developed both fixed effect and random effect panel regression models by the day of the week, whether it is weekday or weekend. After excluding two airports area—Midway International Airport and O'Hare International Airport—with no sociodemographic information available, the final models used a total of 111 census tracts in the City of Chicago with data for 16 months for each tract. Table 18 reports the panel regression results of the impact of TNCs on overall rail transit ridership during the weekday. Both outcome variables and explanatory variables are all measured in logs, so the coefficients represent the percent increase in rail transit ridership that accompanies percent changes in total trips of TNC. The significant result of the Hausman test suggests that a fixed effect model is preferred over a random effect model.

As shown in Table 18, the analysis reveals that TNC trips are statistically associated with rail transit ridership for the weekday. However, the direction of association depends on the type of trip, whether it is an exclusive or shared ride. Specifically, exclusive TNC trips are positively associated with the demand for rail transit, while shared TNC trips have a negative impact on the rail transit demand. The coefficients can be interpreted as elasticity since both TNC trips and rail transit ridership are log-transformed in the model. The results show that the relationship between the two modes is inelastic. A 1% increase in exclusive TNC trips is expected to have approximately a 0.05% increase in rail transit ridership. On the contrary, a 1% increase in shared TNC trips is associated with a decrease in rail transit ridership by 0.01%. These findings support both complementary and substitutionary effects of TNCs on rail transit demand. The percentage of TNC trips out of rail transit service hours is not statistically significant in the weekday rail transit demand model.

In examining sociodemographic characteristics, I found population density, income level, and households with zero vehicle account for rail transit ridership. Specifically, population density and the percentage of households with zero vehicle are positively correlated with rail transit ridership. The level of household income is negatively associated with rail transit ridership. Contrary to my expectation, the percentages of households with income below $50,000 are negatively associated with rail transit ridership.

Table 18 – Fixed effect panel regression model for weekday rail transit demand

Variables	Coefficient		Std. error	t-statistics
ln(Single TNC trips)	0.048	***	0.014	3.36
ln(Shared TNC trips)	-0.011	***	0.003	-3.18
Percentage of TNC trips occurring out of rail transit service hour	-0.074		0.066	-1.13
ln(Population density)	0.156	**	0.093	1.67
ln(Employment density)	0.001		0.003	0.32
Percentage of households with income below $25k	-0.144	***	0.301	-3.46
Percentage of households with income between $25-50k	-0.720	***	0.324	-2.22
Percentage of households with income between $50-75k	-0.667	**	0.321	-2.08
Percentage of households with income between $75-100k	-0.582	***	0.278	-2.09
Percentage of households with zero vehicle	0.555	**	0.271	2.05
Intercept	10.230	***	0.962	10.63
Number of observations	1,776			
Number of entities	111			
Number of time periods	16			
R-squared	0.37			

*Significant at 90% **Significant at 95% ***Significant at 99%

Table 19 presents the panel regression results of the impact of TNCs on demand for rail transit, measured by the number of ridership during the weekend. The significant result of the Hausman test suggests that a fixed effect model is preferred over a random effect model. Similar to the previous model for weekday rail transit ridership, an exclusive TNC trip is positively correlated with rail transit ridership, while a shared TNC trip is negatively associated with rail transit ridership. Specifically, a 1% increase in exclusive TNC trips results in an increase in rail transit ridership by 0.15%, while a 1% increase in shared TNC trips is associated with a 0.01% decrease in rail transit ridership. The percentage of TNC trips occurring out of rail transit service hours is insignificant in the weekend model.

For the weekend model, I found that higher population and employment densities account for the increase in rail transit ridership. For a 1% increase in population density, I expect rail transit ridership to increase by 0.38%. The size of the coefficient is larger than that in the weekday model. Contrary to the weekday model, employment density is statistically significant. A 1% increase in employment density is associated with a 0.02% increase in rail transit ridership. Interestingly, socio-demographic attributes including household income and zero vehicle households become insignificant in the weekend model.

Table 19 – Fixed effect panel regression model for weekend rail transit demand

Variables	Coefficient		Std. error	t-statistics
ln(Single TNC trips)	0.145	***	0.028	5.14
ln(Shared TNC trips)	-0.014	**	0.007	-2.07
Percentage of TNC trips occurring out of rail transit service hour	-0.119		0.099	-1.20
ln(Population density)	0.382	***	0.136	2.81
ln(Employment density)	0.017	*	0.009	1.94
Percentage of households with income below $25k	0.876		0.570	1.54
Percentage of households with income between $25-50k	0.762		0.531	1.44
Percentage of households with income between $50-75k	0.470		0.427	1.10
Percentage of households with income between $75-100k	-0.396		0.476	-0.83
Percentage of households with zero vehicle	-0.234		0.420	-0.56
Intercept	4.941	***	1.437	3.44
Number of observations	1,776			
Number of entities	111			
Number of time periods	16			
R-squared	0.23			

*Significant at 90% **Significant at 95% ***Significant at 99%

4.5 Discussion and Conclusion

As the sharing economy has grown rapidly in the world economy, new mobility services such as Uber and Lyft have begun to challenge traditional transportation providers. The growing popularity and exponential growth of such services have led to concerns that they will compete and eventually replace existing public transportation systems. In fact, major U.S. cities have experienced the growing popularity of TNC services and the stagnant or declined transit ridership at the same time. While previous studies have attempted to examine the relationship between two services, the effect of ridesourcing varies considerably by context. Also, they often failed to consider different types of ridesourcing services by treating them as a single homogenous mode. In particular, the two types of ridesourcing services—an exclusive and shared ride—may attract distinctive groups of travelers. Thus, this study conducts a longitudinal analysis of the determinants of rail transit ridership in Chicago to capture different impacts of ridesourcing on transit demand based on the type of services, whether exclusive or shared rides. Specifically, this study developed two panel regression models for different ridesourcing service types (i.e., exclusive vs. shared ride) and different service times (i.e., weekdays vs. weekends).

The findings suggest that TNCs show heterogeneous impacts on the demand for rail transit. TNC trips with exclusive riders (ridehailing service) are positively associated with rail transit ridership, while TNC trips with multiple riders (ridesharing) are negatively correlated with rail transit ridership. The heterogeneous effects by the type of TNC services appear both in the weekday and weekend models. The results indicate that ridehailing service complements rail transit, increasing average rail transit ridership by 0.05% for weekdays and 0.15% for weekends. Based on the average values of monthly rail transit ridership in Table 17, a 1% increase in ridehailing trips during weekdays (134 trips) is related to the increase in rail transit ridership by 0.05% (53 trips). During weekend, an increase in ridehailing trips by 1% (62 trips) is associated with the rise in rail transit ridership by 0.15% (33 trips). This finding supports a hypothesis that TNC tends to increase the reach and flexibility of fixed-route and fixed-schedule rail transit. Also, one reason ridehailing service is a complement rather than a substitute for rail transit may be that rail transit is still much cheaper to use. The average fare of TNC for exclusive riders is $7.50, while the regular fare of rail transit is $2.50.

Conversely, ridesharing seems to compete with rail transit, reducing rail transit ridership for both weekdays and weekends by 0.01%. A 1% increase in ridehailing trips during weekdays (25 trips) reduces rail transit ridership by 0.01% (11 trips). During weekend, an increase in ridehailing trips by 1% (9 trips) is related to the decrease in rail transit ridership by 0.01% (3 trips). The two modes share the similarity because they are based on sharing at a lower cost of travel. Since ridesharing offers relatively lower fare than ridehailing and more flexibility than public transit, travelers who can afford may choose ridesharing services over rail transit to take advantage of such benefits. Thus, the results support a hypothesis that TNC can substitute for public transit. Particularly, ridesharing service may induce existing transit riders to switch their mode of travel from rail transit to TNC.

Although the results present the heterogeneous impacts of TNCs on rail transit demand, the findings suggest that the complementing effect of ridehailing is better than the substituting effect of ridesharing. This highlights the need for planners and practitioners to craft policies and strategies that encourage and maximize the complementary effect of ridehailing. I believe there is an opportunity to solve the first and last mile problem by connecting rail transit with TNC service. People assume that they can get to their destination faster with TNC service than with public transit. However, that is not always true. Rail transit can be faster than TNC in particular times and areas such as peak hours in downtown with heavy traffic. Thus, planners and policymakers may promote a multimodal trip that connects TNC service with rail transit. The integration of ridehailing service with the transit system has occurred in several cities where transit agencies collaborate with

Uber to subsidize rides within a city boundary. The partnership's success may depend largely on its ability to find the right targets and locations.

It should be noted that this study has some limitations that further research could address. First, it is difficult to generalize the results of this study for other locations because it focused on the City of Chicago due to the data availability. Since ridesourcing is a growing global phenomenon, more fruitful cross-sectional comparisons of transit demand between multiple cities or regions around the world can be made in future research to better understand the factors affecting transit ridership and estimate the effects of ridesourcing on transit demand if data is available.

Second, this study should be extended to examine the effects of TNC on each mode of public transit, such as bus and rail, rather than focusing on rail transit demand. Since two modes have different frequencies and service coverage, their demand may be differently affected by ridesourcing services. The prior studies found some empirical evidence on the heterogeneity in the effects of ridesourcing on different modes of public transit. Barbar and Burtch (2020) found that ridesourcing substitutes for the use of city buses, while it complements the use of commuter rails over the 12 months following the entry of ridesourcing. Clewlow and Mishra (2017) claim that ridesourcing tends to draw people away from both light rail services (a 3 % reduction in use) and bus services (a 6% reduction in use, while ridesourcing complements commuter rail services (a 3% increase in use). Currently, the City of Chicago provides ridership data for rail and bus systems. While rail transit ridership is available at a station level, bus ridership is given at a route level. It is possible to distinguish census tracts with high demand for bus systems from the others with less accessibility. However, it is difficult to identify the exact volume of trips that are generated from each census tract. Thus, future research would provide a more in-depth understanding of how TNCs affect public transit ridership by investigating each mode of transit, which may serve different routes and passengers.

Third, this study did not include rail transit-related factors such as transit capacity and quality of service. In fact, such variables were collected and tested in the panel regression models. However, these factors tend to be invariant over a short period. Thus, the panel regression models did not estimate the effect of those factors on rail transit ridership. Extending the research period to multiple years or other analytical methods estimating the effect of time-invariant variables may help resolve this issue.

Finally, future research should investigate the spatial and temporal variations of both ridesourcing and transit ridership. While this study separately estimated models for weekday and weekend trips, detailed information on the spatial (e.g., ridership in central areas vs. less urban areas) and temporal patterns (e.g., peak vs. non-peak time) of travel demand should be considered in the model estimations. The effect of ridesourcing on public transit could also be different depending on the purpose of trips. For example, Murphy and Felgon (2016) suggest that travelers tend to use ridesourcing for social and recreational purposes while they often take public transit for work purposes. Thus, the differentiating of trip purposes in analytic models would effectively estimate the relationship between two services. Also, ridesourcing would generate new trips—induced travel—that did not exist in the absence of ridesourcine service. This impact could be neutral for rail transit because it neither increases nor decreases transit ridership, representing more travel in the aggregate. Further research needs to separate the neutrality out as a third possible impact on transit use.

CHAPTER 5. CONCLUSIONS AND FUTURE WORK

5.1 Research Summary

As a popular planning strategy, TOD has been applied to reduce automobile trips and diversify travel modes due to easy access to transit, better walkability, and proximity to diverse amenities. Currently, more than 4,400 fixed-route rail transit stations exist in the U.S. While some rail transit stations are already established as TODs, others offer the potential for various forms of TOD practice. The variation in transit areas may lead to different impacts on transportation outcomes, and the degree to which components of TODs influence travel behavior is still debatable. Also, TNCs have disrupted traditional modes of travel in recent years. However, the relationship between TNCs and public transit, whether they compete with or complement each other, is still unclear. Thus, this dissertation addresses some critical questions regarding travel behavior in station areas: 1) do people walk more in TOD? 2) are residents more multimodal in TOD? and 3) what is the potential impact of TNCs on transit demand?

The different data and methodologies are used to address these research questions. In Chapter 2, this dissertation examined the role of rail transit access in promoting walking trips for purposes other than transit use in TOD areas. This study employed a propensity matching score (PSM) to identify two groups of TAZs that have similar built environment attributes, making rail transit access the key differentiator. Then I compared walking behavior between TOD with rail access and non-TOD areas without rail access by employing multi-level logistic regression models. After controlling for sociodemographic and travel characteristics, the results show that rail transit access is strongly associated with the prevalence of walking trips for both commuting and non-commuting purposes. There are two theoretical propositions to explain why a higher level of walking trips for purposes other than transit use is observed in TOD areas. The behavioral spillover theory suggests that a high volume of people who walk to and from transit stations in TOD areas may attract more walking trips to other destinations. The social interaction theory indicates that people tend to increase their propensity to walk when there is a high volume of pedestrians in TODs.

Chapter 3 investigated the effects of land use attributes in TOD on residents' sustainable and multimodal travel behavior in different station areas. By employing factor- and cluster-analysis techniques on various land use characteristics around station areas, this study classified 4,400 fixed-guideway transit station areas in the U.S. into four types: business district, town center, neighborhood, and suburban SAs. For residents in each type of station area, this study measured multiple indicators that capture sustainable and multimodal travel behavior with five categories of travel modes, including driving, walking, cycling, public transit, and ridesourcing. Each type of station area presents different patterns of multimodality of residents. Compared to those in all other types, residents in the business district station areas show the highest level of multimodality, having relatively equal distribution in mode shares at the day, week, and tour levels of the trip. On the other hand, residents in suburban station areas are less multimodal due to their large share of driving ranging from 71 to 84% at all trip levels. To separate out the effect of various TOD attributes on residents' multimodal travel behavior, this study developed multiple regression models by station types and tour purposes. Also, the models controlled for residential self-selection by including the probability of individuals living the other types of TOD except one's own. The results of multiple regression models present that not all land use attributes promote sustainable and multimodal travel behavior of those who live in station areas. Specifically, the results show heterogeneity across station area types. Although not always, higher food/drink service density and job to population balance account for more multimodality and sustainable modes of travel of residents. However, activity density and regional job accessibility tend to increase multimodality of residents only in neighborhood station areas, and significant association. Intersection density is statistically significant only in business district station areas. These findings suggest the need to develop different strategies for promoting multimodality by the type of TODs. Since business district stations are mainly located in the most intensely developed area in the central city, improving existing station areas or further development of TOD construction may be limited. Specific policy incentives, such as density bonuses, alleviating land cost, optimizing land use patterns, and improving street networks, are recommended.

Chapter 4 explored the heterogeneous impacts of TNCs on transit demand whether two types of services—ridehailing and ridesharing—account for changes in rail transit ridership in Chicago. By employing panel regression models at the census tract level, this study found both complementing and substituting effects of TNCs on rail transit ridership. While ridehailing service for single riders is positively associated with rail transit, increasing the ridership by 0.05% for weekdays and 0.15% for the weekend, ridesharing service for multiple riders is negatively related to rail transit, reducing the ridership by 0.01% for both weekday and weekend. These findings support two hypotheses: 1) ridesourcing service supplements public transit with fixed-route and fixed-schedule, and

2) ridesharing replaces public transit because of relatively lower cost than ridehailing and more convenient and flexible service than transit. This study highlights the need for planners and practitioners to craft strategies that encourage the complementary effect of ridehailing. The integration of ridesourcing services with transit systems can be achieved through the partnership between transit agencies and TNCs.

5.2 Dissertation Contributions

This dissertation extends the literature on TOD and multimodal travel behavior from the lens of theoretical contribution. While previous studies have investigated either TOD or multimodality, none have attempted to examine both topics. Also, most of the prior studies on multimodality have focused on identifying travelers who present multimodal travel behavior and analyzing which sociodemographic attributes determine their multimodality. This dissertation combined two topics by analyzing the level of multimodality of residents in different TOD areas. This dissertation found evidence of a heterogeneity effect of land use attributes around transit areas in the level of multimodality after controlling for socio-demographics and residential self-selection. The finding is meaningful because it informs practitioners to develop different policies and strategies by the type of TODs for promoting residents' sustainable and multimodal travel behavior.

With respect to new mobility services, this dissertation contributes a theoretical advance by separating different impacts of TNCs by the type of services, whether it is exclusive or shared services, on transit demand. To date, the explosive growth of TNCs has stimulated a debate on whether they complement or substitute for public transit. Two mechanisms are often used to explain how TNCs affect public transit (Hall et al., 2018).

Theoretically, TNCs could take travelers away from public transit by providing convenient and flexible services. On the other hand, TNCs could complement public transit by filling the spatial and temporal gap in fixed routes and fixed schedules of public transit. The growing body of literature on the relationship between ridesourcing and public transit has produced inconsistent results that largely depend on context. Also, most studies consider TNCs as a single homogeneous mode rather than separating different types of TNC services (i.e., an exclusive or shared service). Since two types of TNC services may serve distinctive groups of travelers and geographical areas, this dissertation identified different types of TNCs. By applying longitudinal data, this dissertation found both complementarity and substitution effects of TNCs. The results suggest that a shared TNC service substitutes rail transit but complements rail transit during both weekday and weekend. This finding is important for transportation planners in integrating on-demand mobility services into transit systems.

The empirical contributions of this dissertation primarily come from the case being analyzed and the dataset that enables the empirical estimations. In the existing TOD literature, a limited number of studies have attempted to investigate transit areas in the entire U.S. This dissertation offers a robust classification of TOD types by capturing various conditions of existing TODs from more than 20 land use attributes around station areas and identifying their heterogeneous outcomes on sustainable and multimodal travel behavior. As a result, this dissertation identified four distinctive TOD types (i.e., business district, town center, neighborhood, and suburban) and the variation across the types. The typology of TODs and their heterogeneous effects on travel behavior shows how much difference in multimodal travel behavior planners can expect by converting a transit area from one TOD type to another.

Furthermore, the use of large-scale datasets with the help of open-source data resources highlights the empirical contributions of this dissertation. In Chapter 3, this dissertation used three sets of data, including the location of all fixed-guideway transit facilities across the U.S., the land use attributes around station areas, and the national household travel survey. Notably, this dissertation measured land use attributes by employing a variety of datasets from ACS, LEHD, OSM, Google Places API, and Walkscore. In Chapter 4, this dissertation employed a large dataset of ridesourcing trip records, including about 105 million trips from November 2018 to February 2020. While previous studies conducted a cross-sectional analysis on this dataset, this dissertation employed a longitudinal analysis using fixed effect panel regression models to prove the impacts of TNCs on rail transit demand.

5.3 Research Limitations and Directions for Future Research

The dissertation has several limitations that future research could address. First, it is difficult to generalize the findings from the first and third studies for the nation or other geographies because the study areas are limited to Atlanta and Chicago. Therefore, more fruitful cross-sectional comparisons of travel behavior between multiple cities or regions can be made in future research to understand better the factors affecting travel behavior in TOD areas and estimate the effects of TNCs on transit demand. However, the cross-sectional analysis does not prove causal relations, and future studies may employ other adequate models, such as structural equation modeling, to estimate the causality.

Second, future research may better understand how sustainable travel can be attained by increasing multimodality with more complex multimodality measures with more detailed categories of trip purpose. Also, future research can test group membership which categorizes individuals into several traveler types as well as multimodality indicators by employing latent class cluster analysis. In future research, the list of variables needs to be extended by considering the following variables: parking supplies and prices, quality of transit service, or public safety. Furthermore, the methodology of the second study can be improved when it investigates changes in multimodality over time. Although both the 2009 and 2017 NHTS are cross-sectional data that do not trace changes in the travel behavior of the same person over time, the comparison of the two data may provide a better understanding of the change in multimodal travel behavior changes around station areas.

Third, future research should extend the type of public transit by examining each mode rather than treating transit as a single mode. Specifically, the third study did not examine bus transit ridership which is not publicly available at the transit stop level. Since two modes have different levels of service and capacity, their demand may be differently associated with TNC services. Also, the spatial and temporal variations of both ridesourcing and transit ridership should be thoroughly examined in future research to provide a more in-depth understanding of how TNCs affect public transit ridership. The relationship between TNCs and public transit could also be different depending on the purpose of trips. Furthermore, the regression models did not include transit capacity and quality of service variables. Since these variables tend to be time-invariant over a short period, extending the research period to multiple years may help reflect changes in such factors.

Finally, this dissertation used the data collected before the outbreak of COVID -19 does not examine how the pandemic affects residents' travel behavior in TOD areas. Since the COVID -19 pandemic is a global health crisis, governments have imposed various rules, such as social distancing, curfew, and lockdown, to prevent the spread of the virus. With the restrictions imposed by governments, the COVID-19 led to unprecedented changes in people's lives worldwide. First, the pandemic has accelerated the trend toward telecommuting, but it is uncertain whether the shift to working from home is permanent or temporary. Second, the pandemic has disrupted individuals' mobility options. TNCs have suspended their shared ride service, and transit agencies have modified public transit systems by reducing service and capacity limits and following the social distancing policy and other safety measures. With limited transport sharing options, travelers have increased their use of private automobiles, leading to traffic congestion and air pollution. At the same time, the pandemic has increased the demand for more active and sustainable options, such as walking and cycling. Since the COVID-19 pandemic has caused a severe impact than past pandemics and is still going on by threatening global health, future research should explore the impact of COVID-19 on individuals' travel behavior by revisiting the topics in this dissertation.

APPENDICES

A.1 Appendix – Chapter 4:

Table 20 – Descriptive statistics of variables that included in the final fixed effect panel models

Variable	Description	Mean	Standard deviation
Dependent variable			
Monthly rail ridership	Natural log of weekday rail ridership	11.07	0.90
	Natural log of weekend rail ridership	9.48	0.95
Independent variables			
Monthly TNC trips	Natural log of weekday single trips	7.23	2.99
	Natural log of weekend single trips	6.34	3.46
	Natural log of weekday shared trips	5.98	3.06
	Natural log of weekend shared trips	5.04	3.22
	Percentage of weekday trips occurring out of rail transit service hour	0.00	0.03
	Percentage of weekend trips occurring out of rail transit service hour	0.02	0.06
Population density	Natural log of the total number of populations per square mile	9.53	0.79
Employment density	Natural log of the total number of the primary jobs per square mile	8.16	2.48
Household income	Percentage of households with income below $25k	0.27	0.17
	Percentage of households with income between $25-50k	0.18	0.09
	Percentage of households with income between $50-75k	0.14	0.05
	Percentage of households with income between $75-100k	0.10	0.05
	Percentage of households with income above $100k	0.30	0.19
Zero-vehicle households	Percentage of zero-vehicle households	0.31	0.15
Transit services	Transit score (out of 100) capturing the quality of all public transit offerings	70.62	16.28

Table 21 – Census tract specific constant terms in the fixed effect panel regression for weekday and weekend rail transit demand

Variable	Weekday			Weekend		
	Coefficient		Robust S.E.	Coefficient		Robust S.E.
Census Tract 2	-1.381	***	0.070	-1.332	***	0.104
Census Tract 3	-0.014		0.085	-0.153		0.122
Census Tract 4	-0.405	***	0.058	-0.374	***	0.097
Census Tract 5	-0.818	***	0.129	-1.169	***	0.192
Census Tract 6	-0.595	***	0.109	-0.676	***	0.162
Census Tract 7	-0.423	***	0.105	-0.537	***	0.158
Census Tract 8	-0.726	***	0.074	-0.706	***	0.109
Census Tract 9	-0.102		0.095	-0.301	**	0.143
Census Tract 10	-0.445	***	0.082	-0.180		0.136
Census Tract 11	-1.282	***	0.080	-1.514	***	0.177
Census Tract 12	-0.812	***	0.106	-0.691	***	0.165
Census Tract 13	-1.045	***	0.112	-0.986	***	0.181
Census Tract 14	-0.171		0.130	0.413	*	0.239
Census Tract 15	-0.800	***	0.123	-0.583	***	0.196
Census Tract 16	0.887	***	0.094	1.097	***	0.178
Census Tract 17	0.617	***	0.131	0.841	***	0.201
Census Tract 18	-0.154		0.119	0.090		0.200
Census Tract 19	-0.614	***	0.104	-0.897	***	0.147
Census Tract 20	0.970	***	0.124	1.211	***	0.217
Census Tract 21	-0.209	*	0.122	-1.041	***	0.224
Census Tract 22	-0.364	**	0.141	-0.728	***	0.233
Census Tract 23	-0.514	***	0.129	-0.431	**	0.198
Census Tract 24	0.159	**	0.077	0.118		0.115
Census Tract 25	0.010		0.085	-0.058		0.174
Census Tract 26	-1.299	***	0.078	-1.557	***	0.183
Census Tract 27	-0.576	***	0.108	-0.665	***	0.158
Census Tract 28	-0.592	***	0.113	-0.526	***	0.170
Census Tract 29	-0.285	***	0.102	-0.030		0.145
Census Tract 30	-0.216	***	0.080	-0.372	**	0.158
Census Tract 31	0.185	***	0.056	-0.116		0.092
Census Tract 32	-0.222	***	0.080	-0.204		0.133
Census Tract 33	-0.373	***	0.132	-0.275		0.214
Census Tract 34	-0.230	**	0.113	-0.092		0.182
Census Tract 35	-0.289	**	0.121	-0.570	***	0.182
Census Tract 36	-0.638	***	0.115	-0.629	***	0.192
Census Tract 37	-0.208	***	0.052	-0.198	**	0.096
Census Tract 38	-0.621	***	0.066	-0.694	***	0.118
Census Tract 39	-0.964	***	0.141	-0.393		0.260
Census Tract 40	-0.843	***	0.079	-0.462	***	0.148
Census Tract 41	-0.660	***	0.123	-0.062		0.248
Census Tract 42	-0.809	***	0.131	-1.293	***	0.235
Census Tract 43	-1.217	***	0.090	-0.995	***	0.142
Census Tract 44	-1.466	***	0.053	-1.289	***	0.099
Census Tract 45	-2.152	***	0.084	-2.035	***	0.138
Census Tract 46	-1.234	***	0.065	-1.409	***	0.103
Census Tract 47	-1.277	***	0.064	-1.714	***	0.089
Census Tract 48	1.463	***	0.129	1.421	***	0.224
Census Tract 49	1.180	***	0.135	1.210	***	0.231
Census Tract 50	-0.815	***	0.099	-0.303	*	0.177
Census Tract 51	-1.450	***	0.112	-1.401	***	0.199

Table 21 continued						
Census Tract 52	0.560	***	0.092	1.120	***	0.144
Census Tract 53	-0.043		0.097	0.093		0.154
Census Tract 54	-1.377	***	0.071	-1.418	***	0.120
Census Tract 55	-1.533	***	0.083	-1.212	***	0.144
Census Tract 56	-1.277	***	0.220	-0.966	***	0.333
Census Tract 57	-2.093	***	0.099	-2.358	***	0.178
Census Tract 58	-1.515	***	0.061	-1.607	***	0.109
Census Tract 59	0.149		0.092	-0.179		0.140
Census Tract 60	-0.311	***	0.112	-0.202		0.169
Census Tract 61	0.182		0.170	0.577	**	0.286
Census Tract 62	-1.096	***	0.137	-1.159	***	0.239
Census Tract 63	-2.056	***	0.128	-2.304	***	0.206
Census Tract 64	0.092		0.087	0.265	*	0.154
Census Tract 65	-0.074		0.108	-0.337	**	0.162
Census Tract 66	0.631	**	0.264	1.080	***	0.387
Census Tract 67	-1.574	***	0.207	-0.439		0.339
Census Tract 68	-0.655	***	0.148	-0.121		0.274
Census Tract 69	-1.253	***	0.158	-1.002	***	0.266
Census Tract 70	-0.551	***	0.103	-0.335	**	0.164
Census Tract 71	-0.263	***	0.079	0.034		0.126
Census Tract 72	-0.922	***	0.125	-0.277		0.196
Census Tract 73	-0.045		0.096	0.399	***	0.152
Census Tract 74	-1.142	***	0.132	-1.017	***	0.231
Census Tract 75	-0.937	***	0.148	-0.343		0.252
Census Tract 76	-0.995	***	0.172	-0.586	**	0.271
Census Tract 77	-0.613	***	0.139	-0.162		0.222
Census Tract 78	-1.234	***	0.131	-0.553	***	0.196
Census Tract 79	-0.184		0.168	0.341		0.254
Census Tract 80	-1.010	***	0.229	-0.345		0.349
Census Tract 81	0.152	**	0.073	0.236	**	0.114
Census Tract 82	-0.235	***	0.077	-0.119		0.137
Census Tract 83	-1.032	***	0.165	-0.583	**	0.257
Census Tract 84	-1.023	***	0.112	-0.837	***	0.183
Census Tract 85	-0.457	***	0.092	-0.418	***	0.153
Census Tract 86	-0.968	***	0.070	-1.038	***	0.100
Census Tract 87	-0.185		0.140	-0.098		0.219
Census Tract 88	0.834	***	0.103	0.975	***	0.163
Census Tract 89	-0.149		0.196	0.412		0.304
Census Tract 90	-0.264		0.171	0.209		0.263
Census Tract 91	-0.148		0.217	0.431		0.336
Census Tract 92	-1.599	***	0.088	-1.611	***	0.125
Census Tract 93	-1.787	***	0.083	-2.037	***	0.132
Census Tract 94	-1.462	***	0.148	-1.421	***	0.246
Census Tract 95	-0.492	***	0.152	-0.447	*	0.246
Census Tract 96	-0.565	***	0.160	-1.120	***	0.241
Census Tract 97	-0.476	***	0.151	-1.135	***	0.224
Census Tract 98	2.124	***	0.114	1.710	***	0.202
Census Tract 99	-0.471	***	0.121	-0.556	***	0.178
Census Tract 100	-1.201	***	0.084	-1.397	***	0.137
Census Tract 101	-0.309	***	0.061	0.030		0.103
Census Tract 102	-1.577	***	0.063	-1.758	***	0.090
Census Tract 103	-1.275	***	0.092	-0.975	***	0.159
Census Tract 104	-1.207	***	0.129	-0.774	***	0.238
Census Tract 105	0.330	**	0.161	0.294		0.247
Census Tract 106	-1.274	***	0.115	-1.126	***	0.177

Table 21 continued						
Census Tract 107	-0.077		0.137	0.451	**	0.210
Census Tract 108	-0.783	***	0.178	-0.286		0.274
Census Tract 109	0.949	***	0.156	1.282	***	0.229
Census Tract 110	-0.119		0.099	-0.576	***	0.158
Census Tract 111	-0.861	***	0.213	-0.582	*	0.310

*Significant at 90% **Significant at 95% ***Significant at 99%

REFERENCES

Alemi, F., Circella, G., Handy, S., & Mokhtarian, P. (2018). What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. *Travel Behaviour and Society*, *13*, 88-104.

Alshalalfah, B., & Shalaby, A. S. (2007). Case study: Relationship of walk access distance to transit with service, travel, and personal characteristics. *Journal of Urban Planning and Development*, *133*(2), 114-118. doi.org/10.1061/(ASCE)0733-9488(2007)133:2(114)

Arrington, G., & Cervero, R. (2008). *Effects of TOD on housing, parking, and travel*. (Transit Cooperative Research Program Report 128). Washington, DC: Transportation Research Board.

Atkinson-Palombo, C., & Kuby, M. J. (2011). The geography of advance transit-oriented development in metropolitan Phoenix, Arizona, 2000–2007. *Journal of Transport Geography*, *19*(2), 189-199. doi.org/10.1016/j.jtrangeo.2010.03.014

Austin, M., Belzer, D., Benedict, A., Esling, P., Haas, P., Miknaitis, G., Zimbabwe, S. (2010). Performance-based transit-oriented development typology guidebook. Retrieved from https://www.cnt.org/sites/default/files/publications/CNT_TODTypologyGuidebook.pdf

Babar, Y., & Burtch, G. (2020). Examining the heterogeneous impact of ride-hailing services on public transit use. *Information Systems Research*, *31*(3), 820-834. doi.org/10.1287/isre.2019.0917

Belzer, D., & Autler, G. (2002). *Transit oriented development: moving from rhetoric to reality* (pp. 06-15). Washington, DC: Brookings Institution Center on Urban and Metropolitan Policy.

Bento, A. M., Cropper, M. L., Mobarak, A. M., & Vinha, K. (2005). The effects of urban spatial structure on travel demand in the United States. *Review of Economics and Statistics*, *87*(3), 466-478. doi.org/10.1162/0034653054638292

Bernick, M., & Cervero, R. (1997). *Transit villages in the 21st century*. New York: McGraw-Hill.

Bertolini, L. (1999). Spatial development patterns and public transport: the application of an analytical model in the Netherlands. *Planning Practice and Research*, *14*(2), 199-210. doi.org/10.1080/02697459915724

Bhat, C. R. (2000). A multi-level cross-classified model for discrete response variables. *Transportation Research Part B: Methodological*, *34*(7), 567-582. doi.org/10.1016/S0191-2615(99)00038-7

Bhat, C. R., & Guo, J. Y. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research Part B: Methodological*, *41*(5), 506-526. doi.org/10.1016/j.trb.2005.12.005

Blumenberg, E., & Pierce, G. (2014). Multimodal travel and the poor: evidence from the 2009 National Household Travel Survey. *Transportation Letters*, *6*(1), 36-45. doi.org/10.1179/1942787513Y.0000000009

Boarnet, M. G., & Crane, R. (1998). Public finance and transit-oriented planning: New evidence from southern California. *Journal of Planning Education and Research*, *17*(3), 206-219. doi.org/10.1177/0739456X9801700302

Boarnet, M., & Crane, R. (2001). The influence of land use on travel behavior: Specification and estimation strategies. *Transportation Research Part A: Policy and Practice*, *35*(9), 823-845. doi.org/10.1016/S0965-8564(00)00019-7

Brazil, N., & Kirk, D. S. (2016). Uber and metropolitan traffic fatalities in the United States. *American journal of epidemiology*, *184*(3), 192-198. doi.org/10.1093/aje/kww062

Brodeur, A., & Nield, K. (2018). An empirical analysis of taxi, Lyft and Uber rides: Evidence from weather shocks in NYC. *Journal of Economic Behavior & Organization*, *152*, 1-16. doi.org/10.1016/j.jebo.2018.06.004

Buehler, R., & Hamre, A. (2015). The multimodal majority? Driving, walking, cycling, and public transportation use among American adults. *Transportation*, *42*(6), 1081-1101. doi.org/10.1007/s11116-014-9556-z

Calthorpe, P. (1993). *The next American metropolis: Ecology, community, and the American dream*. New York: Princeton Architectural Press.

Cao, X., Mokhtarian, P. L., & Handy, S. L. (2009). Examining the impacts of residential self-selection on travel behaviour: A focus on empirical findings. *Transport Reviews*, *29*(3), 359-395. doi.org/10.1080/01441640802539195

Carr, L. J., Dunsiger, S. I., & Marcus, B. H. (2011). Validation of Walk Score for estimating access to walkable amenities. *British journal of sports medicine*, *45*(14), 1144-1148. doi.org/10.1136/bjsm.2009.069609

Carroll, P., Caulfield, B., & Ahern, A. (2017). Examining the potential for car-shedding in the Greater Dublin Area. *Transportation Research Part A: Policy and Practice*, *106*, 440-452. doi.org/10.1016/j.tra.2017.10.019

Cervero, R. (1993). Ridership impacts of transit-focused development in California. Retrieved from https://cloudfront.escholarship.org/dist/prd/content/qt8sr9d86r/qt8sr9d86r.pdf

Cervero, R. (2004). *Transit-oriented development in the United States: Experiences, challenges, and prospects* (Vol. 102). Transportation Research Board.

Cervero, R., Murphy, S., Ferrell, C., Goguts, N. & Tsai, Y. (2004). *Transit-oriented development in the United States: Experiences, challenges, and prospects*. (Transit Cooperative Research Program Report 102). Washington, DC: Transportation Research Board.

Cervero, R., Ferrell, C., & Murphy, S. (2002). Transit-oriented development and joint development in the United States: A literature review. TCRP Research Results Digest No. 52. Retrieved from http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_rrd_52.pdf

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199-219. doi.org/10.1016/S1361-9209(97)00009-6

Chatman, D. (2006). *Transit-oriented development and household travel: A study of California cities*. Los Angeles: Institute of Transportation Studies, School of Public Affairs, University of California, Los Angeles.

Chatman, D. G. (2009). Residential choice, the built environment, and non-work travel: Evidence using new data and methods. *Environment and Planning A*, *41*(5), 1072-1089. doi.org/10.1068/a4114

Chatman, D. G. (2013). Does TOD need the T? On the importance of factors other than rail access. *Journal of the American Planning Association*, *79*(1), 17-31. doi.org/10.1080/01944363.2013.791008

Chatman, D. G., Cervero, R., Moylan, E., Carlton, I., Weissman, D., Zissman, J., ... & Tischler, D. (2014). *Making Effective Fixed-Guideway Transit Investments: Indicators of Success. Volume 1: Handbook*, *Volume 2* (No. Project H-42).

Circella, G., & Alemi, F. (2018). Transport policy in the era of ridehailing and other disruptive transportation technologies. In *Advances in Transport Policy and Planning*, *1*, 119-144. Academic Press. doi.org/10.1016/bs.atpp.2018.08.001

Clauss, T., & Döppe, S. (2016). Why do urban travelers select multimodal travel options: A repertory grid analysis. *Transportation Research Part A: Policy and Practice*, *93*, 93-116. doi.org/10.1016/j.tra.2016.08.021

Clewlow, R. R., & Mishra, G. S. (2017). Disruptive transportation: The adoption, utilization, and impacts of ride-hailing in the United States. Retrieved from https://escholarship.org/uc/item/82w2z91j

Cohen, A., & Shaheen, S. (2018). *Planning for shared mobility*. doi.org/10.7922/G2NV9GDD

Craun, K. J., & Chih-Hung, M. (2017). A comparison of volunteered geographic information (VGI) collected in rural areas to VGI collected in urban and suburban areas of the united states. *Citizen empowered mapping* (pp. 173-197). Springer, Cham.

D'Agostino Jr., R. B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, *17*(19), 2265-2281. doi.org/10.1002/(SICI)1097-0258(19981015)17:19<2265::AID-SIM918>3.0.CO;2-B

Deka, D., & Fei, D. (2019). A comparison of the personal and neighborhood characteristics associated with ridesourcing, transit use, and driving with NHTS data. *Journal of Transport Geography*, *76*, 24-33. doi.org/10.1016/j.jtrangeo.2019.03.001

Dolan, P., & Galizzi, M. M. (2015). Like ripples on a pond: Behavioral spillovers and their implications for research and policy. *Journal of Economic Psychology*, *47*, 1-16. doi.org/10.1016/j.joep.2014.12.003

Dubin, J. A., & McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica: Journal of the Econometric Society*, 345-362. doi.org/10.2307/1911493

Durand, C. P., Tang, X., Gabriel, K. P., Sener, I. N., Oluyomi, A. O., Knell, G., Kohl III, H. W. (2016). The association of trip distance with walking to reach public transit: Data from the California household travel survey. *Journal of Transport & Health*, *3(*2), 154-160. doi.org/10.1016/j.jth.2015.08.007

Evans, I., Pratt, R. H., Stryker, A., & Kuzmyak, J. R. (2007). *Transit-oriented development—Traveler response to transportation system changes*. (Transit Cooperative Research Program Report 95). Washington, DC: Transportation Research Board.

Ewing, R., & Cervero, R. (2010). Travel and the built environment: A meta-analysis. *Journal of the American Planning Association*, *76*(3), 265-294. doi.org/10.1080/01944361003766766

Ewing, R., & Hamidi, S. (2014). Longitudinal analysis of transit's land use multiplier in Portland (OR). *Journal of the American Planning Association*, *80*(2), 123-137. doi.org/10.1080/01944363.2014.949506

Frank, L., Bradley, M., Kavage, S., Chapman, J., & Lawton, T. K. (2008). Urban form, travel time, and cost relationships with tour complexity and mode choice. *Transportation*, *35*(1), 37-54. doi.org/10.1007/s11116-007-9136-6

FTA. (2019). Federal Transit Administration announces $19.2 million for transit planning in communities nationwide. Retrieved from https://www.transit.dot.gov/about/news/federal-transit-administration-announces-192-million-transit-planning-communities

Gallivan, F., Rose, E., Ewing, R., Hamidi, S., & Brown, T. (2015). *Quantifying transit's impact on GHG emissions and energy use—The land use component*. (Transit Cooperative Research Program Report 176). Washington, DC: Transportation Research Board.

Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health Services Research*, *49*(5), 1701-1720. doi.org/10.1111/1475-6773.12182

Gehrke, S. R., Felix, A., & Reardon, T. G. (2019). Substitution of ridehailing services for more sustainable travel options in the greater Boston region. Transportation Research Record: *Journal of the Transportation Research Board*, *2673*(1), 438-446. doi.org/10.1177/0361198118821903

Goetzke, F., & Andrade, P. M. (2010). Walkability as a summary measure in a spatially autoregressive mode choice model: An instrumental variable approach. *Progress in Spatial Analysis* (pp. 217-229): Heidelberg: Springer.

Graehler, M., Mucci, R. A., & Erhardt, G. D. (2019). Understanding the recent transit ridership decline in major US cities: Service cuts or emerging modes. *In 98th Annual Meeting of the Transportation Research Board*, Washington, DC.

Greenwald, M. J., & Boarnet, M. G. (2001). Built environment as determinant of walking behavior: Analyzing non-work pedestrian travel in Portland, Oregon. *Transportation research record, 1780*(1), 33-41. doi.org/10.3141/1780-05

Haas, P., Miknaitis, G., Cooper, H., Young, L., & Benedict, A. (2010). Transit oriented development and the potential for VMT-related greenhouse gas emissions growth reduction. Retrieved from https://community-wealth.org/sites/clone.community-wealth.org/files/downloads/paper-haas-et-al.pdf

Hale, C. (2014). TOD versus TAD: The great debate resolved…(?). *Planning Practice & Research*, *29*(5), 492-507. doi.org/10.1080/02697459.2012.749056

Hall, J. D., Palsson, C., & Price, J. (2018). Is Uber a substitute or complement for public transit?. *Journal of Urban Economics*, *108*, 36-50. doi.org/10.1016/j.jue.2018.09.003

Hawkins, J., & Habib, K. N. (2020). Finding the Subway Disruption Regimes of Switching Subway to Uber in Toronto. *Transportation Research Record*, *2674*(12), 303-311. doi.org/10.1177/0361198120953797

Heinen, E., & Chatterjee, K. (2015). The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey. *Transportation Research Part A: Policy and Practice*, 78, 266-282. doi.org/10.1016/j.tra.2015.05.015

Heinen, E., & Mattioli, G. (2019). Does a high level of multimodality mean less car use? An exploration of multimodality trends in England. *Transportation*, *46*(4), 1093-1126. doi.org/10.1007/s11116-017-9810-2

Higgins, C. D., & Kanaroglou, P. S. (2016). A latent class method for classifying and evaluating the performance of station area transit-oriented development in the Toronto region. *Journal of Transport Geography*, *52*, 61-72. doi.org/10.1016/j.jtrangeo.2016.02.012

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199-236. doi.org/10.1093/pan/mpl013

Kamruzzaman, M., Baker, D., Washington, S., & Turrell, G. (2014). Advance transit oriented development typology: Case study in Brisbane, Australia. *Journal of Transport Geography*, *34*, 54-70. doi.org/10.1016/j.jtrangeo.2013.11.002

Kong, H., Zhang, X., & Zhao, J. (2020). How does ridesourcing substitute for public transit? A geospatial perspective in Chengdu, China. *Journal of Transport Geography*, *86*, 102769. doi.org/10.1016/j.jtrangeo.2020.102769

Kuhnimhof, T., Buehler, R., Wirtz, M., & Kalinowska, D. (2012). Travel trends among young adults in Germany: increasing multimodality and declining car use for men. *Journal of Transport Geography*, *24*, 443-450. doi.org/10.1016/j.jtrangeo.2012.04.018

Kwoka, G. J., Boschmann, E. E., & Goetz, A. R. (2015). The impact of transit station areas on the travel behaviors of workers in Denver, Colorado. *Transportation Research Part A: Policy and Practice*, *80*, 277-287. doi.org/10.1016/j.tra.2015.08.004

Laham, M. L., & Noland, R. B. (2017). Non-work trips associated with transit-oriented development. *Transportation Research Record*, *2606*(1), 46-53. doi.org/10.3141/2606-06

Langlois, M., van Lierop, D., Wasfi, R. A., & El-Geneidy, A. M. (2015). Chasing sustainability: Do new transit-oriented development residents adopt more sustainable modes of transportation? *Transportation Research Record*, *2531*(1), 83-92. doi.org/10.3141/2531-10

Lanzini, P., & Thøgersen, J. (2014). Behavioural spillover in the environmental domain: An intervention study. *Journal of Environmental Psychology*, *40*, 381-390. doi.org/10.1016/j.jenvp.2014.09.006

Lanzini, P., & Khan, S. A. (2017). Shedding light on the psychological and behavioral determinants of travel mode choice: A meta-analysis. *Transportation Research Part F: Traffic Psychology and Behaviour*, *48*, 13-27. doi.org/10.1016/j.trf.2017.04.020

Loa, P., Mashrur, S. M., & Nurul Habib, K. (2021). What influences the substitution of ride-sourcing for public transit and taxi services in Toronto? An exploratory structural equation model-based study. *International Journal of Sustainable Transportation*, 1-14. doi.org/10.1080/15568318.2021.1978018

Loutzenheiser, D. R. (1997). Pedestrian access to transit: Model of walk trips and their design and urban form determinants around Bay area rapid transit stations. *Transportation Research Record*, *1604*(1), 40-49. doi.org/10.3141/1604-06

Manaugh, K., & El-Geneidy, A. (2011). Validating walkability indices: How do different households respond to the walkability of their neighborhood?. *Transportation research part D: transport and environment*, *16*(4), 309-315. doi.org/10.1016/j.trd.2011.01.009

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531-542. doi.org/10.2307/2298123

Manski, C. F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives*, *14*(3), 115-136. doi.org/10.3386/w7580

Molin, E., Mokhtarian, P., & Kroesen, M. (2016). Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers. *Transportation Research Part A: Policy and Practice*, *83*, 14-29. doi.org/10.1016/j.tra.2015.11.001

Moran, M. R., Rodríguez, D. A., & Corburn, J. (2018). Examining the role of trip destination and neighborhood attributes in shaping environmental influences on children's route choice. *Transportation Research Part D: Transport and Environment*, *65*, 63-81. doi.org/10.1016/j.trd.2018.08.001

Mostofi, H., Masoumi, H., & Dienel, H. L. (2020). The Relationship between Regular Use of Ridesourcing and Frequency of Public Transport Use in the MENA Region (Tehran and Cairo). *Sustainability*, *12*(19), 8134. doi.org/10.3390/su12198134

Murphy, C. M., & Felgon, S. F. (2016). *Shared Mobility and the Transformation of Public Transit*. (Transit Cooperative Research Program Report 188). Washington, DC: Transportation Research Board. doi.org/10.17226/23578

Nasri, A., Carrion, C., Zhang, L., & Baghaei, B. (2018). Using propensity score matching technique to address self-selection in transit-oriented development (TOD) areas. *Transportation*, *47*(1), 359-371. doi.org/10.1007/s11116-018-9887-2

Nasri, A., & Zhang, L. (2014). The analysis of transit-oriented development (TOD) in Washington, DC and Baltimore metropolitan areas. *Transport Policy*, *32*, 172-179. doi.org/10.1016/j.tranpol.2013.12.009

Nobis, C. (2007). Multimodality: facets and causes of sustainable mobility behavior. *Transportation Research Record: Journal of the Transportation Research Board*, (2010), 35-44. doi.org/10.3141/2010-05

Oviedo, D., Granada, I., & Perez-Jaramillo, D. (2020). Ridesourcing and travel demand: Potential effects of transportation network companies in Bogotá. *Sustainability*, *12*(5), 1732. doi.org/10.3390/su12051732

Park, K., Ewing, R., Scheer, B. C., & Tian, G. (2018). The impacts of built environment characteristics of rail station areas on household travel behavior. *Cities*, *74*, 277-283. doi.org/10.1016/j.cities.2017.12.015

Parker, T. (2002). Statewide transit-oriented development study: Factors for success in California. Retrieved from http://www.dot.ca.gov/hq/MassTrans/Docs-Pdfs/TOD-Study-Final-Rpt.pdf

Pushkarev, B., & Zupan, J. M. (1977). *Public transportation and land use policy*. Bloomington: Indiana University Press.

Ralph, K., Voulgaris, C. T., Taylor, B. D., Blumenberg, E., & Brown, A. E. (2016). Millennials, built form, and travel insights from a nationwide typology of U.S. neighborhoods. *Journal of Transport Geography*, *57*, 218-226. doi.org/10.1016/j.jtrangeo.2016.10.007

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Newbury Park, CA: Sage.

Rayle, L., Dai, D., Chan, N., Cervero, R., & Shaheen, S. (2016). Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transport Policy*, *45*, 168-178. doi.org/10.1016/j.tranpol.2015.10.004

Reusser, D. E., Loukopoulos, P., Stauffacher, M., & Scholz, R. W. (2008). Classifying railway stations for sustainable transitions–balancing node and place functions. *Journal of Transport Geography*, *16*(3), 191-202. doi.org/10.1016/j.jtrangeo.2007.05.004

Renne, J. L., Hamidi, S., & Ewing, R. (2016). Transit commuting, the network accessibility effect, and the built environment in station areas across the United States. *Research in Transportation Economics*, *60*, 35-43. doi.org/10.1016/j.retrec.2017.02.003

Rodríguez, D. A., & Joo, J. (2004). The relationship between non-motorized mode choice and the local physical environment. *Transportation Research Part D: Transport and Environment*, *9*(2), 151-173. doi.org/10.1016/j.trd.2003.11.001

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55. doi.org/10.1093/biomet/70.1.41

Rubin, D. B. (2001). Using propensity scores to help design observational studies: Application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, *2*, 169-188. doi.org/10.1023/A:1020363010465

Salon, D. (2015). Heterogeneity in the relationship between the built environment and driving: Focus on neighborhood type and travel purpose. *Research in Transportation Economics*, *52*, 34-45. doi.org/10.1016/j.retrec.2015.10.008

Scheiner, J., Chatterjee, K., & Heinen, E. (2016). Key events and multimodality: A life course approach. *Transportation Research Part A: Policy and Practice*, *91*, 148-165. doi.org/10.1016/j.tra.2016.06.028

Schwanen, T., & Mokhtarian, P. L. (2005). What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography*, *13*(1), 83-99. doi.org/10.1016/j.jtrangeo.2004.11.001

Sidharthan, R., Bhat, C. R., Pendyala, R. M., & Goulias, K. G. (2011). Model for children's school travel mode choice: Accounting for effects of spatial and social interaction. *Transportation Research Record*, *2213*(1), 78-86. doi.org/10.3141/2213-11

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science*, *25*(1), 1-21. doi.org/10.1214/09-STS313

Suzuki, H., Cervero, R., & Iuchi, K. (2013). *Transforming cities with transit: Transit and land-use integration for sustainable urban development*. Washington, DC: The World Bank.

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioral Research*, *46*(1), 90-118. doi.org/10.1080/00273171.2011.540475

Uber (Feb 29, 2020). Uber is now in over 10,000 cities globally. Retrieved from https://www.uber.com/newsroom/10000-cities/

U.S. Department of Transportation (2017). Summary of travel trends: 2017 National Household Travel Survey. Retrieved from https://nhts.ornl.gov/assets/2017_nhts_summary_travel_trends.pdf

Vale, D. S., & Pereira, M. (2016). Influence on pedestrian commuting behavior of the built environment surrounding destinations: A structural equations modeling approach. *International Journal of Sustainable Transportation*, *10*(8), 730-741. doi.org/10.1080/15568318.2016.1144836

Vij, A., Carrel, A., & Walker, J. L. (2013). Incorporating the influence of latent modal preferences on travel mode choice behavior. *Transportation Research Part A: Policy and Practice*, *54*, 164-178. doi.org/10.1016/j.tra.2013.07.008

Young, M., & Farber, S. (2019). The who, why, and when of Uber and other ride-hailing trips: An examination of a large sample household travel survey. *Transportation Research Part A: Policy and Practice*, *119*, 383-392. doi.org/10.1016/j.tra.2018.11.018

Young, H. P. (1996). The economics of convention. *Journal of Economic Perspectives*, *10*(2), 105-122. doi.org/10.1257/jep.10.2.105

Zemp, S., Stauffacher, M., Lang, D. J., & Scholz, R. W. (2011). Classifying railway stations for strategic transport and land use planning: Context matters! *Journal of transport geography*, *19*(4), 670-679.

Zhang, Y., & Zhang, Y. (2018). Exploring the relationship between ridesharing and public transit use in the United States. *International journal of environmental research and public health*, *15*(8), 1763. doi.org/10.3390/ijerph15081763

Zielstra, D., Hochmair, H. H., & Neis, P. (2013). Assessing the Effect of Data Imports on the Completeness of OpenStreetMap–A United States Case Study. *Transactions in GIS*, *17*(3), 315-334. doi.org/10.1111/tgis.12037