

July 2019

EQUITY AND EFFICIENCY IN MULTI-MODAL TRANSPORTATION SYSTEMS

Nicholas M. Fournier
University of Massachusetts - Amherst

Follow this and additional works at: https://scholarworks.umass.edu/dissertations_2



Part of the [Transportation Engineering Commons](#)

Recommended Citation

Fournier, Nicholas M., "EQUITY AND EFFICIENCY IN MULTI-MODAL TRANSPORTATION SYSTEMS" (2019).
Doctoral Dissertations. 1570.
https://scholarworks.umass.edu/dissertations_2/1570

This Open Access Dissertation is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

**EQUITY AND EFFICIENCY
IN MULTI-MODAL TRANSPORTATION SYSTEMS**

A Dissertation Presented

by

NICHOLAS MARC FOURNIER

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2019

Civil & Environmental Engineering

© Copyright by Nicholas Marc Fournier 2019

All Rights Reserved

EQUITY AND EFFICIENCY IN MULTI-MODAL TRANSPORTATION SYSTEMS

A Dissertation Presented

by

NICHOLAS MARC FOURNIER

Approved as to style and content by:

Eleni Christofa, Ph.D., Chair

Eric Gonzales, Ph.D., Member

Song Gao, Ph.D., Member

Hari Balasubramanian, Ph.D., Member

Richard Palmer, Ph.D., Department Head
Civil & Environmental Engineering

DEDICATION

To girl from the city by The Bay. You not only joined me on this adventure, but encouraged it from the beginning. None of this would be possible without you, your infinite patience, support, and understanding while I worked on “our” doctorate.

To this adventure, and many more.

ACKNOWLEDGMENTS

I would like to thank my advisor Eleni Christofa, who plucked me from academic obscurity and not only taught me to conduct research, but to *love* it. She had complete faith in my abilities and always supportive, never once undercutting my ideas, but instead helped formulate my ideas into research concepts. Her dedication to my success is unparalleled and I am truly indebted to her, I hope to match this dedication in my own career, so that perhaps someday pass it onto someone else.

I would like to thank the committee for giving me their time and expertise. Specifically, Eric Gonzales for his uncanny mathematical mind in helping me construct, formulate, and solve what eventually became my dissertation. I am grateful to Song Gao and her wealth of knowledge and expertise on the subject of behavioral modeling. I am appreciative of Hari Balasubramanian for helping me develop the non-negative deviation method for population synthesis in his linear programming course. In addition to the committee, I am grateful to Michael Knodler, who advised me in the very beginning of my academic career and from whom I am still learning.

I would also like to give credit to the planners, with their bleeding hearts for urbanity and humanity that I share. Their fundamentally different thoughts and ideas on societal systems fostered my own deviation from prescribed methodologies. This interdisciplinary experience has become one of my greatest assets both professionally and personally.

This research was supported by the Dwight D. Eisenhower Graduate Fellowship Program and the United States Department of Energy's (US DOE) Advanced Research Projects Agency – Energy (ARPA-e). This research was carried out in part

under the US DOE's ARPA-e "Traveler Response Architecture using Novel Signaling for Network Efficiency in Transportation" (TRANSNET) Program, Award Number DE-AR0000611. I would also like to thank the University of Massachusetts Amherst's Transportation Center (UMTC) for their support.

ABSTRACT

EQUITY AND EFFICIENCY IN MULTI-MODAL TRANSPORTATION SYSTEMS

MAY 2019

NICHOLAS MARC FOURNIER

B.S., M.S.C.E., M.R.P., UNIVERSITY OF MASSACHUSETTS AMHERST

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Eleni Christofa, Ph.D.

The land-use pattern for many cities is a central business district surrounded by sprawling suburbs. This pattern can lead to an inefficient and congestion-prone transportation system due to a reliance on automobiles, because high-capacity transit is not efficient in low-density areas where insufficient travelers can access transit. This also poses an equity concern as the monetary cost of faster and more expensive travel disproportionately burdens low income travelers. This dissertation presents a deterministic approximation of a discrete choice model for mixed access and mainline transportation modes, meaning that travelers may use different modes to access a mainline system, such as transit. The purpose is to provide a tractable computationally efficient model to address the first/last mile problem using a system-wide pricing policy that can account for heterogeneous values of time; a problem that is difficult to solve efficiently using a stochastic model. The model is structured for a catchment area around a central access point for a mainline mode, approximating choice

by comparing modal utility costs. The underlying utility model accommodates both fixed prices (e.g., parking, fixed tolls, and fares) and distance-based unit prices (e.g. taxi fare, bike-share, and distance tolls) that may be set in a coordinated way with respect to value of time. Using numerical analysis, the deterministic model achieved results within 4% accuracy of a stochastic logit-based model, and within 6% of measured values. The final model achieved a 57% reduction in generalized travel time and improved the Gini inequity measure from 0.21 to 0.03.

CONTENTS

	Page
ACKNOWLEDGMENTS	v
ABSTRACT	vii
LIST OF TABLES	xi
LIST OF FIGURES	xii
NOMENCLATURE	xiv
 CHAPTER	
1. INTRODUCTION	1
1.1 Motivation	1
1.2 Research Question	3
1.3 Research Contribution	4
1.4 Dissertation Organization	6
2. LITERATURE REVIEW	8
2.1 Transit access, density, and the urban form	10
2.2 System models	12
2.3 Discrete choice	13
2.4 Comparison of system models and discrete choice models	18
2.5 Value of time	19
2.6 Population Synthesis	21
2.7 Summary of Literature	31
3. POPULATION & WORKPLACE SYNTHESIS	34
3.1 Simultaneous fitting of population and workplace	35
3.2 Joint re-weighting	38
3.3 Joint sampling	42
3.4 Data	43

3.5	Evaluation	46
3.6	Summary of findings	54
4.	MULTI-MODAL SYSTEM MODEL	57
4.1	Concept	57
4.2	Mathematical Model	60
4.3	Conditional choice of mainline mode	63
4.4	Congestion	64
4.5	Optimal Mode Split	67
4.6	Pricing	68
4.7	Income, Equity, and the Value of Time	72
4.8	Data	76
4.9	Evaluation	77
4.10	Summary of findings	93
5.	CONCLUSIONS	95
5.1	Dissertation Contributions	95
5.2	Summary of findings	96
5.3	Future work	98
	BIBLIOGRAPHY	102

LIST OF TABLES

Table	Page
2.1 Example IPU joint list	27
3.1 Example linear program joint list	40
3.2 Data used in population synthesis	43
3.3 Population synthesis variables	44
3.4 Computation time comparison of re-weighting methods	50
4.1 Utility model parameters	58
4.2 Worcester–Boston commuter model parameters	77
4.3 Optimal and observed radii	83
4.4 Optimal pricing	85
4.5 Model demand estimation comparison	85

LIST OF FIGURES

Figure	Page
2.1 Red/blue bus nested logit	16
3.1 Modeling framework	35
3.2 Origin-destination-industry conceptualization	36
3.3 Comparison of fit by method (1:1 scale)	48
3.4 Comparison of fit by method (5000:1 scale)	49
3.5 Map of marginal fit for census tracts	51
3.6 Population generation results	52
3.7 Workplace allocation by origin-destination pair	53
4.1 Illustration of Model Concept	57
4.2 Spatial approximation of radii boundaries for access mode choice	59
4.3 Access mode choice model comparison	60
4.4 Nested mode choice	64
4.5 Relationship between flow, density, and travel time	65
4.6 Worcester residents employment density	76
4.7 Value of time and income in MTS	78
4.8 Value of time distribution in PUMS	79
4.9 Comparative fit between deterministic and nested logit models	81
4.10 Effects of value of time on model accuracy	82

4.11	Optimal demand allocation	83
4.12	Mainline travel time with congestion	84
4.13	Optimal logit and dis-utility comparison	87
4.14	Optimal parameters by VOT	88
4.15	Optimal pricing varying by VOT	90
4.16	Comparison between generalized monetary and temporal user costs	92
5.1	Possible model applications	100

NOMENCLATURE

Acronyms

ACS	American Community Survey
AET	All Electronic Tolling
BN	Bayesian Network
CA	Continuum Approximation
CBD	Consolidated Business District
CO	Combinatorial Optimization
CTPS	Central Transportation Planning Staff
DC	Discrete Choice
DI	Destination-Industry
GBA	Greater Boston Area
GIS	Geographic Information System
HMM	Hidden Markov Models
IIA	Independent Irrelevant Alternatives
IPF	Iterative Proportional Fitting
IPU	Iterative Proportional Updating
LAD	Least Absolute Deviation
LEHD	Longitudinal Employer-Household Dynamics
LODES	LEHD Origin-Destination Employment Statistics
MCMC	Monte-Carlo Markov Chain
MFD	Macroscopic Fundamental Diagram
MPO	Metropolitan Planning Organization

MTS Massachusetts Household Travel Survey
NLAD Non-negative Least Absolute Deviation
NNLS Non-negative Least Squares
OD Origin-Destination
ODI Origin-Destination-Industry
OI Origin-Industry
OLD Ordinary Least Squares
PUMA Public Use Microdata Area
PUMS Public Use Microdata
RMSE Root Mean Square Error
RMSN Root Mean Square Normalized Error
TNC Transportation Network Company
VMT Vehicle Miles Travelled
VOT Value of Time

Symbols

α	Congestion model parameter	<i>unitless</i>
β	Scaling parameter	<i>unitless</i>
δ	Population density	pa_x/km^2
γ	Congestion model parameter	<i>unitless</i>
\hat{r}	Optimal radius	<i>km</i>
λ	Demand	pa_x
\bar{r}_a	Average access mode radial distance	<i>km</i>
\bar{v}_a	Average access mode speed	km/hr
ϕ	Congestion model parameter	<i>unitless</i>
π	Mathematical Constant	3.141592653589...
θ	Ratio of highway users to all drivers	<i>unitless</i>
ξ	Value of time	$\$/hr$
C_a^d	Monetary access distance cost	$\$$
C_a^f	Monetary access fixed costs	$\$$
C_m	Monetary fixed mainline cost	$\$$
G	Gini Coefficient	<i>unitless</i>
GT	Generalized Travel Time	<i>hours or minutes</i>
IV	Logsum coefficient	<i>unitless</i>
L	Mainline mode distance	<i>km</i>
R	Maximum access radius	<i>km</i>
r^*	Target tipping point radius	<i>km</i>
r°	Null tipping point radius	<i>km</i>
r'	Effective existing radius	<i>km</i>
r_a	Radial distance for access mode a	<i>km</i>
T_a	Access travel time	<i>hours or minutes</i>
t_a	Fixed access time	<i>hours or minutes</i>

T_m	Mainline travel time	<i>hours or minutes</i>
t_m	Fixed mainline time	<i>hours or minutes</i>
TT	Total Travel Time	<i>hours or minutes</i>
U	Total utility	<i>unitless</i>
v_a	Access mode speed	<i>km/hr</i>
v_m	Mainline mode speed	<i>km/hr</i>
Y	Access mode utility	<i>unitless</i>
Z	Mainline mode utility	<i>unitless</i>

CHAPTER 1

INTRODUCTION

1.1 Motivation

The dream for many is simply a beautiful home and a stress free commute. However, with populations in the United States and around the world continuing to grow and becoming increasingly reliant on urban centers, this dream is transitioning into a nightmare of congested roads, polluted air, and swaths of wasteful parking. With the transportation sector accounting for nearly one-third of global greenhouse gas emissions, the need for more efficient transportation is not only a social concern, but rapidly becoming an environmental imperative. One that must not only address high density urban centers, but the surrounding lower density regions which largely contribute to automobile congestion. Traditional solutions of large scale infrastructure investment are often impossible or impractical due to financial, spatial, or political constraints. Moreover, many existing fixed route transit systems in the United States are underutilized due to low demand, resulting in funding cuts that further undermine ridership. A more holistic solution is one that not only accounts for multiple modes, but align the parallel perspectives of modal demand and optimal system utilization.

For many developed nations, the era of fast paced construction of transportation infrastructure to keep up with demand has largely ended. Yet many cities remain plagued with congestion caused by inefficient utilization of infrastructure, compounded by sprawling suburban regions surrounding central business districts (CBD). This creates a persistent challenge in how to serve a concentrated demand from a spatially distributed source using efficient high-capacity transit. The recent in-

roduction of new forms of transportation under the “mobility as a service” umbrella, namely e-hailing services, transportation network companies (TNCs), and bike-share systems, are becoming a popular choice among travelers and have brought additional complexities in how multi-modal networks operate and travelers behave. While such services provide additional alternatives to users and are said to have increased the “car-free by choice” population, they still largely contribute to unsustainable congestion in urban multi-modal networks. For example, average traffic speeds in midtown Manhattan, NY have slowed since the advent of TNCs due to the potential for an inefficient abundance of empty e-hailing taxis circling for passengers (Schaller, 2017). Furthermore, it is still unclear whether e-hailing services complement or compete with transit services.

In low population density areas, the number of travelers within walking distance of transit often does not support fixed route transit. Colloquially called the “first/last mile” problem, it is the challenge in getting travelers to and from a transit access point. Although travelers may access transit by automobile, either by driving to a park-n-ride or using an e-hailing service operated by a transportation network company (TNC), there is little incentive to transfer to transit when completing the whole trip by automobile is, or is perceived as, more convenient. This often results in overly congested roads and underutilized transit capacity (Schaller, 2017; Gan and Ye, 2018). Alternatively, bicycling is often cited as an excellent alternative access mode serving populations between walking and driving. However, bicycling has its own set of social and infrastructural challenges limiting it, such as physical impedance, discomfort, safety, and social stigmas (Boarnet et al., 2017).

The complications in modeling multi-modal systems come from the need to realistically represent both the supply and demand, and to do so at the scales of the local neighborhood and broader urban region. Travel demand depends on the behavioral characteristics of the population that are difficult to model or predict due to lack of

data and the uncertainty associated with behavior and choice. The demand patterns, in turn, impact available supply. Thus, mode choice should be considered in the context of equilibrium models that represent the inter-related nature of transportation supply and demand.

Setting policies to influence demand while considering supply is not only a political challenge, but also a modeling challenge. Realistic stochastic demand models introduce mathematical hindrances in finding an optimal solution for the system. Conversely, generalized deterministic system models can more easily find a system optimum, but lack user heterogeneity (e.g. assume uniform travelers with a single value of time (VOT)). Although the lack of heterogeneity may be justified by using a mean value, the resulting pricing policies may exacerbate, rather than mitigate, existing inequities in the population. It is essential that equity is accounted for when developing policy to ensure that system improvements do not come at the cost of equity.

1.2 Research Question

The task of shifting low-density populations to more efficient and sustainable modes, such as bicycling, transit, or ride-hailing services, is extremely difficult. Transit can be efficient, but inherently relies on high population density within walking distance to support operation. Bicycles provide excellent cost and environmental benefits, but are impractical for most long-distance commutes. Emerging on-demand ride hailing services offer great potential for eliminating the need for parking, but are currently too costly to serve as a daily commute mode and still absorb the same road space as single occupant vehicles. Although each mode alone is impractical, there are many cases where modes can be integrated in new ways to provide a feasible optimal solution. This elicits the question:

*Can an **access mode congestion pricing model** improve system performance in terms of **time, cost, and equity**?*

1.3 Research Contribution

This dissertation presents a hybrid mode choice system model using a deterministic model for mixed access (e.g., walk, bike, and drive) and mainline modes (e.g. highway, in other words drive on the highway, and rail). The proposed model is based on comparing modal choice utility over a continuous catchment area around an access point. Unlike a stochastic based model, the deterministic global optimum can be found quickly, providing a system-wide optimal pricing policy to be set by the operator. This pricing policy incentivizes mode-shift to achieve a certain objective, such as minimum average travel time. Furthermore, the pricing inputs are provided as both fixed and distance based prices independently for each mode, providing an opportunity to integrate Transportation Network Companies (TNC) (e.g., Uber and Lyft), bike-share systems, and distance based tolling.

The purpose is to provide a tractable and computationally lightweight model to address the first/last mile problem using a system-wide pricing policy that can account for heterogeneous values of time; a problem that is difficult to solve efficiently using a stochastic model. The computationally simple deterministic model makes it possible to more easily explore outcomes. Unlike most system models, which use a single value of time for an entire population, this model accounts for a varying value of time (VOT), optimizing pricing for the entire system across multiple value of time groups. This accounting for all value of time groups is essential as congestion is experienced collectively, thus optimizing demand must account for all groups. Furthermore, accounting for heterogeneous populations is not just beneficial to model accuracy, but enables the results to be evaluated for equity implications resulting from policy decisions.

To apply the model, a test case of Worcester to Boston, Massachusetts is used. Although the pricing optimization itself is performed using a simple VOT distribution and aggregated values (e.g., commuter population density), a synthetic population

provides a more realistic joint distribution of study variables. In particular, the precise number of residents commuting to Boston from the Worcester region, their existing mode choices, and their household income which becomes the basis for VOT. This synthetic population is generated for the entire population of 106,166 households with 282,142 total individual persons, each with workplace destination assigned using industry sector as the linking distribution. The population synthesis makes two contributions: First is workplace assignment is integrated into the population synthesis process itself to generate a more accurate result. Second is a much more efficient person-household re-weighting algorithm based on non-negative least deviation fitting (NLAD). This efficient algorithm enables many more population features to be included in the synthesis, specifically the fourteen additional industry sectors used to generate the accurate workplace assignment.

In summary, the dissertation makes the following contributions:

1. Population Synthesis

- (a) **Integrated population synthesis and fixed work-place assignment**
- (b) **Faster non-negative least-deviation (NLAD) method** for fitting joint population of persons and households

2. Mixed-access Model

- (a) A congestion pricing model with **multiple access modes**
- (b) Allows for **value of time to vary across both income and mode**
- (c) Model framework allows **flexible objective function** (e.g., optimize equity directly)
- (d) Measured impacts of transport **system efficiency gains on user equity**
- (e) **Computationally lightweight**, meaning it requires little computational resources and quickly reaches a solution

1.4 Dissertation Organization

Within this section, an abstract discussion of the proposed methodology is described in order to provide a skeletal framework for subsequent chapters. Chapter 2 is an in-depth literature review and discussion of previous research efforts and modeling approaches to understand their contributions and gaps. Within the literature review, there is first a broad discussion of urban planning and transit access in Section 2.1. Then the literature review shifts to a more modelling focus with a review of system models in Section 2.2 and a detailed review of discrete choice models in Section 2.3 to provide an understanding of behavioral modelling. A comparison of system models and discrete choice model is discussed in Section 2.4. Then in Section 2.5, previous theoretical and applied studies on value of time and its estimation methods is discussed. Lastly, a thorough discussion on population synthesis methods is provided in Section 2.6 before a final summary that identifies gaps in Section 2.7.

Chapter 3 describes the population synthesis process used in this dissertation, the data required, and the resulting population in six sections. Section 3.1 describes the process where population synthesis and workplace assignment is integrated using Iterative Proportional Fitting. Section 3.2 discusses the joint re-weighting of person-household joint populations and the proposed non-negative least absolute deviation (NLAD) fitting method used to improve this process. Section 3.3 then describes the sampling method used to generate the final realized population. After the population synthesis methodology, Section 3.4 describes the data used to generate the population. Finally, Section 3.5 presents and evaluates the resulting synthetic population generated, which is summarized in Section 3.6.

Chapter 4 first introduces the multi-modal model for a simple single VOT case. The proposed multi-modal methodology is described in seven subsections: The basic concept of the model as it relates to access mode and discrete choice utility is presented in Section 4.1. Then in Section 4.2 the model is formulated to minimize travel time,

given the necessary constraints to ensure the model is mathematically feasible. Next the mainline mode choice is introduced as a spatial interpretation of conditional choice depending upon access mode in Section 4.3. Building upon the mainline mode, congestion is introduced into the model in Section 4.4 which is then compiled into a final continuous objective function in Section 4.5. This objective function provides the optimal allocation of demand to achieve minimum travel time, which is then realized through an optimal pricing policy determined in Section 4.6. This pricing is presented separately for access mode and mainline mode in Sections 4.6.1 and 4.6.2. A brief discussion on pricing implementation is presented in Section 4.6.3. Access mode pricing yields the deficit cost in monetary prices needed to adjust the observed demand to become optimal demand. The mainline mode pricing is a decomposition of a nested logit model, which extracts the pricing necessary to achieve the desired optimal mode split given the access mode split determined by the optimization in Section 4.5. The model is then extended in Section 4.7 to account for a varying VOT distribution as well as a means of measuring the model's impacts on equity using a Gini coefficient as a measure of equity.

Chapter 5 concludes the dissertation by highlighting key contributions in Section 5.1, followed by a summary of findings presented in Section 5.2. Section 5.2 summarizes numerical findings, but also highlights key research findings and contributions. Section 5.3 then presents possible future research work and applications of the multi-modal system model and integrated population synthesis.

CHAPTER 2

LITERATURE REVIEW

Transportation modeling is the attempt to create an abstract representation of a real system or situation in the realm of transportation. A larger transport “model” typically incorporates several smaller models to address the various aspects of human transport. In the most general sense, transportation modeling approaches tend to fall into one of two categories: *system models* and *data-driven models*. These two schools of thought can be defined in parallel to Kant’s 1781 definition of *a priori* and *a posteriori* knowledge. *A priori* models are classical mathematical formulations for entire systems. Their strength comes from their elegant simplicity and mathematical form, but often lack empirical validation (Daganzo, 2010). Their counterpart, *a posteriori* models are statistically fit directly from empirical data. Their strength comes from truth in data, but in an age of data-driven decision making, this *laissez faire* approach can often lead to the data driving itself in a feedback loop of perverse incentives (Train, 1978, 1986; Ben-Akiva and Lerman, 1985; McFadden, 1978).

At a macroscopic level, transportation modeling is often divided into the two sides of supply and demand. In general the typical approach is to first estimate transportation demand (e.g. the number of people who choose each mode) and to then meet that demand with supply (e.g. build bigger roads or run more trains and buses). Demand is often treated as a fixed monolith from which we must work around. However, any economist would argue that demand is merely the reverse perspective of the dynamic problem.

As mentioned, this one sided approach can lead to perverse incentives and induced demand. In addition to this one sided approach, transport models often lack flexibility in combinatorial mode choices. At most they are capable of a fixed set of bundled options, assuming travelers will rely entirely on one or another mode for travel. As we enter an era of increasingly complex systems with evolving mode varieties, such as mobile ride-hailing services (e.g., Uber & Lyft), dock-less bike shares, and shared automobiles; it is important to take full advantage of these emerging systems with models that can account for multi-modal systems.

Rather than continually adjusting the supply system to meet a dynamic demand, the opposite stance may be taken to adjust demand through pricing policies to optimize the system. Recent research has demonstrated the potential for personalized and optimized incentive programs to improve system performance (Azevedo et al., 2018). In order to support effective decision making for policies related to infrastructure investment, operations planning, and demand management (e.g., through incentives), there is a need for simplified models that consider the variations across different travelers, e.g., varying values of time to account for modal preferences and socio-demographic factors. This includes considering the effect of distance from transit stations on mode choice decisions as well effects of other types of heterogeneity of travel preferences across the population.

The following literature review begins with an introduction to transit access in Section 2.1, discussing its importance and contextual relevance to density and urban planning. The following three sections are structured around the two modeling paradigms of system models and discrete choice models in Sections 2.2 and 2.3, respectively, followed by a comparison in Section 2.4. The nexus bridging the two methodological approaches is the “value of time” measure, discussed in Section 2.5. The value of time measure is given empirical value in this dissertation through a data synthesis process called population synthesis, which is discussed in Section 2.6.

2.1 Transit access, density, and the urban form

It is no great revelation that transit access is an essential aspect of transit itself. Either from a quantitative technical engineering perspective or a planner’s holistic qualitative perspective. Perhaps planners have intuitively recognized how integral accessibility is to transit, for it has been exhaustively studied and examined in countless ways. One of the most critical aspects of transit access is density (Cervero and Kockelman, 1997). However, for much of 20th century, planners saw dirty and congested city centers as the source of social woes. As a result, planners often shifted towards decentralization and low-density design by taking advantage of individual mobility made possible by an affordable automobile and a burgeoning late-20th century middle class (Fishman, 1987).

Perhaps since Jane Jacobs’ seminal work *The Death and Life of Great American Cities* in her fight against “urban renewal” and the destructive interstate highway projects (Jacobs, 1961), a generation of contemporary planners now realize the negative effects this had on urban form, communities, economics, environment, equity, and overall transportation effectiveness (Cervero and Kockelman, 1997). Today there are a plethora of revivalist initiatives as ubiquitous as “New Urbanish” that bring both hope to urbanism, but also concern over gentrification and the erasure of organic culture (Grant, 2005; Smith, 2002). Moreover, there are of course dissenting opinions arguing that transit projects are a financial failure and merely a planner’s fantasy of a bygone era; that the superior mobility of automobiles should be embraced (Pickrell, 1992). Looking into the future, the relative effects of emerging technologies (e.g. ride sharing and autonomous vehicles) may no doubt incite a further debate.

In a purely utilitarian economic sense, affordable transportation provides the necessary nexus between housing and productive employment (Sandoval et al., 2011; Pathak et al., 2017). Of course, many planners recognize transportation’s importance in life outside work and the impact it has on a community (Cervero and Guerra, 2011;

Cervero and Duncan, 2003). For example, classical commuter focuses park-and-ride models are being grown to foster communities around the transit access centers in what is called transit oriented development (Turnbull et al., 2004; Caltrans, 2010; Center for Neighborhood Technology, 2006). Furthermore, there is also a growing trend to consider bicycle access in these models, for their low cost, small footprint, and environmentally friendly benefits (Flamm and Rivasplata, 2014). In analyzing the fare-box, many transit researchers found that subsidizing lower fares can have surprising societal and economic impacts beyond a marginal increase in ridership (Farber et al., 2014; Cervero, 1982; Cervero and Kockelman, 1997; Sandoval et al., 2011). Moreover, mass transportation is inherently more environmentally friendly than individualized transport. Although governmental air quality regulation began in the late 1960's with a focus on regional air pollution (e.g., smog), environmental concern recently has become a global concern with transportation being the second largest source of greenhouse gases (Hoehne and Chester, 2017).

Although studies from a planning perspective provide a broad insight into systemic issues, they are largely high-level and focus on qualitative measures, ultimately lacking quantitative solutions. From an engineering perspective, transportation can be quantified in many ways. As a fundamental example, Hägerstrand (1970) posited a measure of transportation access as the number of destinations reachable within a given distance and time. More recently, data-driven econometric approaches quantify difficult to measure factors, such as ease of access and “taste” preferences of individuals (Dong et al., 2006; Tong et al., 2015; El-Geneidy et al., 2016; Ben-Akiva et al., 1996). In effect, quantifying qualitative aspects of transportation. To date an ever growing number of transportation models are being developed for a variety of applications, ranging from simplistic geometric models to complex computer simulations.

2.2 System models

The desire to improve multi-modal transportation system efficiency through pricing is a decades old ambition. Since Vickrey’s model (1969), many have extended this work to include competing modes (Tabuchi, 1993; Huang, 2000; Kraus, 2003; Gonzales and Daganzo, 2012, 2013), and to include population heterogeneity through discrete choice models (Huang, 2002). Much of this work, especially in transportation economics, has focused on a single corridor or a simplified representation of a city as a linear system. For example, recognizing that transportation infrastructure requires space itself, Solow and Vickrey (1971) analyzed traffic patterns in an idealized linear city to find an economic equilibrium allocation of space to balance land value and congestion costs. These models have been developed to account for the costs and externalities of transportation (Solow, 1972, 1973; Wheaton, 1998) as well as equilibrium and optimum urban land use patterns (Anas et al., 1998; Anas and Xu, 1999; Rossi-Hansberg, 2004).

A significant body of literature has focused on equilibrium traffic assignment problems that consider detailed mode and route choices on specific links in a network. Equilibrium traffic models are necessary for estimating demand to be used as input in pricing optimization. These include stochastic equilibrium (Daganzo and Sheffi, 1977), dynamic user equilibrium (Ben-Akiva et al., 1986; de Palma et al., 1983), experience-based models (Iida et al., 1992; Fujii and Kitamura, 2000; Polak, 1998), boundedly rational user equilibrium (Lou et al., 2010), late arrival penalized user equilibrium (Watling, 2006), and multi-objective equilibrium models (Dial, 1996; Chen et al., 2010; Zhang et al., 2013; Levinson and Zhu, 2013; Wang et al., 2014).

A separate body of literature has been developed around modeling transit networks in regions. Analytical continuum approximation models date back to Holroyd (1967) and have been recently used to compare the performance of radial, grid, and hybrid network structures, as well as feeder-bus transit systems (Daganzo, 2010; Ba-

dia et al., 2014, 2017; Chen et al., 2015; Sivakumaran et al., 2012). The advantages of these models is that the costs for users and operating agencies can be associated with the performance of a transit system based on a small number of design variables (e.g. route spacing, stop spacing, and service headway) and the number of people using the system. This aggregated approach of modeling transportation fits well with the network-level perspective of the Macroscopic Fundamental Diagram (MFD), because the characteristics of the traveling public and region can be described in terms of high-level, aggregate quantities.

In regards to transit access models in addressing the first-last mile problem, to date there is a growing body of research focused on developing park-and-ride system models. These models seek to better understand station choice of users (Mahmoud et al., 2014), find optimal locations (Wang et al., 2014), optimize pricing (Zhu et al., 2013; Li et al., 2014; Kono et al., 2014), and even commoditized park-and-rides with parking permit trading schemes (Zhang et al., 2011). Although planners have adeptly focused analysis and policy around mixed-modal access (e.g., walk, bike, and drive), the quantitative transportation system models are lacking. Moreover, there is not only a lack of multiple access modes, but a lack of any access modes other than walking and driving.

2.3 Discrete choice

A discrete choice model is for the most part an umbrella term for any model which attempts to predict or explain the choices between discrete alternatives from empirical data. These models are generally regression-based and include such discrete choice model forms as Logistic Regression (Logit), Multinomial Logit, Probit, Nested Logit, Generalized Extreme Value, and Latent Class. Models used for the discrete choice of *mode* will be discussed in this review. Ben-Akiva and Bierlaire (1999) describe

a framework for discrete choice models as having four components: *decision-maker*, *alternatives*, *attributes*, and *decision rule*.

The *decision-maker* is any decision-making entity and their characteristics. This could be an individual person, a collection of individuals in a household, or an organization. Decision-makers are defined using characteristics (i.e., variables) for the disaggregated set of individuals. The variables are usually socio-economic and demographic attributes (e.g., age, income, gender, education, occupation, etc.). The *alternatives* are the possible choices available to the decision-maker. Alternatives could be which mode to choose, route to take, or location to reside in. The *attributes* are costs and benefits associated with each alternative (e.g., cost, travel-time, mode features, etc.). These attributes can be generic or alternative-specific. The *decision rule* is the process by which the decision-maker considers their choices, such as maximum utility. The output of discrete choice models can be deterministic with a single result, or probabilistic with a probability associated with each choice.

2.3.1 The logit model

Most choice models are based on *utility theory* in which each alternative provides a certain value, or utility, to the decision-maker. A deterministic decision rule is made for whichever option offers the highest utility to the decision-maker (Fishburn, 1970). For an alternative n and decision-maker i , utility is given by equation 2.1 as

$$U_{ni} = V_{ni} + \epsilon_{ni} \tag{2.1}$$

where U_{ni} is the utility and V_{ni} is the representative utility. V_{ni} is referred to as the “representative utility” because $V_{ni} \neq U_{ni}$. This is due to unobservable attributes affecting uncertainty in the decision-maker’s choice, captured by the error term, ϵ_{ni} .

In most cases, representative utility V is estimated using linear parameters fitted with regression. For example such as

$$V_{ni} = \beta x_{ni} \tag{2.2}$$

where β is the linear coefficient for variable x_{ni} . This simple form enables a very flexible platform from which to estimate utility using any number of desired variables. From this estimated utility, the probability of each alternative is then given by equation 2.3 as

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_{j=1}^J e^{V_{nj}}} \tag{2.3}$$

where P_{ni} is the probability of alternative n being chosen by decision-maker i , V_{ni} is the representative utility for alternative n , and $\sum_{j=1}^J e^{V_{nj}}$ is the sum of utility for all alternatives. This final expression provides an elegant and flexible platform of discrete choice and has led to a growing family of logit based choice models. Logit models are a very popular estimation method in transportation, but is also widely used in other fields, such as economics, psychology, and sociology.

2.3.2 Nested logit

In standard logit models, alternatives are treated as independent of each other, such as the choice between driving or riding the bus. This is referred to as Independent Irrelevant Alternatives (IIA) property (Ben-Akiva and Lerman, 1985). This becomes problematic when similar options are presented. Such is the case in Ben-Akiva's (1973) red/blue bus paradox where three alternatives are given: drive, blue bus, and red bus. Since the red and blue buses are identical, one would expect the modal split to be 50%–25%–25% for drive, red, and blue. However, a standard logit model treats them independently, thus yielding an even split of 33%–33%–33%. This overestimates

net mode-share of 66% for transit, despite red and blue buses essentially providing identical utility. Ben-Akiva (1973) proposed a Nested Logit Model to resolve this issue by providing predefined model partitioning of the alternatives into “nests”. An example of this is shown in Figure 2.1.

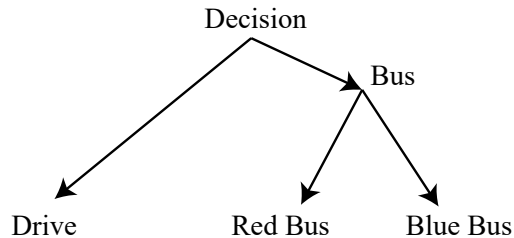


Figure 2.1: Red/blue bus nested logit

This can be accomplished mathematically by creating a utility function for the desired nests as shown in equation 2.5

$$U_{nj} = W_{nk} + Y_{nj} + \epsilon_{nj} + \epsilon_{nk} \quad (2.4)$$

$$j \in C_k \quad (2.5)$$

where C_k is the set of nested alternatives and W_{ik} and Y_{ij} are representative utilities for variables in nests j and k . The error terms ϵ are assumed to be independent and Gumbel distributed, scaled by the parameter μ_m . A composite utility for W_{ik} then is as in equation 2.6:

$$W_{C_k} = W_{C_k} + \frac{1}{\mu_m} \ln \left(\sum_{j \in C_k} e^{\mu_m W_{ij}} \right) \quad (2.6)$$

where μ_m is a unitless correlation term for the nest, often called the “logsum” term. Equation 2.6 is then nested in equation 2.3 to provide a nested probability function. An example of this nesting for the red/blue bus scenario is shown in equation 2.7.

$$P_{bus} = \left(\frac{e^{\beta x_{red}}}{e^{\beta x_{red}} + e^{\beta x_{blue}}} \right) \frac{e^{\frac{\mu}{\mu_m} \ln(e^{\beta x_{red}} + e^{\beta x_{blue}})}}{e^{\frac{\mu}{\mu_m} \ln(e^{\beta x_{red}} + e^{\beta x_{blue}})} + e^{\frac{\mu}{\mu_m} \ln(e^{\beta x_{drive}})}} \quad (2.7)$$

2.3.3 Mixed logit

In addition to the problem of IIA, there are two other limitations: (1) the population is assumed to have uniform preferences, (2) the independent irrelevant alternatives (IIA) property restricts substitution of similar choices, and (3) there is no account for unobserved factors over time. First, in a standard logit all decision-makers are treated as uniform with only a single β in the utility function of $V_{ni} = \beta x_{ni} + \epsilon_{ni}$. This inability to account for a heterogeneous population severely impacts the consideration of alternative modes (e.g., bicyclists), which typically have varying degrees of ability, confidence, and comfort as users. Second, the IIA property means that any introduction of new modes is severely limited if they have similar utility to other modes (e.g., introducing electric cars) (Brownstone and Train, 1998). Lastly, the standard logit does not consider repeated or sequential decisions over time; thus, assuming that each decision is a new one and is not affected by previous decisions.

To surmount the limitations of the standard logit model, McFadden and Train (2000) proposed the Mixed Logit Model. This was achieved by incorporating a distributed β_n in the utility function $V_{ni}(\beta_n) = \beta_n x_{ni} + \epsilon_{ni}$. The probability can then be estimated as the standard logit randomly distributed across $f(\beta)$ given as

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta \quad (2.8)$$

where L_{ni} is the standard logit model given as

$$L_{ni} = \frac{e^{\beta_n x_{ni}}}{\sum_j e^{\beta_n x_{nj}}} \quad (2.9)$$

In its complete form, the mixed logit is given as

$$P_{ni} = \int \left(\frac{e^{\beta_n x_{ni}}}{\sum_j e^{\beta_n x_{nj}}} \right) f(\beta) d\beta \quad (2.10)$$

The mixed logit essentially is a logit that is integrated across a continuous density function that can account for heterogeneous decision-makers.

2.3.4 Latent class logit

It is often that population data is only available in a discrete distribution, not a continuous function. To address this, the mixed logit can be discretized to represent segments of the population in what is called the Latent Class Logit (Kamakura and Russell, 1989; Chintagunta et al., 1991). The probability P_{ni} , of alternative n and individual i , is estimated as

$$P_{ni} = \sum_{m=1}^M s_m \left(\frac{e^{\beta_m x_{ni}}}{\sum_j e^{\beta_m x_{nj}}} \right) \quad (2.11)$$

where s_m is the share of population segment m of M total segments.

2.4 Comparison of system models and discrete choice models

With the continued advancement and availability of high performance computing, computationally intensive stochastic behavioral models have been pushed to the forefront of transportation research. Stochastic models (e.g., discrete choice) utilize abundant data for empirically fit models based on random utility theory, providing generalizable and realistic results (Ben-Akiva and Lerman, 1985; McFadden,

1978; Train, 1978). Stochastic behavioral models are well suited for user-side demand estimation models, but often lack a closed-form solution, obscuring parameter relationships and requiring heuristic solutions that do not guarantee global optimality (Sivakumaran et al., 2012). Alternatively, continuum approximation models can provide robust closed-form solutions, revealing the relationships between input variables and their optimal values (Daganzo and Sheffi, 1977; Newell, 1973; Yang and Bell, 1997). Continuum approximation models tend to be more general and their simplified nature typically only allows for a select few competing modes. Moreover, accounting for a mixture of access modes has yet to be accounted for in a pricing optimization model.

2.5 Value of time

“Value of time” is a measure often used in transportation models as a way to estimate the additional cost imposed on users due to the time spent traveling. Particularly since the proliferation of random utility and disaggregate data based models (i.e. logit), a quantifiable value of time can be extracted from the fitted utility estimations. The value of time is in itself a quantification of an abstraction. It is the value, measured either in utility or monetarily, that a person holds for a particular increment of time given some activity or purpose. Value of time is not entirely understood nor fully quantified as it varies across individuals and purpose, but also in time as well. For example, a rushed morning commute compared to a more relaxed return trip. Many studies have been conducted on the subject, but tend to be either very general with little detail, or too detailed and case specific.

Value of time is something long understood in the realm of labor and work as the value extracted from time spent working. However, outside of wages, pay scale, and productivity, it is also a concept that has become of particular importance in the field of transportation. That is, how much value does a person associate with time spent

traveling in addition to the actual monetary cost? This concept enables researchers and practitioners to convert from the abstract realm of utility into the real world monetary value.

Becker (1965) and Beesley (1965) provided the first theoretical accounts for value of time in transportation. Becker's theory of time and money assumed that an individual optimizes their exchange between leisure and work time. Meaning that time spent traveling to work is essentially unearned income; thus, their value of time can be derived as the marginal between travel time and money, i.e. the person's after-tax wage rate.

DeSerpa (1971) expanded upon this by formulating a utility function of U that is affected by the goods consumption G of time T_k an individual spends on K activities, which can include work time T_w , and leisure time T_ℓ . Assuming there is a time that must be spent on each activity and a total time constraint T° (e.g., hours not sleeping) $\sum_k T_k \leq T^\circ$, the individual will optimize their utility given their budget $G \leq Y + wT_w$. Where Y is unearned income and w is wage. The problem can be solved using a Lagrangian function

$$v_T^k \equiv \left(\frac{dY}{dT_k} \right)_V = - \frac{\partial V / \partial \bar{T}_k}{\partial V / \partial Y} = \frac{\phi_k}{\lambda} \quad (2.12)$$

$$v_T^K \equiv \frac{\mu}{\lambda} - \frac{U_{TK}}{\lambda} = w + \frac{U_{T_w} - U_{TK}}{\lambda} \quad (2.13)$$

with the multiplier constraints of λ for budget, μ for time, and ϕ_K for activity k . Within this formulation, we can extract the value of time as

$$\frac{\mu}{\lambda} = \frac{\partial V / \partial \bar{T}_k}{\partial V / \partial Y} \quad (2.14)$$

In terms of a logit model’s utility estimation, the value of time can be extracted from the various estimation coefficients for cost and time components (Ben-Akiva et al., 1993; Train, 2009). Utility in the form $U_{nj} = \alpha c_{nj}/w_n + \beta t_{nj} + \dots$ where c is cost, w is wages, t is time, and the estimation coefficients are α and β . The value of time then becomes

$$dc/dt = -(\beta/\alpha)w_n \tag{2.15}$$

which is essentially a ratio of that activity’s value to the person’s wage rate.

From this basis, a plethora of empirical studies have been conducted to study the many possible time valuations and methodological varieties (Hensher, 1976, 2001). Abrantes and Wardman (2011) provides an empirical estimation of value of travel time, depending upon the various types of time spent (e.g. in vehicle time, wait time, congestion time, etc.). Since value of time is a marginal benefit (i.e. a partial derivative) studies often focus on travel time savings from choosing some time saving alternative, e.g. toll roads (Shires and de Jong, 2009; Li and Hensher, 2012; Steimetz and Brownstone, 2005).

Small (2012) provides a detailed review of the literature, noting that many studies find the value of time to be about half the gross wage rate, which is consistent with earlier findings by Lave (1969). However, the value of time varies with income (Johnson, 1966) and tends to increase over longer times, suggesting travel fatigue. This is particularly important for active-transportation modes such as walking and biking, which value of time can vary across as well (Jara-Díaz et al., 2008; Uchida, 2014; Wichman and Cunningham, 2017).

2.6 Population Synthesis

While the field of transportation demand modeling is departing from classical trip-based approaches and shifting toward purely activity-based models, it is still

necessary to possess data with fixed workplace locations (e.g., home and work-based origin destination) as well as the socio-demographic characteristics (e.g., age, gender, income, person-household relationships, and household attributes) of the target population. However, such a complete joint set of data with socio-demographic attributes and workplace location is not available. Instead it must be synthesized using available census or sample data.

Population synthesis is a process by which a complete population data set (typically disaggregated data) is synthesized using partial data that are available (e.g., population samples or total census counts). Meaning that the missing portions of a data set are re-created by expanding or estimating for the unknown portions using existing data. Traditionally, population synthesis relied upon a well-matured deterministic fitting method called Iterative Proportional Fitting (IPF), where a multidimensional joint frequency matrix (i.e., a population sample) is fitted to known marginal totals (i.e., census totals). An alternative deterministic method is Combinatorial Optimization (CO), which re-frames it as an optimization problem to find a set of joint frequency matrix weights that minimize error with respect to the target marginal totals. More recently, population synthesis has shifted towards using more probabilistic approaches, such as Bayesian Networks (BN) and Monte-Carlo Markov Chain simulation. The benefit of these probabilistic approaches is that they require smaller sample sizes and free the researcher to use variable types other than binned frequencies (e.g., continuous age variables instead of bins of 10-14, 15-19, 20-24, etc.).

Workplace location models may be estimated using disaggregated data (e.g. discrete choice models or BN), or using aggregated data (e.g., with classical fitting methods, such as IPF and CO). The former is more flexible, capable of including a large number and variety of socio-demographic variables with accurate results, but the spatial precision is limited by sample size. The latter is typically provided at

a very fine grain spatial level (e.g., census blocks), but with fewer and less flexible variables (e.g., categorical frequencies only, not continuous variables).

In general, population synthesis methods can be categorized into three broad groups: (1) Iterative Proportional Fitting (IPF), (2) Combinatorial Optimization (CO), and (3) Probabilistic Simulation Synthesis (PSS) based approaches. The following literature review of population synthesis methods is structured around these three groups, but focusing mainly on IPF as it is most pertinent to the method developed in this dissertation. The following population synthesis literature review is as outlined as follows. First in Section 2.6.1 presents IPF along with the associated method of Iterative Proportional Updating (IPU) for synthesizing multilevel populations in Subsection 2.6.1.1. Next a brief overview of CO is presented in Section 2.6.2 before presenting PSS approaches in Section 2.6.3. Section 2.6.4 then discusses the advancement of workplace assignment models to date.

2.6.1 Iterative Proportional Fitting (IPF)

Data can be cleaved into two distinct types, aggregated and disaggregated data. Aggregated data are the totals of a particular subject or variable (e.g., total number of men or women), referred to as *marginal* data. Aggregated population data in the U.S. is generally available from the U.S. Census Bureau (U.S. Census Bureau, 2010, 2015), which provides tabulated totals for variables, such as totals by age, sex, occupation, etc. Disaggregated data in contrast, are comprised of individual persons in the population and their characteristics, referred to as *microdata*.

For decades the backbone of most population synthesizers has been IPF, a method for expanding a small microdata sample (called a seed) to match marginal totals through an iterative fitting process (Deming et al., 1940; Stephan, 1942; Choupani and Mamdoohi, 2016; Pritchard and Miller, 2012). Introduced by Deming et al. (1940), IPF is an iterative process used to fit joint distribution cells in an n -dimensional

contingency table when the marginal totals are known. The process begins by setting initial “seed” values in the cells $\hat{\pi}^0$ equal to the sample frequencies p . Then for each dimension (e.g. rows i and columns j) the cells are proportionally adjusted to fit the marginal totals for the respective dimension. This is expressed as

$$\hat{\pi}^{(0)} = p \quad (2.16a)$$

$$\hat{\pi}_{ij}^{(\eta-1)} = \hat{\pi}_i^{(\eta-2)} \frac{p_{i+}}{\sum_j \hat{\pi}_j^{(\eta-2)}} \quad (2.16b)$$

$$\hat{\pi}_j^{(\eta)} = \hat{\pi}_i^{(\eta-1)} \frac{p_{+i}}{\sum_i \hat{\pi}_i^{(\eta-1)}} \quad (2.16c)$$

where η is the iteration number. Once all dimensions are proportionally adjusted, a measure of fit calculated, completing one iteration. This iterative process is repeated until error converges to some desired threshold. Deming et al. (1940) proposed an objective function of minimizing the least square error, also called a chi-squared function, shown in equation (2.17).

$$(X')^2 = \sum_{i,j} \frac{(p_{ij} - \hat{\pi}_{ij})^2}{\hat{\pi}_{ij}} \quad (2.17)$$

where p_{ij} is the sample proportion of cell ij and $\hat{\pi}_{ij}$ is the estimated cell proportion. Mosteller (Mosteller, 1968) advanced the technique by showing with equation (2.18) that cross-product ratios could be used to adjust the table while preserving its structure at each iteration.

$$\frac{n_{ij}n_{hk}}{n_{ik}n_{hj}} = \frac{p_{ij}^{(\eta)} p_{hk}^{(\eta)}}{p_{ik}^{(\eta)} p_{hj}^{(\eta)}} \quad (2.18)$$

where n is sample of observations in two multi-way tables of proportions h, i, j , and k ; with η being the iteration number. Then Ireland and Kullback (1968) further showed that cell probabilities can be estimated for multi-way contingency tables, the importance of this is that IPF can be extended to high dimensional contingency tables.

Wong (1992) then tested the utility of IPF for generating populations for geographers, until finally Beckman et al. (1996) was the first to utilize IPF for population synthesis with disaggregated travel demand modeling.

IPF requires initial seed values to begin proportional fitting. Any zero cells in the seed will remain as a zero during IPF and not be fitted. There are two types of zero cells, “sampling” zeros that occur when there are no representatives captured in the sample (e.g., rare combinations), and “structural” zeros that represent impossible combinations in the data (e.g., a head of household that is under aged). The difficulty in handling zero cells is the need to preserve structural zeros while adding heterogeneity by filling sampling zeros. One solution to the zero cell problem is to simply set a very small arbitrary value (e.g., 0.001) for zero cells (Beckman et al., 1996). This allows the cell to be fitted and helps IPF to converge. However, this also removes any structural zeros in the seed, introducing the potential for impossible combinations to occur. Another solution is to substitute missing cells using values from a larger sample (e.g., the entire study area rather than a sub region). In order to ensure proportional unity, the borrowed values are adjusted proportionally by the ratio of sub-sample size to the total sample size (Ye et al., 2009; Guo and Bhat, 2007).

Alternatively, sample-less populations may be generated using structured marginals (Barthelemy and Toint, 2013) with IPF. The weights are then integerized and replicated to form a near perfect disaggregate population (Lovelace et al., 2014; Ballas et al., 2005a,b). However, this destroys the intricate household-person relationships that can be extracted organically from a joint sample. Demand models often rely on decisions made at the household level (Guo and Bhat, 2007). For this reason it is often necessary to synthesize a multilevel population (i.e., persons and households). Multi-level populations must then be reconstructed using an algorithm, but this often comes with a loss of accuracy (Lovelace and Dumont, 2016). Sample-based approaches tend to be preferred, largely because public use microdata are typically available in

most countries where population syntheses are performed. For example, Public Use Microdata Sample (PUMS) in the United States (U.S. Census Bureau American Community Survey, 2015), Public Use Microdata Files (PUMFs) in Canada, and Samples of Anonymised Records (SARs) in the United Kingdom.

2.6.1.1 Iterative Proportional Updating (IPU)

Generating multilevel populations tends to be one of the most challenging problems in population synthesis. In general, multilevel populations are synthesized by drawing households from a joint microdata sample of persons and households. The sampled households along with their associated persons are replicated into a pool of joint persons and households (Beckman et al., 1996; Auld and Mohammadian, 2010).

Beckman et al. (1996) estimated joint populations by fitting households using IPF, then used the IPF weights to draw from a joint sample. However, using only households leaves person characteristics uncontrolled, therefore, introducing error. Error was partially mitigated by incorporating broad person level variables into households (e.g., number of workers, children, or adults). This was further improved through sampling algorithms, relation matrices, multiple IPF steps, or improved classification and regression trees (Le et al., 2016; Zhu and Ferreira, 2014; Guo and Bhat, 2007; Arentze et al., 2007; Arentze and Timmermans, 2004).

Ye et al. (2009) provided a major breakthrough by proposing a novel fitting algorithm called Iterative Proportional Updating (IPU). IPU re-weights households in a microdata sample using separate IPF weights for persons and households as marginal constraints in the subsequent IPU step. This yields a single joint weight that accounts for both persons and households simultaneously. It is based on the basic principle of IPF, but instead of individual households and persons, IPU re-weights a *joint* sample of household and person types using separate IPF estimates as marginals in a subsequent fitting process. The results yield a re-weighted value for each household that

accounts for both persons and household attributes. The algorithm is performed by first structuring the joint person-household sample data (e.g., Public Use Microdata Sample) into a joint list, such as the example in Table 2.1.

Table 2.1: Example IPU joint list

Sample ID	Weight	Households		Persons		j
		Type 1	Type 2	Type 1	Type 2	
1	w_1	1	0	1	1	d_{ij}
2	w_2	0	1	0	2	
3	w_3	0	1	1	1	
4	w_4	1	0	3	0	
5	w_5	0	1	0	1	
6	w_6	1	0	1	0	
i	w_i					
Marginal constraint		35	45	124	137	c_j

The household and person types are combinatorial types, meaning that there is a unique type for each possible combination of household or person variables. The values in the columns under person and household type represent the number of persons or household types in each household sample. Thus, the number of columns in the joint table is equal to the total number of all possible person variable combinations plus all possible household variable combinations. The number of rows in the joint table is equal to the number of household samples in the data. The iterative process begins by proportionally adjusting the weights w_i to match the marginal constraint in each column. Once all columns are adjusted, completing one iteration, a measure of fit is checked using the objective function”

$$\delta = \frac{\sum_j \left[\frac{\left| \left(\sum_i d_{ij} w_i - c_j \right) \right|}{c_j} \right]}{m} \quad (2.19)$$

where δ is the absolute relative difference value, d_{ij} is the frequency of sample population characteristic j in household i , w_i is the weight for household i , c_j is the

marginal constraint for the population characteristic, and m is the total number of marginal constraints in the problem. The process repeats iteratively until a specified convergence criteria for δ is met. However, depending on the sample size and possible combinations, the resulting table can quickly become an extremely large matrix creating massive memory requirements. Although it is possible to reduce memory requirements using sparse matrices, the dimensions of the matrix itself are still computationally cumbersome, greatly slowing computation for a process that could require many iterations to converge.

2.6.2 Combinatorial Optimization (CO)

Though popular, IPF is not the only technique used in population synthesis. Another classical deterministic approach is Combinatorial Optimization (CO) (Owens and Rao, 1995; Voas and Williamson, 2000; Abraham et al., 2012). CO treats population synthesis as an optimization problem, where the number of representatives in the joint sample (i.e., sample weight) is optimized to match the marginal totals. CO also offers the possibility of integer optimization, eliminating the need for probabilistic sampling or decimal “integerization” (Lovelace and Ballas, 2013). However, a major weakness of using CO is the inherent disregard of attribute association and weight (i.e., the frequency a combination of attributes) (Pritchard and Miller, 2012). While IPF will preserve patterns in a microdata sample based on frequency, CO will minimize error even if it means setting unrealistic weights (e.g. zero). This potentially leads to over-fitting or loss of heterogeneity. In general, CO is less common, but can provide precise and computationally efficient results (Hermes and Poulsen, 2012).

2.6.3 Probabilistic Simulation Synthesis (PSS)

IPF and CO rely on classical fitting and re-weighting methods for populations, but more recently a pure simulation based probabilistic approach has proven superior

in many regards. Rather than determining household weights using IPF and then drawing, simulation-based approaches effectively fit and draw samples simultaneously by sampling directly with a conditional Monte-Carlo Markov Chain (MCMC).

Farooq et al. (2013) used a Gibbs sampler to draw from a person level population sample, checking the fit against marginals to achieve a near perfect fit. Casati et al. (2015) improved upon this by proposing a two-step method using a Gibbs sampler followed by a re-weighting step to satisfy both individual and household margins.

A major weakness of simulation-based methods is the lack of heterogeneity in the sample, meaning that persons or households cannot be synthesized in the population if they are not represented in the sample (Farooq et al., 2013). Sun and Erath (2015) proposed a new approach using Bayesian Networks (BN) to map and reconstruct the joint conditional probabilities one pair of variables at a time from their marginals in the population; in effect, reintroducing heterogeneity into the population that may have been lost by solely relying on full joint conditionals. This ability to reconstruct populations also means that the method requires smaller sample sizes than IPF to achieve a satisfactory level of accuracy. A further improvement by Saadi et al. (2016) using Hidden Markov Models (HMM) helps capture hidden correlations between the diversity of variables in subgroups of the population.

These probabilistic approaches are state-of-the-art in population synthesis providing accurate results, but unlike IPF or CO that rely on discrete categorical frequencies, probabilistic approaches have the flexibility to handle continuous variables. This flexibility has spurred the interest in integrating advanced computation methods with population synthesis methods, such as machine learning, to further refine results (Borysov et al., 2018). Although PSS methods do not forbid integration of workplace assignment and population generation per se, but no examples have been found in the literature yet. Furthermore, PSS approaches are limited in synthesizing joint populations, requiring that household relationship structures (e.g., spouse,

child, non-relative etc.) be defined by researchers (Sun and Erath, 2015). This is problematic as there are an unimaginable variety of family and non-family household structures.

2.6.4 Workplace assignment

Traditional trip-based models allocate aggregated travelers from origins to destinations using an origin-destination (OD) assignment matrix fitted with aggregated trip generation data. For example, the number of workers that live in each origin and the total number of workers that work in each destination. To fit the OD matrix, the cells in the matrix (i.e., OD pairs) are given an initial weight based on some weighting scheme, such as the common “gravity model” (Voorhees, 1956). Most aggregated trip-based models fall into this classical model of iterative fitting, but vary by their weighting procedures (Abdel-Aal, 2014), such as the maximum entropy (Wilson, 2011), intervening opportunities (Stouffer, 1940), or radiation laws (Simini et al., 2012). These models make alternative assumptions (e.g., distance based attraction) or add complexity in order to account for a variety of socioeconomic factors. However, these aggregated approaches all rely on IPF and are confined to a single trip purpose at a time (e.g., work trips).

With the development of discrete choice models and the ability to break free from single purpose OD matrices, transport modeling has largely shifted away from rigid deterministically fit assignment models (McFadden, 1978; Train, 1986; Ben-Akiva and Lerman, 1985). An ever growing family of increasingly complex models are being developed to model individual decisions (e.g., for mode, purpose, time of day, and destination) (Bowman and Ben-Akiva, 2001; Bowman et al., 1998; Dong et al., 2006; Recker, 2001). The challenge with these types of models is that as the desired spatial resolution increases, ever larger sets of sample data are required to build these models. While methodologies to deal with the limitations of discrete spatial

choice modeling have been proposed (Guevara, 2010), data limitations still poses a problem to fine grain destination choice models as sample data can become too sparse for accurate estimation.

2.7 Summary of Literature

Demand models typically rely on a family of discrete choice models, estimated using previous travel behavior or stated preferences. Discrete choice models treat each alternative as a singular entity (e.g. walk, drive, or train). As a practical solution, multi-modal combinations are typically modeled by creating “bundles” of alternatives (e.g. walk to train). The utility of bundles are then fitted as a unique alternative and the choice probability is estimated as such. In most cases this is a perfectly acceptable practice because “feeder” modes are often relatively uniform (e.g. walk distance around a transit stop). However, this poses as a limitation to more complex multi-modal combinations due to the lack of control for individual modal utility and the ability to predict for hypothetical new combinations. The purpose of combining new modes is to investigate the potential for improved overall travel benefit by manipulating not merely mainline modes, but both access and mainline modes.

Multi-modality might be solved by chaining decisions together through estimation of the probability for each mode segment of the journey separately. However, this is a partially flawed approach as a decision-maker will not make each choice independently, but instead will consider the total benefit in a single decision. Moreover, this risks underestimating costs associated with transferring modes. Thus, by splitting the modes into decisions, one violates the fundamental assumption that a decision maker is logical and will choose the combination of alternatives which provides the greatest utility. A more appropriate and parsimonious model would be to incorporate and account for the multiple utilities at the decision point.

To accomplish this task, a demand model must be developed in a way that not only accounts for system costs, but formulated in such a way that is easily optimizable. Furthermore, the model should also account for population heterogeneity, as expressed by value of time variation. Although value of time is an already extensively researched subject, a varying value of time has yet to be included in a system-demand model of this kind. Moreover, the relationship between value of time and income is often cited across the literature Small (2012), yet little research has been undertaken to investigate it specifically. The aforementioned research may be distilled into the following critical research gaps:

1. *System models*: System models provide a mathematical approximation for an entire system, revealing parameter relationships and easily finding a robust global optimum. However, these models have yet to include multiple access modes and user heterogeneity. Moreover, demand is often considered a fixed input, or assumed to conform to supply.
2. *Demand models*: Contemporary stochastic demand models (e.g., logit-based) can achieve a high degree of empirical accuracy and flexibility with respect to heterogeneity, but fundamentally do not possess a closed-form. This makes them impossible to solve analytically and obscuring parameter relationships. Optimization of such models requires heuristic solutions that can become computationally intensive and do not guarantee global optimality.
3. *Economic models*: A multitude of economic modelling and theory has been published surrounding mono-centric or poly-centric cities and economic equilibrium. However, these models often oversimplify travel time to a single mode, or disregard congestion in the system. Moreover, these models are closed systems in that they do not consider movement between cities, only within. Lastly, these

models purely seek to find efficient equilibriums and fail to consider population heterogeneity (i.e., value of time).

4. *Equity*: Although equity is often raised as a social issue resulting from transportation policy, it is largely researched *a posteriori*. Transportation models rarely, if ever, address equity as a factor in the model itself and not merely a latent variable or external consideration.

The proposed research seeks to address these research gaps with a new hybrid deterministic model that accounts for multiple access modes, varying values of time, and can be optimized with computational efficiency. Such a model can be used to not only provide an optimal pricing policy to achieve desired demand, but to explore policy results more easily and possibly address equity concerns.

CHAPTER 3

POPULATION & WORKPLACE SYNTHESIS

Both population synthesis and workplace destination have classically relied on IPF, yet to date the two fitting processes have not been integrated into a single process despite sharing a common algorithm. The benefits of such an integration provides a much more efficient and seamless generation and assignment process. This is done by first fitting a workplace assignment model using aggregated employment data using aggregated population data as a constraint (i.e., known home location as origin totals). The joint distribution of individuals in a synthetic population is then fitted, also using IPF, with workplace destination as a variable using the workplace assignment model as a marginal constraint. Thus, as the synthetic population is fitted with respect to persons and households, it is also fitted with respect to workplace. The result yields a synthetic population already fitted with home origin and workplace destination for each individual.

This proposed unified process makes two contributions; first by integrating population synthesis and workplace assignment, and second by developing a more efficient joint person-household matching approach to handle the additional population attributes added. The joint person-household matching approach is based on Combinatorial Optimization (CO) with linear optimization of non-negative least absolute deviation (NLAD). The new re-weighting method is capable of handling larger matrices more efficiently than typical methods, a necessary consideration when many additional features are incorporated for workplace assignment. This not only makes integrated workplace assignment possible, but also provides inherently robust results

while exploiting the computational efficiency of linear optimization. The proposed method is compared against the more conventional algorithm of Iterative Proportional Updating (IPU) as well as a non-negative least squares (NNLS) approach.

The proposed integrated population synthesis and workplace assignment process is displayed visually through a schematic process flowchart in Figure 3.1. In general the process is divided in to four steps: (1) origin-destination fitting, (2) person and household IPF, (3) joint re-weighting, and (4) joint sampling.

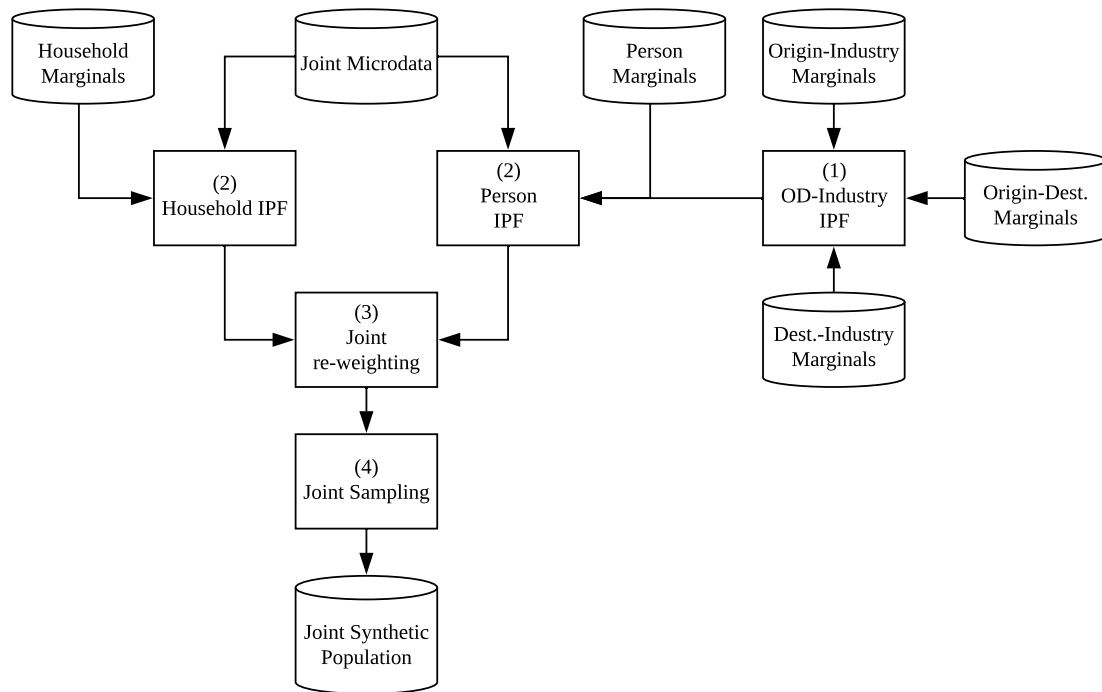


Figure 3.1: Modeling framework

3.1 Simultaneous fitting of population and workplace

Within the proposed process, IPF is used at three separate instances: persons, households, and origin-destination. The origin-destination IPF is performed in step labeled (1) as a pre-processing step for aggregated origin-destination data. The purpose of the pre-processing step is to obtain the joint distribution for origin, destina-

tion, and industry from the simple “flat” two-dimensional tables. The resulting joint distribution is then used as marginals in the person level IPF of step (2).

Despite being classically fitted using IPF, the origin-destination-industry (ODI) matrix is weighted using observed OD totals, not an assumed model as the gravity or similar model. The fitting process merely extends the OD table across multiple dimensions using origin by industry (OI) and destination by industry (DI) as marginals, meaning that the resulting matrix is empirically fit. In this dissertation a three-dimensional matrix was generated for origin, destination, and industry; however, the process is flexible in that it can accommodate higher dimensional matrices by incorporating additional socio-demographic stratification. The three-dimensional matrix is formed by the three marginal tables of origin-industry, destination-industry, and origin-destination (see Figure 3.2). Industry by origin or destination make the vertical faces and origin-destination totals make the cube base. The cube interior is the unknown joint distribution of home-workplace pairs, which is fitted using multi-dimensional IPF.

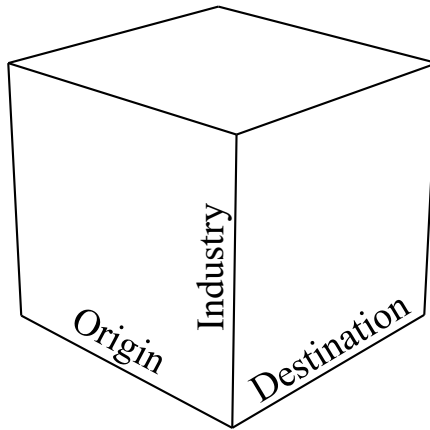


Figure 3.2: Origin-destination-industry conceptualization

3.1.1 Marginal consistency

Before the multidimensional matrix is constructed, the marginals must be checked for consistency between the origin, destinations, and person marginals (i.e., the marginal totals are equal). It is likely that the tables will not match perfectly due to sampling error, shifts in the population over time, unemployed persons, or persons that enter/leave the study region. Although the differences may be minor, IPF requires perfect consistency between marginals. The minor differences between the OD, OI, and DI marginals can be corrected by proportionally adjusting each marginal to match each other.

The adjustment process begins by treating the population marginals as the OI marginal. The OD and DI tables will be adjusted to match the OI table (i.e., population marginal). The OD table is adjusted first to match the OI table, then the DI table is adjusted to match the OD; effectively using the OD table as a bridge between origins and destinations. At this point, non-working persons are excluded because the aggregated employment data only accounts for employed persons. Once the tables are adjusted, the missing portion of non-working persons are added back to the OD, DI, and OI tables using the original total of unemployed persons in the population marginals. Since the aggregated data reflects *workplace* only, the non-working persons are their origin (i.e., home zone) as also their destination to ensure the totals are consistent.

3.1.2 Origin-Destination-Industry seed

Although it is possible to perform IPF using only the OI and DI marginals with an assumed *a priori* seed model, such as the gravity or radiation models, the OD marginal itself may be used to seed the ODI matrix. This provides the important OD distribution that is then fitted to the other two OI and DI marginals. Moreover, since all marginals are contingent upon each other (e.g., the origins in OI overlap with

origins in the OD, and the destinations in the OD then overlap with the destinations in DI), this adds structure to the IPF process helping to ensure accurate fitting (Barthelemy and Toint, 2013; Lovelace and Dumont, 2016). This means that the OD marginal table itself may be used as seed data, building an empirically fit OD assignment model.

3.1.3 Integrated person fitting

Once the origin-destination-industry (ODI) matrix is fitted, the joint distribution of destinations can be treated as marginal totals for destinations of persons. This is performed in a zone-by-zone basis where a slice of the three-dimensional cube is taken for the respective origin zone. This yields a joint table of total workers at each destination by origin zone and by industry sector. The joint table is then used as a marginal constraint in IPF along with the conventional demographic variables for persons (e.g., age, gender, industry, etc.). The joint destination table by industry, as opposed to destination only, provides additional structural constraint to the IPF, helping to ensure a good fit (Lovelace and Dumont, 2016). IPF for households is also performed at this step in tandem with the person IPF.

3.2 Joint re-weighting

To generate a joint population of households and persons in step (3) (see Figure 3.1), the separate IPF weights must be jointly re-weighted using the joint person and household sample. A common re-weighting method is Iterative Proportional Updating (IPU) (Ye et al., 2009). However, IPU is computationally intensive and given a large number of zones and person level variables, IPU requires a very long time to compute (e.g. approximately over 100 hours in this case). To improve upon the computational efficiency of IPU, the joint re-weighting problem is recast as an optimization problem. The optimization problem is then solvable by minimizing error

through either non-negative least squares (NNLS) or non-negative least absolute deviation (NLAD) objective functions. For the overall synthetic population generation, only NLAD is used due its superior performance and the excessive time required to compute a full population using the other methods (e.g., 720 days for NNLS). However, the performance of NNLS and IPU is compared to NLAD using a smaller set of zones and by treating the entire GBA as a single zone.

3.2.1 Non-negative least squares (NNLS)

The problem is formulated as an optimization problem that minimizes an error by changing joint person-household population weights. The optimal solution can then be found more directly through efficient optimization algorithms, rather than iterating until convergence as is the case with IPU. The problem may be easily formulated into the familiar $Ax = b$ format, treating each joint sample weight as a decision variable x , and the sample values as constraints in an A matrix, and the IPF marginal cells as the right hand side target value of b . The objective is to minimize the fitting error by finding a combination of samples x that best fit the IPF marginals b . This optimization approach is similar to CO in the sense that it uses optimization to find weights, but unlike conventional CO it does not use individual population variable totals as the target. Instead this approach using IPF contingency table results, rather than individual variables. This not only provides a joint person-household weight, but helps to preserve the attribute associations in the population that may be lost with conventional CO (Pritchard and Miller, 2012). For demonstration, the generic data is reformatted into Table 3.1 for an optimization approach. As with IPU, the person and household types are combinatorial, resulting in a table with as many rows as total possible combinations of person and household variables, and as many columns as household samples. Each column contains a household sample, specifying which household type it is and then the number of person types in each household sample.

Table 3.1: Example linear program joint list

Sample ID	Weight	x_1	x_2	x_3	x_4	x_5	x_6	b
Households	Type 1	1	0	0	1	0	1	35
	Type 2	0	1	1	0	1	0	45
Persons	Type 1	1	0	1	3	0	1	124
	Type 2	1	2	3	0	1	0	137

Although this problem can be solved analytically with ordinary least squares (OLS) regression, this poses a problem in the context of a population for which there cannot be negative weights (e.g., cannot have negative households or persons). Thus, this problem requires a non-negativity constraint and cannot be solved analytically. The non-negative OLS problem, called non-negative least squares (NNLS) is presented in Equation (3.1) and can be solved using quadratic programming.

$$\min \|b - Ax\|^2 \tag{3.1a}$$

$$\text{s.t. } x \geq 0 \tag{3.1b}$$

This NNLS problem is often solved using an algorithm developed by Lawson and Hanson (1995). However, for large scale problems, such as the one posed in this dissertation, this algorithm becomes computationally inefficient. Alternatively, a linear approach such as NLAD, may actually yield superior computational performance in this case.

3.2.2 Non-negative least absolute deviation (NLAD)

To avoid the quadratic complication, the problem can be made linear by instead optimizing for least absolute deviation with Equation (3.2). Least absolute deviation (LAD) is a well established sibling to OLS regression often used as a “robust” alternative less sensitive to outliers (Bloomfield and Steiger, 1984; Davis and Dunsmuir,

1997). In typical regression applications, LAD is less attractive than OLS because it cannot be solved analytically. However, since the non-negativity constraint already makes the problem unsolvable analytically, the linear form proves more efficient in this case.

$$\min |b - Ax| \tag{3.2a}$$

$$\text{s.t. } x \geq 0 \tag{3.2b}$$

The single constraint of $x \geq 0$, again is for the simple reason that there cannot be a negative number of households or persons. The absolute value is needed in order to account for both positive and negative deviations, which was otherwise handled in OLS by the squared term. This can be further recast as a linear program as follows:

$$\min \sum \epsilon \tag{3.3a}$$

$$\text{s.t. } Ax - \epsilon \leq b \tag{3.3b}$$

$$Ax - \epsilon \geq -b \tag{3.3c}$$

$$x \geq 0 \tag{3.3d}$$

where ϵ is an artificial variable for error (i.e., the deviation) to be minimized in the objective function (3.3a). Absolute deviation is provided by the mirrored inequality constraints (3.3b) and (3.3c), ensuring that both positive and negative error is minimized. This linear mathematical program can be solved by the simplex method, using almost any optimization package. However, the matrix created by the problem is quite large with tens of thousands of decisions variables and potentially hundreds of thousands of constraints (i.e., one per household sample), and would benefit from a well coded optimization package.

The project work flow and data handling is written in *R*, but the optimization algorithms are coded as dedicated functions using more efficient programming languages and simply executed with *R*. In all cases, the joint sample is stored as a sparse matrix in *R* before being passed to the respective algorithms, greatly reducing the required memory and improving overall performance for all methods. NLAD utilizes an open source linear programming package written in C++ called “Clp”, developed and maintained by Computational Infrastructure for Operations Research (COIN-OR) Foundation (2017). NNLS also utilizes open source software package called “nnls”, which is based on the Lawson and Hanson (1995) algorithm and written in Fortran (Katharine M. Mullen and Ivo H. M. van Stokkum, 2015). IPU is not readily available as an *R* package, but was coded as a custom *R* package in C++ by the authors to provide a competitive performance comparison.

3.3 Joint sampling

The results for all re-weighting methods are decimal weights for each joint record in the microdata. The joint weights can then be used as weighted probabilities to generate the final population with Monte-Carlo sampling in step (4) of the process (see Figure 3.1). This sampling process is no different than with existing methods (e.g., IPU). However, microdata typically does not contain OD information and cannot be re-weighted with respect to OD. Instead a simple two step random sampling procedure is used to generate the final population. First, joint household-persons are generated by sampling from the microdata using the new joint weights. Then from this joint sample, the destination is drawn using the person-destination IPF weights as proportional probabilities for each person given their person type.

3.4 Data

Data utilized for population synthesis consists of aggregated marginal totals, disaggregated microdata samples, and aggregated OD totals by industry. All data are publicly available from the United States Census Bureau, and are summarized in Table 3.2. The marginal tables are provided by the United States Census Bureau’s American Community Survey (ACS) (U.S. Census Bureau, 2015). As opposed to the decennial census, which is a full census collected only every 10 years, the ACS is a program that performs ongoing data collection used to estimate adjusted tables for more recent years between decennial census years. The microdata are also managed by the ACS program of the United States Census Bureau, referred to as Public Use Microdata Samples (PUMS). The PUMS are provided as roughly a five percent sample of the households and persons in the population.

Table 3.2: Data used in population synthesis

Table/Dataset Name	Year	Program	Description
<i>Marginal data</i>			
B19001	2015	ACS 5-year	Household income
C24050	2015	ACS 5-year	Industry & occupation
B01001	2015	ACS 5-year	Age & sex
<i>Microdata</i>			
ss15pma	2011-2015	PUMS	Disaggregate persons sample
ss15hma	2011-2015	PUMS	Disaggregate households sample
<i>Origin-Destination data</i>			
ma_wac_S000_JT00	2015	LODES	Workplace destination by industry
ma_rac_S000_JT00	2015	LODES	Workplace origin totals by industry
ma_od_main_JT00	2015	LODES	Workplace origin-destination totals

The OD totals are managed by the Center for Economic Studies of the United States Census Bureau under the Longitudinal Employer Household Dynamics (LEHD) program. This program also collects home and work locations of individuals, with origins and destinations aggregated by various stratification (e.g., industry sector), called the LEHD Origin-Destination Employment Statistics (LODES). LODES pro-

vides aggregated OD pair totals for census blocks in a set of data tables stratified by demographics. The demographic data stratification are provided for origins or destinations separately, not simultaneously. For example, the total number of workers for each origin-destination pair are provided in one table with two separate tables stratified for origin by industry and destination by industry.

Since both the PUMS and census tables are managed and provided by the U.S. Census Bureau, they largely share the same variables and data structure, requiring very little adjustment to make them compatible. In some cases however, continuous variables (e.g., income and age) in the disaggregated PUMS needs to be binned as discrete variables to match the grouping used in the aggregated census tables. Table 3.3 summarizes the overall variables used for person and household IPF.

Table 3.3: Population synthesis variables

Household income	Sex	Age	Mode	Industry	Occupation	Home Workplace
≤\$10,000	Male	0-9	Drive	None	None	(965x965 tracts)
\$10,000-14,999	Female	10-14	Transit	Natural resources	Service	
\$15,000-19,999		15-19	Other	Transportation and utilities	Production and transportation	
\$20,000-24,999		20-24	None	Finance and real-estate	Sales, office, and administration	
\$25,000-29,999		25-44		Educational and social-work	Management, business, scientific, and arts	
\$30,000-34,999		45-54		Professional, scientific, and management	Natural resources, construction, and maintenance	
\$35,000-39,999		55-64		Retail trade		
\$40,000-44,999		>65		Wholesale trade		
\$45,000-49,999				Information		
\$50,000-59,999				Construction		
\$60,000-74,999				Manufacturing		
\$75,000-99,999				Arts and accommodation		
\$100,000-124,999				Other		
\$125,000-149,999						
\$150,000-199,999						
≥\$200,000						

3.4.1 Marginal data

The census tables contain data for all census tracts in the Commonwealth of Massachusetts. These tables are provided as either the decennial census or the American Community Survey (ACS) program. The decennial census is a complete census completed every 10 years and the ACS is an ongoing sampling program which publishes estimated tables every year. A summary of the tables used are shown in Table 3.2 under *Marginal Data*. Census tracts that exist outside of the study area were removed using GIS shapefiles. It should be noted that some of the tables contain joint contingency tables, such as SF-DP1 which contains both age and sex. In this case we can have a joint count for the number of men and women within an age group. The benefit of this is it eliminates the need to fit for these variables jointly, improving IPF accuracy and performance. To make these tables compatible with IPF, they are restructured to be multidimensional, rather than tabular. For example, SF1-DP1 would become a three dimensional matrix with each dimension being age, sex, and census tract.

3.4.2 Micro-data

The microdata samples used in this dissertation are Public Use Microdata Samples (PUMS) obtained from the U.S. Census Bureau (U.S. Census Bureau American Community Survey, 2015). The PUMS are five percent disaggregate population samples collected as part of the ACS. The samples are compiled and anonymized for public use. The smallest spatially allocated area available in the PUMS are Public Use Microdata Areas (PUMA). PUMAs are areas that contain at least 100k persons in the population, which are relatively large compared to census tracts used for synthesis, but smaller than most counties.

3.4.3 Longitudinal Origin-Destination Employment Statistics

The origin-destination totals are managed by the Center for Economic Studies of the United States Census Bureau under the Longitudinal Employer-Household Dynamics (LEHD) program. This program also collects home and work locations of individuals, with origins and destinations aggregated by various stratifications (e.g. industry sector), called the LEHD Origin-Destination Employment Statistics (LODES) (U.S. Census Bureau Center for Economic Studies, 2017). The LODES tables provides aggregated origin-destination pair totals for census blocks, but the demographic stratifications are only provided for origins or destinations, not simultaneously. For example, the total number of trips from each origin-destination pair are provided in one table with two separate tables stratified for origin by industry and destination by industry. In this case, the aggregated data is stratified by industry sector.

3.5 Evaluation

The results are described in three subsections: joint re-weighting method comparison, population generation results, and workplace assignment results. Joint re-weighting results present the accuracy and computational performance comparison between IPU, NNLS, and NLAD when performed for a single zone. The subsequent sections then demonstrate the final population and workplace assignment results using the NLAD re-weighting method. Only NLAD was used to generate a full population due to the excessive computation time required the synthesis all 965 census tracts with the other methods.

The results are validated using Root Mean Square Error (RMSE) and Root Mean Square Normalized Error (RMSN). RMSE is calculated as

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{b}_i - b_i)^2}{n}} \quad (3.4)$$

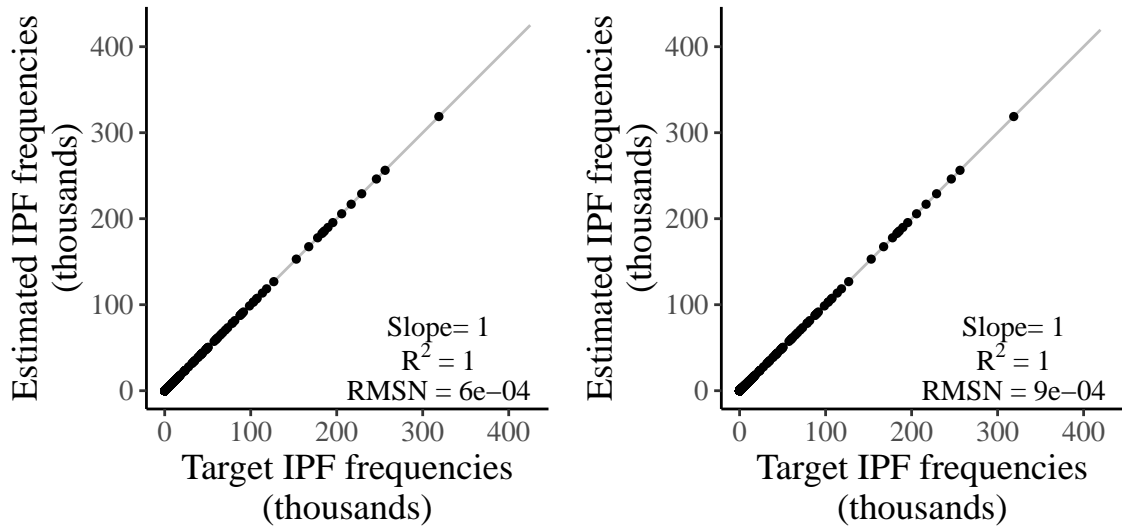
where n is the number of values, \hat{b}_i is the estimated value of variable i , and b_i is the actual values. For example, b_i is the frequency of person or household type i . A good fit will yield a small RMSE. However, since the following comparisons contain a wide range of values (e.g., between tracts, total region, and ODs) a normalized RMSE value is used in order to make the errors more comparable across tests. A commonly used alternative is to normalize the RMSE value by the mean \bar{b} to account for relative error between differently sized values, further calculated as

$$RMSN = \frac{RMSE}{\bar{b}} \quad (3.5)$$

3.5.1 Joint re-weighting results

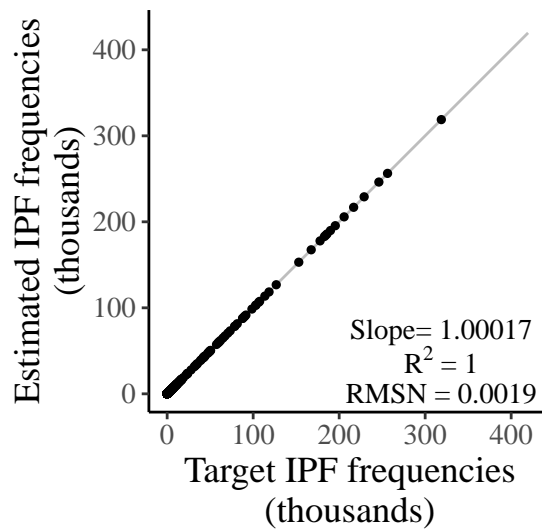
As a general comparison of fitting accuracy, persons and households are jointly re-weighted using the three re-weighting methods of (1) NNLS, (2) NLAD, and (3) IPU for the entire Greater Boston Area treated as a single zone. Figure 3.3 is a comparison of the fit results for the three methods. The target values from IPF for persons and households (i.e., the b values) are shown on the horizontal axes and the vertical axes are the fit results when the joint weights are multiplied by the joint sample matrix (i.e., the Ax result). A good fit will be along the diagonal, meaning that the correct number of both persons and households are fitted when $Ax = b$ is evaluated. Note that the weights at this point are decimal values, which is why the results are near perfect. Error will be introduced when weights are sampled as discrete persons and households, but as a measure of fitting performance that fact is irrelevant.

NNLS and NLAD both achieve much better fit results than IPU with RMSNs of 0.0318, 0.0824, and 0.1787, respectively. Figure 3.4 presents the same results as in



(a) NNLS

(b) NLAD



(c) IPU

Figure 3.3: Comparison of fit by method (1:1 scale)

Figure 3.3, but zoomed in at 100:1 scale. This closer perspective reveals interesting insights into the inherent properties of NNLS, NLAD, and IPU. Least square based NNLS will find the mean fit, fitting the points near the diagonal target line. In contrast, least deviation based NLAD finds the median fit, placing the points either

perfectly on the line or almost not at all. IPU is neither, resulting a combination of near perfect and approximately fit results.

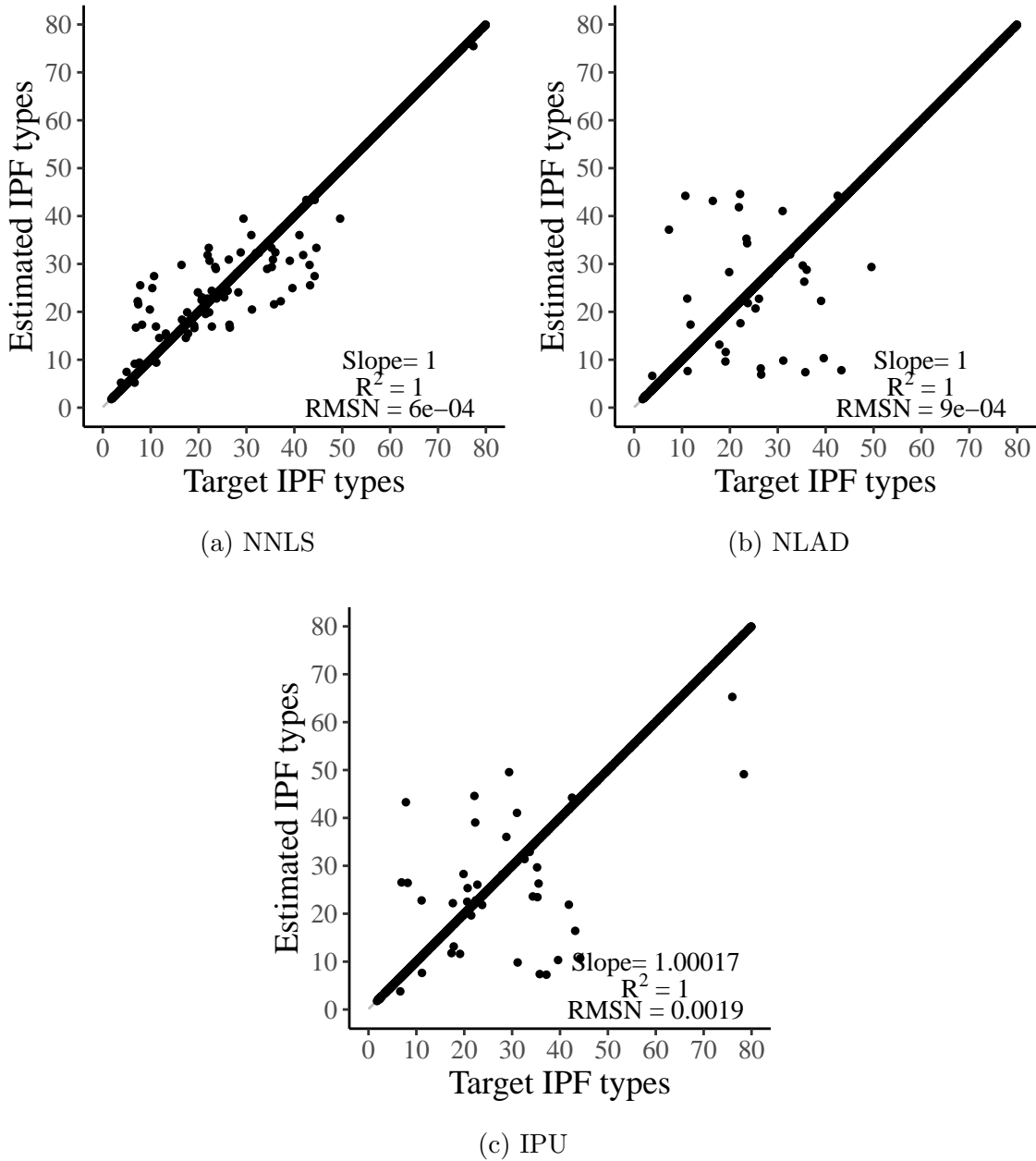


Figure 3.4: Comparison of fit by method (5000:1 scale)

Although both NNLS and NLAD outperformed IPU in terms of goodness of fit, NLAD substantially outperformed both NNLS and IPU in terms of computational

time. The computational time for each of the three methods to fit a single zone are shown in Table 3.4. Although NNLS yielded the most accurate results, a computation time nearly 1000 times longer than NLAD does not justify the relatively small gain in accuracy over NLAD. Overall, NLAD achieved the greatest results in computation time while still providing an excellent fit.

Table 3.4: Computation time comparison of re-weighting methods

Method	RMSN	Computation time
NNLS	0.6e-4	5.89 hours
IPU	1.9e-3	2.37 minutes
NLAD	9e-4	19.75 seconds

IPU did achieve a substantial improvement in computation time over NNLS, but even IPU is still much slower than NLAD. The performance success of NLAD may be attributed to the linear form, allowing it to be solved with an efficient optimization algorithm (e.g., simplex), whereas IPU must fit *all* cells iteratively until convergence. Furthermore, the IPU computation time can greatly vary depending on the relative difference between the microdata sample and the target distribution. For example, if the proportional distribution of a microdata sample poorly fits the target marginal distribution then IPU will require additional iterations to achieve convergence; whereas NLAD will always use a similar number of steps to find a solution regardless of initial condition. The total estimated computation time for 965 census tracts is approximately 16 hours for NLAD, 105 hours for IPU, and 17,300 hours for NNLS on a single computer core. This of course can be reduced significantly if census tracts are processed on parallel computer cores, but a full run using IPU and NNLS were not performed as they would have required a substantially high computation time.

3.5.2 Person-household population generation results

Validation of the synthetic population generated using NLAD is performed at two levels: marginal totals and cells proportions. The marginal comparison measures how well the aggregated variable totals in the synthetic population fit the actual control variables in the census. The cell level comparison is a much more in depth comparison, comparing combinatorial person and household type frequencies in the synthetic population to the PUMS. This helps ensure that the actual individual person and household types (i.e., joint distributions) are properly synthesized. However, since the PUMS is only a sample, this comparison must be performed proportionally where the distribution percentages are compared. An overall marginal level comparison for individual census tracts is shown in Figure 3.5.

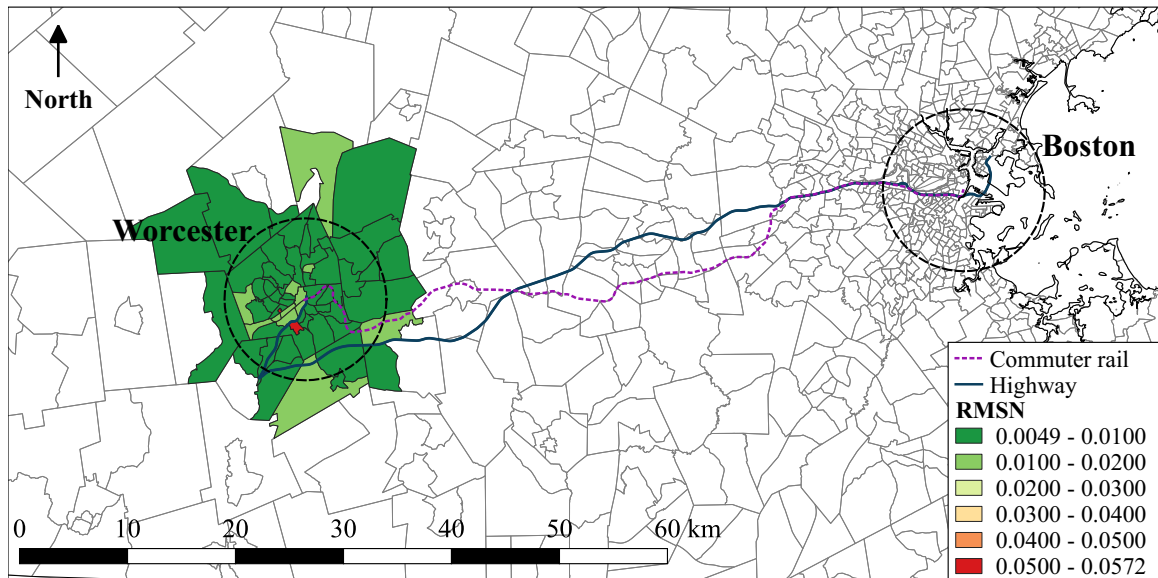


Figure 3.5: Map of marginal fit for census tracts

The marginal comparison is performed across all individual census tracts, achieving an RMSN of 0.0715 (see Figure 3.6a). These comparisons show the final realized population results (i.e., not just weighted fit) on the vertical axes, against the expected census totals shown on the horizontal axes. The cell level comparison achieved an

RMSN of 6.5779 (see Figure 3.6b). Although the cell level RMSN is much higher than at the marginal level, this is to be expected since it is a very sparse comparison with many small values magnifying error. Moreover, any error in the sample itself will be evident against the synthesized data.

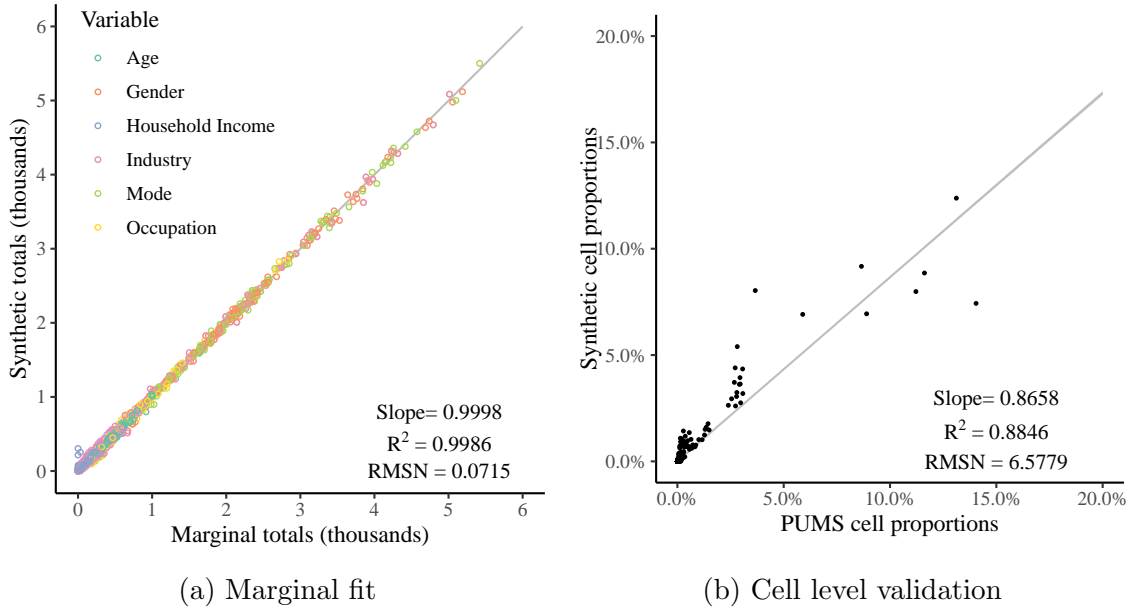


Figure 3.6: Population generation results

Using R^2 and the estimated slope as an additional measure of it, all comparisons yielded very good fit results of 0.999 and 0.8846 for the marginal and cell levels, respectively.

3.5.3 Origin-destination results

Results up to this point only considered demographic variables, not workplace assignment. A final check is to cross-validate the allocation of synthesized persons to origins and destinations. This is performed at the aggregated level by comparing the aggregated totals in the synthetic population to the actual totals in the LODES marginals. This comparison is similar to the validation for the synthetic population, but can only be performed at the aggregated level because OD microdata at this fine

grain resolution is not available. The results are well fit, with RMSN values of 0.3889 for OD pair totals. These applied workplace assignment results are visually displayed in Figure 3.7.

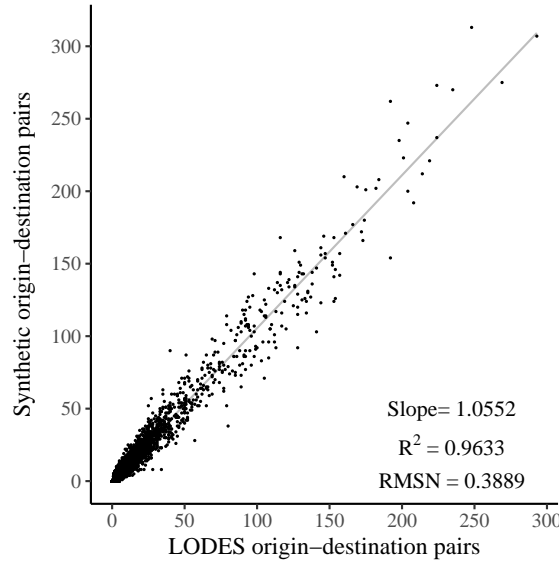


Figure 3.7: Workplace allocation by origin-destination pair

The reason that the figures are plotted on different scales is due to the variation between origin and destination totals. This is a byproduct of census tracts being delineated roughly by population size, but not by employment size; in other words, while residential location is dispersed fairly evenly, it is likely that certain census tracts (e.g., downtown) will attract a high concentration of workers and others very few. Regardless, the results show a very good fit, though it is impossible to validate further using only aggregated data. However, since industry sector is a shared variable between both the population and workplace assignment, it provides an additional level of cross validation, meaning that at the very least, persons are properly allocated to destinations based on industry. Further stratification of workplace assignment by additional variables, such as age and gender, would strengthen the validation results.

3.6 Summary of findings

As travel demand models shift towards pure activity-based models, workplace assignment is still an important input for activity-generation in state-of-the-art microscopic travel demand models. For example, many travel related activities take place in conjunction with work trips, such as shopping trips on the way home from work or picking up school-age children. Although discrete choice spatial models are possible to use, aggregated employment data is often readily available at a higher spatial resolution than in disaggregated samples, making the use of classically fit models attractive. This dissertation presents and applies an integrated population synthesis and workplace assignment method using aggregated employment data and an efficient person-housing matching method based on non-negative least deviation fitting. Such an integrated approach can be easily integrated in current common practice in existing models in the United States and elsewhere. The specific application described in this dissertation synthesized a population of 4.6-million people and 1.7-million households in the Greater Boston Area, which is ultimately utilized for an energy assessment simulation of an activity-based demand and multi-modal supply simulation (Fournier et al., 2018). The resulting population achieved a marginal level fit RMSN of 0.0715, cell level fit RMSN of 6.5779, and a workplace assignment fit RMSN of 0.3889.

The overall application for the population synthesis, workplace assignment, and person-household matching are achieved good fit results. However, there are several areas of potential refinement. The first is in including additional OD stratification variables. This synthesis presents only a single stratification variable for industry sector, but the workplace assignment can be easily modified for any number of stratifications (e.g., age and gender). Additional stratification variables would not only make the workplace assignment more disaggregated, but effectively it would add more constraints to the integrated IPF process, further improving the accuracy of the final

workplace assignment. A second area of improvement is the error dispersion noticeable for infrequent persons and household types incurred during the joint sampling process. This dispersion is visible in Figure 3.6a where the fitting points for infrequent types are more dispersed than their more common counterparts. This may be refined by using an improved sampling algorithm that does not over-penalize rare sample weights Lovelace and Ballas (2013).

An area worth further investigation is the impact of using an optimization based re-weighting approach (i.e., NLAD), as opposed to traditional proportional fitting (i.e., IPU). A well known weakness of Combinatorial Optimization (CO) population generation is that it fits sample attributes to the marginal variables, ignoring attribute associations (Pritchard and Miller, 2012). A similar limitation may be true for NLAD where an optimal solution is found using some combination of samples, but potentially eliminates redundant sample weights to zero. This raises the concern that NLAD may eliminate households that should exist in the population. However, unlike conventional CO that uses individual variables as marginals directly, the marginals in the proposed NLAD method uses combinatorial joint weights from IPF, preserving attribute association through the initial IPF step.

The proposed integrated process makes two contributions. First by exploiting the common IPF algorithm used in population synthesis and workplace assignment for an integrated method. This minimizes errors that would be introduced through independently estimated models. Second, this dissertation develops an efficient joint person-household re-weighting technique based on non-negative least absolute deviation (NLAD) fitting, substantially reducing computation time by one-sixth compare to the conventional iterative proportional updating (IPU) method. This new re-weighting technique makes the integrated process possible by being able to efficiently handle additional shared attributes in the population and workplace data (e.g., employment). The proposed technique outperforms both the conventional IPU method

and non-negative least squares (NNLS) in terms of computation time, while providing a final joint person-household population and workplace assignment.

CHAPTER 4

MULTI-MODAL SYSTEM MODEL

4.1 Concept

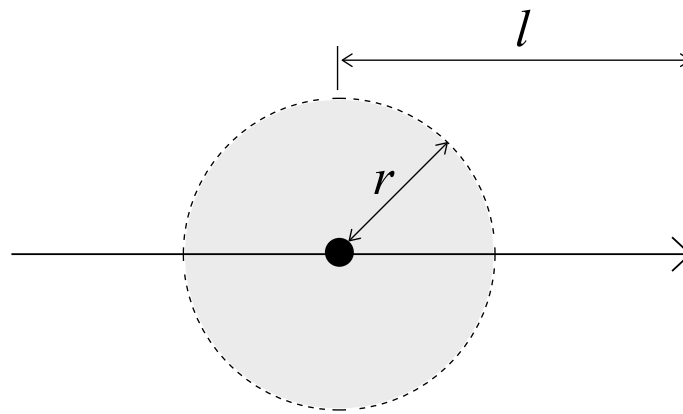


Figure 4.1: Illustration of Model Concept

The proposed model is structured for a catchment area around a central access point, which serves a mainline mode, as a simple catchment zone around an access point; see Figure 4.1. This mainline mode may be part of a larger more complex network system, but in a simplistic case it can be represented as a direct link to a central point, such as the central business district (CBD). Mode choice can then be broken down into two basic sub-models for access mode and mainline mode. The respective utility cost, or dis-utility since it is a cost not a benefit, may be formulated as

$$-U = Y + Z = \xi(t_a + \frac{r}{v_a}) + rC_a^d + C_a^f + \xi(t_m + \frac{L}{v_m}) + C_m \quad (4.1a)$$

$$-Y = \xi(t_a + \frac{r}{v_a}) + rC_a^d + C_a^f \quad (4.1b)$$

$$-Z = \xi(t_m + \frac{L}{v_m}) + C_m \quad (4.1c)$$

$a \in \{w, b, d\}$ for walk, bike, drive

$m \in \{p, h\}$ for park-n-ride and highway

where U is the joint dis-utility of the combination of an access and mainline mode, Y is the access mode dis-utility, and Z is the mainline mode dis-utility. The negativity in each case ensures that the function is a dis-utility, where a user will choose the alternative of least cost. The parameters used are as described in Table 4.1.

Table 4.1: Utility model parameters

Parameter/Variable	Unit	Description
ξ	$\$/time$	value of time (VOT)
r	$distance$	access distance
L	$distance$	mainline distance
v_a	$distance/time$	access speed
v_m	$distance/time$	mainline speed
t_a	$time$	access mode startup time
t_m	$time$	mainline mode delay time
C_a^d	$\$/distance$	unit cost per distance for access mode
C_a^f	$\$$	fixed cost for access mode
C_m	$\$$	fixed cost for mainline mode

This basic dis-utility function may be applied to independent access modes, such as walk, bike, and drive; as well as mainline modes of highway and train. A conventional approach to mode choice is to enter these dis-utility functions into the general logit function in Equation (4.2) to yield the probability P_j of the discrete mode choice $j \in J$, where J is the set of available modes, and β is an additional scaling parameter.

$$P_j = \frac{e^{-\beta U_j}}{\sum_J e^{-\beta U_j}} \quad (4.2)$$

By varying the radial distance, the linear dis-utility and choice probability may be represented visually in Figures 4.3a and 4.3b, respectively. The total demand for each mode $\lambda_a^{(s)}$ from a stochastic solution (s), is then the product of population density δ , and the area under the probability curve P_a (see Figure 4.3b). This can be achieved radially by integrating probability as a function of r from 0 to the maximum radius R using the shell integration method, expressed as

$$\lambda_a^{(s)} = 2\pi\delta \int_0^R P_a(r) dr \quad (4.3)$$

The problem, however, is that the logit function is not a closed form expression, making the optimization of pricing to incentivize demand difficult. Alternatively, the basic geometry of the model in Figure 4.2 may be exploited to approximate the logit model, while remaining in an analytical form that can be optimized efficiently.

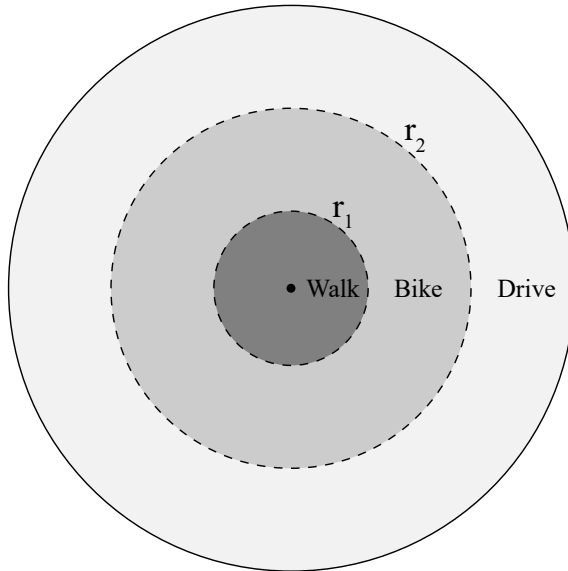


Figure 4.2: Spatial approximation of radii boundaries for access mode choice

Assuming that travelers choose the mode with the least dis-utility, all travelers between radius 0 and r_1 will walk, all travelers between r_1 and r_2 will bike, and all travelers between r_2 and R will drive in the deterministic case as shown in Figure 4.3a. Demand $\lambda_a^{(d)}$ for each access mode in the deterministic case (d) is simply the product of the population density δ , and the ring areas between the tipping points at radii r_1 and r_2 . In other words, the area under each probability curve in Figure 4.3b is approximately equal to the area formed between the respective radii that determine the boundaries of choosing that mode. For example, the sum of probabilities under the walk curve is equal to the area formed between 0 and r_1 .

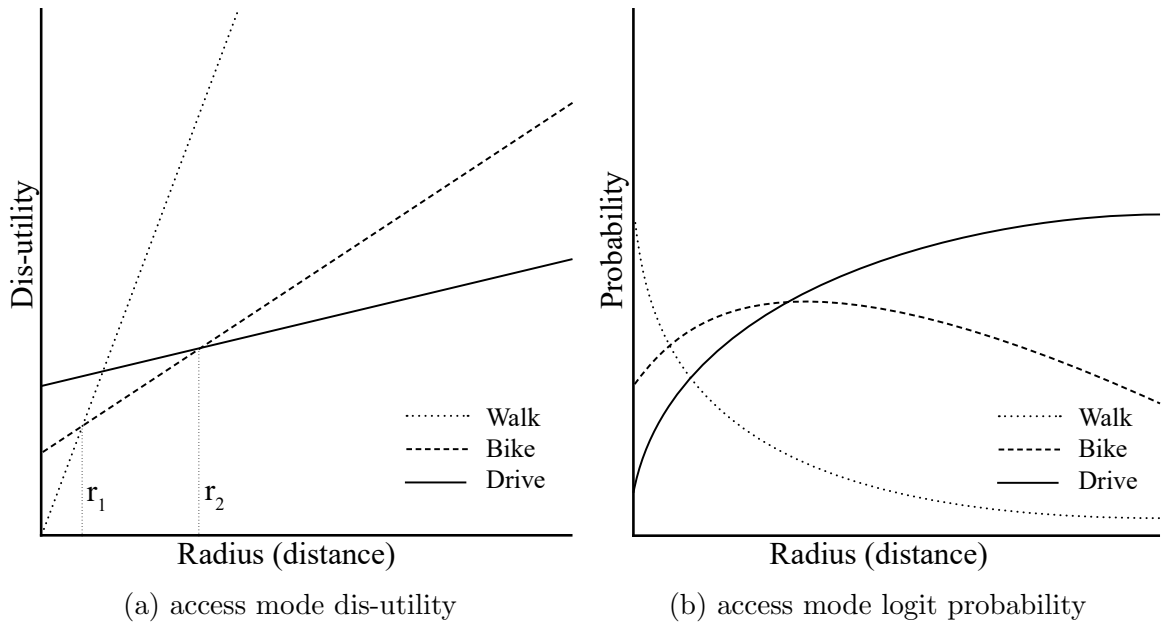


Figure 4.3: Access mode choice model comparison

4.2 Mathematical Model

Various objectives may be formulated to include other costs, such as agency cost; however, a simple case of minimizing total travel time TT , will be used for demonstration. Generically this may be formulated in Equation (4.4) as the sum of the products

of demand in number of travelers $\lambda_a^{(d)}$ and average travel time T_a , respectively, for each access mode $a \in A$, where A is the set of all access modes.

$$TT = \sum_A T_a \lambda_a^{(d)} \quad (4.4)$$

The deterministic travel demand $\lambda_a^{(d)}$, between two radii r_i and r_{i+1} , is derived from the area of a circle multiplied by the population density, as follows

$$\lambda_a = \pi \delta (r_{i+1}^2 - r_i^2) \quad (4.5)$$

The average travel time T_a is the average radial distance \bar{r}_a divided by average speed \bar{v}_a , which is assumed to be constant across the network. The average speed \bar{v}_a is calculated as the average distance divided by the sum of startup time t_a (e.g., time to unlock bicycle or start and un-park automobile) and travel time \bar{r}_a/v_a .

$$\bar{v}_a = \frac{\bar{r}_a}{t_a + \frac{\bar{r}_a}{v_a}} \quad (4.6)$$

The average radial distance \bar{r}_a may be calculated using the population's probability distribution function $f(r)$. Assuming a uniformly distributed population, the probability density function increases linearly with r , $f(r) = kr$. The value of k is determined using the law of total probability to solve $\int_{r_{min}}^{r_{max}} kr \, dr = 1$, implying that the probability density function is

$$f(r) = \frac{2r}{r_{max}^2 - r_{min}^2} \quad \text{for } r \in [r_{min}, r_{max}] \quad (4.7)$$

The expected radius is $E(r) = \int r f(r) dr = \int k r^2 dr$, which yields a formula for average radius, \bar{r}_a , for points within a ring assuming uniform density. For access mode a serving the population between r_i and r_{i+1} , the average distance is

$$\bar{r}_a = \frac{2(r_{i+1}^3 - r_i^3)}{3(r_{i+1}^2 - r_i^2)} \quad (4.8)$$

Recalling that average travel time is $T_a = \bar{r}_a/\bar{v}_a$, and substituting for \bar{r}_a and \bar{v}_a , a formula for travel time as a function of r_i and r_{i+1} is expressed as

$$T_a = \frac{\bar{r}_a}{\bar{v}_a} = \left[t_a + \frac{2(r_{i+1}^3 - r_i^3)}{3v_a(r_{i+1}^2 - r_i^2)} \right] \quad (4.9)$$

which can then be used to formulate an objective function for minimizing total travel time of all users in the system

$$TT_a(r_i, r_{i+1}) = \lambda_a T_a = \pi\delta \left[t_f(r_{i+1}^2 - r_i^2) + \frac{2}{3v_a}(r_{i+1}^3 - r_i^3) \right] \quad (4.10)$$

where t_f is the fixed additional travel time specific to the combination of access and mainline mode (i.e., startup time t_a plus mainline travel time T_m), and r_i and r_{i+1} are the decision variables. The radii subscripts i and $i+1$ reflect whether the relevant distances are 0 to r_1 for walk, r_1 to r_2 for bike, or r_2 to R for drive. Global monetary costs (e.g., tolls, fares, parking, etc.) experienced by the user may also be included into the model to express the generalized cost. This can be done by converting monetary costs to generalized time (GT) cost using the VOT as in the expression

$$GT_a(r_i, r_{i+1}) = \lambda_a \left[T_a + \frac{\bar{r}C_a^d + C_a^f}{\xi_a} \right] \quad (4.11)$$

$$= \pi\delta \left[t_f(r_{i+1}^2 - r_i^2) + \frac{2}{3}(r_{i+1}^3 - r_i^3) \left(\frac{1}{v_a} + \frac{C_a^d}{\xi} \right) + \frac{C_a^f}{\xi} \right] \quad (4.12)$$

where C_a^d and C_a^f are the existing fixed or distance-based monetary costs. A final caveat of this model is that a basic hierarchical constraint must be imposed such that

$$\text{Vehicle speed: } C_W + \frac{1}{v_W} \geq C_B^d + \frac{1}{v_B} \geq C_D^d + \frac{1}{v_D} \quad (4.13a)$$

$$\text{Vehicle startup time: } C_W^f + t_W \leq C_B^f + t_B \leq C_D^f + t_D \quad (4.13b)$$

meaning that the distance *cost* of walking is greater than that of biking, which is greater than that of driving, and the fixed *cost* of driving is greater than that of biking, which is greater than that of walking. With the given linear constraints, Equations (4.10) and (4.11) are convex and can be efficiently optimized to obtain a global optimum solution.

4.3 Conditional choice of mainline mode

Since highway travel time ultimately depends upon the number of drivers that use the highway λ_H rather than park-and-ride to the train λ_D , the mode-share $\theta = \frac{\lambda_H}{\lambda_D}$ of highway drivers to all drivers must also be determined. This may be conceptualized spatially in Figure 4.4a where the portion of walk, bike, and drive are determined by the radii r_1 and r_2 , but highway and park-and-ride are subsequently split within the drive portion. This can be thought of as analogous to the nested logit illustrated in Figure 4.4b.

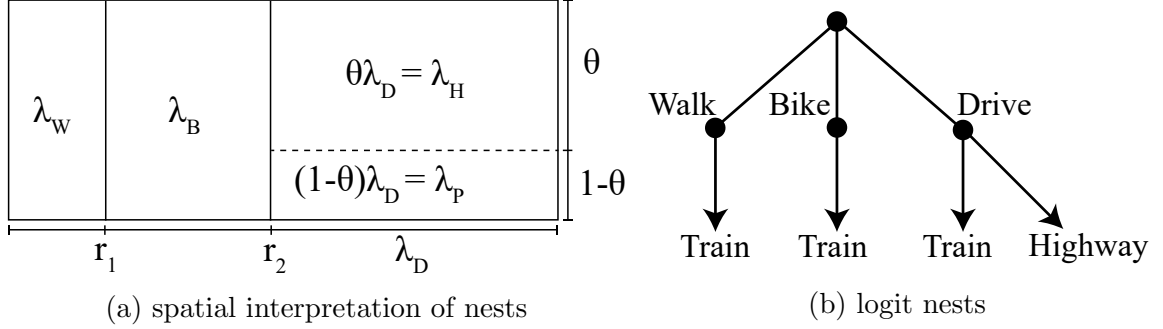


Figure 4.4: Nested mode choice

With T_H as highway travel time, and T_T as the train travel time, the total travel time for driving is estimated as in Equation (4.14). It is essentially the product of the total drive demand λ_D and the sum of the proportional mainline travel times θT_H , $(1 - \theta)T_T$, and drive access time T_D .

$$TT_D = \lambda_D [\theta T_H + (1 - \theta)T_T + T_D] \quad (4.14)$$

4.4 Congestion

Thus far in the dissertation, congestion has been absent from the model. To be more realistic, once the mainline mode is incorporated, a model for congested mainline travel time must be considered. The following section describes this formulation and justification for utilizing a simpler continuous function for congested travel time.

Mainline dis-utility is a fixed constant time cost that may be added to t_a in Equation (4.10). The effect of the mainline mode is a vertical shift in dis-utility while the slope remains unchanged. This provides a generic input for the mainline travel time, allowing for flexibility in mainline travel time estimation. Assuming that train transit is reliable and does not experience delay due to congestion, then the train

travel time T_T in Equation (4.15), depends on the distance, L ; speed, v_T ; and any fixed delay, t_T , such as the waiting time associated with transit headway.

$$T_T = t_T + \frac{L}{v_T} \quad (4.15)$$

The travel time for driving, which is a congestible mode, requires accounting for the network's capacity to move car traffic and can account for the effect of traffic volumes on speed. The Macroscopic Fundamental Diagram (MFD) (see Figure 4.5a), provides an aggregate representation of network traffic conditions that can be used to characterize traffic speeds as a function of traffic flow q in vehicles per time, and density k in vehicles distance.

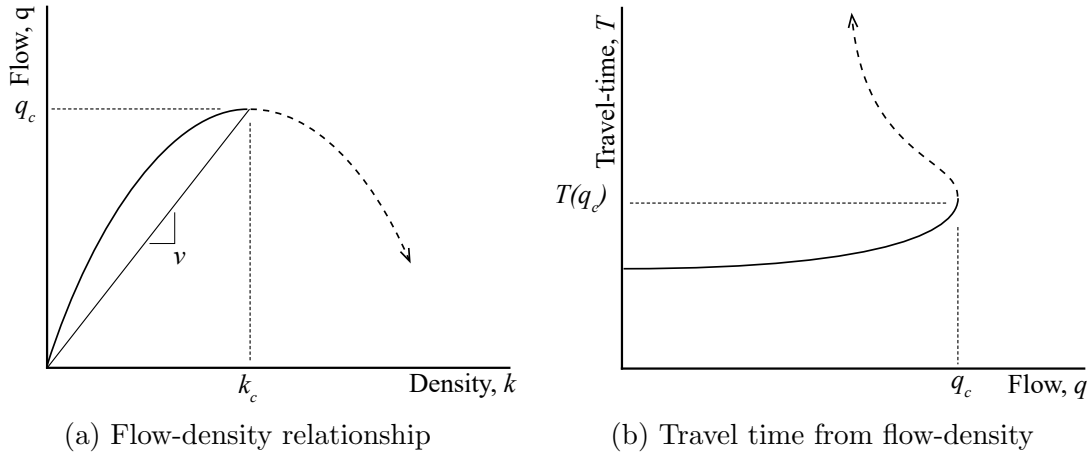


Figure 4.5: Relationship between flow, density, and travel time

A generic mathematical model for the flow-density relationship shown in Figure 4.5a is a function $q(k)$, where a and b are model fitting parameters subject to boundary constraints. The average travel time, including the effects of congestion, can then be derived by solving for k and substituting into the inverse of speed $\frac{1}{v} = \frac{k}{q}$ as

$$T_H = f(q) = \frac{1}{v} = \frac{k(q)}{q} \quad (4.16)$$

where T_H is in units of time per distance per lane (e.g., hours/km/lane). However, the resulting travel time function is multi-valued for each flow (see Figure 4.5b), complicating the optimization. Steady-state conditions can only be maintained for highway demand that does not exceed the network capacity, $\lambda_H \leq q_c$, but each flow is associated with an “uncongested” state with density less than the critical density associated with capacity, k_c , and a “congested” state with density exceeding k_c , as shown in Figure 4.5a.¹

Since the overall objective is to identify system optimum pricing strategies, travel times can be modeled using the bottom (solid) part of the $T_H(q)$ curve, because uncongested conditions use fewer resources to serve traffic flow and are therefore always preferable to congestion. Pricing strategies should aim at maintaining uncongested conditions. Thus, travel time for highway users can be modeled using a continuous non-decreasing function of demand $T_H(\lambda_H)$ for cases when pricing or demand management will keep demand from exceeding the highway’s capacity.

For the purposes of demonstrating the application of the model, the travel time function for the highway between Worcester and Boston is characterized by Equation (4.17),

$$T_H(\lambda_H) = t_H + \frac{L}{v_H} \left[1 + \alpha \left(\frac{\lambda_H}{\gamma} \right)^\phi \right] \quad (4.17)$$

where λ_H is the highway demand, and α , ϕ , and γ are model parameters. Although a hard constraint for capacity is not set, it is assumed that the model results will never

¹these regions are referred to as “congested” and “hypercongested,” respectively, in the economics literature (Gonzales, 2015).

exceed capacity. This expression represents only the uncongested lower part of the travel time vs. flow relation, illustrated in Figure 4.5b.

4.5 Optimal Mode Split

Mainline travel times for train and highway in Equations (4.15) and $T_H(\lambda_H)$ can be incorporated into the respective access travel time functions in Equations (4.11) and (4.14). Once expanded to include mainline travel time, the walk, bike, and drive travel times are calculated as in Equations (4.18a-4.18c), respectively.

$$GT_W(0, r_1) = \pi \delta \left[(r_1^2 - 0^2) \left(t_W + t_T + \frac{L}{v_T} \right) + \frac{2}{3} (r_1^3 - 0^3) \left(\frac{1}{v_W} + \frac{C_W^d}{\xi} \right) + \frac{C_W^f + C_T^f}{\xi} \right] \quad (4.18a)$$

$$GT_B(r_1, r_2) = \pi \delta \left[(r_2^2 - r_1^2) \left(t_B + t_T + \frac{L}{v_T} \right) + \frac{2}{3} (r_2^3 - r_1^3) \left(\frac{1}{v_B} + \frac{C_B^d}{\xi} \right) + \frac{C_B^f + C_T^f}{\xi} \right] \quad (4.18b)$$

$$GT_D(r_2, \theta) = \pi \delta \left[(R^2 - r_2^2) \left(t_D + \theta T_H(\lambda_H) + (1 - \theta) \left(t_T + \frac{L}{v_T} \right) \right) + \frac{2}{3} (R^3 - r_2^3) \left(\frac{1}{v_D} + \frac{C_D^d}{\xi} \right) + \frac{C_D^f + \theta C_H^f + (1 - \theta) C_T^f}{\xi} \right] \quad (4.18c)$$

where the highway travel time T_H is a function of highway demand λ_H . Highway demand is calculated as the portion of total drive access users taking the highway given by the θ ratio, calculated as

$$\lambda_H = \theta \lambda_D = \theta \pi \delta (R^2 - r_2^2) \quad (4.19)$$

The final objective function then becomes the summation of the three generalized travel time functions with three decision variables of r_1 , r_2 , and θ in Equation (4.20).

$$\min GT(r_1, r_2, \theta) = GT_W(0, r_1) + GT_B(r_1, r_2) + GT_D(r_2, \theta) \quad (4.20)$$

The optimal parameters, \hat{r}_1 , \hat{r}_2 , and $\hat{\theta}$ yield the allocation of demand to access and mainline modes that minimizes the generalized total travel time while accounting for congestion. All components TT_W , TT_B , and TT_D are convex, thus the summation is

also convex, providing an objective function that can be efficiently solved to obtain the global optimum.

4.6 Pricing

To achieve the optimal mode share determined by Equation (4.20), society might need to impose additional policies to incentivize users through the intervention price parameters \hat{C}_a^f , \hat{C}_a^d , and \hat{C}_m . This would adjust the dis-utility functions shown in Figure 4.3a such that the intersections for r_1 and r_2 are shifted from their observed position to match the optimal values (i.e., \hat{r}_1 and \hat{r}_2). First the observed “effective” radii r' and highway ratio θ' must be determined from an observed mode share P' . This can be done using the proportions in Figure 4.4a from the proportional demand P_{am} and the known total radius R to calculate the effective proportional radii r'_1 and r'_2 , expressed more fully as

$$r'_1 = P'_{WT} \cdot R \quad (4.21a)$$

$$r'_2 = (P'_{WT} + P'_{BT}) \cdot R \quad (4.21b)$$

$$\theta' = \frac{P'_{DH}}{P'_{DT} + P'_{DH}} \quad (4.21c)$$

4.6.1 Access Mode Pricing

The point of intersection can be solved when the costs are known by setting Equation (4.1b) for the access modes' dis-utilities equal to each other (i.e., $Y_a = Y_{a+1}$) and solving for r_i . This yields Equation (4.24). Without any intervention, the given parameters yields a null tipping point radii r_i° calculated as

$$r_i^\circ = \frac{\xi (t_{a+1} - t_a) + C_{a+1}^f - C_a^f}{\xi \left(\frac{1}{v_a} - \frac{1}{v_{a+1}} \right) + C_a^d - C_{a+1}^d} \quad (4.22)$$

where i is the tipping point radius between incremental access modes a and $a+1$ (i.e., r_1 is tipping point between walk and bike and r_2 is the tipping point between bike and drive). Incorporating intervention prices \hat{C}_a^f and \hat{C}_a^d yields the cost *differential* needed to shift from the null radii r_i° , to some target radii r_i^* . However, we seek to find the differential from the optimal to the existing, not from the null. This total differential is simply calculated as

$$r_i^* = r_i' + r_i^\circ - \hat{r}_i \quad (4.23)$$

where $\hat{r}_i - r_i^\circ$ accounts for the difference between the null and optimal condition, which can then be used to find the differential intervention prices with

$$r_i^* = \frac{\xi(t_{a+1} - t_a) + C_{a+1}^f - C_a^f + \hat{C}_{a+1}^f - \hat{C}_a^f}{\xi\left(\frac{1}{v_a} - \frac{1}{v_{a+1}}\right) + C_a^d - C_{a+1}^d + \hat{C}_a^d - \hat{C}_{a+1}^d} \quad (4.24)$$

However, when the intervention price variables are not known, an “optimal” set of prices must be found to achieve the desired radii. This can be done by reformulating Equation (4.24) for r_1 and r_2 as constraints in an optimization problem as shown in Equations (4.25b-4.25c). An infinite number of solutions may exist, thus a reasonable objective function in Equation (4.25a) might be to minimize the square magnitude of costs to shift the radii as in Equation (4.25a). From a social perspective, this would minimize the quantity of money changing hands. Equations (4.25d-4.25e) impose the hierarchical constraints required for the deterministic model from Equation (4.13). An additional constraint could also be set to maintain revenue neutrality (Equation 4.25f).

$$\min \sum_J C_j^2 \quad (4.25a)$$

$$\text{s.t. } r_1^* \hat{C}_W^d - r_1^* \hat{C}_B^d + \hat{C}_W^f - \hat{C}_B^f = \xi \left[t_B - t_W + r_1^* \left(\frac{1}{v_B} - \frac{1}{v_W} \right) \right] + r_1^* (C_W^d - C_B^d) + C_B^f - C_W^f \quad (4.25b)$$

$$r_2^* \hat{C}_B^d - r_2^* \hat{C}_D^d + \hat{C}_B^f - \hat{C}_D^f = \xi \left[t_D - t_B + r_2^* \left(\frac{1}{v_D} - \frac{1}{v_B} \right) \right] + r_2^* (C_B^d - C_D^d) C_D^f - C_B^f \quad (4.25c)$$

$$C_W^d + \hat{C}_W^d + \frac{\xi}{v_W} \geq C_B^d + \hat{C}_B^d + \frac{\xi}{v_B} \geq C_D^d + \hat{C}_D^d + \frac{\xi}{v_D} \quad (4.25d)$$

$$C_W^f + \hat{C}_W^f + \xi t_W \leq C_B^f + \hat{C}_B^f + \xi t_B \leq C_D^f + \hat{C}_D^f + \xi t_D \quad (4.25e)$$

$$\sum C_j = 0 \quad (4.25f)$$

4.6.2 Mainline Mode Pricing

The optimal prices from solving Equation (4.25) yield only the access mode costs, the equation does not yet consider the two mainline modes. Recalling the conditional mode choice model described in Section 4.3 where $\theta = \lambda_H/\lambda_D$, the analogous nested logit may be decomposed to extract the cost differential necessary to achieve the desired mode split θ . Assuming revenue neutrality, the individual mainline mode prices for highway C_H , and train C_T , can be analytically determined. The following section describes this decomposition to yield highway and train prices. Although these prices may be implemented as a distance based cost as well as a fixed cost from a mainline perspective (e.g., distance-based tolling), only a fixed cost parameter is necessary, because the total cost associated with the mainline part of the trip will always be the same for mainline distance L .

Walk and bike access users are limited to only the train for mainline mode, so determination of the relevant radii is sufficient to determine the number traveling by each mode. Drive access users have an additional choice between train or highway, so in addition to determining the number of drivers by radius, they must be further distinguished by mainline mode. A deterministic model would assign all trips to the

lowest cost mainline mode unless the costs were exactly equal. For the model to predict non-zero numbers of highway and transit users in uncongested conditions, a stochastic logit model is needed. Unlike the access mode, everyone travels the same mainline distance, so this does not introduce the same computational complexity. In the logit model, this introduces the problem of independence of irrelevant alternatives (IIA). A common solution to this problem is to nest the choices using a nested logit model as in Equations (4.26).

$$P_m = P_a \cdot P_{m|a} \quad (4.26a)$$

$$P_a = \frac{e^{-Y_a}}{\sum_J e^{-Y_j}} \quad (4.26b)$$

$$P_{m|a} = \frac{e^{-Z_m + \mu IV_m}}{\sum_K e^{-Z_k + \mu IV_k}} \quad (4.26c)$$

$$IV_m = \ln \left(\sum e^{-Y_a} \right) \quad (4.26d)$$

where P_m is the probability of the mainline nest (i.e. highway or train), $P_{a|m}$ is the probability of access mode a given a mainline nest m , IV_m is the logsum for nest m , and μ is the logsum coefficient. When $\mu = 1$, there is no correlation, and when $\mu = 0$, there is perfect correlation among alternatives. It can be assumed that the choices are perfectly correlated because highway is exclusive to driving, reducing the conditional probability of choosing highway given driving as the access mode to $P_{H|D} = \frac{e^{-Z_H}}{e^{-Z_H} + e^{-Z_T}}$ and the conditional probability of choosing train given walk or bike access modes to $P_{W|T} = P_{B|T} = 1$, because users that user walk or bike as their access mode must take the train. Thus, $P_H = P_D \cdot P_{H|D} = \frac{\lambda_H}{\lambda} = \frac{\lambda_D}{\lambda} \cdot \frac{e^{-Z_H}}{e^{-Z_H} + e^{-Z_T}}$ the logit can be solved to extract the cost differential $\Delta C_m = C_H - C_T$ as in Equation (4.27a). Subsequently if revenue neutrality is maintained across the mainline modes as well, then $\lambda_H C_H + \lambda_T C_T = 0$ may be used to solve for the specific cost values.

$$\hat{C}_H - \hat{C}_T = \xi(t_T - t_H) - \ln\left(\frac{\lambda_H}{\lambda_D - \lambda_H}\right) + C_T - C_H \quad (4.27a)$$

$$0 = \lambda_H(C_H + \hat{C}_H) + \lambda_T(C_T + \hat{C}_T) \quad (4.27b)$$

$$\hat{C}_H = \left[\xi(t_T - t_H) - \ln\left(\frac{\lambda_H}{\lambda_D - \lambda_H}\right) + 2C_T - C_H \left(\frac{\lambda_H}{\lambda_T} + 1\right) \right] \left(\frac{\lambda_T}{\lambda_T + \lambda_H}\right) \quad (4.27c)$$

4.6.3 Pricing Implementation

The flexibility of the model's pricing inputs as both fixed and distance-based prices provides potential for multi-modal integration. Fixed costs may be attributed to more classical features, such as parking, transit fare, bridge tolls, etc., whereas distance based costs can help integrate future technologies, such as gantry mounted all electronic tolling (AET), Vehicle Miles Traveled (VMT) fees, or even bike-share and e-hailing taxi fares. Moreover, with increasingly inter-operable payment systems, it is possible to promote an efficient multi-modal pricing system through price collaboration, such as providing a discount when transferring to transit.

4.7 Income, Equity, and the Value of Time

Thus far the fundamental mode choice model uses a homogeneous population with a single VOT. This is not only unrealistic, but poses equity concerns considering that VOT is often correlated with income (Small et al., 2005; Hensher, 1976; Li and Hensher, 2012; Hensher, 2001). To ensure that the transportation system is both efficient and equitable, the VOT must also be a parameter to account for varying distributions of income or preferences. Since Equation (4.20) is convex, further summation remains convex, enabling the population to be segmented into any number of discrete population groups optimized for the total generalized travel time in Equation (4.28).

$$\min \sum_K GT_k(r_{1k}, r_{2k}, \theta_k) \quad (4.28)$$

The population may then be segmented by VOT ξ_k for K discrete segments with a population density of δ_k , modeled by some population probability density function. Such a function might follow a log-normal distribution commonly used to model income distributions. Similarly, the pricing policy can also be optimized by extending Equation (4.25) with a set of access pricing variables C_{ak}^d , C_{ak}^f , and the mainline prices of C_{Hk} and C_{Tk} , associated with each population segment k with Equation (4.29). The same constraints hold, but with a corresponding set for each k .

$$\min \sum_K \sum_J C_{jk}^2 \quad (4.29a)$$

$$\text{s.t. } r_{1k}^* \hat{C}_{Wk}^d - r_{1k}^* \hat{C}_{Bk}^d + \hat{C}_{Wk}^f - \hat{C}_{Bk}^f = \xi_k \left[t_B - t_W + r_{1k}^* \left(\frac{1}{v_B} - \frac{1}{v_W} \right) \right] + r_{1k}^* \left(C_{Wk}^d - C_{Bk}^d \right) + C_{Bk}^f - C_{Wk}^f \quad (4.29b)$$

$$r_{2k}^* \hat{C}_{Bk}^d - r_{2k}^* \hat{C}_{Dk}^d + \hat{C}_{Bk}^f - \hat{C}_{Dk}^f = \xi_k \left[t_D - t_B + r_{2k}^* \left(\frac{1}{v_D} - \frac{1}{v_B} \right) \right] + r_{2k}^* \left(C_{Bk}^d - C_{Dk}^d \right) + C_{Dk}^f - C_{Bk}^f \quad (4.29c)$$

$$C_{Wk}^d + \hat{C}_{Wk}^d + \frac{\xi_k}{v_W} \geq C_{Bk}^d + \hat{C}_{Bk}^d + \frac{\xi_k}{v_B} \geq C_{Dk}^d + \hat{C}_{Dk}^d + \frac{\xi_k}{v_D} \quad (4.29d)$$

$$C_{Wk}^f + \hat{C}_{Wk}^f + \xi_k t_W \leq C_{Bk}^f + \hat{C}_{Bk}^f + \xi_k t_B \leq C_{Dk}^f + \hat{C}_{Dk}^f + \xi_k t_D \quad (4.29e)$$

$$\sum C_{jk} = 0 \quad (4.29f)$$

4.7.1 Estimating value of time

Empirical estimation of the VOT in transport is typically achieved by comparing the coefficients β , for time and cost fitted in a discrete choice model (Hensher, 1976; Li and Hensher, 2012; Hensher, 2001; Small et al., 2005). Estimating VOT varying across income is accomplished by estimating a multinomial logit choice model with income as fixed effects,

$$V_{ijk} = \alpha_j + \beta_k^{time} t_{ik} + \beta_k^{\$} C_{ik} \quad (4.30)$$

where V_{ijk} is the mode choice utility, α_j is the intercept for choice alternative j , and C_{ik} and t_{ik} are the costs imposed on the individual i in money and time, respectively; with income group k as the fixed effect in the model. From this, discrete VOT can be calculated for each income group k as

$$\xi_k = \frac{\beta_k^{time}}{\beta_k^{\$}} \quad (4.31)$$

4.7.2 Extrapolating value of time to full population

In order to provide any number of discrete segment breaks in a VOT distribution beyond what is available in sample data, a smooth continuous distribution function can be fitted to a sample. However, such a sample distribution does not exist and must be estimated. This is done by further modeling the correlation between income and VOT itself. A simple linear model would follow the form $\xi = \alpha + \beta Income$, where the estimated VOT is merely a scaled model of income with a fixed intercept. Applying the linear VOT-income model to a sufficiently large sample distribution of income, the distribution can be fitted using the appropriate parametric model, such as a log-normal distribution in Equation (4.32).

$$f(k) = \frac{1}{k\sigma\sqrt{2\pi}} e^{-\frac{(\ln(k) - \mu)^2}{2\sigma^2}} \quad (4.32)$$

where μ is the mean and σ is the standard deviation for the probability density function $f(k)$, for discrete segment k . Once fitted, any scale and number of discrete segments may be created from the distribution for analysis.

The value of time (VOT) was modeled by first estimating the VOT for eight discrete household income groups using data from the Massachusetts Household Travel Survey (MTS) (Massachusetts Department of Transportation, 2012). The monetary costs for the respective modes and their alternatives were determined by either im-

puting known prices based on location for commuter rail and transit fares, or by using travel distance with an average cost of \$0.40 per km (\approx \$0.64 per mile) for driving. The travel times were obtained using the estimated travel time matrices from the Boston Region Metropolitan Planning Organization (MPO)/Central Transportation Planning Staff (CTPS) (Massachusetts Central Transportation Planning Staff, 2018). Once a VOT was determined for each of the income groups, a linear model was constructed using the median income value of each group.

This VOT model was then applied to a full synthetic population for the Worcester region. This larger distribution was then used to fit a continuous log-normal distribution used to model population density δ_k for k income groups. The purpose of using a synthetic population is to not only create a value of time distribution, but to determine the joint distribution of travelers commuting to Boston, their value of time, and their current mode choice.

4.7.3 Gini equity coefficient measure

Accounting for varying income groups presents the prospect of evaluating equity in the system, in addition to efficiency. Though there are many different methods to measure equity, a very common and universally accepted measure is the Gini coefficient (Gini, 1912; Litchfield, 1999). In addition to its popularity, the Gini coefficient can be easily calculated for a discrete probability distribution directly, rather than aggregating individuals in a population. The Gini coefficient is calculated with Equations (4.33).

$$G = \frac{1}{2\mu} \sum_{k=1}^n \sum_{h=1}^n f(y_k)f(y_h)|y_k - y_h| \quad (4.33a)$$

$$\mu = \sum_{i=1}^n y_k f(y_k) \quad (4.33b)$$

where G is the Gini coefficient which ranges from 0 to 1, 0 being perfectly equitable and 1 being perfectly inequitable. $f(y_k)$ is the discrete probability function for the percent of the population with income y_k . The Gini coefficient is calculated as one half the total absolute relative difference in income normalized by sum of all costs in the population, μ . Instead of income, average generalized cost \overline{GT}_k is used for each VOT group k .

4.8 Data

As an application of the proposed model, an ideal case study is the Worcester to Boston commute in Massachusetts, USA; shown in Figure 4.6. Worcester is a satellite city to Boston with competing mainline transportation modes of tolled highway and a commuter rail line. Both the highway and commuter rail line originate at relatively central locations in Worcester and Boston. A summary of the basic parameters used for this numerical example are detailed in Table 4.2, which includes both the general system parameters, as well as vehicle specific parameters. All parameter values are approximate for demonstration purposes.

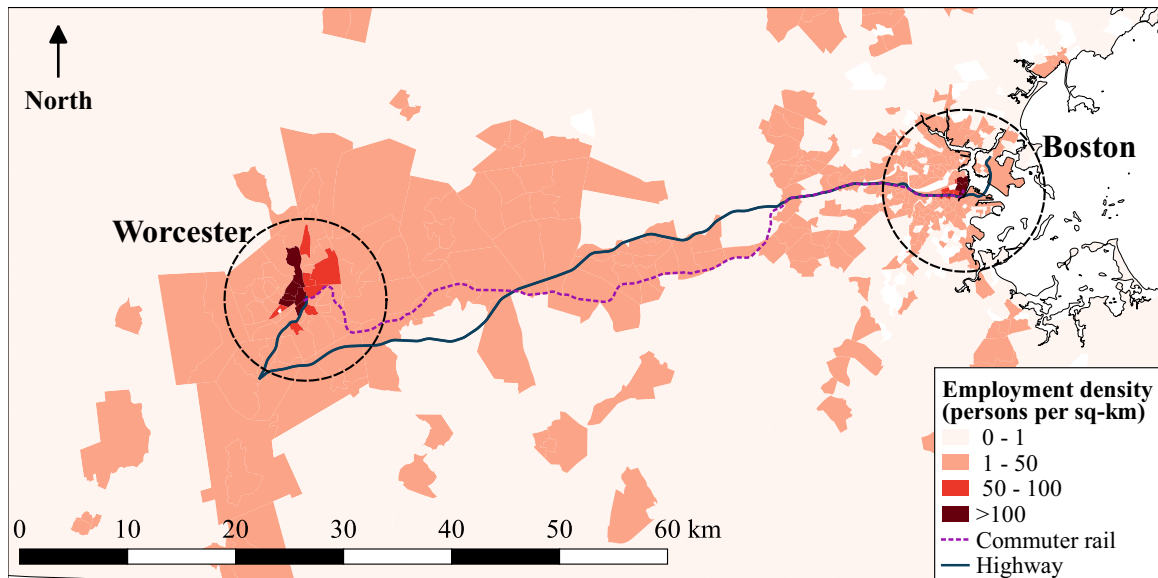


Figure 4.6: Worcester residents employment density

Table 4.2: Worcester–Boston commuter model parameters

System Parameters								
Parameter	Value	Units	Description					
δ	49.14	<i>pax/km²</i>	Average commuting population density					
ξ	17.21	<i>\$/hour</i>	Mean value of time (VOT)					
R	7.5	<i>km</i>	Approximate access radius					
L	70	<i>km</i>	Approximate distance to Boston					
α	0.1	–	Congestion model parameter					
ϕ	4	–	Congestion model parameter					
γ	4,400	–	Highway throughput parameter					
Commuters	3%	–	Share of population that travels from Worcester to Boston					
Boardings	1,000	<i>pax</i>	Average daily passenger boardings at Worcester					
Auto parking	700	<i>pax</i>	Number of auto parking spaces available					
Bike parking	30	<i>pax</i>	Number of bike parking spaces					
Walk access	270	<i>pax</i>	Remaining number of passengers boarding minus parking					
Vehicle Specific Parameters								
Parameter		Units	Description	Walk	Bike	Drive	Train	Highway
v		<i>km/hour</i>	Speed	5	15	40	85	100
t_0		hour	Startup time	0	$\frac{5}{60}$	$\frac{10}{60}$	0	0
C^d		<i>\$/km</i>	existing distance prices	0	0.1	0.4	0	0.4
C^f		\$	existing fixed prices	0	0	5	11.5	2.9

The value of time (VOT) was modeled by first estimating the VOT for eight discrete household income groups using data from the Massachusetts Household Travel Survey (MTS) containing 61,777 trips reported by 15,828 travelers (Massachusetts Department of Transportation, 2012). The monetary costs for the respective modes and their alternatives were determined by either imputing known prices based on location for commuter rail and transit fares, or by using travel distance with an average cost of \$0.40 per km (\approx \$0.64 per mile) for driving. The travel times were obtained using the estimated travel time matrices from the Boston Region Metropolitan Planning Organization (MPO)/Central Transportation Planning Staff (CTPS) (Massachusetts Central Transportation Planning Staff, 2018).

4.9 Evaluation

This findings section is split into two main parts using data for the Worcester–Boston commute example case. First, the deterministic model itself is validated

against a logit-based counterpart for a simplified case with a single average VOT. Second, the deterministic model is extended to include multiple values of time across the distribution.

4.9.1 Value of time

The value of time (VOT) was modeled by first estimating the VOT for eight discrete household income groups using data from the Massachusetts Household Travel Survey (MTS) containing 61,777 trips reported by 15,828 travelers. The monetary costs for the respective modes and their alternatives were determined by either imputing known prices based on location for commuter rail and transit fares, or by using travel distance with an average cost of \$0.40 per km (\approx \$0.64 per mile) for driving. The travel times were obtained using the estimated travel time matrices from the Boston Region Metropolitan Planning Organization (MPO)/Central Transportation Planning Staff (CTPS). Once a VOT was determined for each of the income groups, a linear model was constructed using the median income value of each group, shown in Figure 4.7.

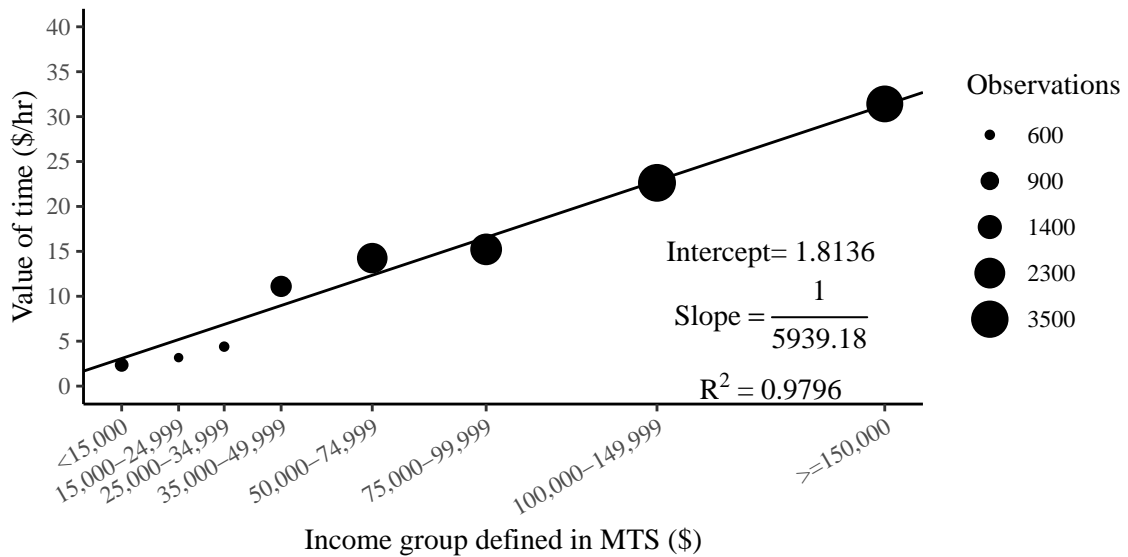


Figure 4.7: Value of time and income in MTS

This VOT model was then applied to a much larger data set using the 2016 Massachusetts Public Use Microdata Sample (PUMS) containing 343,615 individuals, provided by the U.S. Census (U.S. Census Bureau American Community Survey, 2015). This larger distribution was then used to fit a continuous log-normal distribution used to model population density δ_k for k income groups, shown in Figure 4.8. The mean VOT was determined to be \$17.21 per hour, with a standard deviation of \$1.93 per hour. This mean value is used as the VOT for the single VOT validation example.

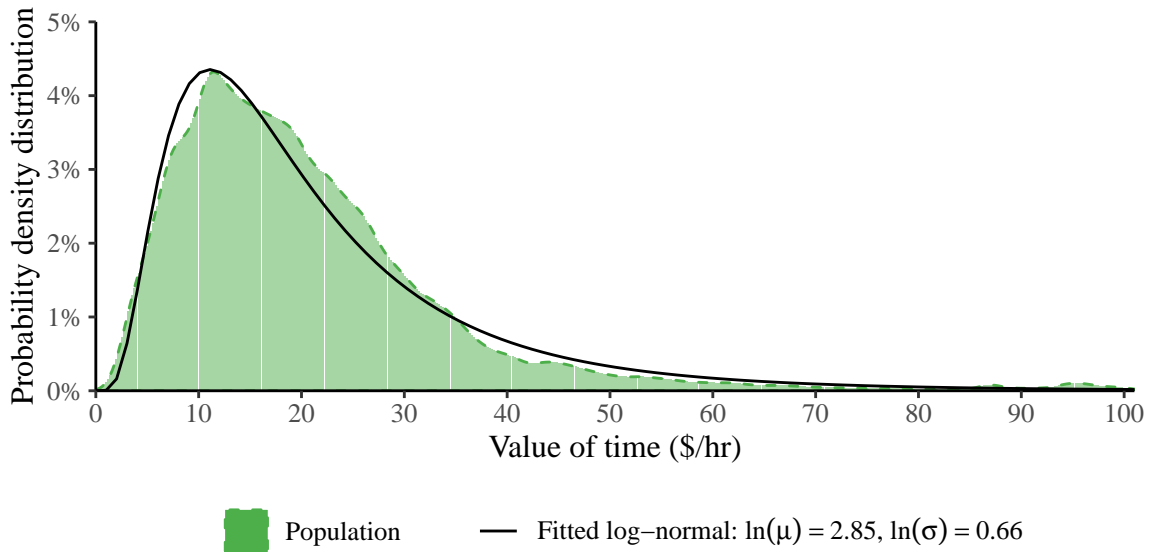


Figure 4.8: Value of time distribution in PUMS

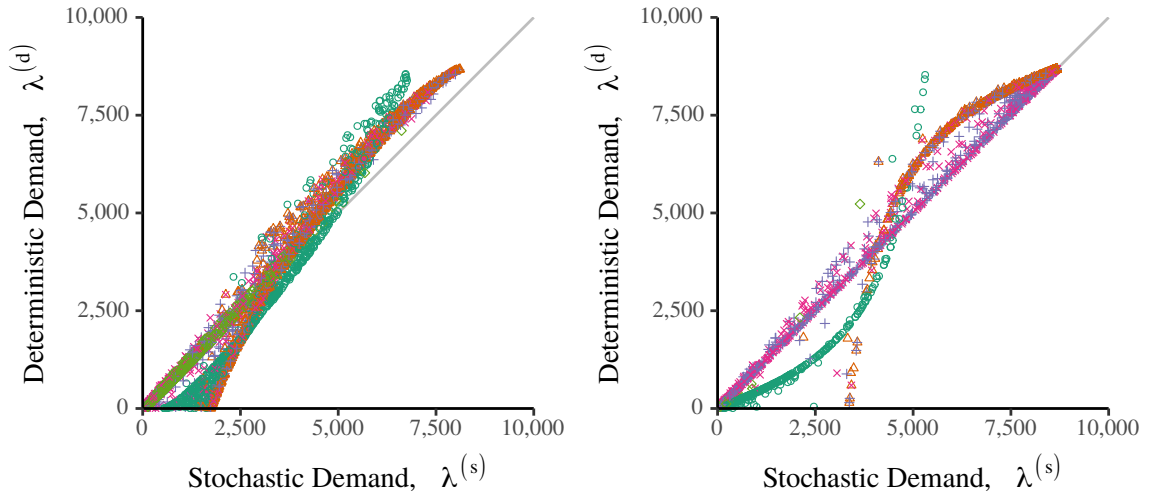
4.9.2 Validation

The deterministic model is validated by directly comparing estimated demand for each mode using the deterministic and logit methods, $\lambda^{(d)}$ and $\lambda^{(s)}$, respectively. Since the logit model does not provide a closed form solution for the access mode choice, a comparison must be evaluated numerically. This is done by generating 10,000 random price combinations ranging between -\$10 to \$10 and comparing the demand resulting from the deterministic model and the stochastic model. Figure 4.9

is a numerical comparison between the deterministic approximation and the stochastic logit model. The plot is generated by drawing 10,000 random price combinations ranging between -\$10 to \$10 to reveal the functional relationship. To determine the effect of fixed prices, C_a^f , and distance prices, C_a^d , on the model, the numerical comparison is performed three times. Once with only fixed prices varying (Figure 4.9a), another with only distance prices varying (Figure 4.9b), and a third with both prices varying (Figure 4.9c). To determine the effects of VOT on the model, this analysis is then repeated across a range of VOTs, ranging from 0 to 50 with 5 $\$/hour$ increments. Percent root-mean square error (RMSE) is used as an overall measure of fit for each of the price and VOT sets, expressed in Equation (4.34) as

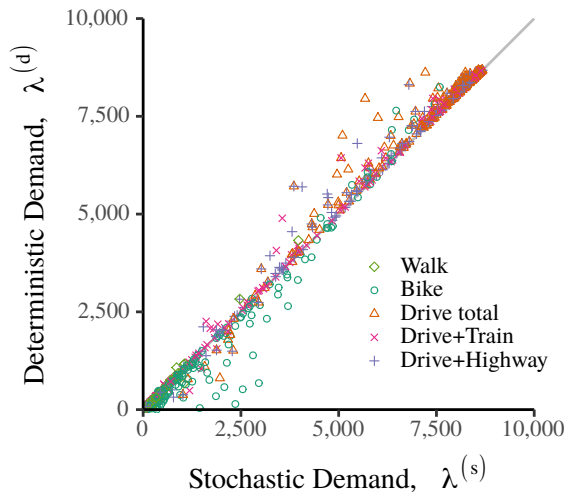
$$\%RMSE = \frac{100}{\lambda_{total}} \cdot \sqrt{\frac{\sum_{i=1}^N (\lambda^{(d)} - \lambda^{(s)})^2}{n}} \quad (4.34)$$

where the stochastic and deterministic estimation for demand are $\lambda^{(s)}$ and $\lambda^{(d)}$, respectively. Alternatively, the comparison could be using measured demand versus modeled demand.



(a) Fixed price only

(b) Distance price only



(c) Both prices vary

Figure 4.9: Comparative fit between deterministic and nested logit models

The fixed price only model in appears to provide the stable results, but tends to have larger errors towards the ends (i.e. at origin and maximum radius) for biking and driving. This may be problematic when the radii for biking is often in this range. Conversely, the distance price only model provides better fit at the ends of the

range, but quickly degrades as biking and driving demand move towards each other. However, in practice it is unlikely that the demand for biking will exceed that of driving, so it may prove beneficial to retain distance pricing in a model. As expected, the combination of two price inputs, shown in Figure 4.9c, yields a hybrid of the two models, providing a more reliable fit overall.

The overall analysis results in Figure 4.10 show that the greatest loss in accuracy occurs at lower VOTs, but the magnitude of the loss differs when using the different price sets. With fixed costs only, the model accuracy increases from 17% to 2%RMSE over a range of 5 to 50^{\$/hour}. The same is true for distance prices only, but to a lesser degree, ranging from 6% to 2%RMSE over the same VOT range. However, when using a combination of pricing schemes, different values of time have little effect on the model’s accuracy, ranging from 4% to 2%RMSE, with an average of 2.56%RMSE.

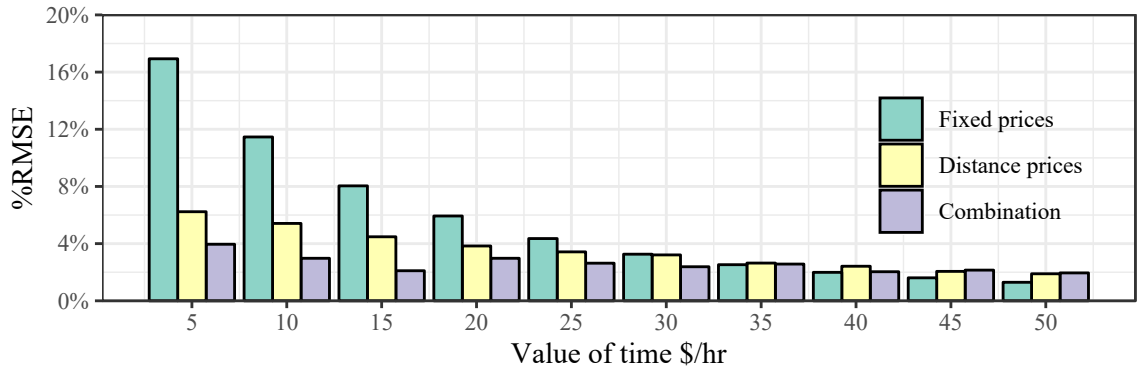


Figure 4.10: Effects of value of time on model accuracy

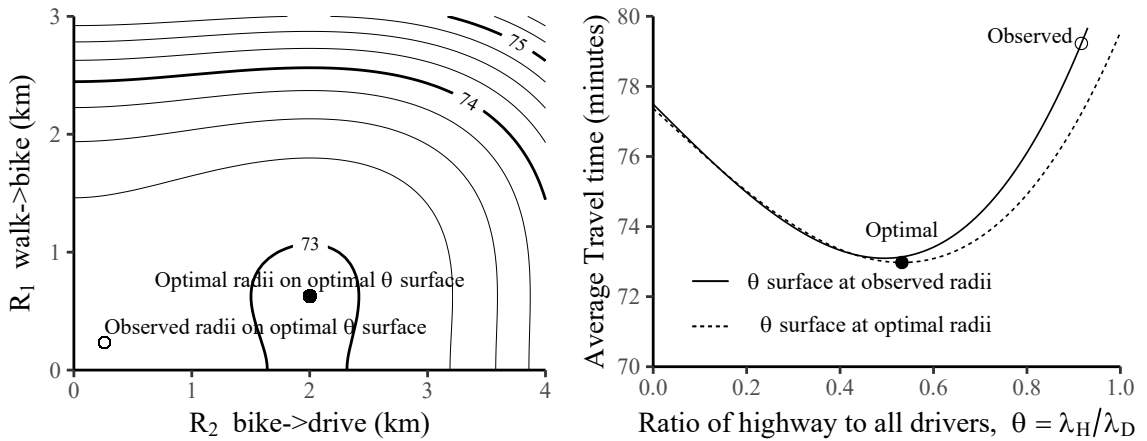
4.9.3 Single Value of Time – Optimal Mode Split

Using the observed mode-share in Table 4.2, the effective “observed” radii and θ are calculated using Equation set (4.21). The optimal radii and θ are then determined through optimization of Equation (4.20). The calculated observed radii and optimal radii are presented in Table 4.3.

Table 4.3: Optimal and observed radii

	r_1	r_2	θ	Average travel time
Observed	0.23	0.26	0.92	79.23 minutes
Optimal	0.63	2.00	0.53	72.97 minutes

The shape of the objective function is shown in Figure 4.11. In Figure 4.11a the gradient contours for the travel time are plotted by varying the access mode radii while holding the third term θ as constant. A unique optimal point is visible at the bottom of the distorted bowl shape. Conversely, in Figure 4.11b the travel time function is plotted by varying mainline θ and holding the radii constant. Two surfaces are present in Figure 4.11b because the optimal radii are used for the optimal case and the observed effective radii are used for the observed case. Regardless of radii, there appears to be a unique optimum at the bottom of the curve.



(a) Travel time across radii with optimal θ (b) Travel time across θ with optimal radii

Figure 4.11: Optimal demand allocation

Another perspective of the results is the highway travel time as users are loaded onto the highway, as shown in Figure 4.12. One might assume that the optimal highway utilization would be when the highway travel time is equal to the train

travel time. However, a more global optimal point actually leaves the highway with excess capacity. This is because the access mode is being considered in the proposed model. System performance is improved by maintaining a free-flowing highway by having more drivers take the train, as well as by leveraging the shorter start up times and access distances for walk and bike.

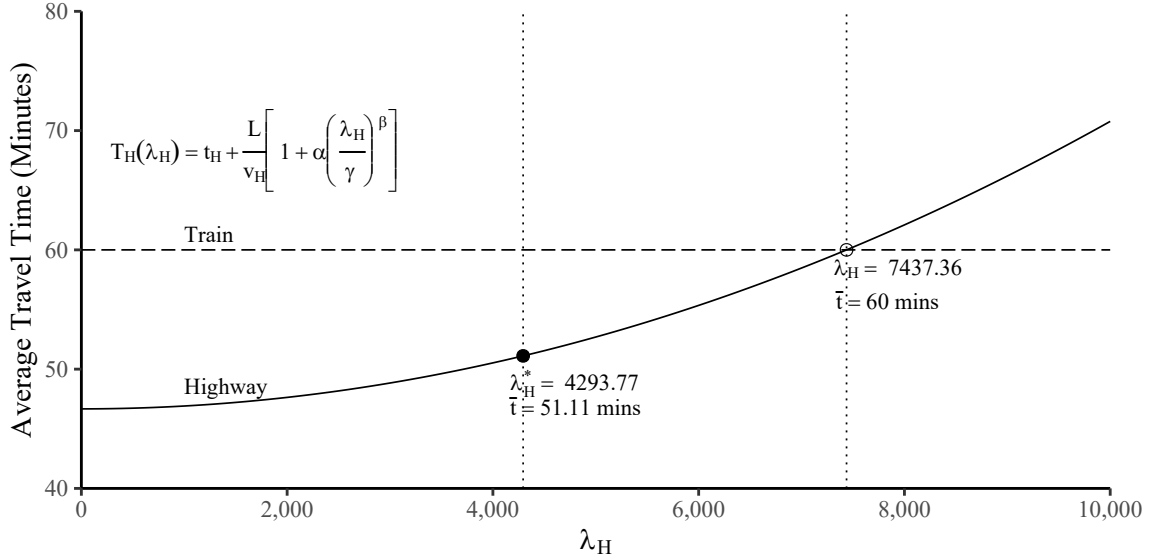


Figure 4.12: Mainline travel time with congestion

4.9.4 Single Value of Time – Optimal Prices

Using the proposed optimization approach in Equation (4.25), optimal access mode prices are determined for the three price sets of fixed prices only, distance prices only, and the combination of the two. Each of the three price sets are optimized using a revenue neutral scheme. The mainline prices are calculated using the nesting approach proposed in Equation (4.27), also using a revenue neutral scheme. The optimal prices are presented in Table 4.4 for the three sets of prices. The table also presents the mainline prices.

Table 4.4: Optimal pricing

Price set	Access prices						Mainline prices	
	C_W^d	C_B^d	C_D^d	C_W^f	C_B^f	C_D^f	C_H	C_T
Fixed only	–	–	–	-\$1.02	-\$0.12	\$1.13		
Distance only	-\$4.18	-\$0.32	\$4.50	–	–	–	\$1.22	-\$1.20
Combination	-\$0.22	-\$0.05	\$0.28	-\$0.96	-\$0.10	\$1.06		

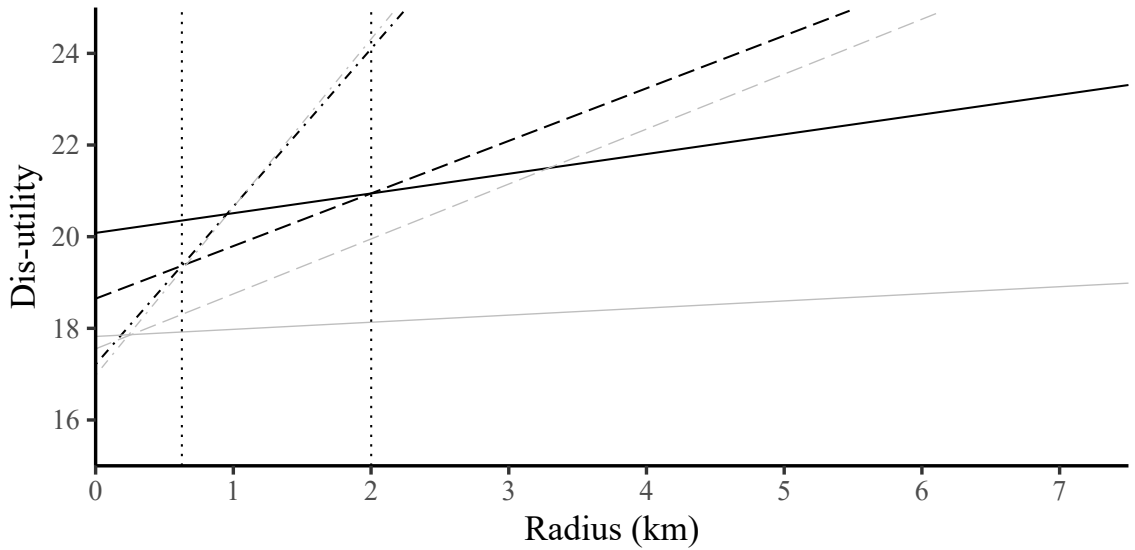
The effect on results for each pricing set are compared by modeling the demand for the observed condition using stochastic and deterministic methods as well as for the optimal condition without additional pricing. The results of this are presented in Table 4.5. A percent root-mean square error (%RMSE) is calculated for each model’s fit against the observed condition, as well as between model fit (i.e., difference between deterministic and stochastic). The reason that the observed condition is modeled with pricing is because the optimal point is when costs are zero, the additional pricing costs reflect the unobserved utility “deficit” required to move the observed condition to the optimal condition. Modeling the observed condition also provides a useful measure of overall accuracy, not just relative precision between models.

Table 4.5: Model demand estimation comparison

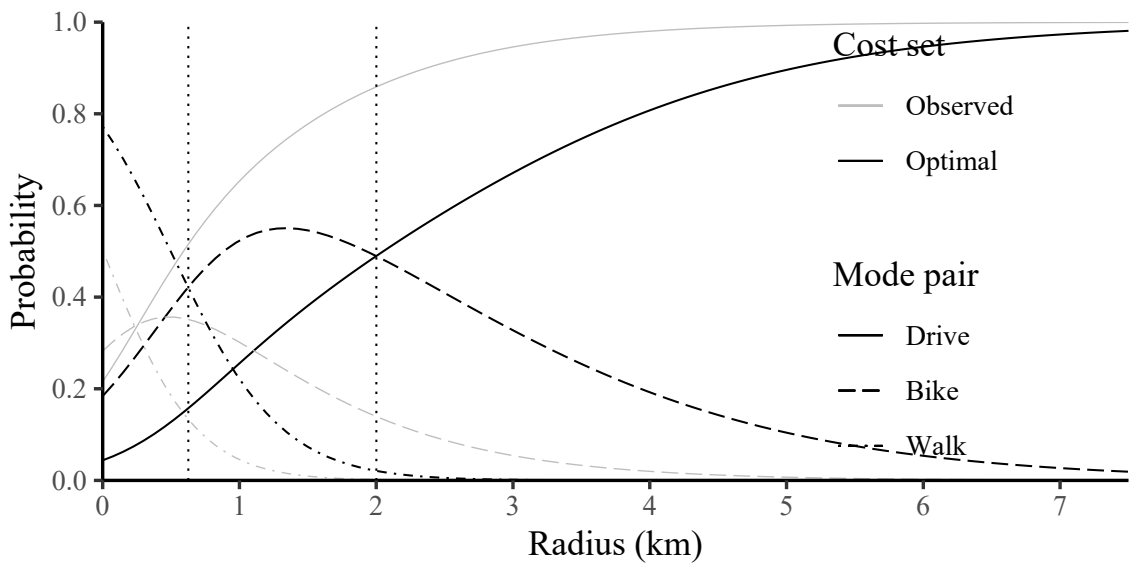
Price scheme	Model	λ_W	λ_B	λ_D	$\lambda_{D T}$	$\lambda_{D H}$	TT	%RMSE	
								Deterministic vs. Stochastic	Observed vs. model
Matching observed condition									
	Measured	270	30	8,383	700	7,683	79.23	–	–
Combination	Deterministic	8	2	8,673	628	8,045	81.43	2.53%	2.77%
	Stochastic	27	280	8,376	607	7,770	78.66		
Distance	Deterministic	8	2	8,673	185	8,045	81.43	0.17%	5.32%
	Stochastic	9	20	8,654	185	8,469	81.34		
Fixed	Deterministic	8	2	8,673	628	8,045	81.43	4.85%	2.77%
	Stochastic	33	542	8,108	587	7,521	77.91		
Optimal condition									
–	Deterministic	60	557	8,066	3,771	4,294	74.97	6.48%	–
	Stochastic	104	1,326	7,253	3,391	3,862	71.48		

As a demonstration of how the dis-utility functions are modified by the prices, Figure 4.13 displays both the original and optimized functions for dis-utility and

subsequent logit curves using a combination price set. In both Figures 4.13a and 4.13b the original functions without any additional costs are shown in gray, and the optimized functions are shown in black. The dis-utility plot in Figure 4.13a shows how the functions are adjusted to intersect at the optimal radii, shown as the vertical dotted lines. The original dis-utility functions all nearly intersect at the same point, reflecting the small number of bicyclists. In the logit plot in Figure 4.13b, the curves for walk and bike grow to represent a much larger share while the driving curve shrinks.



(a) Optimized access mode dis-utility



(b) Optimized access mode logit

Figure 4.13: Optimal logit and dis-utility comparison

An artifact of the model is that the dis-utility functions intersect exactly where the logit functions intersect. However, unlike the all-or-nothing deterministic case, the intersection merely represents where the probabilities intersect. This corresponds to the results in Table 4.5 where the deterministic model always yields the same result,

but the stochastic model can vary. This is because the deterministic model simply finds the radii where dis-utility intersects to approximate demand. Thus, as long as dis-utility intersects at the desired radii the result is the same, regardless of slope or intercept. However, the stochastic model determines demand by integrating the area under the non-linear probability curve across the radii. The steepness of the curves, or the magnitude of the difference in dis-utility, will then affect the results.

4.9.5 Distributed Value of Time – Optimal Mode Split

Using the continuous log-normal distribution function for VOT, a range of discrete probability distribution segments are created. The product of this probability distribution and the population density δ provides a set of densities δ_k , for each VOT segment in the population. From this, a set of optimal parameters for r_{1k} , r_{2k} , θ_k , are determined for each segment k in the initial parameter optimization step. In this dissertation, a discrete distribution size of thirty equal-length segments were chosen for optimization. Although this is a fairly large number of segments, it was chosen in order to provide a reasonably smooth result for analysis and graphical representation. The results of the parameter optimization step are shown in Figure 4.14.

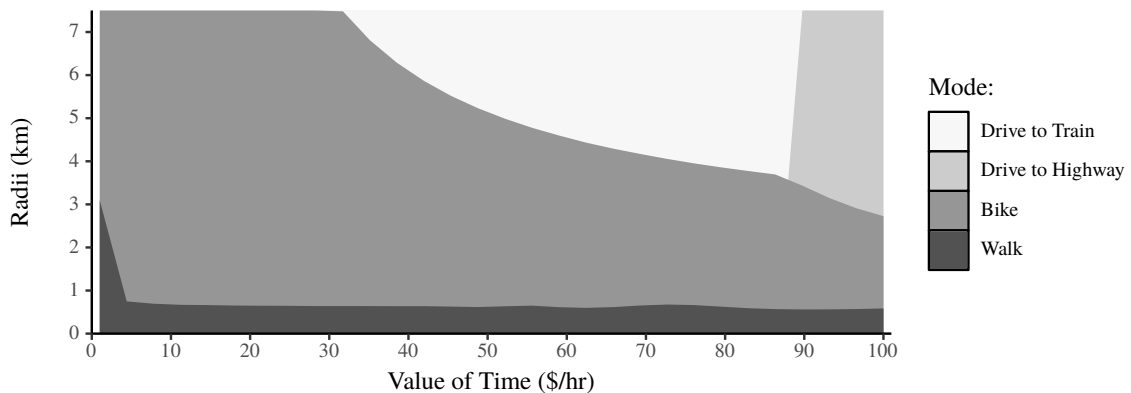
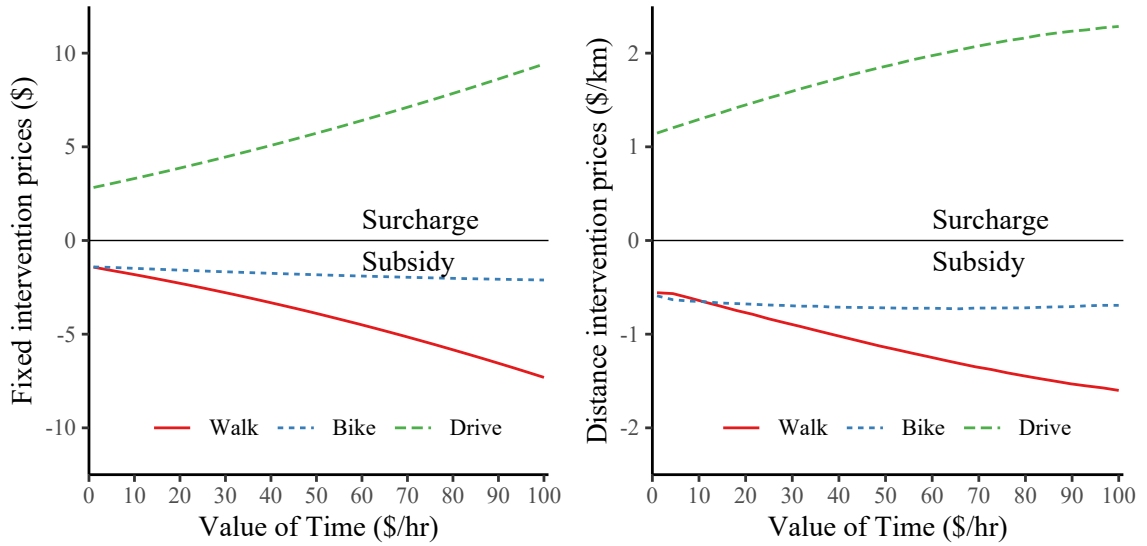


Figure 4.14: Optimal parameters by VOT

The horizontal axis in Figure 4.14 is the range of VOT and the vertical axis shows the tipping point radii for the corresponding VOT (i.e., the point at which users change mode choice). The gradual descent of bike users makes intuitive sense, because as the users' VOT increases, the time savings of the much faster drive mode begins to outweigh the cheaper but slower bike mode. This is also true for the walk mode, but at a much steeper slope, indicating that all but the poorest users will choose another mode if they travel further than approximately $\frac{2}{3}$ of a kilometer. A particularly interesting result is the severe change of drive users from train to highway. Unlike the single VOT case where θ is between zero and one, here we see that the shift from train to highway is almost binary depending upon VOT. This is because from the perspective of an individual VOT segment, the choice between train and highway doesn't vary by distance and only depends on whether the VOT is high enough to justify the travel time savings. A θ between zero and one would occur when a VOT segment spans the range of VOT where this choice occurs, as is the case when the mean VOT is used for the entire system.

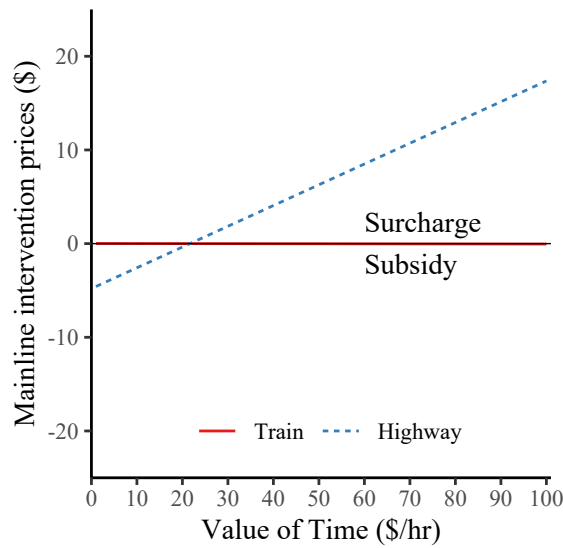
4.9.6 Distributed Value of Time – Optimal Pricing

Although the results of a single mean VOT system optimization may simply be the aggregated result of a distributed VOT system, the pricing policy and generalized cost experienced does not affect users equally. A set of prices for all VOT groups was determined simultaneously using Equation (4.29), minimizing the total magnitude of money changing hands while achieving the desired radii and θ parameters for all k . The results of this are presented in Figure 4.15, displaying the costs of fixed only, distance only, and combination across VOT.



(a) Fixed prices by VOT

(b) Distance prices by VOT



(c) Mainline prices by VOT

Figure 4.15: Optimal pricing varying by VOT

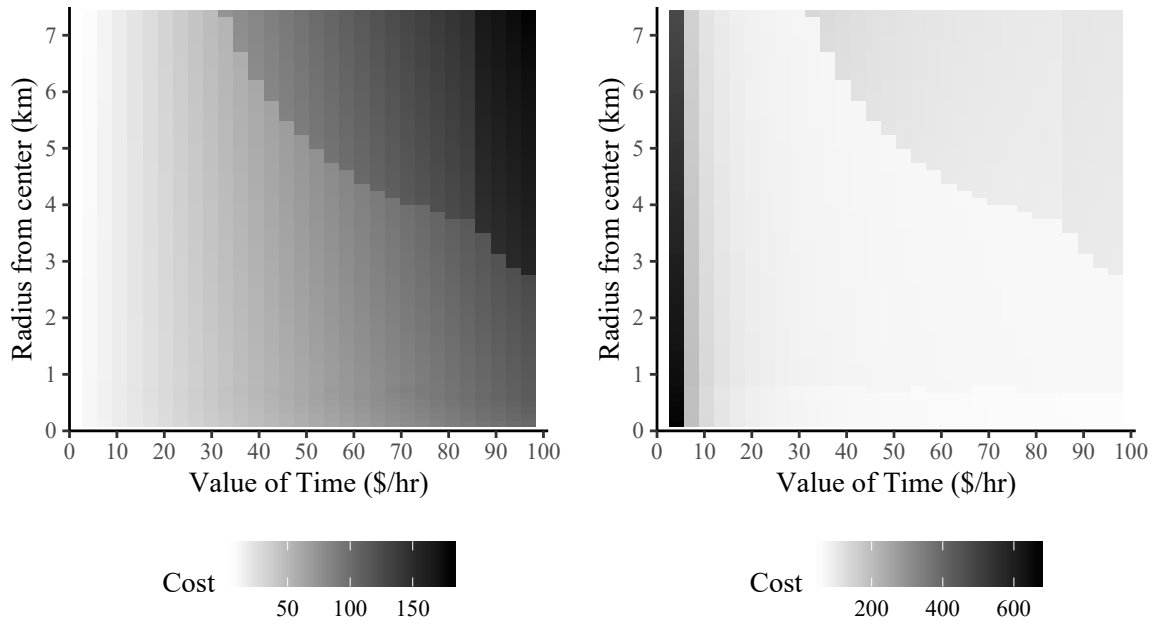
Similarly to the pricing determined in the single VOT case, the walk, bike, and train modes are all subsidized while drive and highway have a surcharge. The difference in this case is the magnitude of the surcharge and subsidies increases with VOT. This makes intuitive sense, as a greater monetary magnitude is needed to sway the

users with higher values of time. More interesting is the non-linear pricing for access modes, with fixed pricing increasing at an increasing rate and distance based prices increasing at a decreasing rate. With access mode, the magnitude of the mainline mode also increased with VOT, but unlike access mode it primarily affected one mode, highway users. Moreover, the prices for mainline modes were not only a subsidy or surcharge for all users, but instead is a subsidy for low VOT users and a surcharge for high VOT users. Interestingly, the transition from subsidy to surcharge occurs approximately at the mean VOT value.

4.9.7 Distributed Value of Time – Equity

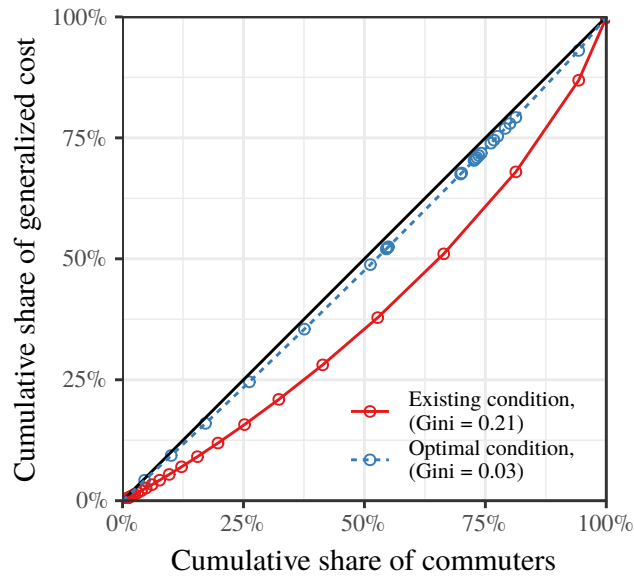
A fundamental disadvantage of a system optimum is the potential for users to experience inequitable costs. This cost can be experienced across two user variables, VOT or radii, and is measured either as the generalized time cost or as the generalized monetary costs. These costs are presented graphically by VOT and radii with generalized monetary cost in Figure 4.16a and generalized time cost in Figure 4.16b. When measuring generalized cost as a monetary cost, the wealthier VOT users appear to experience the greatest cost burden, but when viewed from a temporal lens, the poorest travelers experience the greatest burden. This is not just because wealthier users experience greater pricing costs, but because the poorer users' VOT is so low their burden appears low from a monetary perspective. Time is a non-transferable and an equally finite resource, because once expended time cannot be redeemed nor can it be redistributed, and all users possess the same “wealth” of time by sharing a 24-hour day. For these reasons, time will be used as the measure of equity, rather than the monetary cost. Furthermore, it makes moral sense to use the measure that exposes the cost burden on lower income users, not the inverse. The equity measure, the Gini coefficient, is calculated using Equation (4.33) across VOT. It is calculated across VOT for two reasons. First, because users can feasibly choose their location in

the radii but they cannot so easily choose their VOT. Second, radii is a physical cost based on location, it would make little sense to subsidize users to live further away.



(a) Generalized monetary cost (\$)

(b) Generalized time cost (minutes)



(c) Gini Lorenz curves

Figure 4.16: Comparison between generalized monetary and temporal user costs

The overall results of the distributed VOT optimization achieved a 57% reduction in generalized travel time from 256 minutes to 110 minutes. The resulting average generalized time costs are higher than the single VOT validation example in Table 4.3. This is because now the average cost accounts for the much higher costs experienced by low VOT users whereas before it only considered the average VOT and the average travel time. Despite this, the Gini equity measure improved by 86% from 0.21 to 0.03. Figure 4.16c is a graphical representation of the Gini coefficient as the ratio of the area under the Lorenz curve to the area under the diagonal. A perfectly equitable society (i.e., Gini = 1) would be a diagonal line and a perfectly inequitable society (i.e., Gini = 0) would be along the bottom edge. This is interesting result considering that neither objective function directly contains any equity factor. This is because when minimizing the total generalized *time* cost, the burden on low VOT users is taken into account effectively using money is used as a transferable resource to achieve this minimized overall cost. This is apparent in Figure 4.15c where low VOT users are actually subsidized as an incentive to take the faster highway mode.

4.10 Summary of findings

The deterministic model yielded a good fit against both a logit-based counterpart and the measured demand with an accuracy of 4% and 6%, respectively. The deterministic model also can account for both distance-based and fixed monetary costs, which when combined yield a more reliably accurate model regardless of value of time magnitude. The importance of this is that the deterministic model can be deployed more easily to explore policy decisions and their impacts, such as equity. Results of which can then be further refined with the use of a stochastic model.

Through empirical analysis it was found that value of time is linearly proportional to income. Thus, making the income distribution a realistic proxy for a value of time distribution. The deterministic model is then extended to include varying values of

time, resulting in a 57% reduction in generalized travel time and improved the Gini inequity measure from 0.21 to 0.03. This shows that improving system efficiency does not necessarily come at the cost of equity. In particular, transferable monetary costs can be used to offset the non-transferable temporal cost burden experienced by users, particularly by lower income.

CHAPTER 5

CONCLUSIONS

This dissertation presents a deterministic approximation of a discrete choice model for multi-modal access and mainline mode choice that can account for varying values of time to allow for an evaluation of equity. The objective is to develop a choice model that can be efficiently optimized through pricing schemes; a current challenge using stochastic logit-based models. This is achieved by first optimizing demand using a spatial model, then subsequently determining the pricing necessary to achieve the optimal demand. The added complexity of nested mode choice is modeled using an auxiliary spatial model to determine the mainline cost differential. In addition to the model's efficient optimization, it also possesses the ability to account for varying values of time and the flexibility to set both fixed and distance-based pricing policies for specific modes. This value of time heterogeneity is essential for assessing the impacts of pricing intervention on equity, this heterogeneity was made possible by the integrated population synthesis.

5.1 Dissertation Contributions

In summary, the dissertation makes the following contributions:

1. Population Synthesis
 - (a) **Integrated population synthesis and fixed work-place assignment**
 - (b) **Faster non-negative least-deviation (NLAD) method** for fitting joint population of persons and households

2. Mixed-access Model

- (a) A congestion pricing model with **multiple access modes**
- (b) Allows for **value of time to vary across both income and mode**
- (c) Model framework allows **flexible objective function** (e.g., optimize equity directly)
- (d) Measured impacts of transport **system efficiency gains on user equity**
- (e) **Computationally lightweight**, meaning it requires little computational resources and quickly reaches a solution

5.2 Summary of findings

The model was applied for a case study of commuting between Worcester and Boston, Massachusetts, USA. The application utilized a synthetic population to generate a realistic household income distribution of travelers as the basis for the value of time distributions. Through an empirical analysis of a travel survey it was found that the value of time varies proportionally with income, making it possible to model a value of time distribution for the synthetic population. The synthetic population achieved accurate results with root mean squared normalized (RMSN) values of 0.0715 for marginals, 6.5779 for joint distribution cells, and 0.3889 for OD totals. Although the joint distribution cell fit is less accurate, it is of less concern for this particular application as population is re-aggregated into the VOT distribution.

The proposed integrated population synthesis process makes two contributions. First by exploiting the common Iterative Proportional Fitting (IPF) algorithm commonly used in population synthesis and workplace assignment for an integrated method for both. This minimizes errors that would be introduced through independently estimated models. Second, the efficient joint person-household re-weighting technique based on non-negative least absolute deviation (NLAD) fitting substan-

tially reduces computation time by one-sixth compare to the conventional iterative proportional updating (IPU) method. This new re-weighting technique makes the integrated process possible by being able to efficiently handle additional shared attributes in the population and workplace data (e.g., employment). The proposed technique outperforms both the conventional IPU and non-negative least squares (NNLS) methods in terms of computation time with result of similar accuracy.

Before the multi-modal model was applied to a varying value of time distribution, it is first validated against a stochastic counterpart using numerical analysis for a single value of time. The deterministic model achieved results within 4% accuracy of the stochastic logit-based model, and within 6% of measured values. Once applied to a varying value of time distribution, the model results show great potential, achieving a 57% reduction in generalized travel time and improves the Gini inequity measure from 0.21 to 0.03.

A summary of the major research findings are as follows:

- 1. The deterministic model provides an accurate approximation of a stochastic logit-based model. Meaning that the model can be deployed to easily explore policy decisions and their impacts; results of which can then be further refined with the use of a stochastic model if so desired.*
- 2. Overall system travel times can be improved by manipulating access mode demand. That is, mainline mode choice can be optimized by adjusting prices for access modes. Mainline prices can then be used to achieve the optimal split for drive-access users specifically.*
- 3. Improving system efficiency does not necessarily come at the cost of equity. In particular, transferable monetary costs can be used to offset the non-transferable temporal cost burden experienced by users, particularly of lower income.*

5.3 Future work

Possible future work is focused primarily on the multi-modal system model as the main thrust of the dissertation, but the population synthesis also possesses potential for future work.

5.3.1 Population synthesis & workplace assignment

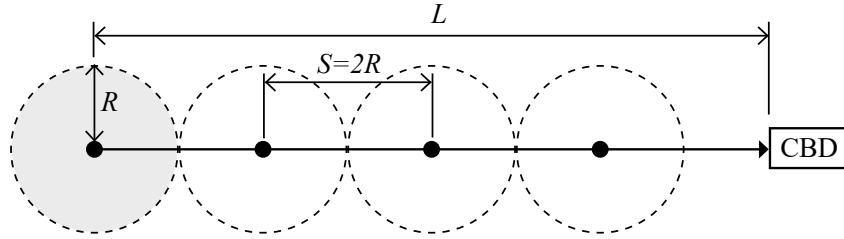
Regarding the population synthesis, the results a good fit, but there are two main areas of potential refinement.

1. *Expand origin-destination stratification:* This synthesis presents only a single stratification variable that is for industry sector. Although using fourteen industry sectors is robust on its own, the workplace assignment can be modified for any number of stratification variables (e.g., age and gender). Additional stratification variables would effectively add more constraints to the integrated process, improving the accuracy of the final workplace assignment.
2. *Apply the integrated framework to probabilistic synthesis:* A second, and more worthwhile improvement would be to investigate the potential fusion of high-resolution aggregated workplace data (i.e., LODES) with disaggregated probabilistic simulation synthesis (PSS) methods (e.g., Bayesian Networks). If possible, this would likely provide a significant improvement to workplace assignment while still generating high quality socio-demographic synthetic populations using advanced PSS methods. However, the feasibility of integrating PSS with aggregated location data is uncertain as no examples have been found in the literature.

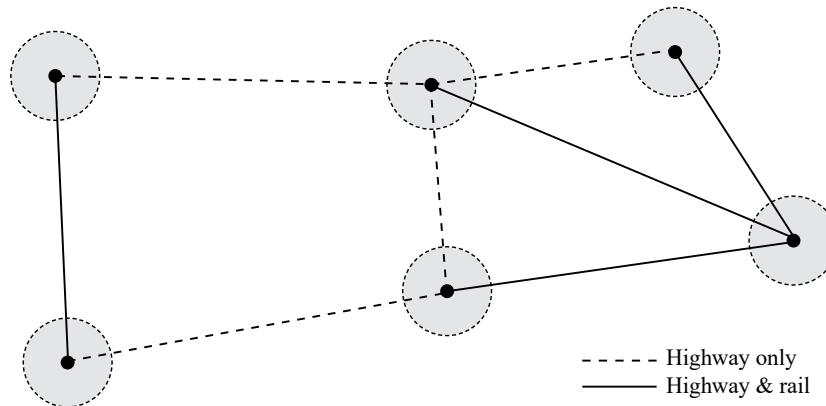
5.3.2 Multi-modal system model

There are several extensions or potential applications of the multi-modal model that are worthy of future research.

1. *Account for systematic error:* Overall, the deterministic model provides a reasonable approximation of its logit-based counterpart to within 4% error. However, much of this error is systematic when dissected by fixed and distance-based pricing. It may be possible to correct this error for a more reliable solution. Alternatively, the stochastic and deterministic models could be used together where the efficient deterministic model is used to quickly find an approximate deterministic solution that is then refined using a more computationally intensive, but possibly more precise, stochastic model.
2. *Network application:* A major benefit of the deterministic model is that it requires very little computational resources, allowing for future application in a larger network. One such application could include a corridor of stations as in Figure 5.1a that feed into a central business district in a commuter rail type model. At a more macroscopic level, the model could be applied to a network of cities, as in Figure 5.1b, and optimized to provide pricing schemes for inter-city travel.



(a) Corridor application



(b) Network application

Figure 5.1: Possible model applications

3. *Agency costs:* The current model does not account for the agency costs in the generalized cost function used. If the model is to provide robust insights into system operations, as shown in Figures 5.1a and 5.1b, it is essential that the model accounts for the operational costs associated with each mode. Fortunately, this can be achieved relatively easily due to the deterministic form of the presented model. Agency costs need only be formulated and incorporated into the generalized cost function, such as a marginal operating cost per user for a distance or time in operation (e.g., $\frac{\$}{\text{pax-km}}$).
4. *Local congestion:* The current model does not account for local congestion. Although this is justifiable for walk and bike modes, this is not entirely valid

for driving, especially if applying the model to a much larger scale (e.g., inter-city travel). An additional sub-model will need to be formulated to account for local travel time and the impact congestion has on it, much like it was for mainline travel.

5. *Equity as the objective:* In the current model, pricing is determined by simply minimizing the total magnitude of money changing hands. Although this inadvertently resulted in an improvement of equity, it is not necessary guaranteed. It may be possible to more intelligently improve equity by directly optimizing prices using equity as an objective function. The Gini coefficient calculation is ill-suited as an objective function due to the permutations and subtraction required. However, it is possible to reformulate or construct a proxy function that can be optimized easily to minimize inequity.
6. *Exploring trade off between efficiency and equity:* Assuming that an objective function is formulated to directly express some measure of equity, it would be possible to explore whether a relationship exists between optimal efficiency and optimal equity. If so, a bi-objective mathematical function could be developed to explore whether the two are mutually exclusive and to what extent. Meaning, to determine if there is a trade-off between efficiency and equity, and if perhaps some compromise can be achieved (e.g., pareto).

BIBLIOGRAPHY

- Abdel-Aal, M. M. M. Calibrating a trip distribution gravity model stratified by the trip purposes for the city of Alexandria. *Alexandria Engineering Journal*, 53(3): 677–689, 9 2014.
- Abraham, J. E., Stefan, K. J., and Hunt, J. D. Population synthesis using combinatorial optimization at multiple levels. *Transportation Research Record*, 17, 2012.
- Abrantes, P. A. and Wardman, M. R. Meta-analysis of UK values of travel time: An update. *Transportation Research Part A: Policy and Practice*, 45(1):1–17, 1 2011.
- Anas, A. and Xu, R. Congestion, Land Use, and Job Dispersion: A General Equilibrium Model. *Journal of Urban Economics*, 45(3):451–473, 5 1999.
- Anas, A., Arnott, R., and Small, K. A. Urban Spatial Structure. *Journal of Economic Literature*, 36(3):1426–1464, 1998.
- Arentze, T., Timmermans, H., and Hofman, F. Creating Synthetic Household Populations: Problems and Approach. *Transportation Research Record: Journal of the Transportation Research Board*, 2014:85–91, 12 2007.
- Arentze, T. A. and Timmermans, H. J. A learning-based transportation oriented simulation system. *Transportation Research Part B: Methodological*, 38(7):613–633, 8 2004.
- Auld, J. and Mohammadian, A. Efficient Methodology for Generating Synthetic Populations with Multiple Control Levels. *Transportation Research Record: Journal of the Transportation Research Board*, 2175(1):138–147, 1 2010.
- Azevedo, C. L., Seshadri, R., Gao, S., Atasoy, B., Akkinepally, A. P., Christofa, E., Zhao, F., Trancik, J., and Ben-Akiva, M. Tripod: Sustainable Travel Incentives with Prediction, Optimization, and Personalization. *97th Annual Meeting of the Transportation Research Board*, 1:19, 2018.
- Badia, H., Estrada, M., and Robusté, F. Competitive transit network design in cities with radial street patterns. *Transportation Research Part B: Methodological*, 59: 161–181, 2014.
- Badia, H., Argote-Cabanero, J., and Daganzo, C. F. How network structure can boost and shape the demand for bus transit. *Transportation Research Part A: Policy and Practice*, 103:83–94, 2017.

- Ballas, D., Clarke, G., Dorling, D., Eyre, H., Thomas, B., and Rossiter, D. SimBritain: A spatial microsimulation approach to population dynamics. *Population, Space and Place*, 11(1):13–34, 1 2005a.
- Ballas, D., Clarke, G. P., and Wiemers, E. Building a dynamic spatial microsimulation model for Ireland. *Population, Space and Place*, 11(3):157–172, 5 2005b.
- Barthelemy, J. and Toint, P. L. Synthetic population generation without a sample. *Transportation Science*, 47(2):266–279, 2013.
- Becker, G. S. A Theory of the Allocation of Time. *The Economic Journal*, 75(299):493, 9 1965.
- Beckman, R. J., Baggerly, K. A., and McKay, M. D. Creating synthetic baseline populations. *Transportation Research Part A: Policy and Practice*, 30(6):415–429, 1996.
- Beesley, M. E. The Value of Time Spent in Travelling: Some New Evidence. *Economica*, 32(126):174, 5 1965.
- Ben-Akiva, M. and Bierlaire, M. Discrete choice methods and their applications to short term travel decisions. *Handbook of transportation science*, 23:5–33, 1999.
- Ben-Akiva, M., De Palma, A., and Kanaroglou, P. Dynamic model of peak period traffic congestion with elastic arrival rates. *Transportation Science*, 20(3):164–181, 1986.
- Ben-Akiva, M., Bolduc, D., and Bradley, M. *Estimation of travel choice models with randomly distributed values of time*. Number 1413. 1993.
- Ben-Akiva, M., Bowman, J., and Gopinath, D. Travel demand model system for the information era. *Transportation*, 23(3), 8 1996.
- Ben-Akiva, M. E. *Structure of passenger travel demand models*. PhD thesis, Massachusetts Institute of Technology, 1973.
- Ben-Akiva, M. E. and Lerman, S. R. *Discrete choice analysis: theory and application to travel demand*, volume 9. MIT press, 1985.
- Bloomfield, P. and Steiger, W. L. *Least Absolute Deviations*. Birkhäuser Boston, Boston, MA, 1984. ISBN 978-1-4684-8576-9.
- Boarnet, M. G., Giuliano, G., Hou, Y., and Shin, E. J. First/last mile transit access as an equity planning issue. *Transportation Research Part A: Policy and Practice*, 103:296–310, 9 2017.
- Borysov, S. S., Rich, J., and Pereira, F. C. Scalable Population Synthesis with Deep Generative Modeling. *arXiv preprint arXiv:1808.06910*, 2018.

- Bowman, J. and Ben-Akiva, M. Activity-based disaggregate travel demand model system with activity schedules. *Transportation Research Part A: Policy and Practice*, 35(1):1–28, 1 2001.
- Bowman, J. L., Bradley, M., Shiftan, Y., Lawton, T. K., and Ben-Akiva, M. E. Demonstration of an activity based model system for Portland. In *8th World Conference on Transport Research, Antwerp, Belgium*, 1998.
- Brownstone, D. and Train, K. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89(1-2):109–129, 11 1998.
- Caltrans. Park and Ride Program Resource Guide. Technical report, 2010.
- Casati, D., Müller, K., Fourie, P. J., Erath, A., and Axhausen, K. W. Synthetic Population Generation by Combining a Hierarchical, Simulation-Based Approach with Reweighting by Generalized Raking. *Transportation Research Record: Journal of the Transportation Research Board*, 2493:107–116, 1 2015.
- Center for Neighborhood Technology. Paved Over: Surface Parking Lots or Opportunities for Tax-Generating, Sustainable Development? Technical Report November, 2006.
- Cervero, R. The transit pricing evaluation model: A tool for exploring fare policy options. *Transportation Research Part A: General*, 16(4):313–323, 7 1982.
- Cervero, R. and Duncan, M. Walking, Bicycling, and Urban Landscapes: Evidence From the San Francisco Bay Area. *American Journal of Public Health*, 93(9): 1478–1483, 2003.
- Cervero, R. and Guerra, E. E. Urban Densities and Transit : A Multi-dimensional Perspective. *Institute of Transportation Studies University of California, Berkely*, (September):15, 2011.
- Cervero, R. and Kockelman, K. Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3): 199–219, 9 1997.
- Chen, A., Oh, J.-S., Park, D., and Recker, W. Solving the bicriteria traffic equilibrium problem with variable demand and nonlinear path costs. *Applied Mathematics and Computation*, 217(7):3020–3031, 2010.
- Chen, H., Gu, W., Cassidy, M. J., and Daganzo, C. F. Optimal transit service atop ring-radial and grid street networks: A continuum approximation design method and comparisons. *Transportation Research Part B: Methodological*, 81:755–774, 11 2015.
- Chintagunta, P. K., Jain, D. C., and Vilcassim, N. J. Investigating heterogeneity in brand preferences in logit models for panel data. *Journal of Marketing Research*, pages 417–428, 1991.

- Choupani, A.-A. and Mamdoohi, A. R. Population Synthesis Using Iterative Proportional Fitting (IPF): A Review and Future Research. *Transportation Research Procedia*, 17:223–233, 2016.
- Computational Infrastructure for Operations Research (COIN-OR) Foundation. Clp, 2017.
- Daganzo, C. *Public Transportation Systems: Basic Principles of System Design , Operations Planning and Real-Time Control*. Number October. 2010.
- Daganzo, C. F. and Sheffi, Y. On stochastic models of traffic assignment. *Transportation science*, 11(3):253–274, 1977.
- Davis, R. A. and Dunsmuir, W. T. M. Least Absolute Deviation Estimation for Regression with ARMA Errors. *Journal of Theoretical Probability*, 10(2):481–497, 4 1997.
- de Palma, A., Ben-Akiva, M., Lefevre, C., and Litinas, N. Stochastic equilibrium model of peak period traffic congestion. *Transportation Science*, 17(4):430–453, 1983.
- Deming, W. E., Stephan, F. F., and Frederick F. Stephan. On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. *The Annals of Mathematical Statistics*, 11(4):427–444, 1940.
- DeSerpa, A. C. A Theory of the Economics of Time. *The Economic Journal*, 81(324): 828, 12 1971.
- Dial, R. B. Bicriterion traffic assignment: basic theory and elementary algorithms. *Transportation science*, 30(2):93–111, 1996.
- Dong, X., Ben-Akiva, M. E., Bowman, J. L., and Walker, J. L. Moving from trip-based to activity-based measures of accessibility. *Transportation Research Part A: Policy and Practice*, 40(2):163–180, 2 2006.
- El-Geneidy, A., Levinson, D., Diab, E., Boisjoly, G., Verbich, D., and Loong, C. The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transportation Research Part A: Policy and Practice*, 91:302–316, 9 2016.
- Farber, S., Bartholomew, K., Li, X., Páez, A., and Nurul Habib, K. M. Assessing social equity in distance based transit fares using a model of travel behavior. *Transportation Research Part A: Policy and Practice*, 67:291–303, 9 2014.
- Farooq, B., Bierlaire, M., Hurtubia, R., and Flötteröd, G. Simulation based population synthesis. *Transportation Research Part B: Methodological*, 58:243–263, 12 2013.
- Fishburn, P. C. Utility theory for decision making. Technical report, Research Analysis Corp Mclean va, 1970.

- Fishman, R. The End of Suburbia: A New Kind of City Is Emerging—the “Technoburb”. *Los Angeles Times*, 1987.
- Flamm, B. and Rivasplata, C. Perceptions of Bicycle-Friendly Policy Impacts on Accessibility to Transit Services: The First and Last Mile Bridge. Technical report, 2014.
- Fournier, N., Chen, S., Needell, Z., Lima, I. V. D., Deliali, K., Araldo, A., Prakash, A. A., Azevedo, C. M. L., Christofa, E., Trancik, J., Ben-Akiva, M., and Akkinapally, A. Integrated Simulation of Activity-Based Demand and Multi-Modal Dynamic Supply Simulation for Energy Assessment. In *21st IEEE International Conference on Intelligent Transportation Systems*, 2018.
- Fujii, S. and Kitamura, R. Anticipated travel time, information acquisition, and actual experience: Hanshin expressway route closure, Osaka-Sakai, Japan. *Transportation Research Record: Journal of the Transportation Research Board*, (1725): 79–85, 2000.
- Gan, H. and Ye, X. Will commute drivers switch to park-and-ride under the influence of multimodal traveler information? A stated preference investigation. *Transportation Research Part F: Traffic Psychology and Behaviour*, 56:354–361, 7 2018.
- Gini, C. Variabilità e mutabilità. *Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi*, 1912.
- Gonzales, E. J. Coordinated pricing for cars and transit in cities with hypercongestion. *Economics of Transportation*, 4(1-2):64–81, 3 2015.
- Gonzales, E. J. and Daganzo, C. F. Morning commute with competing modes and distributed demand: user equilibrium, system optimum, and pricing. *Transportation Research Part B: Methodological*, 46(10):1519–1534, 2012.
- Gonzales, E. J. and Daganzo, C. F. The evening commute with cars and transit: duality results and user equilibrium for the combined morning and evening peaks. *Transportation Research Part B: Methodological*, 57:286–299, 2013.
- Grant, J. *Planning the Good Community*. Routledge, 11 2005. ISBN 9780203479469.
- Guevara, C. A. *Endogeneity and sampling of alternatives in spatial choice models*. PhD thesis, Massachusetts Institute of Technology, 2010.
- Guo, J. and Bhat, C. Population synthesis for microsimulating travel behavior. *Transportation Research Record: Journal of the Transportation Research Board*, (2014):92–101, 12 2007.
- Hägerstrand, T. What about people in regional science? *Papers in regional science*, 24(1):7–24, 1970.

- Hensher, D. The value of commuter travel time savings: Empirical estimation using an alternative valuation model. *Journal of Transport Economics and Policy*, pages 167–176, 1976.
- Hensher, D. A. The valuation of commuter travel time savings for car drivers: Evaluating alternative model specifications. *Transportation*, 28(2):101–118, 2001.
- Hermes, K. and Poulsen, M. A review of current methods to generate synthetic spatial microdata using reweighting and future directions. *Computers, Environment and Urban Systems*, 36(4):281–290, 7 2012.
- Hoehne, C. G. and Chester, M. V. Greenhouse gas and air quality effects of auto first-last mile use with transit. *Transportation Research Part D: Transport and Environment*, 53:306–320, 6 2017.
- Holroyd, E. M. The optimum bus service: a theoretical model for a large uniform urban area. In *Proceedings of the Third International Symposium on the Theory of Traffic Flow Operations Research Society of America*, 1967.
- Huang, H.-J. Fares and tolls in a competitive system with transit and highway: the case with two groups of commuters. *Transportation Research Part E: Logistics and Transportation Review*, 36(4):267–284, 2000.
- Huang, H.-J. Pricing and logit-based mode choice models of a transit and highway system with elastic demand. *European Journal of Operational Research*, 140(3): 562–570, 8 2002.
- Iida, Y., Akiyama, T., and Uchida, T. Experimental analysis of dynamic route choice behavior. *Transportation Research Part B: Methodological*, 26(1):17–32, 1992.
- Ireland, C. T. and Kullback, S. Contingency tables with given marginals. *Biometrika*, 55(1):179–188, 1968.
- Jacobs, J. *The death and life of great American cities*. 1961. ISBN 0394421590 9780394421599 067974195X 9780679741954.
- Jara-Díaz, S. R., Munizaga, M. A., Greeven, P., Guerra, R., and Axhausen, K. Estimating the value of leisure from a time allocation model. *Transportation Research Part B: Methodological*, 42(10):946–957, 12 2008.
- Johnson, M. B. Travel time and the price of leisure. *Economic Inquiry*, 4(2):135–145, 1966.
- Kamakura, W. A. and Russell, G. A probabilistic choice model for market segmentation and elasticity structure. 1989.
- Kant, I. Critique of pure reason. *Modern Classical Philosophers, Cambridge, MA: Houghton Mifflin*, pages 370–456, 1781.

- Katharine M. Mullen and Ivo H. M. van Stokkum. The Lawson-Hanson algorithm for non-negative least squares, 2015.
- Kono, Y., Uchida, K., and Andrade, K. Economical welfare maximisation analysis: Assessing the use of existing Park-and-Ride services. *Transportation*, 2014.
- Kraus, M. A new look at the two-mode problem. *Journal of Urban Economics*, 54(3):511–530, 2003.
- Lave, C. A. A behavioral approach to modal split forecasting. *Transportation Research*, 3(4):463–480, 12 1969.
- Lawson, C. L. and Hanson, R. J. *Solving Least Squares Problems*. Society for Industrial and Applied Mathematics, 1 1995. ISBN 978-0-89871-356-5.
- Le, D.-t., Cernicchiaro, G., Zegras, C., and Ferreira, J. Constructing a Synthetic Population of Establishments for the Simmobility Microsimulation Platform. *Transportation Research Procedia*, 19:81–93, 2016.
- Levinson, D. and Zhu, S. A portfolio theory of route choice. *Transportation Research Part C: Emerging Technologies*, 35:232–243, 2013.
- Li, Z. and Hensher, D. A. Estimating values of travel time savings for toll roads: Avoiding a common error. *Transport Policy*, 24:60–66, 11 2012.
- Li, Z.-C., Wang, Y.-D., Lam, W. H. K., Sumalee, A., and Choi, K. Design of Sustainable Cordon Toll Pricing Schemes in a Monocentric City. *Networks and Spatial Economics*, 14(2):133–158, 6 2014.
- Litchfield, J. A. Inequality: Methods and Tools. Text for World Bank’s website on Inequality, Poverty and Socio-economic performance. 1999.
- Lou, Y., Yin, Y., and Lawphongpanich, S. Robust congestion pricing under boundedly rational user equilibrium. *Transportation Research Part B: Methodological*, 44(1): 15–28, 2010.
- Lovelace, R. and Ballas, D. Truncate, replicate, sample: A method for creating integer weights for spatial microsimulation. *Computers, Environment and Urban Systems*, 41:1–11, 2013.
- Lovelace, R. and Dumont, M. *Spatial microsimulation with R*. CRC Press, 1 edition, 2016.
- Lovelace, R., Ballas, D., and Watson, M. A spatial microsimulation approach for the analysis of commuter patterns: from individual to regional levels. *Journal of Transport Geography*, 34:282–296, 1 2014.
- Mahmoud, M. S., Habib, K. N., and Shalaby, A. Park-and-Ride access station choice model for cross-regional commuting: case study of Greater Toronto and Hamilton Area, Canada. *Transportation Research Record*, 2014.

- Massachusetts Central Transportation Planning Staff. Massachusetts Travel Demand Model [data], 2018.
- Massachusetts Department of Transportation. Massachusetts Travel Survey [data file]. Technical report, Massachusetts Department of Transportation, 2012.
- McFadden, D. Modelling the choice of residential location. *Spatial Interaction Theory and Planning Models*, 673(477):75–96, 1978.
- McFadden, D. and Train, K. Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, 15(5):447–470, 2000.
- Mosteller, F. Association and Estimation in Contingency Tables. *Journal of the American Statistical Association*, 63(321):1, 3 1968.
- Newell, G. F. Scheduling, Location, Transportation, and Continuum Mechanics: Some Simple Approximations to Optimization Problems. *SIAM Journal on Applied Mathematics*, 25(3):346–360, 11 1973.
- Openshaw, S. and Rao, L. Algorithms for reengineering 1991 Census geography. *Environment and planning A*, 27(3):425–446, 1995.
- Pathak, R., Wyczalkowski, C. K., and Huang, X. Public transit access and the changing spatial distribution of poverty. *Regional Science and Urban Economics*, 66:198–212, 9 2017.
- Pickrell, D. A Desire Named Streetcar: Fantasy and Fact in Rail Transit Planning. *Journal of the American Planning Association*, 58(February 2015):158–176, 1992.
- Polak, J. W. The influence of alternative traveller learning mechanisms on the dynamics of transport systems. In *Transportation planning methods. Proceedings of seminar D held at AET European Transport Conference, Loughborough University, U.K.*, volume P423, pages 14–18, 9 1998.
- Pritchard, D. R. and Miller, E. J. Advances in population synthesis: fitting many attributes per agent and fitting to household and person margins simultaneously. *Transportation*, 39(3):685–704, 5 2012.
- Recker, W. W. A bridge between travel demand modeling and activity-based travel analysis. *Transportation Research Part B: Methodological*, 35(5):481–506, 6 2001.
- Rossi-Hansberg, E. Optimal urban land use and zoning. *Review of Economic Dynamics*, 7(1):69–106, 1 2004.
- Saadi, I., Mustafa, A., Teller, J., Farooq, B., and Cools, M. Hidden Markov Model-based population synthesis. *Transportation Research Part B: Methodological*, 90: 1–21, 8 2016.

- Sandoval, J. O., Cervero, R., and Landis, J. The transition from welfare-to-work: How cars and human capital facilitate employment for welfare recipients. *Applied Geography*, 31(1):352–362, 1 2011.
- Schaller, B. Empty Seats, Full Streets. Fixing Manhattan’s Traffic Problem. *Schaller Consulting, December*, 1(3), 2017.
- Shires, J. D. and de Jong, G. C. An international meta-analysis of values of travel time savings. *Evaluation and Program Planning*, 32(4):315–325, 11 2009.
- Simini, F., González, M. C., Maritan, A., and Barabási, A.-L. A universal model for mobility and migration patterns. *Nature*, 484(7392):96–100, 2012.
- Sivakumaran, K., Li, Y., Cassidy, M. J., and Madanat, S. Cost-saving properties of schedule coordination in a simple trunk-and-feeder transit system. *Transportation Research Part A: Policy and Practice*, 46(1):131–139, 1 2012.
- Small, K. A. Valuation of travel time. *Economics of Transportation*, 1(1-2):2–14, 2012.
- Small, K. A., Winston, C., and Yan, J. Uncovering the Distribution of Motorists’ Preferences for Travel Time and Reliability. *Econometrica*, 73(4):1367–1382, 7 2005.
- Smith, N. New Globalism, New Urbanism: Gentrification as Global Urban Strategy. *Antipode*, 34(3):427–450, 6 2002.
- Solow, R. M. Congestion, Density and the Use of Land in Transportation. *The Swedish Journal of Economics*, 74(1):161, 3 1972.
- Solow, R. M. Congestion Cost and the Use of Land for Streets. *The Bell Journal of Economics and Management Science*, 4(2):602, 1973.
- Solow, R. M. and Vickrey, W. S. Land use in a long narrow city. *Journal of Economic Theory*, 3(4):430–447, 12 1971.
- Steimetz, S. S. and Brownstone, D. Estimating commuters’ “value of time” with noisy data: a multiple imputation approach. *Transportation Research Part B: Methodological*, 39(10):865–889, 12 2005.
- Stephan, F. F. An Iterative Method of Adjusting Sample Frequency Tables When Expected Marginal Totals are Known. *The Annals of Mathematical Statistics*, 13 (2):166–178, 1942.
- Stouffer, S. A. Intervening opportunities: a theory relating mobility and distance. *American sociological review*, 5(6):845–867, 1940.
- Sun, L. and Erath, A. A Bayesian network approach for population synthesis. *Transportation Research Part C: Emerging Technologies*, 61:49–62, 12 2015.

- Tabuchi, T. Bottleneck Congestion and Modal Split. *Journal of Urban Economics*, 34(3):414–431, 11 1993.
- Tong, L., Zhou, X., and Miller, H. J. Transportation network design for maximizing space–time accessibility. *Transportation Research Part B: Methodological*, 81:555–576, 11 2015.
- Train, K. A validation test of a disaggregate mode choice model. *Transportation Research*, 12(3):167–174, 6 1978.
- Train, K. *Qualitative choice analysis: Theory, econometrics, and an application to automobile demand*, volume 10. MIT press, 1986.
- Train, K. E. *Discrete choice methods with simulation*. Cambridge university press, 2009.
- Turnbull, K. F., Pratt, R. H., Evans, I. V., Levinson, H. S., and others. *Traveler Response to Transportation System Changes Handbook, Third Edition: Chapter 3, Park-and-Ride/Pool*. Transportation Research Board, Washington, D.C., 8 2004. ISBN 978-0-309-37513-9.
- Uchida, K. Estimating the value of travel time and of travel time reliability in road networks. *Transportation Research Part B: Methodological*, 66:129–147, 8 2014.
- U.S. Census Bureau. 2010 Decennial Census Tables, 2010.
- U.S. Census Bureau. 2011-2015 5-year American Community Survey Public Use Microdata Sample, 2015.
- U.S. Census Bureau American Community Survey. 2011-2015 ACS 5-year PUMS, 2015.
- U.S. Census Bureau Center for Economic Studies. Longitudinal Employer-Household Dynamics Origin-Destination Employer Statistics, 2017.
- Vickrey, W. S. Congestion Theory and Transport Investment. *The American Economic Review*, 59(2):251–260, 1969.
- Voas, D. and Williamson, P. An evaluation of the combinatorial optimisation approach to the creation of synthetic microdata. *Population, Space and Place*, 6(5): 349–366, 2000.
- Voorhees, A. M. A general theory of traffic movement. *Transportation*, 40(6):1105–1116, 11 1956.
- Wang, G., Jia, N., Ma, S., and Qi, H. A rank-dependent bi-criterion equilibrium model for stochastic transportation environment. *European Journal of Operational Research*, 235(3):511–529, 2014.

- Watling, D. User equilibrium traffic network assignment with stochastic travel times and late arrival penalty. *European journal of operational research*, 175(3):1539–1556, 2006.
- Wheaton, W. C. Land Use and Density in Cities with Congestion. *Journal of Urban Economics*, 43(2):258–272, 3 1998.
- Wichman, C. J. and Cunningham, B. Notching for Free : Do Cyclists Reveal the Value of Time?*. (September), 2017.
- Wilson, A. G. *Entropy in urban and regional modelling*, volume 1. Routledge, 2011.
- Wong, D. W. S. The Reliability of Using the Iterative Proportional Fitting Procedure. *The Professional Geographer*, 44(3):340–348, 8 1992.
- Yang, H. and Bell, M. G. H. Traffic restraint, road pricing and network equilibrium. *Transportation Research Part B: Methodological*, 31(4):303–314, 8 1997.
- Ye, X., Konduri, K., Pendyala, R. M., Sana, B., and Waddell, P. A methodology to match distributions of both household and person attributes in the generation of synthetic populations. In *88th Annual Meeting of the Transportation Research Board, Washington, DC*, 2009.
- Zhang, K., Mahmassani, H. S., and Lu, C.-C. Dynamic pricing, heterogeneous users and perception error: Probit-based bi-criterion dynamic stochastic user equilibrium assignment. *Transportation Research Part C: Emerging Technologies*, 27:189–204, 2013.
- Zhang, X., Yang, H., and Huang, H.-J. Improving travel efficiency by parking permits distribution and trading. *Transportation Research Part B: Methodological*, 45(7): 1018–1034, 8 2011.
- Zhu, C., Jia, B., Han, L., and Gao, Z. Commuting Pattern with Park-and-Ride Option for Heterogeneous Commuters. *Discrete Dynamics in Nature and Society*, 2013:1–8, 2013.
- Zhu, Y. and Ferreira, J. Synthetic Population Generation at Disaggregated Spatial Scales for Land Use and Transportation Microsimulation. *Transportation Research Record: Journal of the Transportation Research Board*, 2429:168–177, 2014.