Spring 4-4-2023

Urban Public Transportation Planning with Endogenous Passenger Demand

Yifei Sun
Dartmouth College, yifei.sun.th@dartmouth.edu

Urban Public Transportation Planning with Endogenous Passenger Demand

A Thesis
Submitted to the Faculty
in partial fulfillment of the requirements for the
degree of

Doctor of Philosophy

by Yifei Sun

Thayer School of Engineering
Dartmouth College
Hanover, New Hampshire

Sep 2022

Examining Committee:

Chairman _____
Vikrant S. Vaze

Member _____
Geoffrey G. Parker

Member _____
Robert A. Shumsky

Member _____
António Pais Antunes

_____
F. Jon Kull, Ph.D.
Dean of Guarini School of Graduate and Advanced Studies

Abstract

An effective and efficient public transportation system is crucial to people's mobility, economic production, and social activities. The Operations Research community has been studying transit system optimization for the past decades. With disruptions from the private sector, especially the parking operators, ride-sharing platforms, and micro-mobility services, new challenges and opportunities have emerged. This thesis contributes to investigating the interaction of the public transportation systems with significant private sector players considering endogenous passenger choice. To be more specific, this thesis aims to optimize public transportation systems considering the interaction with parking operators, competition and collaboration from ride-sharing platforms and micro-mobility platforms. Optimization models, algorithms and heuristic solution approaches are developed to design the transportation systems. Parking operator plays an important role in determining the passenger travel mode. The capacity and pricing decisions of parking and transit operators are investigated under a game-theoretic framework. A mixed-integer non-linear programming (MINLP) model is formulated to simulate the player's strategy to maximize profits considering endogenous passenger mode choice. A three-step solution heuristic is developed to solve the large-scale MINLP problem. With emerging transportation modes like ride-sharing services and micro-mobility platforms, this thesis aims to co-optimize the integrated transportation system. To improve the mobility for residents in the transit desert regions, we co-optimize the public transit and ride-sharing services to provide a more environment-friendly and equitable system. Similarly, we design an integrated system

of public transit and micro-mobility services to provide a more sustainable transportation system in the post-pandemic world.

Acknowledgements

First and foremost, I would like to give my sincere gratitude to my advisor Professor Vikrant Vaze. It was my honor to have him as my first supervisor when I started my academic journey. Having just switched my major from Economics to Operations Research, I was both confused and anxious about exploring this new area. Professor Vaze was a patient and intelligent guide with immense knowledge. I still remember the time we sat together for three hours to figure out the formulation of a model and the time when he spent two hours helping me debug the code. He celebrated my progress and gave me guidance when research came to stagnation. He managed to "infect" me with his contagious passion for research. I earnestly appreciate you being my advisor during my five years of Ph.D.

I would also like to thank Professor António Pais Antunes for his excellent mentoring and guidance. It was my great pleasure to work with you on the Parking project. That was one of my first projects in urban transportation planning. You always provided me with informative resources and insightful suggestions. It was an unforgettable experience to collaborate during the darkest days of the Pandemic across the Atlantic Ocean.

Thanks should also go to Professor Geoffrey Parker and Professor Robert Shumsky. Professor Parker was a committee member on both my Ph.D. qualifying exam and thesis defense. I still remember his questions regarding business and social impacts on my Ph.D. qualifying exam. These questions enable me to focus more on the greater influence of my research. Professor Shumsky provided invaluable insights into the post-pandemic transportation system which became an important part of my thesis.

I thank my fellow labmates for stimulating great research ideas and discussions. I remember the days when I was working with Navid to finish the quarterly report before the deadline. I remember the numerous research discussions I had with Pushpendra, Pheobe, and Lilly. I hope all of you enjoy the rest of your Ph.D. and wish you the best for your future career.

It was my honor to work with the Thayer staff. I would like to thank everyone in Thayer Computing, Mark, Matt, Jan, Dan, and Zack in particular. Sorry about the trouble I caused when I crashed the servers. And thank you Mark for helping me set up my first parallel computing code. I would also like to thank Amy and Liz in Thayer Career Service for providing invaluable interview tips and career advice. I would also like to thank Candace and Holly for providing support for the International student organization.

I would like to thank Sameed Ali, the soon-to-be Doctor Ali, for giving me constant support and love during the past five years. Thanks for being my go-to CS handbook when I have any questions remotely related to computers. Thanks for taking care of my frustration when I was screaming at my screen searching for a bug. Thanks for being my cheerleader when I am going through 5 interviews in a week. I could not imagine going through my Ph.D. at Dartmouth without you. It wouldn't be the same experience and I wouldn't be the same person I am today.

Finally, I would like to give my deepest appreciation to my parents. Thanks for your unconditional love and support. I haven't met them in person for the past three years. Hopefully, we will see each other in the near future.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Designing a sustainable and equitable transportation system is crucial to society. The public transportation system plays an essential role in providing mobility to residents, especially to socio-demographic minorities and marginalized communities. An efficient public transit system can attract people away from using private vehicles, and taxis — modes that produce more greenhouse gas emissions. The urban transportation system is a complex system including multiple decision-makers including transit agencies, passengers, parking agencies, and emerging traveling service operators. With the increasing adoption of emerging traveling modes like ride-sharing and micro-mobility, it is increasingly important for the city planner to design a public transportation system considering the interaction of different operators, traveling modes, and passenger choices.

In 2010, the transportation sector accounted for 13.5% of global greenhouse gas, and road transportation emitted 10% of global greenhouse gas (Burke, 2010). The transportation sector is predicted to be responsible for 30–50% of CO2 emissions by 2050 (Nebojsa, 2000). Public transportation is an essential segment of building a sustainable transportation system. Metro only generates less than 10g CO2eq/pkt compared to 230g CO2eq/pkt for

gasoline sedan (de Bortoli, 2021). An efficient design of public transportation is essential to countering climate change.

A better urban transportation system is also important to build a more equitable society. The lack of access to public transportation poses a severe threat to people's mobility and financial well-being. A recent study by (Jiao and Bischak, 2018) shows that over 1 in 8 residents are living in transit desert regions in the most severely affected cities. Residents in the transit deserts are forced to travel using private vehicles, taxis, or ride-hailing services which produce more greenhouse gas emissions. Lack of access to public transportation often coincides with a lack of access to high-quality healthcare and fresh food which makes the inequality problem even worse.

The pandemic has forced city planners to rethink the urban planning process. Transit ridership decreased significantly during the pandemic which made the financial condition of transit agencies even worse EBP (2021). The possible continuation of remote-working practices may prolong the situation. Therefore, the design of an efficient public transportation system is of great importance for the post-pandemic world.

This thesis aims to design an efficient urban transportation system by investigating the interaction between public transit operators and other key operators, co-optimizing the transportation system with other emerging travel modes considering endogenous passenger demand. To be more specific, we analyze the capacity-pricing interaction of public transit and parking operators. The impacts of different policies are evaluated in terms of the reduction of vehicle travel times and congestion in the city center. We also design a transportation system where the transit network, frequencies and on-demand fleet size are optimized to minimize the social cost in the transit desert regions. We analyze the equity improvement of our transportation system by analyzing the disutility decrease for people with no private vehicles and lower-income communities. We provide a framework to co-design the public transit as well as micro-mobility system by determining the transit network, bike lane, and micro-mobility deployment to cater to commuters navigating the post-pandemic world.

1.2 Interaction of Transit and Parking Operators

An efficient public transit system is important to alleviate environmental problems like greenhouse gas emissions and increase social mobility. The availability of affordable parking spaces can incentivize people to use private vehicles instead of using public transportation. When parking operators lower the parking fees or increase parking capacities in the central business district, it will attract more people away from taking public transit and toward using private vehicles. Similarly, when transit increases its frequencies or lowers its fare, people are more likely to take public transit rather than drive to work. Therefore, the interaction between the pricing and capacity decisions of transit and parking operators is worth investigating in order to design a better transportation system.

We contribute by proposing a two-stage game-theoretic model and algorithmic framework to capture the interactions between the decisions of transit and parking operators. Both operators maximize their own profits under various forms and degrees of government regulation. In the first stage of the game, they decide their respective capacities in terms of transit frequencies and parking capacities. In the second stage, both operators make pricing decisions, namely transit fares and parking fees. The passengers make their travel mode decisions based on various costs, including notably the extra travel time induced by congestion. A novel heuristic solution method is developed using a semi-approximate approach to make the two-stage model tractable and solvable. We develop an original end-to-end computational framework to simulate, solve and evaluate the solutions of two-stage games for study networks with realistic characteristics. The impacts of various capacity decisions by the operators as well as various incentive-generation and regulatory strategies by the city planners are evaluated.

1.3 Co-optimization of Transit and Ride-sharing System

Transit deserts pose a significant hurdle to local mobility and social equity. In the worst affected regions, more than 10% of residents don't have sufficient access to public transportation. The problem also has a disproportionate impact on people with no private vehicles and lower-income households since they have to pay for the more expensive taxis or purchase private vehicles. However, transit agencies do not appear to be capable of solving this problem by simply expanding their network since they already suffer from severe financial deficits. Similarly, this problem cannot be solved simply by offering more on-demand services in the transit desert regions. On-demand services may introduce more congestion to the city center and possibly even more greenhouse gas emissions than driving a private vehicle. Therefore, we consider designing an integrated transportation system including fixed-route transit and on-demand services.

We provide a comprehensive framework to address the transit desert problem by co-optimizing transit networks and frequencies as well as the fleet size of on-demand services. A mixed-integer non-linear programming model is developed to optimize transit services (including transit network design and frequencies for each line) and on-demand fleet size while considering the interaction with on-demand services and passenger mode choice. In order to solve the large-scale non-linear non-convex model, we design an original two-step heuristic solution approach. In the first step, we apply a rounding heuristic method to select transit lines from the set of potential transit lines. In the second step, we determine the frequency for each transit line and the fleet size for on-demand services by iteratively running a mixed-integer second-order conic programming (MISOCP) model. The framework is applied to those parts of the greater Boston area that the MBTA has identified as transit desert regions. The results indicate that we are able to reduce the total systemwide cost by 3%. Low-income commuters and communities with no private vehicles enjoy a much greater disutility reduction of 14% and 21% respectively. It demonstrates that our framework is able to alleviate the transit desert problem and provides a more equitable transportation

system.

1.4 Co-optimization of Transit and Micro-mobility System

The pandemic has revealed both challenges and opportunities within the current transportation system. Due to social distancing requirements and remote-working practices, transit ridership was reduced by 90% at the beginning of the pandemic. However, it also boosted the growth of micro-mobility traveling where it is easier to comply with the social distancing regulations. Governments built bike lanes to accommodate increasing micro-mobility trips. Bike-sharing and scooter-sharing platforms are also taking steps to expand their services. Life may never be the same in the post-pandemic world. Work-from-home is likely to continue in the form of hybrid working practices after the pandemic. An integrated design of public transportation and micro-mobility systems is essential to the city planner in the post-pandemic world.

We aim to co-optimize the system of fixed-route transit, bike lanes, and micro-mobility service operations considering endogenous passenger demand. The city planner decides the transit network, frequencies, bike lane network, and micro-mobility vehicle deployment. The passengers can choose to take any alternative out of fixed-route transit, bike-sharing services, multi-modal services, or outside options based on the attractiveness of each mode. A mixed-integer non-linear programming model is formulated to minimize the weighted cost of passenger disutility, operational costs, and environmental costs.

1.5 Thesis Outline

The contents of the thesis are organized as follows. In **Chapter 2**, we focus on analyzing the interaction of capacity-pricing decisions of public transit and parking operators. A two-

stage game-theoretic model is developed to formulate the problem and solved using a novel semi-approximate heuristic. **Chapter 3** aims to generate an optimal transportation system of fixed-route transit and on-demand services considering endogenous passenger demand. Transit network, frequencies, and on-demand fleet size are decided using an MINLP model and a two-step heuristic solution approach. **Chapter 4** focuses on optimizing an integrated transportation network of public transit, bike lane network, and micro-mobility operations in response to challenges in the post-pandemic world. Chapter 5 summarizes the main conclusions and lists the directions for future research.

Chapter 2

Comprehensive Public Transit Design Considering Parking Operator's Response Using a Tractable Two-stage Framework

2.1 Introduction

Transit systems are advocated both by international organizations (such as the United Nations and the European Union) and by the governments of virtually every country as a beneficial alternative to private automobiles for satisfying the transportation needs of urban commuters. Indeed, such systems can contribute significantly to alleviating the current environmental problems (fossil fuel consumption, greenhouse gas, pollutant emissions, and traffic congestion) and social exclusion concerns associated with urban mobility (Schiller et al., 2010; Miller, 2014).

However, transit usage in the United States is relatively low compared to other developed countries. To make things worse, transit ridership has been declining for the past

decades. According to a report published by National Geographic (2009), transit usage in the United States ranked the lowest among the 17 countries investigated. Freemark (2021) concluded that the transit is struggling to retain ridership over the past half-century. Only 5% of workers get to work by bus or train, compared to 9% half a century ago. An efficient and accessible transit network is essential to a sustainable urban transportation system (Arbex and da Cunha, 2015; Cipriani et al., 2012a). Therefore, many studies have focused on improving the transit network design (Arbex and da Cunha, 2015; Cipriani et al., 2012a; Liang et al., 2019; Feng et al., 2019; Canca et al., 2019), conducting the optimal frequency level (Wei et al., 2021; Jiang et al., 2022; Sun and Szeto, 2019), and setting the optimal transit fare (Zhou et al., 2019; Guo et al., 2021; Yang and Tang, 2018; Li et al., 2009).

However, previous studies rarely considered external effects outside the public transit system. The change in the transit network design will not only have an impact on the utilities of those passengers taking transit, but it also exerts an impact on traveler mode choice, congestion, and the decisions of other operators like taxis, ride-hailing operators, and parking operators. As the data from U.S. Department of Transportation (2016) shows the major alternative to taking transit is driving private vehicles. According to U.S. Department of Transportation (2016), 76.6% of commuters drive alone to work in the United States, while only 5.2% of commuters use public transit. Parking plays a significant role when commuters make their decisions on whether to drive to work (Franco, 2017). Limited parking capacities and higher parking fees can deter people from driving to adopting other modes of commuting (Proulx et al., 2014; Hess, 2001; Yan et al., 2019). On the contrary, free parking and ample parking spots will discourage people to switch from driving. When transit operators design the public transit system, it is also important to consider how parking operators will respond to transit frequencies and fare changes. Therefore, in this paper, we want to investigate the optimal transit network design considering the response of the parking operators.

In this paper, we contribute toward bridging that literature gap by presenting a versatile

model and algorithm framework to optimize the transit frequency setting by considering endogenous passenger choice, road congestion, and the interactions with the decisions of parking operators. Transit operators minimize systemwide passenger travel costs. The parking operators are assumed to maximize their own profits. We provide a model framework where we investigate different scenarios where transit operators can either respond to parking operators' decisions flexibly or transit operators will not respond after making their decisions. The parking operators can either not respond, respond by adjusting the parking fees alone or adjusting both parking fees and parking capacities. We model endogenous passenger choice using the Sales Based Linear Programming (SBLP) model from Gallego et al. (2015). The model is a nonlinear mixed-integer programming model. In order to make it tractable, we develop a semi-approximate heuristic algorithm to solve this complicated model. We evaluate the parking operators' responses under various combinations of transit decisions. Then we use second-order regression to approximate transit revenue and systemwide passenger travel cost. Finally, we obtain the optimal frequency levels by minimizing systemwide passenger travel costs while satisfying a certain level of farebox recovery rate. This framework is tested using the data from Boston Proper (including Back Bay, Bay Village, Beacon Hill, Chinatown, Downtown Crossing, Fenway-Kenmore, Financial District, Government Center, Leather District, North End, South End, West End, and the Waterfront), Massachusetts. We obtained the optimal transit frequencies and transit fares considering the parking operator's response. Furthermore, we evaluate the impact of various government policies and provide valuable policy insights.

2.2 Related Literature

2.2.1 Transit Pricing and Network Design Literature

A review of the existing literature on optimizing transit pricing and transit network design indicates that the societal importance of these systems plays a central role in these stud-

ies. The first stream is focused on establishing the optimal transit pricing and subsidization rules. Li et al. (2009) investigate the optimal transit pricing structures under monopoly market and oligopoly market respectively. They developed a network-based model to take the uncertainty and transit arrival reliability into account. Yang and Tang (2018) investigated how to use fare rewards to incentivize passengers from traveling during peak hours. The total revenue was kept unchanged, while the total passenger traveling cost was reduced by 25%. However, they did not consider an endogenous demand where passengers can switch to other travel modes when the price becomes higher. Zhou et al. (2019) studies the equity aspect of transit fare. The proposed methods such as "trajectory rebuilding" and "fare matching" demonstrate the impact of transit fares on equity. Guo et al. (2021) applied microeconomic models to optimize time-dependent transit fares. They concluded that time-dependent pricing can avoid cross-subsidization among travelers in different time periods.

Another stream of literature focuses on the transit network design and frequency setting problem (TNDFSP). TNDFSP aims to find a set of transit lines and their associated frequencies for a transit system. It is a combinatorial problem that is computationally difficult to solve (Arbex and da Cunha, 2015). Cipriani et al. (2012a,b); Arbex and da Cunha (2015); Feng et al. (2019); Liang et al. (2019); Canca et al. (2019); Szeto and Jiang (2014a); Asadi Bagloee and Ceder (2011) solved the TNDFSP using Genetic Algorithm, column generation, branch and cut, bee colony algorithm, and metaheuristic method. However, the studies mentioned above did not consider optimizing the transit price along with the transit network and frequency setting. Bertsimas et al. (2020) demonstrated the benefit of optimizing transit fare and transit frequencies jointly in a multimodal transportation network. Zhang et al. (2018) compared the frequencies and prices under profit maximization schemes and social surplus maximization schemes. They also investigated the impacts of fare regulation, frequency regulation, goal regulation, and fiscal regulation. However, none of these papers take road congestion into account.

Capacity and pricing are among the most important decisions by transit and other operators, with numerous studies focusing on optimizing these decisions (e.g., García and Marín, 2002; Savage, 2010; Abdelfatah and Taha, 2014; Huang et al., 2016). High transit frequency and affordable fares make transit more desirable to passengers while limited and expensive parking spaces deter trips by private vehicles. Moreover, changes in the capacities and prices of one operator impact the other operator by changing the passengers' travel choice decisions. Prices are much easier to adjust and are often decided much later, than physical capacity allocation decisions. Pricing in service industries is becoming increasingly dynamic with many service industries moving away from fixed pricing strategies (e.g., McGill and van Ryzin, 1999; Kimes, 2003; Nicas, 2015; Schechner, 2017). Under dynamic pricing, service firms first decide facility locations, schedules, and capacities, and then prices are decided through complex interactions between the competing firms and the customers. A two-stage, rather than single-stage, model is a suitable way of modeling such competition. Additionally, the insights and results from the single-stage models are known to be quite different from those from two-stage models, and the latter has much better alignment with empirical data (Hansen, 1990; Vaze and Harder, 2017; Harder and Vaze, 2017). Furthermore, two-stage optimization is widely used in urban transportation settings (Lindsey, 2012; Wan and Zhang, 2013; de Rus and Socorro, 2019).

In our paper, the transit operator optimizes both transit frequencies and transit fares. One important issue worth discussing is the order of decision-making for transit fares and transit frequencies. In practice, both transit fares and transit frequencies have changed over time. In 2016, Single-ride bus fares increased from \$1.60 to \$1.70 and a single ride on the subway increased from \$2.10 to \$2.25 (DeCosta-Klipa, 2016). In July 2021, The Massachusetts Bay Transportation Authority (MBTA) Express Bus routes Outer Express were reduced from \$5.25 to \$4.25 per ride (MBTA, 2021b). Metropolitan Transportation Authority (MTA) increased their base transit fare in 2003, 2009, 2013, and 2015 from \$2 to \$2.25, \$2.5, and \$2.75 (Wikipedia, 2022). Chicago Transit Authority (CTA) increased

the fare by $0.25 in 2009 and increased it again in 2019 (CBS, 2018). CTA is proposing a reduction in transit fares to boost ridership after the hit of the pandemic (Ford, 2021). Regarding changes in the transit frequencies and schedule, MTA reduced the public transit services by $100 million with bus services taking the biggest hit. Over 150 subway stations were affected and 21 local and 12 express bus services were discontinued (MTA, 2010). Some bus services were partially restored in 2013 based on customers' feedback (CBS, 2013). In 2008, services were increased on subway line 3 in response to increasing Harlem ridership. From a practical standpoint, it seems like transit fare changes are more often than transit frequency changes.

2.2.2 Response From Other Operators

Besides transit operators, there are other major stakeholders in the urban transportation system. The customer will respond to the changes in transit scheduling and transit fare by changing their travel modes. Taxi operators may consider changing their prices. Ride-hailing platforms, which adopted a much more flexible pricing strategy, can modify the two-sided market prices to adjust the supply of drivers and passenger demand. Parking operators are able to adjust their prices and capacities to maximize their profit. Among all these stakeholders, we focus on the response from the parking operator. Driving private vehicles is the major alternative to taking public transit in the United States. There are more than 70% of commuters drive to work nationally. In Boston city, 48.1% of commuters drive to work while 33% of commuters take transit (Boston Transportation Department, 2017). Moreover, parking plays a significant role in the decision of commuter travel mode. Therefore, it is important to take the parking operators' responses into account when deciding the optimal transit frequencies and fares.

Separately, there is a considerable amount of literature on the impact of parking pricing on travel mode choice and congestion. Arnott and Rowse (1999), as well as Arnott and Inci (2006), presented a downtown parking model which investigates the interaction

between traffic congestion and saturated on-street parking. Arnott (2006) examined the optimal parking policies under equilibrium in the parking garage market. Liu and Geroliminis (2016) investigated how cruising for parking reshapes the morning commute and developed a dynamic model of pricing to reduce total social cost, including cruising time cost, moving time cost, and schedule delay cost. Eftekhari and Ghatee (2017) developed models to estimate the impact of dynamic parking prices on the central business district travel demand. Nourinejad and Roorda (2017) applied a variational inequality model to investigate the impact of hourly parking pricing on travel demand.

Besides parking fees, parking operators can also adjust their parking capacities. In 2001, there were 13830 parking spaces under construction or approved and 11890 parking spaces under review and proposed (Boston Transportation Department, 2017). In East Boston, there was a reduction in parking inventory of 190 parking spaces in 2001 (Boston Transportation Department, 2017). It is nonetheless harder to adjust than parking fares, parking operators can still manage to adjust parking capacities in the long term. In order to ameliorate congestion in the city center and discourage people from driving, city planners have passed regulations to limit parking capacities. The Boston Air Pollution Control Commission administers caps for all or part of the parking supply in Boston Proper, South Boston, and East Boston (City of Boston, 2017). New York City eliminates mandatory parking minimums which require new buildings to include a fixed number of off-street parking spaces (Cuba, 2022). Therefore, government regulations also play a significant role in deciding parking capacities.

An integrated approach to transit and parking planning was considered by Cavadas and Antunes (2018). They developed an optimization model to minimize the joint operating deficit of the transit and parking operators while ensuring required minimum levels of mobility in a city. However, to the best of our knowledge, the topic of transit and parking interactions under competition has not been explored in any part of the existing literature.

2.2.3 Literature Gaps, Contributions

Through the literature review, we found that the interaction of transit and parking operators is not well studied. When optimizing transit frequency or fares, most studies use private vehicles as an alternative travel mode without considering the reaction of parking operators to transit capacity and pricing changes (e.g., Liu et al., 2009; Basso and Jara-Díaz, 2012; Li et al., 2012; Wang et al., 2016). Numerous studies have investigated the impact of transit frequency and fares on reducing city congestion from a transit operator's perspective or the impact of parking fee regulations on reducing city congestion from the parking operator's perspective (e.g., Basso and Jara-Díaz, 2012; Azari et al., 2013). Our study bridges the gap between the currently separate literature streams of transit and parking optimization by optimizing transit decisions by considering the response from the parking operators.

Table 2.1 Model framework

			Upfront Pricing	Flexible Second Stage Pricing	
No parking response			TF+TP (S1)		
Parking Responds	PP only	Stage I	TF + TP (S2)	TF	(S4)
		Stage II	PP	PP & TP	
	PP and PC	Stage I	TF + TP (S3)	TF	(S5)
		Stage II	PC + PP	PC + PP & TP	

*TF denotes transit frequencies; TP denotes transit price; PC denotes parking capacity; PP denotes parking price; S1-S5 denote scenario 1 to scenario 5

We propose a model framework to provide optimal decisions under different assumptions and strategies. For simplicity, we only consider the case where there are only one transit operator and one parking operator. Regarding the parking operator, We evaluate the scenarios under three different assumptions: (1) the parking operator does not respond to the decisions of the transit operator, (2) it responds only by adjusting parking fees, (3) or it responds by adjusting both parking fees and parking capacities. For the transit operator, we evaluate two different strategies: (1) Upfront pricing strategy, which means the transit operator does not change its decision in response to the parking operator, and (2) Flexible

second-stage pricing strategy, which means the transit operator adjust its price in response to parking operator's decision.

With different combinations of the assumptions and strategies aforementioned, we developed a model framework as presented in Table 2.1. The rows indicate whether the parking operator will respond to the transit operator's decision and how they will respond. The columns indicate whether transit would respond to the parking operator's decision in Stage II. TF, TP, PC, and PP denote transit frequencies, transit price, parking capacities, and parking price respectively. S1-S5 denotes Scenario 1 to Scenario 5.

In Scenario 1, we assume that the parking operator doesn't respond and the transit operator optimizes both frequencies and transit fare. In Scenario 2, transit frequencies and transit fare are optimized in the first stage and parking price is optimized in the second stage given transit decisions. In Scenario 3, the transit operator decides frequencies in the first stage. Parking prices and transit fare will reach an equilibrium in the second stage. Scenario 4 and 5 follow similar logic as Scenario 2 and 3 except that the parking operator can also adjust parking capacities in Stage II.

We use the solution concept called subgame perfect pure strategy Nash equilibrium (SPPSNE) for solving the flexible second-stage pricing scenarios. This solution concept offers multiple advantages. First, it is a more intuitive refinement of the general notion of a pure strategy Nash equilibrium (PSNE) for extensive form (i.e., multi-stage) games. Second, in a recent study involving transportation capacity and pricing competition within a different context, this solution concept has been shown to have promising mathematical, computational, and empirical properties (Harder and Vaze, 2017; Vaze and Harder, 2017).

Our goal is to develop a realistic and tractable computational framework for optimizing transit frequencies and transit fares considering the response from parking operators. An improved and more sustainable urban transportation system is typically associated with lower levels of traffic congestion, fossil fuel consumption, and greenhouse gas emissions, enhanced transit ridership, financially sustainable transit systems, and improved access

to urban mobility for passengers. This paper makes three major contributions. 1) We develop an original modeling framework to optimize transit frequencies and transit fares considering interactions with parking operators. The results in Section 2.6 demonstrate the significance of the parking operator's impact on the transit operator's decision. To the best of our knowledge, ours is the first study to model the transit decision considering the parking operator's response. 2) After highlighting the complexity associated with the challenge of solving the proposed model, we develop a new solution method that uses a semi-approximate approach to make the two-stage model tractable and solvable. To the best of our knowledge, ours is the first study to present any tractable approach for solving realistic-sized multi-stage games in any urban transportation setting. 3) Finally, using this framework, we quantitatively assess the impacts under various scenarios mentioned above in Boston city. We demonstrated the accuracy and scalability of our approach. We are also able to draw managerial insights in a realistic city setting.

This paper is organized as follows. In Section 2.3, the model framework is described. The semi-approximate solution approach, and various solution-time acceleration heuristics to enhance model tractability, are detailed in Section 2.4. The setup of the computational studies is presented in Section 2.5, while the computational results are described in Section 2.6. Section 2.7 concludes with a summary of our main findings and future research directions.

2.3 Model

We model a transportation system where there is only one transit operator and one parking operator. The transit operator aims to minimize systemwide passenger travel costs and the parking operator is a profit-maximizing entity. The transit operator charges a flat transit fare for all transit trips while the parking operator can charge different parking fees in different regions. Even though there may be multiple competing parking operators in the reality, we

only consider the case where there is only one monopolistic parking operator in this paper.

We aim to optimize transit frequencies and transit fare under different scenarios as stated in Table 2.1. Section 2.3.1 describes the base model (Scenario 1) where we assume the parking operator does not respond to transit operator's decision. This is an assumption that is widely used in conventional literature. Scenario 1 provides a baseline for the other four scenarios. In Section 2.3.2, we assume the transit operator decides the pricing strategy up front; thus the transit frequencies and fare are decided in Stage I and the parking operator responds in Stage II (corresponding to Scenarios 2 and 3). In Section 2.3.3, we assume that the parking operator responds to the transit operator's decision and the transit operator also adjusts its price leading to an equilibrium in Stage II (scenarios 4 and 5).

2.3.1 Base Model

In the base model, the transit operator optimizes both frequencies and fares while the parking operator's decisions remain unchanged. We consider endogenous demand where travelers can choose between taking public transportation, driving private vehicles, and outside options. Road congestion is also taken into account for private vehicle traveling. This allows us to analyze how an improvement of the transit design will have an impact on ameliorating traffic congestion which is not well-studied in the previous literature.

In most cities, the transit operator is not a merely profit-maximizing agency, it also shoulders the responsibility to provide accessibility and mobility to local travelers. However, the transit operator also has financial incentives to recover the cost of investment and daily operation. Therefore, in our model, the transit operator minimizes passenger traveling costs while satisfying a certain farebox recovery rate by optimizing transit frequencies and transit fare.

2.3.1.1 Objective Function

In most cities, the transit operators are not profit-maximizing agencies. They are heavily subsidized by the government to provide sufficient accessibility and mobility to travelers. However, the financial condition is also of great importance to the transit agency. Therefore, this is a bi-objective problem where we want to minimize both systemwide passenger travel costs and financial loss. We solve this bi-objective problem by minimizing systemwide passenger travel cost (as shown in the Objective Function 2.1) while having farebox recovery rate constraint to control the losses (as shown in constraint 2.2).

$\mathcal{OD}$ denotes the set of OD pairs. $\mathcal{J}$ denotes the set of parking destinations. $\mathcal{R}^D$ denotes the set of routes for driving a private vehicle. C_{od}^O is the passenger travel cost for the outside options in OD pair $od \in \mathcal{OD}$, and D_{od} represents the demand for the OD pair $od \in \mathcal{OD}$.

q_{od}^T	the market share of public transit for OD pair $od \in \mathcal{OD}$
q_{od}^O	the market share of the outside option for $od \in \mathcal{OD}$
q_{od}^D	the market share of driving private vehicles for $od \in \mathcal{OD}$
c_{od}^T	passenger travel cost when using public transit for OD pair $od \in \mathcal{OD}$
c_{od}^D	passenger travel cost when driving private vehicle for OD pair $od \in \mathcal{OD}$

$$Min \sum_{od \in \mathcal{OD}} D_{od}(c_{od}^T q_{od}^T + c_{od}^D q_{od}^D + C_{od}^O q_{od}^O) \tag{2.1}$$

2.3.1.2 Farebox Recovery Constraint

Besides passenger travel cost, another important consideration is the financial conditions. We also have to ensure that the transit operator can recover a certain level of operational cost through passenger fare revenues. Therefore, we add a constraint ensuring a minimum farebox recovery rate. $\mathcal{L}$ is the set of transit lines. C_l is the operating cost for each transit vehicle operating on line $l \in \mathcal{L}$ per unit time. R denotes farebox recovery rate in constraint (2.2).

p^T	transit fare
f_l	transit frequency for transit line $l \in \mathcal{L}$
s_l	transit fleet size for transit line $l \in \mathcal{L}$

$$\sum_{od \in \mathcal{OD}} q_{od}^T p^T \geq R \sum_{l \in \mathcal{L}} C_l s_l \tag{2.2}$$

Transit fleet size is defined in constraints (2.3). f_l is the frequency for transit line l. L_l is the length of the route for transit line l and V^T is the travel speed for transit. The multiplier 2 accounts for the fact that the total length of the line in both directions is $2L_l$.

$$s_l \geq \frac{2L_l f_l}{V^T} \quad \forall l \in L \tag{2.3}$$

2.3.1.3 Travel Mode Attractiveness

The utility of traveling for OD pair od using transit is given by Equation (2.4). This cost is the weighted sum of the in-vehicle time T_{od}^V, transit fare p^T, access utility u_{od}^{AC}. β^V, β^C and β^T are the coefficients for in-vehicle traveling time, transit fare cost, and the alternative-specific constant coefficient for taking transit, respectively. The attractiveness of taking transit (a_{od}^T) is the exponential of the utility u_{od}^T as shown in constraints (2.5).

$$u_{od}^T = \beta^V T_{od}^V + \beta^C p^T + u_{od}^{AC} + \beta^T \quad \forall od \in \mathcal{OD} \tag{2.4}$$

$$a_{od}^T = e^{u_{od}^T} \quad \forall od \in \mathcal{OD} \tag{2.5}$$

The in-vehicle time cost is obtained by multiplying the cost per unit time β^V, and the total in-vehicle time (T_{od}^T) that users need to spend in OD pair od by transit. The in-vehicle

u_r^{DR}	the utility of traveling by driving a private vehicle on route $r \in \mathcal{R}^D$
u_{od}^D	the overall utility for driving private vehicle on OD pair $od \in \mathcal{OD}$
u_{od}^T	the utility for taking public transit for OD pair $od \in \mathcal{OD}$
p_j^D	parking fee at destination $j \in \mathcal{J}$
a_{od}^D	the attractiveness for driving private vehicle for OD pair $od \in \mathcal{OD}$
a_{od}^T	the attractiveness for taking public transit for OD pair $od \in \mathcal{OD}$
u_{od}^{AC}	the access utility for taking public transit for OD pair $od \in \mathcal{OD}$
w_{od}	the waiting time for taking public transit for OD pair $od \in \mathcal{OD}$
t_r^D	the travel time for driving private vehicle for driving route $r \in \mathcal{R}$
z_r^D	binary variable, 1 if route r is taken by commuters, 0 otherwise

time calculation assumes that the transit users always take the shortest path (defined as the one with the minimum travel time), with at most one transfer, along the transit network. Such a shortest path may involve one or zero transfers between different transit lines. Here, a transit line is defined as the sequence of stops served by a transit vehicle. Typically, a particular transit line will be served multiple times a day. The transit network can be thought of as a collection of all such transit lines.

The access utility is the sum of utilities associated with the walking time to and from the transit stops, and the total waiting time at the transit stops. The first component of this access utility function corresponds to the average walking time utility, which is the product of the walking time coefficient, β^K, and the average walking time, K_{od}^T. The second component of the access utility function in Constraints (2.6) is the waiting time utility, which is the product of the waiting time coefficient, β^W, and the average waiting time w_{od}. Constraint (2.7) demonstrated how the waiting time is calculated. l_{od}^1 (l_{od}^2) indicates the index of the transit line for the first (second) segment of OD pair od. If no transfer is needed for the trip, $N_{od}^2 = 0$ and $l_{od}^2 = l_{od}^1$. Otherwise, $N_{od}^2 = 1$.

$$u_{od}^{AC} = \beta^K K_{od}^T + \beta^W w_{od} \quad \forall od \in \mathcal{OD} \tag{2.6}$$

$$w_{od} = \frac{1}{2f_{l_{od}^1}} + \frac{N_{od}^2}{2f_{l_{od}^2}} \quad \forall od \in \mathcal{OD} \tag{2.7}$$

For commuters of OD pair od who choose to drive, they can choose from a set of routes $\mathcal{R}^D_{od}$. Let t^D_r be the car travel time using car route r, β^D be the vehicle depreciation, fuel, and maintenance coefficient per unit distance, and let D^D_r be the total distance traveled by car using route r. Then the generalized travel utility (u^{DR}_r) of driving a car using route r includes the travel time utility, and the vehicle depreciation, fuel, and maintenance utility as expressed in equation (2.8). This utility formulation can also be easily extended to include other utilities, like road tolls for road pricing analysis, by simply adding the road tolls to the end of the equation as an extra term τ_r where τ_r is the toll per car on route r.

β^P denotes the coefficient for parking fees, β^{DR} is the constant coefficient for driving private vehicle and p^D_j is the parking fee for driving destination $j \in \mathcal{J}$. j_{od} is a parameter indicating the driving destination for OD pair od. z^D_r is a binary variable indicating whether route r is taken. u^D_{od} is a variable representing the utility for driving private vehicles for OD pair $od \in \mathcal{OD}$. M is a large number. Constraints (2.9) and (2.10) ensure that the travel utility for od is the largest among all possible routes for OD pair od. Constraints (2.11) calculate the driving attractiveness (a^D_{od}) for OD pair od.

$$u^{DR}_r = \beta^V t^D_r + \beta^D D^D_r \quad \forall r \in \mathcal{R}^D \tag{2.8}$$

$$u^D_{od} \geq u^{DR}_r + \beta^P p^D_{j_{od}} + \beta^{DR} \quad \forall r \in \mathcal{R}^D_{od}, od \in \mathcal{OD} \tag{2.9}$$

$$u^D_{od} \leq u^{DR}_r + \beta^P p^D_{j_{od}} + \beta^{DR} + M(1 - z^D_r) \quad \forall r \in \mathcal{R}^D_{od}, od \in \mathcal{OD} \tag{2.10}$$

$$a^D_{od} = e^{u^D_{od}} \quad \forall od \in \mathcal{OD} \tag{2.11}$$

2.3.1.4 Congestion

We incorporate congestion into the endogenous passenger choice by modeling car travel time t^D_r as a function of the traffic flow as defined in Equation (2.12). The road network consists of a set of segments, denoted by $\mathcal{S}$, and each car route includes a subset of these segments. We adopt the widely used congestion formulation given by the Bureau of Public

Roads (Bureau of Public Roads, 1964). F_s^D is the free flow travel time on segment s. α and β are constant parameters of the congestion model. M_{rs}^D is a binary parameter that equals 1 if car route r travels on segment s, and 0 otherwise. $\mathcal{S}_r^D$ is the set of segments traveled by route r. q_r^{DR} is the fraction of passengers taking car route r. $q_r^{DR} D_{od}$ (where $r \in \mathcal{R}_{od}^D$) is the number of passengers taking route r. The term $\sum_{od \in \mathcal{OD}} \sum_{r \in \mathcal{R}_{od}^D} M_{rs}^D q_r^{DR} D_{od}$ equals the number of vehicles on segment s. C_s^R denotes road capacity for segment s.

q_r^{DR}	the market share of driving private vehicles for route $r \in \mathcal{R}^D$

$$t_r^D = \sum_{s \in \mathcal{S}_r^D} F_s^D \left(1 + \alpha \left(\frac{\sum_{od \in \mathcal{OD}} \sum_{r \in \mathcal{R}_{od}^D} M_{rs}^D q_r^{DR} D_{od}}{C_s^R}\right)^{\beta}\right) \quad \forall r \in \mathcal{R}^D \tag{2.12}$$

The following constraints are used to model user equilibrium traffic assignment problem. The user equilibrium assumes that all drivers for the same OD pair will choose the path that takes the shortest travel time and all the drivers should experience the same travel time (Sheffi, 1985; Patriksson, 2015). The widely used formulation presented and proved by Sheffi (1985) is to minimize the sum of all arcs of the integral between 0 and the segment flow of the travel time function. However, we want to incorporate user equilibrium traffic assignment problem into our model as a set of constraints. Therefore, we come up with a novel method to model traffic assignment problem as shown in constraints (2.13)-(2.17).

$$q_r^{DR} \leq z_r^D \quad \forall r \in \mathcal{R}^D \tag{2.13}$$

$$Mq_r^{DR} \geq z_r^D \quad \forall r \in \mathcal{R}^D \tag{2.14}$$

$$u_r^{DR} \geq u_{r'}^{DR} + M(z_r^D - 1) \quad \forall r, r' \in \mathcal{R}_{od}^D, od \in \mathcal{OD} \tag{2.15}$$

$$\sum_{r \in \mathcal{R}_{od}^D} z_r^D \geq 1 \quad \forall od \in \mathcal{OD} \tag{2.16}$$

$$q_{od}^D = \sum_{r \in \mathcal{R}_{od}^D} q_r^{DR} \quad \forall od \in \mathcal{OD} \tag{2.17}$$

z_r^D is a binary variable indicating whether any passenger takes route r. Therefore, constraints (2.13) and (2.14) ensure that when market share of route r is non-zero ($q_r^{DR} > 0$), $z_r^D = 1$ and when $q_r^{DR} = 0$, $z_r^D = 0$. Constraints (2.15) guarantee that when $u_r^{DR} < u_{r'}^{DR}$, then $z_r^D = 0$. Therefore, no commuter will take a path that has a lower traveling utility. Constraints (2.16) ensures that at least 1 route will be taken by the commuters. Constraints (2.17) makes sure that the market shares individual routes add up to the total market share for the OD pair.

The generalized cost of the outside option is estimated as a function of the generalized cost of making the trip by transit or car. We assume this cost to be OD-dependent so that, similar to the generalized costs of making the trip by transit or car, the generalized cost of the outside option is also a linear function of the OD distance.

2.3.1.5 Passenger Traveling Cost

Passenger traveling cost constraints for driving private vehicle and taking public transit are defined in Constraints (2.18) and (2.19) respectively. Passenger traveling cost for driving is composed of parking fee, in-vehicle traveling time cost as well as driving distance cost. The cost of taking public transit is composed of transit fare, in-vehicle travel time cost, waiting time cost as well as walking time cost. C^{DV}, C^{DD}, C^{TV}, C^{TW}, and C^{TK} are

cost coefficients for driving travel time, driving travel distance, transit travel time, transit waiting time and transit walking time.

$$c_{od}^{D} = (p_{j}^{D} + C^{DV} t_{r}^{D} + C^{DD} D_{r}^{D}) q_{od}^{D} D_{od} \quad \forall od \in \mathcal{OD} \tag{2.18}$$

$$c_{od}^{T} = (p^{T} + C^{TV} T_{od}^{T} + C^{TW} w_{od}^{T} + C^{TK} K_{od}^{T}) q_{od}^{T} D_{od} \quad \forall od \in \mathcal{OD} \tag{2.19}$$

2.3.1.6 Passenger Choice

We adopt Sales Based Linear Programming (SBLP) model from Gallego et al. (2015) to formulate the market share for each travel mode in Constraints (2.20) - (2.22). The market share of each travel mode is proportional to its corresponding attractiveness unless the capacity constraints are violated. Constraints (2.22) ensure that the total market share is 1. We adopted SBLP instead of the widely used multinomial logistic (MNL) method because SBLP can be easily incorporated with the capacity constrained as presented in Section 2.3.1.7, while MNL is not compatible with the capacity constraints.

$$q_{od}^{D} a_{od}^{O} \leq q_{od}^{O} a_{od}^{D} \quad \forall od \in \mathcal{OD} \tag{2.20}$$

$$q_{od}^{T} a_{od}^{O} \leq q_{od}^{O} a_{od}^{T} \quad \forall od \in \mathcal{OD} \tag{2.21}$$

$$q_{od}^{T} + q_{od}^{D} + q_{od}^{O} = 1 \quad \forall od \in \mathcal{OD} \tag{2.22}$$

2.3.1.7 Capacity Constraints

The capacity constraints are demonstrated in Constraints (2.23) and (2.24). $\mathcal{OD}_j$ indicates the set of OD pairs going to destination j. $\mathcal{S}^T$ is the set of segments traveled by public transit. γ_{odls}^{T} is a binary parameter, which equals 1 when the transit passengers in OD pair od travel using transit line l passing through segment s. S_B is the transit vehicle capacity. Constraints (2.23) ensure that the number of demand traveling to destination j using car

does not exceed the parking capacity of j. Constraints (2.24) ensure that for each transit segment, the number of passengers onboard does not exceed the transit vehicle's passenger carrying capacity.

c_j^P	parking capacity at destination $j \in \mathcal{J}$

$$\sum_{od \in \mathcal{OD}_j} D_{od} q_{od}^D \leq c_j^P \quad \forall j \in \mathcal{J} \tag{2.23}$$

$$\sum_{od \in \mathcal{OD}} D_{od} \gamma_{odls}^T q_{od}^T \leq S_B f_l \quad \forall s \in \mathcal{S}^T, l \in \mathcal{L} \tag{2.24}$$

2.3.1.8 Fix Parking Decision

Since we aim to optimize the transit operator's capacity and pricing decisions, the parking capacities and parking fees are given as constants in the optimization model.

$$p_j^D = P_j^D \quad \forall j \in \mathcal{J} \tag{2.25}$$

$$c_j^D = C_j^D \quad \forall j \in \mathcal{J} \tag{2.26}$$

2.3.1.9 Scenario 1

In Scenario 1, the transit operator optimizes both transit frequencies and transit fare with objective function (2.1) and constraints (2.2) - (2.26).

2.3.2 Models for Upfront Pricing

In this section, we provide models where the transit operator makes frequency and pricing decisions in the first stage, and the parking operator responds to the transit operator's decision in the second stage. Once the transit operator makes its decision, it will not adjust

either transit frequencies or transit fare later. This scenario is a true reflection of reality. Since the parking operators usually have more flexibility to adjust their price and capacity than the transit operator. It is usually difficult for transit to adjust either its price or frequencies once the decision has been approved. For example, a 2013 state law bars the MBTA from increasing the fares more often than every two years and limits the increase to 5 percent per year—or 10 percent every two years (DeCosta-Klipa, 2016). However, the parking operator has greater flexibility when deciding its price (ITS, 2017; Milligan, 2020; Zheng and Geroliminis, 2016).

In this section, we talk about how we model the parking operator's response and how we incorporate it when optimizing the first-stage decisions made by the transit operator.

2.3.2.1 Parking Operator's response

The objective of the parking operator is to maximize its profit. The objective function is shown in Objective Function (2.27). j_{od} is a parameter indicating the destination for OD pair od. q^D_{od} is the percentage of demand driving private vehicles for OD pair od and p^D_j is the parking fee for destination j. C^{PO} is the maintenance and operating costs for each parking space per unit time. c^P_j is the parking capacity for destination j. Therefore, the first term calculates the revenue for the parking operator, while the second term is the cost.

$$Max \sum_{od \in \mathcal{OD}} q^D_{od} \cdot p^D_{j_{od}} \cdot D_{od} - \sum_{j \in \mathcal{J}} C^{PO} * c^p_j \tag{2.27}$$

$$f_l = F_l, \quad \forall l \in \mathcal{L} \tag{2.28}$$

$$p^T = P^T \tag{2.29}$$

The parking operator is subject to constraints (2.4)-(2.24) and constraints (2.28) and (2.29). P^t and F_l are parameters denoting transit fare and transit frequencies. Therefore, given a set of transit frequencies F_l and transit fare P^t, we obtain an optimal set of parking

decisions. Depending on whether parking can change its capacities, we can obtain the function as in Constraints (2.30) and (2.31).

$$\overrightarrow{p}^D = f_p(\overrightarrow{f}, p^T) \tag{2.30}$$

$$\overrightarrow{c^P}, \overrightarrow{p}^D = f_{pc}(\overrightarrow{f}, p^T) \tag{2.31}$$

2.3.2.2 Scenario 2

In Scenario 2, we consider pricing adjustment from the parking operator. The objective function is Function (2.1). The Constraints are Constraints (2.2) - (2.24), (2.26), and (2.30).

2.3.2.3 Scenario 3

In Scenario 3, the parking operator responds by adjusting both its capacities and parking fees. Constraint (2.31) is added to the optimization model and Constraints (2.26) are removed. Therefore, the Constraints are (2.2) - (2.24), and (2.31).

2.3.3 Models for Flexible Second Stage Pricing

In this section, we assume that the transit operator can also adjust its price flexibly. The transit operator decides the transit frequencies in the first stage. In the second stage, there will be an equilibrium where both the transit operator and the parking operator do not have incentives to deviate from the current decisions.

We present how we model the second stage equilibrium and how we incorporate it into the first stage optimization in this section.

2.3.3.1 Scenario 4

In Scenario 4, we consider a pricing equilibrium in the second stage between the parking and transit operator. We introduce function (2.32) which is a function computing the op-

timal transit fare given the transit frequencies, parking fees, and parking capacities. It is the result of the optimization model with Objective Function (2.1) with Constraints (2.2) - (2.26), and (2.28).

$$p^T = f_t(\overrightarrow{f}, \overrightarrow{p}^D, \overrightarrow{c}^P) \tag{2.32}$$

Given the transit frequencies, Function (2.30) and (2.32) should be satisfied at the same time. Function (2.33) is introduced to depict the price equilibrium for given transit frequencies.

$$p^T, \overrightarrow{p}^D = f_2(\overrightarrow{f}, \overrightarrow{c}^P) \tag{2.33}$$

Therefore, the final set of constraints for scenario 5 is Constraints (2.2) - (2.24), (2.26), and (2.33).

2.3.3.2 Scenario 5

Similar to Scenario 4, there is a second stage equilibrium between the parking and transit operator. However, the parking operator can also adjust its capacities in scenario 5. Functions (2.31) and (2.32) should be satisfied at the same time.

Correspondingly, Function (2.33) is replaced by Function (2.34).

$$p^T, \overrightarrow{p}^D, \overrightarrow{c}^D = f_{2c}(\overrightarrow{f}) \tag{2.34}$$

Therefore, the final set of constraints for scenario 6 is Constraints (2.2) - (2.24), and (2.34).

2.4 Solution Approach

As presented in Section 2.3, this is a complicated model framework with nonlinear and non-convex constraints in the base models (S1), response function from another operator (S2 - S5), and a second stage equilibrium constraint (S4 and S5). To obtain near-optimal solutions with a reasonable runtime for the base models, we adopt the coordinate descent method (Section 2.4.1) and the acceleration method (Section 2.4.2). The coordinate descent method iteratively computes the driving attractiveness under congestion and the driving demand. The acceleration method shortens the runtime by reducing the range of decision variables in each iteration.

Scenario 2, 3, 4, and 5 involve either a second-stage response from another operator or a second-stage equilibrium. To incorporate it into our model, we generate numerous combinations of transit frequencies and fares to evaluate parking response. Then the regression method is applied to approximate the transit revenue (Section 2.4.3). The first-stage decision is optimized based on the approximated function. Regarding obtaining the equilibrium of the second stage. we iteratively optimize transit fare and parking operator's decision until it converges. We will elaborate on this in Section 2.4.4.

2.4.1 Coordinate Descent Method

The major component and contribution of the base models are incorporating road congestion. It provides important insights into how an improved transit network design interacts with traffic congestion. However, it also adds significant complexity to the model. Therefore, we apply the coordinate descent method which separates the traffic congestion user equilibrium component from the original optimization model.

Figure 2.1 demonstrates the iterative process of the coordinate descent method. The top component is the traffic congestion user equilibrium model while the bottom component is the rest of the optimization.

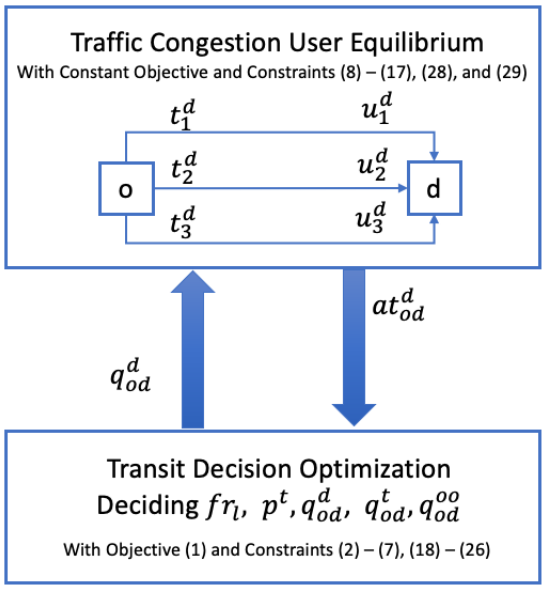


Figure 2.1: Coordinate Descent Method

The top component is a traffic congestion user equilibrium model which has a constant objective function with Constraints (2.12) - (2.17), (2.28), and (2.29) given the driving demand for each OD pair. As a result of this optimization, we can get the travel time for all driving routes (t_r^D) as well as which routes are taken for traveling (z_r^D). Taking these as inputs, we can obtain the attractiveness for driving private vehicles for each OD pair using Function (2.8) - (2.11).

In the bottom component, we take the attractiveness calculated in the traffic congestion user equilibrium component as input to compute the optimal transit decisions and corresponding demands. The constraints in the transit decision optimization model are Constraints (2.2) - (2.7), and (2.18) - (2.26).

The detailed procedure is as follows:

A1. Initialize driving demands as 0 for all OD pairs. Transit frequencies, transit fare, parking capacities, and parking fees are initialized as in the current scenario.

A2. Run the traffic congestion user equilibrium model with Constraints (2.8) - (2.17) and obtain driving attractiveness for all OD pairs.

A3. Use the attractiveness computed from Step A2 as input to the transit decision optimization model with Constraints (2.2) - (2.7), and (2.18) - (2.26).

A4. Repeat Step A2 and Step A3 until convergence.

2.4.2 Acceleration Method

Even after separating the congestion user equilibrium component from the original optimization model, the transit decision optimization is still a nonlinear non-convex model with bilinear and exponential constraints. Even though the Gurobi optimizer is able to solve optimization models with bilinear and exponential models, the runtime increases dramatically with the size of the problem. When running our case study in Boston Proper, we are not able to obtain a solution with a 5% optimality gap within 24 hours. Therefore, we come up with an acceleration method to shorten the runtime.

We observe that the solution doesn't change much after running the optimization model for 20 minutes in most cases. Moreover, the optimality gap reduces much faster when the range for the decision variable is smaller. Therefore, we set a time limit for running the optimization model and use the current solution to create decision variables with tighter lower bounds and upper bounds. We iterate this process until we obtain a solution with an optimality gap of 0.1%.

The detailed procedure is presented below. We use transit fare optimization (Scenario 1) as an example. Similar techniques are applied to all the other Scenarios.

B1. Initialize the lower bound of transit fare as p^T_{min} and the upper bound as p^T_{max}.

B2. Run transit decision optimization model with an optimality gap of 40% and obtain a transit fare solution p^{T*}.

B3. We set the lower bound of transit fare as $max(p^T_{min}, p^{T*} - \frac{p^T_{max} - p^T_{min}}{2})$ and the upper bound as $min(p^T_{max}, p^{T*} + \frac{p^T_{max} - p^T_{min}}{2})$

B4. Update $p^T_{min} = max(p^T_{min}, p^{T*} - \frac{p^T_{max} - p^T_{min}}{2})$ and $p^T_{max} = min(p^T_{max}, p^{T*} + \frac{p^T_{max} - p^T_{min}}{2})$.

B5. Reduce the optimality gap by half and iterative Step B2, B3 and B4 until the optimality gap is smaller than 1%.

We find that the method significantly reduces the runtime. We will also demonstrate in Section 2.5.2.1 that we don't sacrifice much on optimality by adopting the acceleration method.

2.4.3 Regression Method

Scenario 2, 3, 4, 5 are two-stage games. In Scenario 2 and Scenario 3, the transit operator decides transit frequencies and transit fare in the first stage, and the parking operator responds in the second stage. For Scenario 4 and 5, the transit operator decides transit frequencies in the first stage while there is an equilibrium between transit fare and the parking operator's decision in the second stage. The challenge here is how to incorporate the second-stage decisions into the first-stage optimization. Here, we use Scenario 2 as an example.

C1. Generate numerous combinations of transit frequencies and transit fares. The parking operator's decisions are optimized for each set of transit frequencies and transit fare. The optimization model for the parking operator has the Objective Function (2.27) and Constraints (2.4) - (2.24), (2.26), (2.28), and (2.29).

C2. Two linear regressions are run with the dependent variables of systemwide passenger travel cost and Transit revenues. The independent variables are the quadratic functions of transit frequencies and transit fare.

C3. The optimal transit frequencies and fare are obtained by minimizing the systemwide passenger traveling cost (Objective Function (2.37)) while satisfying the farebox recovery constraints (Constraints (2.38) - (2.41)).

The regression method approximates the systemwide passenger travel cost and transit revenue corresponding to Stage II as functions of Stage I decisions. We find that transformed polynomial approximations provide a strong goodness-of-fit to the exact net revenue values calculated at the Stage II PSNEs, and capture the gross properties of the relationship between the Stage I decision variables – capacities and frequencies – and the net revenues at the Stage II PSNE. These transformed polynomials are easy to fit during Step C2 and easy to optimize during Step C3 of the solution method. The independent variables (or predictors) in these regressions are the Stage I decision variables of both operators and the dependent variable (or response) is the net revenue of each player at the Stage II PSNE. Step C2 thus results in a closed-form expression of each operator's Stage I profit.

Our detailed computational experiments showed that a quadratic function of transformations of transit line frequencies and parking capacities provided an appropriate tradeoff between the objectives of achieving a good fit to the exact net revenue values and model tractability. The regression parameters were estimated using an ordinary least squares (OLS) estimation process. Specifically, we set the independent variables in the regressions to be the transit headway and transit fare as well as the interaction terms. The polynomial regression model is given by Equation (2.36). k is the systemwide passenger traveling cost. r^T is the transit revenue. θ_0^K, θ_l^K, ν^K, $\rho_{l_1 l_2}^K$, ζ^K and η_l^K are the regression parameters for systemwide passenger travel cost OLS estimation while θ_0^R, θ_l^R, ν^R, $\rho_{l_1 l_2}^R$, ζ^R and η_l^R are the regression parameters for transit revenue OLS estimation.

$$k = \theta_0^K + \sum_{l \in \mathcal{TR}} \theta_l^K \cdot \frac{1}{f_l} + \nu^K \cdot p^T + \sum_{l_1 \in \mathcal{TR}} \sum_{l_2 \in \mathcal{TR}} \rho_{l_1 l_2}^K \cdot \frac{1}{f_{l_1}} * \frac{1}{f_{l_2}} + \zeta^K (p^T)^2 + \sum_{l \in \mathcal{TR}} \eta_l^K \cdot \frac{1}{f_l} * p^T \tag{2.35}$$

$$r^T = \theta_0^R + \sum_{l \in \mathcal{TR}} \theta_l^R \cdot \frac{1}{f_l} + \nu^R \cdot p^T + \sum_{l_1 \in \mathcal{TR}} \sum_{l_2 \in \mathcal{TR}} \rho_{l_1 l_2}^R \cdot \frac{1}{f_{l_1}} * \frac{1}{f_{l_2}} + \zeta^R (p^T)^2 + \sum_{l \in \mathcal{TR}} \eta_l^R \cdot \frac{1}{f_l} * p^T \tag{2.36}$$

After estimating passenger travel cost and transit revenue, we strive to obtain the optimal first-stage decision for the transit operator. The optimization model for Step C3 is presented below:

$$Min \quad k \tag{2.37}$$

$$k = \theta_0^p + \sum_{l \in \mathcal{TR}} \theta_l^p \cdot f_l^{inv} + \nu^p \cdot p^T + \sum_{l_1 \in \mathcal{TR}} \sum_{l_2 \in \mathcal{TR}} \rho_{l_1 l_2}^p \cdot f_{l_1}^{inv} * f_{l_2}^{inv} + \zeta^p (p^T)^2 + \sum_{l \in \mathcal{TR}} \eta_l^p \cdot f_l^{inv} * p^T \tag{2.38}$$

$$r^T = \alpha_0^T + \sum_{l \in \mathcal{TR}} \alpha_l^t \cdot f_l^{inv} + \beta^t \cdot p^T + \sum_{l_1 \in \mathcal{TR}} \sum_{l_2 \in \mathcal{TR}} \gamma_{l_1 l_2}^t \cdot f_{l_1}^{inv} * f_{l_2}^{inv} + \zeta^t (p^T)^2 + \sum_{l \in \mathcal{TR}} \eta_l^t \cdot f_l^{inv} * p^T \tag{2.39}$$

$$f_l * f_l^{inv} = 1 \quad \forall l \in \mathcal{L} \tag{2.40}$$

$$r^T \geq R * \sum_{l \in \mathcal{L}} C_l * \frac{2L_l \cdot f_l}{V^T} \tag{2.41}$$

2.4.4 Sequential Optimization Heuristic Method

Scenario 4 and Scenario 5 of this approximate solution method involve computationally finding a PSNE for a two-player game at Stage II. This is performed via a sequential optimization heuristic (also known as the best response heuristic), which is a standard method used in prior transportation studies that use computational approaches to solve game-theoretic models (see, e.g., Martín and Román, 2003; Adler, 2001, 2005; Harder and Vaze, 2017). The idea is to implement an iterative response chain, where each player iteratively optimizes its own objective by reacting to the other player's decisions. This response chain will stop when neither player is able to further improve its objective by changing its own decisions. This is when a PSNE is reached. Schematically, this process is as follows:

D1. Initialize with a strategy profile a, where $a = \left(p^T, \overrightarrow{p^D}\right)$. For the transit operator, fix the parking operator's decision and find the optimal response

D2. For the parking operator, fix the transit operator's decision and find the optimal response. Repeat until convergence to PSNE.

Thus, Step C1 requires repeatedly solving combinatorial, non-linear and non-convex optimization formulations described in Section 2.3. These are very challenging optimization problems (Hemmecke et al., 2010; Burer and Letchford, 2012), which only a few software products are capable of handling effectively. Gurobi optimizer (Gurobi Optimization, LLC, 2022a) is one of the most suitable software products to deal with non-linear and combinatorial formulations (Gurobi Optimization, LLC, 2022c). It uses a global optimization approach based on polyhedral branch-and-cut algorithms (Gurobi Optimization, LLC, 2022b), and is the optimization solver software used in this paper.

2.4.5 The Overall Solution Method

We describe the overall solution method is Figure 2.2. This Figure corresponds to the solution method for Scenario 4 and Scenario 5. Scenario 1 - Scenario 3 is partially presented in the Figure. Step C1, C2, and C3 correspond to the regression method discussed in Section 2.4.3. In Step C1, Stage II pure strategy Nash equilibriums are found for a subset of combinations of Stage I decisions. Within Step C1, we run Step D1 and Step D2 iteratively until it reaches an equilibrium (discussed in Section 2.4.4). Due to the nonlinearity and nonconvexity of the optimization model within Step D1 and Step D2, we applied Coordinate Descent Method (discussed in Section 2.4.1) and acceleration method (discussed in Section 2.4.2) to reduce the runtime.

Scenario 1 corresponds to Step D1 in Figure 2 where only the coordinate descent method and acceleration method are being applied. Scenario 2 and Scenario 3 do not have Stage II equilibrium. Therefore, In Step C1, we only run Step D2 for a subset of

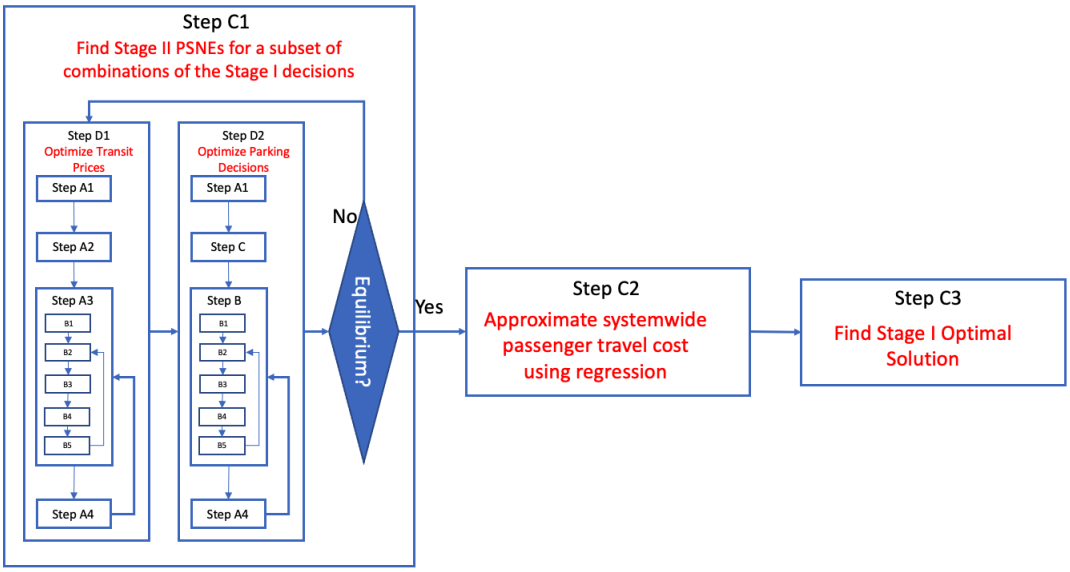


Figure 2.2: Solution method flowchart

combinations of Stage I decisions. Then the regression approximation will be applied.

2.5 Computational Setup and Results

We applied our model and solution approach to Boston proper which includes Back Bay, Bay Village, Beacon Hill, Chinatown, Downtown Crossing, Fenway-Kenmore, Financial District, Government Center, Leather District, North End, South End, West End, and the Waterfront (NeighborhoodX Boston, NA). Since the original model cannot be evaluated in the Boston Proper instance due to its large size. In order to evaluate the efficacy of our method, we apply our algorithm to two smaller instances in South Boston (Waterfront) and Downtown Boston.

2.5.1 Data

Since both private vehicle driving and fixed-route transit services are considered, we obtain both road networks for the Boston Proper region and current subway and bus networks. We are interested in optimizing the transit schedules during the morning rush hours (6:00 a.m.

*Picture from Boston Planning Development Agency

Figure 2.3: Map of Boston city

- 9:00 a.m.).

The demand data is obtained from the 2010 Census LODES database where Workplace Census Block Code, Residence Census Block Code, the total number of jobs, number of jobs under different age brackets, number of jobs under different income levels, and number of jobs in different industries were provided (United States Census Bureau, 2020).

Information on the subway, buses, and commuter rail was obtained from Massachusetts Bay Transportation Authority (MBTA, 2021a) in the format of General Transit Feed Specification (GTFS). It provides the locations of stops, transit lines, lengths, and schedules. The route of the public transit is obtained in the format of shapefile from MassDOT website (MassDOT, 2021).

2.5.2 Computational Results

2.5.2.1 Base Models Method Evaluation

In this section, we evaluate the accuracy and the runtime of our coordinate descent method and acceleration method. Both of these two methods aim to reduce the runtime for the baseline model. However, it is important to evaluate the accuracy by comparing the solution of this method to the original method.

Table 2.2 Optimality Gap and Runtime Comparison of the Orginal Mothod and Our Method

	Optimality Gap	Runtime of the original method	Runtime of our method
Mean	0.03%	2634s	11s
Min	0%	38s	5s
Max	0.1%	5001s	23s

Table 2.2 presented the optimality gap and the runtime of our method and the original method of 40 different combinations of parking fees in the South Boston instance. We find that the optimality gap between our method and the original method is 0.03%, but the average runtime is 250 times smaller. It shows that by adopting the coordinate descent method and the acceleration method, we can maintain high accuracy while reducing the runtime dramatically.

2.5.2.2 Regression Method Evaluation

In this section, we want to evaluate how accurate is the linear regression approximation in Step C2. Table 2.3 presents the R square values for predicting systemwide passenger travel cost and transit revenue under Scenario 2, 3, 4, and 5. The R squares in all cases are above 90% demonstrating the high accuracy of our approximation.

Table 2.3 R square values for Scenario 2, 3, 4, and 5 for predicting systemwide passenger travel cost and transit revenue

R^2	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Systemwide Passenger travel Cost	95%	96%	95%	96%
Transit Revenue	99%	99%	95%	92%

2.6 Model Results

In this section, we present the results from our computational experiments. There are two major questions we want to investigate in this section: (1) whether it is necessary to account for the response from the parking operator when optimizing transit frequencies and transit fare, (2) what's the impact of parking capacity on transit decisions and passenger travel cost, and (3) is upfront pricing strategy or flexible second-stage better in achieving lower passenger travel cost? Section 2.6.1 aims to address the first question by comparing Scenario 1 and Scenario 2. Section 2.6.2 answers the second question by investigating Scenario 2 and 3. Section 2.6.3 focuses on the third question by comparing the results from Scenario 2 and 4 as well as Scenario 3 and 5.

Table 2.4 The Transit Frequencies, Parking Capacities, Transit Fare, Parking Fees and Other Important Metrics for the current scenario and Scenario 1-5

Scenarios		Current Scenario	Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Transit Frequencies	Frequencies 1	3	6	7	7	7	7
	Frequencies 2	15	12	12	12	12	12
	Frequencies 3	8	4	1	1	1	1
	Frequencies 4	3	6	7	7	7	7
	Frequencies 5	3	6	7	7	7	7
	Frequencies 6	7	12	10	10	10	10
	Frequencies 7	9	10	10	10	8	8
	Frequencies 8	6	12	10	10	10	10
	Frequencies 9	10	12	12	12	11	10
	Frequencies 10	10	10	12	12	9	8
	Frequencies 11	8	11	10	10	8	8
Parking Capacities	Capacity 1	7000	7000	7000	5849	7000	6450
	Capacity 2	2520	2520	2520	2463	2520	2386
	Capacity 3	1320	1320	1320	1232	1320	1239
	Capacity 4	60	60	60	39	60	46
	Capacity 5	20	20	20	14	20	12
	Capacity 6	10	10	10	4	10	5
Transit Pricing	Transit fare	1.7	2.1	2.0	1.9	1.7	1.6
Parking Pricing	Parking fee 1 ($)	23	23	22	25	22	24
	Parking fee 2 ($)	26	26	26	25	26	26
	Parking fee 3 ($)	38	38	37	39	38	39
	Parking fee 4 ($)	20	20	19	24	20	21
	Parking fee 5 ($)	20	20	24	22	21	24
	Parking fee 6 ($)	17	17	16	22	16	18
Mode shares	Transit market share	42%	44%	43%	48%	43%	45%
	Driving share	47%	46%	48%	42%	47%	44%
	Outside option share	11%	10%	10%	10%	10%	10%
Other Important Metrics	Passenger travel cost ($ per hour)	212805	205634	203922	203729	205199	207490
	Parking profit ($ per hour)	74148	71694	72062	70748	72160	72259
	Transit profit ($ per hour)	-20175	-26028	-24607	-24182	-21530	-20940
	Transit Revenue ($ per hour)	5075	6522	5993	6418	5320	5160
	Farebox ratio	20%	20%	20%	21%	20%	20%

2.6.1 Impact of the Parking Operator's Response

In this Section, we will analyze whether it is necessary to take the parking operator's response into account and how much impact the parking operator's decision has on the transit operator's decision. Table 2.4 presents the transit frequencies, parking capacities, transit fare, parking fees, travel mode shares, and other important metrics. In the current scenario, the current transit frequencies, transit fare, parking capacities, and parking fees are applied. Under the current scenario, 35% of commuters travel using public transit, 56% of commuters take private vehicles, while the rest 9% travel using outside options like biking. Scenario 1 optimize transit frequencies and transit fare while the parking operator doesn't change the parking capacities and parking fees. To guarantee a fair comparison, we ensure that the farebox recovery rate is the same as the value in the current scenario. Compared to the current scenario, the transit fare increases from 1.7 to 2.1 while there are changes in the transit frequencies as well. The transit market share increases from 35% to 37% and the passenger travel cost decreases by 3.4%.

Scenario 2 assumes that the parking operator will respond to the transit operator's change of decisions. When optimizing the transit frequencies and transit fare, the transit operator also takes the parking operator's response into account. As presented in Table 2.4, the optimal transit frequencies and transit fare in Scenario 2 are different from the ones in Scenario 1. The optimal transit fare is \$2.0 which is slightly smaller than the optimal transit fare in Scenario 1. Only 3 out of the 11 transit lines have the same optimal transit frequencies as in Scenario 1. The difference is due to the fact the parking operator will adjust its parking fees in all regions except for region 2. The difference in transit frequencies and fares as well as parking fees lead to a decrease in passenger travel costs. Compared to Scenario 1, the passenger travel cost only reduces by 3.4% instead of 4.2%.

To provide a fair evaluation of the impact of incorporating the parking operator's response, we further conducted scenario 2, 3, 4, and 5 using the solution from scenario 1. These scenarios provide the results when the transit operator is not considering the park-

ing operator's response, but the parking operator adjusts it decisions in practice. In Table 2.5, scenarios 2', 3', 4' , and 5' are the results of scenario 1 solution evaluated under the assumptions of scenario 2, 3, 4, and 5. By comparing scenario 2' with 2, 3' with 3, 4' with 4, and 5 with 5', we are able to quantify the impact of considering parking responses. In all four comparisons, we find that we are able to reduce the passenger travel cost further when the transit decisions are optimized considering parking responses.

Table 2.5 The Transit Frequencies, Parking Capacities, Transit Fare, Parking Fees and Other Important Metrics for the current scenario and Scenario 1-5

	Scenarios	2'	2	3'	3	4'	4	5'	5
Transit Frequencies	Frequencies 1	6	7	6	7	6	7	6	7
	Frequencies 2	12	12	12	12	12	12	12	12
	Frequencies 3	4	1	4	1	4	1	4	1
	Frequencies 4	6	7	6	7	6	7	6	7
	Frequencies 5	6	7	6	7	6	7	6	7
	Frequencies 6	12	10	12	10	12	10	12	10
	Frequencies 7	10	10	10	10	10	10	10	10
	Frequencies 8	12	10	12	10	12	10	12	10
	Frequencies 9	12	12	12	12	12	12	12	12
	Frequencies 10	10	12	10	12	10	12	10	12
	Frequencies 11	11	10	11	10	11	10	11	10
Parking Capacities	Capacity 1	7000	7000	5767	5849	7000	7000	6953	6450
	Capacity 2	2520	2520	2755	2463	2520	2520	2334	2386
	Capacity 3	1320	1320	1134	1232	1320	1320	1254	1239
	Capacity 4	60	60	28	39	60	60	54	46
	Capacity 5	20	20	13	14	20	20	54	46
	Capacity 6	10	10	6	4	10	10	9	12
Transit Pricing	Transit fare	2.1	2.0	2.1	1.9	2.2	1.7	1.7	1.6
Parking Pricing	Parking fee 1 ($)	22	22	25	25	22	22	23	24
	Parking fee 2 ($)	25	26	23	25	26	26	27	26
	Parking fee 3 ($)	39	37	42	39	41	38	39	39
	Parking fee 4 ($)	25	19	28	24	18	20	19	21
	Parking fee 5 ($)	27	24	24	22	28	21	29	24
	Parking fee 6 ($)	16	16	17	22	16	16	18	18
Mode shares	Transit market share	35%	43%	43%	47%	48%	43%	43%	45%
	Driving share	47%	48%	43%	42%	47%	47%	46%	44%
	Outside option share	9%	10%	10%	10%	10%	10%	10%	10%
Other Important Metrics	Passenger travel cost ($ per hour)	205176	203922	206129	203729	207311	205199	210202	207490
	Parking profit	72019	72062	70375	70748	72068	72160	73238	72259
	Transit Revenue	6331	5993	6962	6418	6661	5320	5188	5160

This is an indication that without considering the response from the parking operator, it is likely that the previous studies are overestimating the social impact of the optimized transit system.

2.6.2 Impact of the Parking Capacity Adjustment

In Scenario 3 and 5, we assume that the parking operator can also flexibly adjust its parking capacities. By taking a closer look at scenario 3, the optimal strategy for the parking operator is to reduce the parking capacities and increase the parking fees. Compared to scenario 2, the average parking fee increases from \$24 to \$26. The total parking capacity decreases from 10930 to 9601. This leads to a decrease in the market share of driving from 48% to 42% while the transit market share increases from 43% to 48%. The transit fare decreased slightly to \$2 to alleviate the negative impact of surging parking fees on passenger travel cost. The results further demonstrate the interaction between the parking operator and the transit operator's decision. It demonstrates the importance of taking considering the parking operator while optimizing the transit system.

Comparing Scenario 5 with Scenario 4, the parking operator similarly reduces its parking capacities and increases its parking fees. Even though the transit operator reduces its transit fare, the overall passenger travel cost increases even further to 207,490.

2.6.3 Analysis of Transit Pricing Strategies

In this section, we will compare the two transit pricing strategies: upfront pricing strategy and flexible second-stage pricing strategy. Comparing Scenario 2 and Scenario 4 as well as Scenario 3 and Scenario 5, we find that the passenger travel cost is lower in Scenario 2 and 3 compared to Scenario 4 and 5. It shows that the upfront pricing strategy provides a better result compared to the flexible second-stage pricing strategy. This is an interesting finding indicating that having the flexibility to adjust its price can lead to worse results.

With closer examination, we find that this finding is not a coincidence. The upfront pricing strategy is always superior to the flexible second-stage pricing strategy. The transit operator enjoys a first-mover advantage with the upfront pricing strategy. We provide proof below:

With the upfront pricing strategy, the transit operator can achieve passenger travel cost

no greater than the cost with the flexible second-stage pricing strategy.

Proof. We assume the optimal results under the flexible second-stage pricing strategy is $(f^*, p^{T*}, c^{P*}, p^{D*})$. Under the upfront pricing strategy, the transit operator can adopt the same transit frequencies and transit fare (f^*, p^{T*}). Since $(f^*, p^{T*}, c^{P*}, p^{D*})$ is a Nash equilibrium solution in the second stage, which means the parking operator has no motivation to deviate from this decision given transit operator's decision (f^*, p^{T*}). Therefore, under the upfront pricing strategy, the parking operator also responds with (c^{P*}, p^{D*}). Therefore, the upfront pricing strategy can always achieve the same passenger cost by adopting the optimal solution from the flexible second-stage pricing strategy. Therefore, with the upfront pricing strategy, the transit operator can achieve passenger travel cost no greater than the cost with flexible second-stage pricing strategy.

2.7 Conclusion

An effective and sustainable urban transportation system should provide residents with sufficient access to affordable urban mobility options while minimizing traffic congestion, greenhouse gas emissions, and reliance on fossil fuels. Nevertheless, the actual performance of most urban transportation systems is the result of complex interactions among various service providers, passengers, and government agencies, whose goals are not completely aligned. Many urban transportation providers are moving toward more dynamic pricing strategies creating further separation between the timelines of capacity and pricing decisions. Therefore, it is important to consider the interaction with other operators when planning the transit network. These multi-stage decisions of the interacting decision-makers motivate using multi-stage game-theoretic models. While these models are good representatives of real-world interactions, they are also notoriously hard to solve. To overcome this challenge, this paper developed a realistic and tractable computational framework for optimizing transit systems considering the response from the parking operator

and evaluating policies to improve the urban transportation system.

We explore the optimal transit frequencies and transit fares under different assumptions and strategies. A comprehensive model framework is developed to optimize transit decisions considering the response from the parking operator. In order to solve the mixed-integer non-linear programming model, we applied coordinate descent, acceleration method, regression method, and sequential optimization heuristic method to provide a near-optimal solution in a reasonable amount of time.

For the case studies used in this paper, we obtained several new insights. First, we found that the capacity and pricing decisions of the parking operator do exert an impact on the optimal transit frequency and pricing decisions. Furthermore, the impact on passenger travel cost reduction can be overestimated when ignoring the paring operator's response. Second, the parking capacity is found to have a significant impact on the overall passenger travel cost. When the parking operator is allowed to adjust parking capacities, they tend to reduce the parking capacities while increasing the parking fees to achieve a higher profit. This adjustment leads to a higher transit market share and sometimes a higher passenger travel cost. Therefore, it is important for the policymaker to implement supplementary regulations while reducing the parking capacities in urban areas. Third, we find that the upfront pricing strategy is better than the flexible second-stage pricing strategy for the transit operator to achieve lower passenger travel cost. The non-flexible pricing strategy can in fact provide the transit operator the first-mover advantage. We also prove theoretically that the upfront pricing strategy is always at least as good as the flexible second-stage pricing strategy.

From a methodological standpoint, a promising future research direction is incorporating the decisions of other competing urban transportation service operators beyond transit and parking. This may include careful modeling of ridehailing and ridesharing operators. From a policy standpoint, while we focused mainly on taxes and subsidies, our framework can be easily used to analyze a variety of other policies and regulatory tools including park-

ing capacity restrictions in certain parts of the city, transit price and frequency regulations on certain routes, modifications to transit fare structures (e.g., flat versus distance-based), congestion tolls in the city center, and many more. We hope that quantification of the costs and benefits of these alternative policy mechanisms using our computational framework provides a path toward enhancing urban transportation systems.

Acknowledgments

The research was financially supported by the Portuguese Science and Technology Foundation (grant SFRH/BD/52359/2013). This material is based upon work supported by the National Science Foundation under Grant No. 1750587.

Chapter 3

Optimizing Transit Network Planning and Local On-demand Services in Transit Desert Regions

3.1 Introduction

Urbanization has brought over 80% of the US population to cities (U.S. Census Bureau, 2010). But this urban and suburban growth has also amplified the prevalence and severity of gaps in public transit accessibility, in and around many major urban centers. The term "Transit desert" has been coined to describe an area of the city where the demand for public transportation exceeds supply (Jiao and Dillivan, 2013). In the most severely affected cities, over 1 in 8 residents were found to live in transit desert regions (Jiao and Bischak, 2018). Transit desert residents are often forced to choose between walking long distances to access the nearest transit stops or using alternative means of transportation like private vehicles, taxis, and ride-hailing services. These alternative means produce more carbon emissions and add more congestion to urban, and suburban streets. Besides, they might not be affordable for some residents leading to financial distress, job losses, and hindered

socioeconomic mobility. Moreover, the transit desert residents are also confronted with a lack of access to fresh food, clean air, and quality medical care (Syed et al., 2013). Transit deserts often coincide with the residential areas of lower-income, underserved, and minority populations exacerbating the income and racial inequalities (Chen, 2019; Williams, 2018).

Transit agencies often lack incentives or are ill-equipped to extend their networks to transit deserts. The public transportation sector has been suffering from falling ridership, especially in the US. Even before the COVID-19 pandemic began, ridership in the previous five years had fallen in 44 of the 50 urban areas (O'Toole, 2020). Many transit agencies are struggling with financial difficulties, with fares often covering under 25% of their costs, and several of them are planning service cuts in the near future. Recently, the Metropolitan Transportation Authority (MTA) in New York has been mulling a 15% cut to subway, bus, and train service, as it tries to adapt to post-pandemic realities (GUSE, 2021), despite receiving $15 billion in federal aid (Meyer, 2021). Therefore, finding ways to improve transit access without a surging budget is key to alleviating the transit desert problem.

The proliferation of ride-hailing services can enhance access for people lacking access to traditional fixed-route transit. Revenue in the Ride-Hailing & Taxi segment in the US is projected to grow at an expected annual growth rate of 4.7%, resulting in a projected market volume of $74.4 billion by 2026 with an average revenue per user of $660 (statista, 2021). Leading ride-hailing platforms | Uber and Lyft | have claimed that ride-hailing services enhance people's mobility and reduce private vehicle ownership (Gordon, 2021). However, some studies indicate that these transportation network companies (TNCs) had an insignificant effect on vehicle ownership, but instead led to a significant decline in public transit ridership, and worsened congestion (Diao et al., 2021; Bliss, 2019; Erhardt et al., 2019; Schaller, 2018; Tarduno, 2021). Given these concerns, governments have applied various regulations to reduce the negative impacts of ride-hailing services. New York City charges TNCs congestion tolls for entering Manhattan's central business district (Guse,

2021). Other collaborative policies are promoting the integration of traditional public transit and on-demand services where TNCs can help solve the first/last-mile problem for fixed-route public transit.

Observing collaboration opportunities between fixed-route transit and on-demand services, multiple transit agencies have announced TNC alliances to fill in the gaps in their services, especially in the transit desert areas, and/or to provide more flexibility to their passengers. St. Louis Metro Transit is partnering with Via to launch multi-modal shared-ride services (Staff, 2020), with multi-modal route options made directly available through their Via app. This new on-demand transit service for the St. Louis region called Via Metro STL charges a flat price of $2 for any ride. Lyft riders in LA could take shared rides to and from select Metro rail stations during weekday rush hours with a $3 flat fare (Metro, 2020). Similar collaborations are in place in many other cities including Miami (Wanek-Libman, 2020), Boston (**?**), Dallas (Fernandez, 2019), Denver (RTD, 2021), Santa Monica (BigBlueBus, 2018), etc. These collaborations could potentially improve accessibility for passengers in transit deserts without exacerbating traffic congestion in city centers. However, the governments and transit agencies currently have limited analytical tools and datasets to estimate the influence of these collaborations (Blodgett et al., 2017). Moreover, few programs have tried finetuning the current systems in light of these collaborations with ride-sharing services.

This paper aims to address these gaps, by providing an analytical framework combining prescriptive models, datasets, and algorithms to alleviate the transit desert problem in a holistic way. It attempts to simultaneously optimize the fixed-route public transit network as well as the fleet size of the on-demand services to serve local demand and/or to act as a feeder to the mainline transit network. This paper is particularly timely, in the context of how the pandemic has brutally exposed the unequal distribution of medical, food, and transportation resources. Our framework can also be seen as a tool for rigorous optimization of the use of transportation infrastructure investments, such as the recent $1 trillion

infrastructure bill in the US of which $39 billion is devoted to expanding the current public transit system. This paper makes the following contributions:

1. We provide a comprehensive analytical framework to address the transit desert problem by co-optimizing the transit network and frequencies as well as the fleet size of on-demand services. The transit desert has been a decades-long problem in the US. With the decreasing transit ridership and with the ongoing plans by transit providers to cut services, the situation is likely to get worse in the foreseeable future. The problem cannot be easily addressed by expanding the current transit network, or by relying solely on ride-hailing/ride-sharing services. Therefore, joint planning of the transit network and ride-sharing service deployment holds the potential to provide a better solution. In this study, we optimize the transit network design, operating frequencies, and on-demand vehicle fleet sizes while accounting for passengers' mode and route choice decisions. Different policies like on-demand services restrictions, road pricing, and transit subsidies can also be evaluated using this framework.

2. We formulate a new optimization model to determine transit network design and frequencies for each line, and on-demand fleet size under passengers' choice decisions. The transit lines are chosen to minimize the social costs. The commuters have the options to travel with public transit, on-demand services alone, multi-modal services (using both on-demand services and fixed-route transit services), and outside options (driving private vehicle or working from home). The discrete choice model is used to model passengers' choices among different travel modes based on expenditure, in-vehicle travel time, waiting time, and walking time. This problem is formulated as a mixed-integer non-linear program. Due to its complexity, we further developed a mixed-integer second-order conic programming model to solve the original problem iteratively. Few papers have investigated public transportation planning with the planning of ride-hailing services (Wei et al., 2021; Liu and Ouyang, 2021). However, these previous studies either didn't consider changing the transit network or were applied to a grid city network. This paper develops a framework to jointly

optimize the public transportation network, frequencies, and on-demand fleet size.

3. We design an original two-step heuristic to solve the MINLP which is non-linear and non-convex. In the first step, we select a subset of transit lines from the full set of potential transit lines by applying the rounding heuristics method. The delayed constraints generation method is applied to reduce the number of constraints. In the second step, we determine the frequencies for each transit line and the fleet size for on-demand services by iteratively running mixed-integer second-order conic programming (MISOCP) model. The non-linear constraints are linearized using first-order Taylor series Expansion. By running the MISOCP model and updating the Taylor Expansion points iteratively, we can obtain a near-optimal solution in less than 2 hours. This approach provides a heuristic method to solve MINLP problems by determining the integer variables in the first step and deciding the continuous variables iteratively in the second step.

4. We provide important policy implications by performing computational experiments in transit desert regions of the greater Boston area. The modeling and solution framework is applied to those parts of the greater Boston area that the MBTA identified as priority regions. The results indicate that we are able to reduce the social cost by more than 3.67%. We are able to get this result within 12 hours. We further analyze the impact of different pricing strategies and find that we can reduce the social cost further (5.67%) if the on-demand service pricing can be adjusted. Moreover, we compare distance-based pricing structures and flat pricing structures and find that the flat pricing structures generate more profit while maintaining a similar service level. The analysis of different commuter groups demonstrates that the improved transit system reduces the commuting disutility for low-income commuters and communities with no private vehicles by 14% and 21% respectively. The results confirm that the framework can provide a more fair and accessible urban public transportation system.

3.2 Related Literature

This paper fits in the literature on improving urban mobility and accessibility combining different types of transportation modes including traditional fixed-route public transit and the fast-growing on-demand services. This paper lies at the intersection of the studies in transit network design, on-demand services operations, and passenger choice modeling. In this section, we review the related research in improving public transportation while considering on-demand services.

3.2.1 Transit Network Design and Frequency Setting Problem

Transit network and frequency design problems are usually seen as a two-stage sequential decision-making process. The transit operators decide on where to build transit lines first and then decide on the frequency for each transit line. This two-stage decision-making process separates the transit network and frequency design problem into transit network design problem and frequency design problem.

Research on transit network design problems has evolved rapidly due to the improvement of search algorithms as well as the advancement of computing technologies. Researchers first generate a pool of potential transit lines to build the network based on demand volumes or expert insights. Then, a number of transit lines will be selected from the potential transit lines as part of the final transit network. The objective of the transit network design problem is usually to maximize ridership under budget constraints. Due to the combinatorial nature of the transit network design problems, directly solving the optimization model would be extremely computationally expensive. Therefore, researchers have applied various heuristic and metaheuristic approaches like Simulated Annealing (Yan et al., 2013), Genetic Algorithm (Liu and Yang, 2007), Tabu search (Fan and Machemehl, 2008), bee colony optimization (Szeto and Jiang, 2014b), and particle swarm optimization (Kechagiopoulos and Beligiannis, 2014) to obtain high-quality solutions. Most early

researchers assumed an exogenously determined transit demand, which means that passengers won't switch between modes of transportation if the service availability and service level in the area changes (Buba and Lee, 2018). In the second stage, the frequency design problem is to determine the optimal frequency for each transit line after the transit lines have been decided. The objective function is usually to minimize the weight cost of passenger waiting time as well as infrastructure and operating costs. However, this problem is highly complicated because we must address transit assignment problems. Transit assignment problems emerge when passengers can take multiple lines to travel from their origin to destination. Researchers have developed multiple models to simulate passengers' decisions and probabilities to take certain transit lines. The most commonly used model is the frequency-based model which assumes passengers arrive randomly at stops and they take whichever bus or train comes first among those that can take them to their destination (Oliker and Bekhor, 2016; Bowman and Turnquist, 1981). Another central theme is the transit congestion effect. If the service is crowded, the waiting, boarding, and dwell time tend to increase, and discomfort is caused to the passenger. If the capacity of the service is lower than the demand, passengers may miss boarding a vehicle (Bowman and Turnquist, 1981).

3.2.2 On-demand Operations

Ride-hailing service platforms like Uber and Lyft facilitate two-sided markets. They don't own cars or hire drivers. Instead, they design algorithms to better match drivers and passengers, manipulate wages to attract or discourage drivers to/from certain areas and adjust prices to profit more during peak hours. Most literature in the ride-hailing field also focuses on matching/dispatching strategies, and optimal surge prices for both drivers and passengers.

Matching/dispatching strategies correspond to the process of finding a driver to serve a passenger's request. Previously, greedy methods, which immediately assign the pas-

senger to the nearest available driver, were widely applied in large taxi companies (Liao, 2003; Zhang and Pavone, 2014). Although greedy methods are easy to implement, they are suboptimal and may lead to longer waiting times for future passengers. An alternative matching approach is batching which collects requests over a short period of time and then optimizes to pair each request to an open driver (Korolko et al., 2018; Mazzuoccolo and Mella, 2021). Batching can provide a better solution than the greedy method since it pools requests information and it is especially helpful for ride-sharing services (e.g. Uber Pool).

Surge price has been widely studied as a solution to solve the wild goose chase problem. Wild goose chase happens when cars are thus sent to pick up distant customers, wasting drivers' time, longer passenger waiting time, and invoking more vehicle mileage and congestion in the city (Castillo et al., 2017). This phenomenon occurs when the demand is highly unbalanced especially during the morning commute when most commuters are commuting from the outskirts to the city center. Surge prices can provide incentives for drivers to exit the city center to outskirt regions serving the demand (Besbes et al., 2021). However, surge prices cannot solve the problem entirely. Due to the demand imbalances, drivers still drive empty cars for a large proportion of time which creates even more pollution than commuting using private vehicles. An alternative solution is to use ride-hailing services as feeders to the transit systems which could solve both accessibility problems and wild goose chase problems. This is the objective of this research.

3.2.3 Interaction of fixed-route and on-demand services

The studies most relevant to this paper are the ones that optimize fixed-route transit while taking on-demand services into account. Wei et al. (2021) developed a mathematical model and a solution heuristic to optimize transit timetables considering ride-hailing competition and traffic congestion. The framework was applied to New York City showing a 0.4%-3% system-wide cost reduction. Compared to this study, we optimize the transit network and frequency setting instead of the timetable, which enables us to decrease the social cost

to a larger extent. Instead of focusing on densely populated areas like New York City, we focus on the transit desert regions and aim to increase mobility for low-income and no-vehicle households. Unlike Wei et al. (2021), we also consider multi-modal services which use on-demand services as feeders into the transit system. There are numerous studies aiming to optimize the design of transit feeder systems. Quadrifoglio and Li (2009) proposed an analytical model and solution framework to determine whether fixed-route services or on-demand services are more effective. They came up with the critical demand density threshold to facilitate the decision-making process for policymakers. Chandra and Quadrifoglio (2013) further investigated the optimal terminal-to-terminal cycle length for a demand-responsive transit feeder system. Wang et al. (2020) studied the optimal feeder transit service area considering multiple types of vehicles. More recent studies focus on hybrid transit optimization which co-optimizes the fixed-route transit network design and demand-responsive services. Chen and Nie investigated the problem of a demand-adaptive paired-line hybrid transit system in different networks (e.g. radial (Chen and Nie, 2018), grid (Chen and Nie, 2017)). Luo and Nie (2019) further analyzed the impact of ride-pooling comparing the results of six distinctive transit systems including two fixed-route transit systems and four hybrid transit systems.

An increasing number of studies are proposing a multi-modal system to tackle first-mile and last-mile problems. Jäppinen et al. (2013) estimated that a bike-sharing system would decrease public transit travel times by 10% in Greater Helsinki. Stiglic et al. (2018) showed that the integration of a ride-sharing system and a public transit system can significantly enhance mobility but driver's willingness to accommodate multiple riders is critical to the efficiency of the system (Stiglic et al., 2018). Lee and Savelsbergh (2017) further improve the multi-modal system by giving flexible dropping-off station options. They found that flexible system offers cost advantages over regional systems.

In this paper, we focus on transit network and frequency design as well as on-demand fleet size optimization. The passengers can choose from different transit modes (public

transit, on-demand services only, multi-modal services, and outside options) based on their corresponding attractiveness. To the best of our knowledge, this is the first paper combining the fields of transit network design, on-demand fleet optimization, multi-modal services and applying it to large-scale real-world scenarios.

3.3 Model Description

In this section, we present our MINLP model to optimize the transit network, its corresponding frequencies and on-demand service fleet size considering the endogenous passenger choice. We also present how we linearize the bilinear and exponential constraints as well as the waiting time constraints in the following sections.

3.3.1 Setting and Assumptions

We consider morning commutes with four different travel modes: public transit, on-demand services, multi-modal services (i.e., commuters taking both fixed-route transit and on-demand service), and other travel options (including private vehicles, biking, not traveling, etc.). Commuters make their decision to travel using a certain mode based on its attractiveness relative to the other modes.

The transit agency plays the role of a central planner who decides the fixed-route transit routes, fixed-route transit frequencies and on-demand ride-sharing service fleet size. The ride-sharing vehicles can pick up and drop off passengers in the local region either by dropping passengers off at their destination or at a transit station. However, it is the commuter's decision whether to use multi-modal services. Traffic congestion is assumed to be not affected in a significant way due to the decisions being made in this study. This is a reasonable assumption, consistent with the expected impacts, as shown in Section XXX.

The objective function of this model is to minimize the total system-wide cost including passenger travel cost, parking cost, fixed-route transit operating cost, on-demand service

operating cost, and environmental emissions cost generated by private vehicles, on-demand vehicles, and buses. The constraints include transit budget constraints, capacity constraints, passenger waiting time, travel time, walking time, attractiveness as well as market share for each travel mode.

The model involves a number of assumptions which are listed below:

1. Passenger mode choice is captured using sales-based linear programming reformulation (Gallego et al., 2015) of a general attraction model, whose special case is the well-known multinomial logit model. The market share of a certain travel mode increases with its attractiveness and the attractiveness is a function of travel time, walking time, waiting time, fare, and parking fee.

2. Passengers choose the shortest-distance route when traveling. It is possible that there are multiple routes to travel in a transit network. There are numerous papers investigating different transit assignment strategies (Li et al., 2018; Sun and Szeto, 2018; Codina and Rosell, 2017; Xie et al., 2020; Wu et al., 2019; Xu et al., 2020) and traffic assignment problems (Feng et al., 2020; Xie et al., 2018; Xie and Xie, 2016). However, our focus is on network design and frequency planning, and therefore, we decided to simplify the transit assignment such that the passenger chooses the shortest-distance path if multiple paths exist.

3. When traveling using fixed-route transit, the passenger will take at most one transfer. Transit agencies generally believe that passengers won't take more than one transfer in a journey (Stern, 1996). Therefore, we assume that the passenger can only take either one local bus directly to the destination, or take two different bus lines with one transfer, or take the subway directly to the city center, or transfer from a bus line to the subway. Similarly, for multi-modal services, we assume the passenger will use on-demand services as the first-mile or last-mile transport or both with one bus/subway journey with no transfer in between.

4. The model simplifies the operations for on-demand services. We approximate the travel time for on-demand services as a function of on-demand fleet size and the number of people taking on-demand services in that region. We adopt the analytical model from Daganzo and Ouyang (2019) which provides an analytical framework to compute the on-demand service detour ratio for non-shared rides, dial-a-ride, and ride-sharing services. We applied the model to our setting and fit a 3-dimensional curve with demand, fleet size and detour ratio and added it to the optimization model as a constraint. It enables us to capture the interaction of on-demand travel time, wait time, fleet size and demand without involving complicated routing and matching details.

5. The model considers exogenous transit fare and on-demand service pricing. We applied the prevalent pricing in the optimization model. For transit fare, we applied a flat price structure which is the status quo in Boston. We borrow Uber's pricing structure for ride-sharing services which has a minimum fare for each trip, a base price for making a request, and an additional price based on travel distance or travel time. In Section 3.7.1, we further analyze the impact of different pricing strategies on social cost and market share.

3.3.2 Model Setup

In our model, we take the perspective of a central planner who makes primary decisions of transit network design, corresponding frequencies and on-demand fleet sizes; auxiliary decisions include fleet size for fixed-route transit, passengers' travel disutility, attractiveness for each travel mode and market share.

3.3.2.1 Transit Network, Frequencies, Fleet Sizing, and Budget

Let $\mathcal{L}$ represent the set of potential transit lines. Let N_{max} and N_{min} be the upper and lower bounds on the total number of allowable lines in the network. The lower and upper bounds on transit frequencies are represented by F_{min} and F_{max} respectively. Let the length (in miles) of each transit line $l \in \mathcal{L}$ be L_l. The travel speed for transit is V^F. The purchasing costs for each fixed-route transit and on-demand vehicle are C^F and C^R respectively. The budget for vehicle purchase is B. We define the following decision variables:

$x_l \in \{0, 1\}$: is 1 if transit line $l \in \mathcal{L}$ is selected, and 0 otherwise
$\xi_{ll'} \in \{0, 1\}$: is 1 if transit lines $l, l' \in \mathcal{L}$ are both selected, and 0 otherwise
$f_l \in \mathcal{R}_+$: frequency of transit line $l \in \mathcal{L}$ if the line is selected, and 0 otherwise
$b^R \in \mathcal{Z}_+$: ride-sharing vehicle fleet size
$b^F \in \mathcal{Z}_+$: fixed-route transit vehicle fleet size

---

Constraints (3.1)-(3.2) ensure that the number of chosen fixed-route lines is within an allowable range, and constraints (3.3) restrict the upper and lower bounds on the fixed-route transit frequencies in the final solution. Constraint (3.4) determines transit fleet size to satisfy the line frequencies while constraint (3.5) enforces the budget (B) spent on purchasing fixed-route and on-demand vehicles.

$$\sum_{l \in \mathcal{L}} x_l \leq N_{max} \tag{3.1}$$

$$\sum_{l \in \mathcal{L}} x_l \geq N_{min} \tag{3.2}$$

$$x_l F_{min} \leq f_l \leq x_l F_{max} \tag{3.3}$$

$$b^F \geq \sum_{l \in \mathcal{L}} \frac{L_l f_l}{V^f} \tag{3.4}$$

$$C^R b^R + C^F b^F \leq B \tag{3.5}$$

3.3.2.2 Path Selection Constraints

Let $\mathcal{P}$ and $\mathcal{OD}$ represent the set of paths and set of OD pairs. $\mathcal{P}^F, \mathcal{P}^M \subseteq \mathcal{P}, \mathcal{P}^F \cap \mathcal{P}^M = \emptyset$ partition the set of paths into those using fixed-route transit and multi-modal trips, respectively. $\mathcal{P}^F_{od} \subseteq \mathcal{P}^F$ and $\mathcal{P}^M_{od} \subseteq \mathcal{P}^M$, respectively, are paths serving fixed-route transit and multi-modal trips for OD pair $od \in \mathcal{OD}$. Also, define $\mathcal{P}^{F2} \subseteq \mathcal{P}^F$ as the subset of transit paths that involve transfers between two transit lines.

$y_p \in \{0, 1\}$	is 1 if all transit lines serving path $p \in \mathcal{P}$ are selected in the final solution, and 0 otherwise
$z_p \in \{0, 1\}$	is 1 if path $p \in \mathcal{P}$ is used to serve passengers, and 0 otherwise. Note that $z_p \leq y_p, \forall p \in \mathcal{P}$

In the preprocessing stage, we generate all possible paths to serve each OD pair assuming all transit lines will be included in the final solution. y_p serves as an indicator of whether a path p is available, that is, whether all transit lines needed to serve the path are included in the final solution. Parameter ϕ^{F1}_{lp} is 1 if the first (or the only) leg of the path $p \in \mathcal{P}^F$ is served by transit line $l \in \mathcal{L}$, and 0 otherwise. Similarly, parameter ϕ^{F2}_{lp} is 1 if path $p \in \mathcal{P}^F$ is served by transit line $l \in \mathcal{L}$ in the last (or the only) leg of the trip, and 0 otherwise. Finally, parameter ϕ^{M}_{lp} is 1 if multi-modal path $p \in \mathcal{P}^M$ is served by transit line $l \in \mathcal{L}$, and 0 otherwise. Recall that we don't consider intra-transit transfers for multi-modal trips; so no parameters are needed to be defined for the second transit leg of a multi-modal trip. L_p is the total length of the path $p \in \mathcal{P}$. M is a large positive number. z_p indicates

whether a path $p \in \mathcal{P}$ will be an actual path traveled by commuters.

$$y_p = \sum_{l \in \mathcal{L}} \sum_{l' \in \mathcal{L}} \phi_{lp}^{F1} \phi_{l'p}^{F2} \xi_{ll'} \quad p \in \mathcal{P}^{F2} \tag{3.6}$$

$$y_p = \sum_{l \in \mathcal{L}} \phi_{lp}^{F1} x_l \quad p \in \mathcal{P}^F \setminus \mathcal{P}^{F2} \tag{3.7}$$

$$y_p = \sum_{l \in \mathcal{L}} \phi_{lp}^{M} x_l \quad \forall p \in \mathcal{P}^M \tag{3.8}$$

$$(L_p - L_{p'}) y_{p'} \leq M(1 - z_p) \quad \forall p, p' \in \mathcal{P}_{od}^F, od \in \mathcal{OD} \tag{3.9}$$

$$\sum_{p \in \mathcal{P}_{od}^F} z_p \leq 1 \quad \forall od \in \mathcal{OD} \tag{3.10}$$

$$M \sum_{p \in \mathcal{P}_{od}^F} z_p \geq \sum_{p \in \mathcal{P}_{od}^F} y_p \quad \forall od \in \mathcal{OD} \tag{3.11}$$

$$z_p \leq y_p \quad \forall p \in \mathcal{P} \tag{3.12}$$

$$(L_p - L_{p'}) y_{p'} \leq M(1 - z_p) \quad \forall p, p' \in \mathcal{P}_{od}^M, od \in \mathcal{OD} \tag{3.13}$$

$$\sum_{p \in \mathcal{P}_{od}^M} z_p \leq 1 \quad \forall od \in \mathcal{OD} \tag{3.14}$$

$$M \sum_{p \in \mathcal{P}_{od}^M} z_p \geq \sum_{p \in \mathcal{P}_{od}^M} y_p \quad \forall od \in \mathcal{OD} \tag{3.15}$$

Constraints (3.6)-(3.8) ensure that y_p equals 1 when all relevant transit lines are selected to serve the path; 0 otherwise. Constraints (3.9)-(3.15) ensure that the shortest-distance path will be selected among the available paths. In constraints (3.9), if the length of a transit-only path p (L_p) is strictly greater than that of another transit-only path p' ($L_{p'}$), and path p' is available (i.e., if $y_{p'}$ = 1), then path p will never be selected (i.e., $z_p = 0$). Constraints (3.10) ensure that at most one transit-only path is selected in each OD pair. Constraints (3.11) make sure that at least one transit-only path will be used to serve passengers in each OD pair if any such path is available for that OD pair. Constraints (3.12) ensure that if a path is used to serve passengers, then it must be available. Constraints (3.13)-(3.15) follow the same logic as constraints (3.9)-(3.11) for multi-modal services. Finally, constraints

(3.16)-(3.18) set $\xi_{ll'}$ to its correct value.

$$\xi_{ll'} \leq x_l \quad \forall l, l' \in \mathcal{L} \tag{3.16}$$

$$\xi_{ll'} \leq x_{l'} \quad \forall l, l' \in \mathcal{L} \tag{3.17}$$

$$\xi_{ll'} \geq x_{l'} + x_l - 1 \quad \forall l, l' \in \mathcal{L} \tag{3.18}$$

3.3.2.3 Waiting Time Constraints

For formulating the waiting time constraints, we define the following decision variables:

$w_p^{F1} \in \mathcal{R}_+$:	passenger waiting time for the first leg (before any transfer) on fixed-route transit path $p \in \mathcal{P}_{od}^F$, and is 0 if path p is not used to serve any passengers.
$w_p^{F2} \in \mathcal{R}_+$:	passenger waiting time for the last leg (after any transfer) on fixed-route transit path $p \in \mathcal{P}_{od}^F$; Note that $w_p^{F2} = 0$ if path p is used to serve any passengers, or if path p involves no transfer.
$w_p^F \in \mathcal{R}_+$:	total passenger waiting time when using fixed-route transit path $p \in \mathcal{P}_{od}^F$
$w_p^M \in \mathcal{R}_+$:	passengers' waiting time for the fixed-route transit phase of multi-modal path $p \in \mathcal{P}_{od}^M$

Constraints (3.19)-(3.22) compute waiting times for the fixed-route phase of each path. Constraints (3.19) ensure that the waiting time is greater or equal to half of the headway if the path is taken, and is 0 otherwise. Constraints (3.20) are the analogous constraints for the last leg of a transit-only path involving a transfer. Constraints (3.21) follow a similar logic for multi-modal services. Constraints (3.22) calculate the total waiting time for the transit-only paths.

$$w_p^{F1} \begin{cases} \geq \frac{1}{2\sum_{l\in\mathcal{L}} f_l z_p \phi_{lp}^{F1}}, & \text{if } z_p \neq 0 \\ = 0, & \text{otherwise} \end{cases} \quad \forall p \in \mathcal{P}^F \tag{3.19}$$

$$w_p^{F2} \begin{cases} \geq \frac{1}{2\sum_{l\in\mathcal{L}} f_l z_p \phi_{lp}^{F2}}, & \text{if } z_p \neq 0 \text{ and } p \in \mathcal{P}^{F2} \\ = 0, & \text{otherwise} \end{cases} \quad \forall p \in \mathcal{P}^F \tag{3.20}$$

$$w_p^{M} \begin{cases} \geq \frac{1}{2\sum_{l\in\mathcal{L}} f_l z_p \phi_{lp}^{F1}}, & if\, z_p \neq 0 \\ = 0, & \text{otherwise} \end{cases} \quad \forall p \in \mathcal{P}^M \tag{3.21}$$

$$w_p^F = w_p^{F1} + w_p^{F2} \quad \forall p \in \mathcal{P}^F \tag{3.22}$$

Constraints (3.19), (3.20) and (3.21) are non-linear on multiple accounts. They include the multiplication of a binary and a continuous variable ($f_l z_p$), a reciprocal function, and also a conditional statement dependent on the value of a binary decision variable. We develop an exact reformulation of these constraints into more tractable mixed-integer second-order conic constraints. First, we tackle the multiplication of a binary and a continuous variable in the denominator, creating additional variables $\zeta_{lp} = f_l z_p$ and adding constraints (3.23)-(3.25) to the optimization model.

$$\zeta_{lp} \leq f_l \quad \forall p \in \mathcal{P}, l \in \mathcal{L} \tag{3.23}$$

$$\zeta_{lp} \leq z_p F_{max} \quad \forall p \in \mathcal{P}, l \in \mathcal{L} \tag{3.24}$$

$$\zeta_{lp} \geq f_l - (1 - z_p) F_{max} \quad \forall p \in \mathcal{P}, l \in \mathcal{L} \tag{3.25}$$

Next, we use these ζ_{lp} variables to convert non-linear constraints (3.19), (3.20) and (3.21) into mixed-integer second-order conic constraints. In constraints (3.26), when z_p equals 0, ζ_{lp} will also be 0 (due to constraints (3.24)) and the left-hand side will be trivially

equal to the right-hand side. When z_p is 1, by moving the first term on the right-hand side to the left, we get $(w_p^{F1} + \sum_{l\in\mathcal{L}} \phi_{lp}^{F1}\zeta_{lp})^2 - (w_p^{F1} - \sum_{l\in\mathcal{L}} \phi_{lp}^{F1}\zeta_{lp})^2 \geq 2$; then by expanding the two quadratic terms on the left-hand side, we get $4w_p^{F1}\sum_{l\in\mathcal{L}} \phi_{lp}^{F1}\zeta_{lp} \geq 2$, and since $\sum_{l\in\mathcal{L}} \phi_{lp}^{F1}\zeta_{lp}$ is strictly positive in this case, this is equivalent to $w_p^{F1} \geq \frac{1}{2\sum_{l\in\mathcal{L}} \phi_{lp}^{F1}\zeta_{lp}}$ which in turn is equivalent to constraints (3.19). To ensure that w_p^{F1} is 0 when z_p is 0, constraints (3.27) are added to the model. Similarly, constraints (3.20)-(3.21) are equivalent to constraints (3.28)-(3.32).

$$(w_p^{F1} + \sum_{l\in\mathcal{L}} \phi_{lp}^{F1}\zeta_{lp})^2 \geq (w_p^{F1} - \sum_{l\in\mathcal{L}} \phi_{lp}^{F1}\zeta_{lp})^2 + 2z_p \quad \forall p \in \mathcal{P}^F \tag{3.26}$$

$$w_p^{F1} \leq \frac{z_p}{2F_{min}} \quad \forall p \in \mathcal{P}^F \tag{3.27}$$

$$(w_p^{F2} + \sum_{l\in\mathcal{L}} \phi_{lp}^{F2}\zeta_{lp})^2 \geq (w_p^{F2} - \sum_{l\in\mathcal{L}} \phi_{lp}^{F2}\zeta_{lp})^2 + 2\sum_{l\in\mathcal{L}} z_p\phi_{lp}^{F2} \quad \forall p \in \mathcal{P}^{F2} \tag{3.28}$$

$$w_p^{F2} \leq \frac{\sum_{l\in\mathcal{L}} z_p\phi_{lp}^{F2}}{2F_{min}} \quad \forall p \in \mathcal{P}^{F2} \tag{3.29}$$

$$w_p^{F2} = 0 \quad \forall p \in \mathcal{P}^F \setminus \mathcal{P}^{F2} \tag{3.30}$$

$$(w_p^{M} + \sum_{l\in\mathcal{L}} \phi^{F1}\zeta_{lp})^2 \geq (w_p^{M} - \sum_{l\in\mathcal{L}} \phi^{F1}\zeta_{lp})^2 + 2z_p \quad \forall p \in \mathcal{P}^M \tag{3.31}$$

$$w_p^{M} \leq \frac{z_p}{2F_{min}} \quad \forall p \in \mathcal{P}^M \tag{3.32}$$

3.3.2.4 Utility Calculation Constraints

For each of the four travel modes for each OD pair, the following constraints calculate the deterministic component of the utility as the weighted sum of travel time, wait time, walking time, and commuting expenditure - their utility coefficients are given by β^T, β^W, β^K, and β^C, respectively. β^F, and β^R are the alternative specific constants for fixed-route transit and ride-sharing, respectively. All β values are negative since the attractiveness of the trip is negatively affected by travel time, wait time, walking time, and commuting

expenditure. Therefore, v_p^F, v_p^R, and v_p^M are also non-positive values. F^F is the (flat) fare for the fixed-route transit. Let F^{R0} be the base fare for ride-sharing services and F^{RT} be the hourly markup. F_{MIN}^R is the minimum fare for the ride-sharing services. $T_p^F \ \forall p \in \mathcal{P}^F$, $T_{od}^R \ \forall od \in \mathcal{OD}^F$, $T_p^{MF} \ \forall p \in \mathcal{P}^M$, $T_p^{MR} \ \forall p \in \mathcal{P}^M$ are the in-vehicle travel times for fixed-route transit alternative, for ride-sharing alternative, for the fixed-route transit phase of the multi-modal alternative, and for the ride-sharing phase of the multi-modal alternative, respectively. Both T_{od}^R and T_p^{MR} ignore the waiting time and detour times associated with ride-sharing services. They are accounted for separately through a multiplier Δ, defined as the detour ratio for ride-sharing services compared to traveling using private vehicles. K_p^F is the walking time for path $p \in \mathcal{P}^F$. The related decision variables are listed below.

$v_{od}^F \in \mathcal{R}_-$	utility of taking the fixed-route transit alternative for $od \in \mathcal{OD}$
$v_{od}^R \in \mathcal{R}_-$	utility of taking ride-sharing alternative for $od \in \mathcal{OD}$
$v_{od}^M \in \mathcal{R}_-$	utility of taking multi-modal alternative for $od \in \mathcal{OD}$
Δ	detour ratio compared to traveling using personal vehicles via the shortest path. For example, $\Delta = 2$ means that the travel time (including wait time) is twice the time required for traveling directly using a personal vehicle.

Constraints (3.33)-(3.35) compute the deterministic utility for fixed-route transit, ride-sharing, and multi-modal alternatives respectively. For fixed-route transit, it equals a weighted sum of wait time, in-vehicle travel time, walking time, transit fare, and an alternative-specific constant. For ride-sharing it equals a weighted sum of in-vehicle travel time, fare, and another alternative specific constant. The ride-sharing fare structure in (3.34) follows the prevalent pricing structure which has a minimum fare for each ride, a base fare, and an hourly markup based on the travel time. For multi-modal services, deterministic utility includes negative contributions from the wait time for the fixed-route transit phase, in-vehicle travel time for both fixed-route transit and ride-sharing phases, and fares for both transit and ride-sharing phases of the multi-modal services.

$$v_p^F = \beta^W w_p^F + \beta^T T_p^F + \beta^K K_p^F + \beta^C F^F + \beta^F \quad \forall p \in \mathcal{P}^F \tag{3.33}$$

$$v_{od}^R = \beta^T T_{od}^R \Delta + \beta^C (max(F^{RT} T_{od}^R \Delta + F^{R0}, F_{MIN}^R)) + \beta^R \quad \forall od \in \mathcal{OD} \tag{3.34}$$

$$\begin{aligned} v_p^M = {} & \beta^W w_p^M + \beta^T T_p^{MR} \Delta + \beta^T T_p^{MF} \\ & + \beta^C (max(F^{RT} T_p^{MR} \Delta + F^{R0}, F_{MIN}^R) + F^F) + \beta^R + \beta^F \quad \forall p \in \mathcal{P}^M \end{aligned} \tag{3.35}$$

3.3.2.5 Market Share Constraints

Let $\mathcal{S}$ represent the set of segments, where a segment is defined as the portion of a fixed-route transit line between consecutive stops. Each segment is associated with a unique transit line. If multiple lines share a pair of consecutive stops, then they correspond to multiple elements in set $\mathcal{S}$. ψ_{ps} is 1 if path $p \in \mathcal{P}$ uses segment $s \in \mathcal{S}$, and 0 otherwise. V_{od}^O is the utility of using the outside option for OD pair od. W_{od}^R is 1 if ride-sharing is available for OD pair od, 0 otherwise. Related decision variables are as below:

$s_{od}^F \in \mathcal{R}_+$	fixed-route transit's market share in OD pair $od \in \mathcal{OD}$
$s_{od}^R \in \mathcal{R}_+$	ride-sharing's market share in OD pair $od \in \mathcal{OD}$
$s_{od}^M \in \mathcal{R}_+$	multi-modal service's market share in OD pair $od \in \mathcal{OD}$
$s_{od}^O \in \mathcal{R}_+$	outside option's market share in OD pair $od \in \mathcal{OD}$
$\delta_s \in \{0, 1\}$	1 if the fixed-route transit vehicle is full on segment $s \in \mathcal{S}$, and 0 otherwise.

Constraints (3.36)-(3.39) compute market shares using the sales-based linear programming model from Gallego et al. (2015). Constraints (3.36) ensure that as long as the capacity constraint is not violated, the market share is proportional to the exponential of the deterministic component of utility for each alternative, and when the capacity constraint is tight, the market share is less than or equal to that in the absence of the capacity constraint. Constraints (3.37) follow a similar logic for multi-modal services. When ride-sharing service is available for this OD pair, the market share is proportional to its attractiveness

(constraints (3.38)), and when it is not available, the market share is 0. Constraints (3.39) ensure that the sum of the market shares for all four travel modes is at least 1 for each OD pair, and they are met with equality at the optimal solution.

$$s_{od}^{F} exp(V_{od}^{O}) \begin{cases} = s_{od}^{O} \sum_{p \in \mathcal{P}_{od}^{F}} z_p exp(v_p^F), & \text{if } \sum_{p \in \mathcal{P}_{od}^{F}} z_p \sum_{s \in \mathcal{S}} \psi_{ps} \delta_s = 0 \\ \leq s_{od}^{O} \sum_{p \in \mathcal{P}_{od}^{F}} z_p exp(v_p^F), & \text{otherwise} \end{cases} \quad \forall od \in \mathcal{OD} \tag{3.36}$$

$$s_{od}^{M} exp(V_{od}^{O}) \begin{cases} = s_{od}^{O} \sum_{p^{\in} \mathcal{P}_{od}^{M}} z_p exp(v_p^M), & \text{if } \sum_{p \in \mathcal{P}_{od}^{M}} z_p \sum_{s \in \mathcal{S}} \psi_{ps} \delta_s = 0 \\ \leq s_{od}^{O} \sum_{p \in \mathcal{P}_{od}^{M}} z_p exp(v_p^M), & \text{otherwise} \end{cases} \quad \forall od \in \mathcal{OD} \tag{3.37}$$

$$s_{od}^{R} exp(V_{od}^{O}) = \begin{cases} s_{od}^{O} exp(v_{od}^{R}), & \text{if } W_{od}^{R} \neq 0 \\ 0, & \text{otherwise} \end{cases} \quad \forall od \in \mathcal{OD} \tag{3.38}$$

$$s_{od}^{F} + s_{od}^{R} + s_{od}^{M} + s_{od}^{O} \geq 1 \quad \forall od \in \mathcal{OD} \tag{3.39}$$

In order to incorporate the capacity constraints, we introduce decision variables g_{od}^{F} and g_{od}^{M}, which convert constraints (3.36)-(3.38) into constraints (3.40)-(3.44). If the vehicle capacity constraint is not violated on any of the fixed-route transit segments associated with the path used to serve passengers in an OD pair $od \in \mathcal{OD}$, then the corresponding g_{od}^{F} and g_{od}^{M} should be set to 0. However, when the vehicle is full on any of the fixed-route transit segments, g_{od}^{F} and g_{od}^{M} can be positive, allowing the market shares for fixed-route transit and multi-modal alternatives to be lower than those under the unlimited capacity setting. We introduce continuous decision variables $\sigma_p^O = z_p s_{od}^O \; \forall p \in \mathcal{P}_{od}, od \in \mathcal{OD}$ (as modeled by constraints (3.45)-(3.47)), as well as binary decision variables η_p, which equal 0 when $z_p \sum_{s \in \mathcal{S}} \psi_{ps} \delta_s = 0$, 1 otherwise (as modeled by constraints (3.48)-(3.50)).

$g^F_{od} \in \mathcal{R}_+$	slack in the fixed-route transit attractiveness-market share constraint in $od \in \mathcal{OD}$
$g^M_{od} \in \mathcal{R}_+$	slack in the multi-modal attractiveness-market share constraint in $od \in \mathcal{OD}$

$$s^F_{od} exp(V^O_{od}) = \sum_{p \in \mathcal{P}^F_{od}} \sigma^O_p exp(v^F_p) - g^F_{od} \quad \forall od \in \mathcal{OD} \tag{3.40}$$

$$s^M_{od} exp(V^O_{od}) = \sum_{p \in \mathcal{P}^M_{od}} \sigma^O_p exp(v^M_p) - g^M_{od} \quad \forall od \in \mathcal{OD} \tag{3.41}$$

$$s^R_{od} exp(V^O_{od}) = W^R_{od} s^O_{od} exp(v^R_{od}) \quad \forall od \in \mathcal{OD} \tag{3.42}$$

$$g^F_{od} \leq M \sum_{p \in \mathcal{P}^F_{od}} \eta_p \quad \forall od \in \mathcal{OD} \tag{3.43}$$

$$g^M_{od} \leq M \sum_{p \in \mathcal{P}^M_{od}} \eta_p \quad \forall od \in \mathcal{OD} \tag{3.44}$$

$$\sigma^O_p \leq z_p \quad \forall p \in \mathcal{P} \tag{3.45}$$

$$\sigma^O_p \leq s^O_{od} \quad \forall p \in \mathcal{P}_{od}, od \in \mathcal{OD} \tag{3.46}$$

$$\sigma^O_p \geq s^O_{od} + z_p - 1 \quad \forall p \in \mathcal{P}_{od}, od \in \mathcal{OD} \tag{3.47}$$

$$\eta_p \leq z_p \quad \forall p \in \mathcal{P} \tag{3.48}$$

$$\eta_p \leq \sum_{s \in \mathcal{S}} \psi_{ps} \delta_s \quad \forall p \in \mathcal{P} \tag{3.49}$$

$$\eta_p \geq z_p + \psi_{ps} \delta_s - 1 \quad \forall s \in \mathcal{S}, p \in \mathcal{P} \tag{3.50}$$

Note that these reformulations have ensured that the only non-linearities remain the first terms of the right hand sides of constraints (3.40)-(3.42), in the form of a decision variable multiplied by an exponential of another decision variable. While these terms make constraints (3.40)-(3.42) highly non-linear and non-convex, some of the most advanced optimization solvers, e.g., Gurobi 9.1, offer quadratic and bilinear solvers which use piece-wise linear approximations as well as cutting planes and branching methods to tackle bilinearity and exponential constraints (Achterberg and Towle, 2020). However, these approaches end

up introducing a significant number of binary integer decision variables which increase the runtime dramatically.

Therefore, we applied an iterative first-order linear approximation to get the solution in a shorter time which we describe further in Section 3.4.3.1. Constraints (3.40)-(3.42) are replaced by first-order Taylor series expansions at points ($\overline{v_p^F}$, $\overline{s_{od}^O}$), ($\overline{v_p^M}$, $\overline{s_{od}^O}$), and ($\overline{v_{od}^R}$, $\overline{s_{od}^O}$) respectively. We also introduce decision variables ν_p^F and ν_p^M representing $z_p v_p^F$ and $z_p v_p^M$ respectively for linearization (as modeled in constraints (3.54)-(3.59)). Constraints (3.40)-(3.42) can be replaced by constraints (3.51)-(3.59).

$$s_{od}^F exp(V_{od}^O) = \sum_{p \in \mathcal{P}_{od}^F} \sigma_p^O exp(\overline{v_p^F}) + \overline{s_{od}^O} \sum_{p \in \mathcal{P}_{od}^F} \nu_p^F exp(\overline{v_p^F}) - \overline{s_{od}^O} \sum_{p \in \mathcal{P}_{od}^F} z_p \overline{v_p^F} exp(\overline{v_p^F}) - g_{od}^F$$

$$\forall od \in \mathcal{OD} \tag{3.51}$$

$$s_{od}^M exp(V_{od}^O) = \sum_{p \in \mathcal{P}_{od}^M} \sigma_p^O exp(\overline{v_p^M}) + \overline{s_{od}^O} \sum_{p \in \mathcal{P}_{od}^M} \nu_p^M exp(\overline{v_p^M}) - \overline{s_{od}^O} \sum_{p \in \mathcal{P}_{od}^M} z_p \overline{v_p^M} exp(\overline{v_p^M})$$

$$- g_{od}^M \quad \forall od \in \mathcal{OD} \tag{3.52}$$

$$s_{od}^R exp(V_{od}^O) = W_{od}^R \overline{s_{od}^O} exp(\overline{v_{od}^R}) + W_{od}^R \overline{s_{od}^O} exp(\overline{v_{od}^R})(v_{od}^R - \overline{v_{od}^R}) + W_{od}^R exp(\overline{v_{od}^R})(s_{od}^O - \overline{s_{od}^O})$$

$$\forall od \in \mathcal{OD} \tag{3.53}$$

$$\nu_p^F \geq -M z_p \quad \forall p \in \mathcal{P}^F \tag{3.54}$$

$$\nu_p^F \geq v_p^F \quad \forall p \in \mathcal{P}^F \tag{3.55}$$

$$\nu_p^F \leq v_p^F + M(1 - z_p) \quad \forall p \in \mathcal{P}^F \tag{3.56}$$

$$\nu_p^M \geq -M z_p \quad \forall p \in \mathcal{P}^M \tag{3.57}$$

$$\nu_p^M \geq v_p^M \quad \forall p \in \mathcal{P}^M \tag{3.58}$$

$$\nu_p^M \leq v_p^M + M(1 - z_p) \quad \forall p \in \mathcal{P}^M \tag{3.59}$$

3.3.2.6 Ride-sharing Travel Time Constraints

An important set of constraints is needed to establish the relationship between ride-sharing detour time (including both travel time and waiting time), ride-sharing fleet size, and the number of customers taking the ride-sharing services. D_{od} represents the demand for OD pair od. Constraint (3.60) model this relationship by means of a new decision variable related to ride-sharing service operations defined as below:

d^R	Number of passengers using ride-sharing services

We adopt the method from Daganzo and Ouyang (2019) to capture this relationship. First, we plotted the data points corresponding to different combinations of fleet sizes and demands, and their estimated detour ratios. Then we fit the data using second-order polynomial regression. For all case studies, the R square was found to be more than 96%. More details are provided in the appendix.

$$\Delta = f(d^R, b^R) \tag{3.60}$$

$$d^R = \sum_{od \in \mathcal{OD}} ((s^R_{od} + s^M_{od}) D_{od}) \tag{3.61}$$

3.3.2.7 Capacity Constraints

K^F represents each fixed-route transit vehicle's passenger carrying capacity. To indicate whether the capacity constraint is binding on each segment, we create new decision variables as below:

g_s	slack in the capacity constraint of fixed-route transit segment $s \in \mathcal{S}$

Define $\mathcal{S}_l$ as the set of segments belonging to transit line $l \in \mathcal{L}$. Constraints (3.62) ensure that the number of passengers on each segment of the fixed-route transit network does not exceed vehicle capacity. Constraints (3.63)-(3.64) ensure that when g_s is 0, δ_s is 1 and when g_s is positive, δ_s is 0.

$$\sum_{od \in \mathcal{OD}} (s_{od}^F \sum_{p \in \mathcal{P}_{od}^F} \psi_{ps} z_p + s_{od}^M \sum_{p \in \mathcal{P}_{od}^M} \psi_{ps} z_p) D_{od} = K^F f_l - g_s \quad \forall s \in \mathcal{S}_l, l \in \mathcal{L} \tag{3.62}$$

$$1 - \delta_s \leq M g_s \quad \forall s \in \mathcal{S} \tag{3.63}$$

$$M(1 - \delta_s) \geq g_s \quad \forall s \in \mathcal{S} \tag{3.64}$$

The multiplication of z_p with s_{od}^F as well as s_{od}^M generates non-linearities in constraints (3.62). We linearize them by defining $\sigma_p^F = z_p s_{od}^F$ and $\sigma_p^M = z_p s_{od}^M$ (as modeled in constraints (3.66)-(3.71)). This converts constraints (3.62) into constraints (3.65).

$$\sum_{od \in \mathcal{OD}} (\sum_{p \in \mathcal{P}_{od}^F} \psi_{ps} \sigma_p^F + \sum_{p \in \mathcal{P}_{od}^M} \psi_{ps} \sigma_p^M) D_{od} = K^F f_l - g_s \quad \forall s \in \mathcal{S}_l, l \in \mathcal{L} \tag{3.65}$$

$$\sigma_p^F \leq z_p \quad \forall p \in \mathcal{P}^F \tag{3.66}$$

$$\sigma_p^F \leq s_{od}^F \quad \forall p \in \mathcal{P}_{od}^F, od \in \mathcal{OD} \tag{3.67}$$

$$\sigma_p^F \geq s_{od}^F + z_p - 1 \quad \forall p \in \mathcal{P}_{od}^F, od \in \mathcal{OD} \tag{3.68}$$

$$\sigma_p^M \leq z_p \quad \forall p \in \mathcal{P}^M \tag{3.69}$$

$$\sigma_p^M \leq s_{od}^M \quad \forall p \in \mathcal{P}_{od}^M, od \in \mathcal{OD} \tag{3.70}$$

$$\sigma_p^M \geq s_{od}^M + z_p - 1 \quad \forall p \in \mathcal{P}_{od}^M, od \in \mathcal{OD} \tag{3.71}$$

3.3.2.8 Objective Function

The model minimizes total systemwide costs including the costs to passengers, costs to operators, and the environmental costs. C^{FW} is the unit cost for walking to the transit station. C^{RT} is the unit cost of in-vehicle travel time using ride-sharing. C_{od}^{OP} is the cost of using the outside option for the passengers in OD pair $od \in \mathcal{OD}$. $C_{od}^{OE} \in \mathcal{OD}$ is the environmental cost of using the outside option in OD pair $od \in \mathcal{OD}$. C^{FE} and C^{RE} are the environmental costs of using each transit/ride-sharing vehicle per unit time. C^{FO} and C^{RO} are the operational costs for each transit/ride-sharing vehicle per unit time. Since the walking cost and in-vehicle travel cost are fixed for each path p, we use the parameter U_p^F to denote these two passenger costs for path p. Similarly, we use parameter U_p^M to denote the cost of the in-vehicle travel time of the fixed-route transit phase of a multi-modal path p.

The objective function is presented next. Its first term is the sum of the walking cost and the in-vehicle travel time cost for fixed-route transit. Similarly, the second term is the in-vehicle travel time cost of the fixed-route phase of the multi-modal alternative. The third and fourth terms are the wait time costs for the fixed-route transit alternative, and for the fixed-route transit phase of the multi-modal alternative, respectively. The fifth and sixth terms are the in-vehicle travel time costs for the ride-sharing alternative and for the ride-sharing phase of the multi-modal alternative, respectively. The seventh term is the cost associated with traveling using outside options which includes parking cost, no travel cost, and travel time cost. The eighth, ninth, and tenth term are the environmental costs for outside option, fixed-route transit, and ride-sharing services, respectively. The last two terms are the operational costs for fixed-route transit and ride-sharing services, respectively.

$$
\begin{aligned}
Min \quad & \sum_{od\in\mathcal{OD}}\left(D_{od}s^F_{od}\sum_{p\in\mathcal{P}^F_{od}} z_p U^F_p\right) + \sum_{od\in\mathcal{OD}}\left(D_{od}s^M_{od}\sum_{p\in\mathcal{P}^M_{od}} z_p U^M_p\right) \\
& + C^{FW}\sum_{od\in\mathcal{OD}}\left(D_{od}s^F_{od}\sum_{p\in\mathcal{P}^F_{od}} w^F_p\right) + C^{FW}\sum_{od\in\mathcal{OD}}\left(D_{od}s^M_{od}\sum_{p\in\mathcal{P}^M_{od}} w^M_p\right) \\
& + C^{RT}\sum_{od\in\mathcal{OD}} D_{od}T^R_{od}\Delta s^R_{od} + C^{RT}\sum_{od\in\mathcal{OD}}\left(D_{od}s^M_{od}\sum_{p\in\mathcal{P}^M_{od}} z_p\Delta T^{MR}_p\right) \\
& + \sum_{od\in\mathcal{OD}} D_{od}(C^{OP}_{od} + C^{OE}_{od})s^O_{od} + (C^{FO} + C^{FE})b^F + (C^{RO} + C^{RE})b^R \qquad (3.72)
\end{aligned}
$$

Note that the first six terms in the objective function are non-linear. The first two terms are the multiplication of a binary variable and a continuous variable, while the rest are the multiplications of two continuous variables. The multiplication of a binary variable and a continuous variable can be linearized by introducing another variable. For the multiplication of two continuous variables, we will use other techniques to handle them.

There are non-linear terms in the objective function including both bilinear terms (the multiplication of two continuous variables) and the multiplication of binary variables and continuous variables. As mentioned in Section 3.3.2.5, we keep the bilinear terms in the first step and introduce decision variables $\mu^M_p = z_p\Delta$ (modeled by constraints (3.74)-(3.76)). The objective function and added constraints for the first step are presented below.

$$
\begin{aligned}
Min \quad & \sum_{od\in\mathcal{OD}}\left(D_{od}\sum_{p\in\mathcal{P}_{od}^{F}}\sigma_p^F U_p^F\right)+\sum_{od\in\mathcal{OD}}\left(D_{od}\sum_{p\in\mathcal{P}_{od}^{M}}\sigma_p^M U_p^M\right) \\
& + C^{FW}\sum_{od\in\mathcal{OD}}\left(D_{od}s_{od}^F\sum_{p\in\mathcal{P}_{od}^{F}}w_p^F\right)+C^{FW}\sum_{od\in\mathcal{OD}}\left(D_{od}s_{od}^M\sum_{p\in\mathcal{P}_{od}^{M}}w_p^M\right) \\
& + C^{RT}\Delta\sum_{od\in\mathcal{OD}}D_{od}T_{od}^R s_{od}^R+C^{RT}\sum_{od\in\mathcal{OD}}\left(D_{od}s_{od}^M\sum_{p\in\mathcal{P}_{od}^{M}}T_p^{MR}\mu_p^M\right) \\
& + \sum_{od\in\mathcal{OD}}D_{od}(C_{od}^{OP}+C_{od}^{OE})s_{od}^O+(C^{FO}+C^{FE})b^F+(C^{RO}+C^{RE})b^R \qquad (3.73)
\end{aligned}
$$

$$
\mu_p^M \le \Delta \quad \forall p\in\mathcal{P}^M \qquad (3.74)
$$

$$
\mu_p^M \le M z_p \quad \forall p\in\mathcal{P}^M \qquad (3.75)
$$

$$
\mu_p^M \ge \Delta - M(1-z_p) \quad \forall p\in\mathcal{P}^M \qquad (3.76)
$$

In Step 2, we completely linearize the model by applying firsr-order Taylor series expansion to the third, fourth, fifth and sixth terms in the objective function (3.73) at points $(\overline{w_p^F}, \overline{s_{od}^F})$, $(\overline{w_p^M}, \overline{s_{od}^M})$, $(\overline{\Delta}, \overline{s_{od}^R})$ and $(\overline{\Delta}, \overline{s_{od}^M})$ respectively. We also introduce decision variables $\rho_p^M = z_p s_{od}^M$ for linearization (modeled by constraints (3.78)-(3.80)). We modify the objective function and add constraints as below.

$$
\begin{aligned}
Min \quad & \sum_{od\in\mathcal{OD}}\Big(D_{od}\sum_{p\in\mathcal{P}_{od}^{F}}\sigma_p^F U_p^F\Big)+\sum_{od\in\mathcal{OD}}\Big(D_{od}\sum_{p\in\mathcal{P}_{od}^{M}}\sigma_p^M U_p^M\Big) \\
& + C^{FW}\sum_{od\in\mathcal{OD}}\Big(D_{od}\sum_{p\in\mathcal{P}_{od}^{F}}(w_p^F\overline{s_{od}^F}+\overline{w_p^F}s_{od}^F-\overline{w_p^F}\,\overline{s_{od}^F})\Big) \\
& + C^{FW}\sum_{od\in\mathcal{OD}}\Big(D_{od}\sum_{p\in\mathcal{P}_{od}^{M}}(w_p^M\overline{s_{od}^M}+\overline{w_p^M}s_{od}^M-\overline{w_p^M}\,\overline{s_{od}^M})\Big) \\
& + C^{RT}\sum_{od\in\mathcal{OD}}\Big(D_{od}T_{od}^R(\Delta\overline{s_{od}^R}+\overline{\Delta}s_{od}^R-\overline{\Delta}\,\overline{s_{od}^R})\Big) \\
& + C^{RT}\sum_{od\in\mathcal{OD}}\Big(D_{od}\sum_{p\in\mathcal{P}_{od}^{M}}(T_p^{MR}\mu_p^M\overline{s_{od}^M}+T_p^{MR}\rho_p^M\overline{\Delta}-z_pT_p^{MR}\overline{s_{od}^M}\,\overline{\Delta})\Big) \\
& + \sum_{od\in\mathcal{OD}}D_{od}(C_{od}^{OP}+C_{od}^{OE})s_{od}^O+(C^{FO}+C^{FE})b^F+(C^{RO}+C^{RE})b^R \qquad (3.77)
\end{aligned}
$$

$$\rho_p^M \le s_{od}^M \quad \forall p\in\mathcal{P}_{od}^M, od\in\mathcal{OD} \qquad (3.78)$$

$$\rho_p^M \le Mz_p^M \quad \forall p\in\mathcal{P}^M \qquad (3.79)$$

$$\rho_p^M \ge s_{od}^M - M(1-z_p^M) \quad \forall p\in\mathcal{P}_{od}^M, od\in\mathcal{OD} \qquad (3.80)$$

3.4 Solution Approach

The model presented in Section 3 aims to optimize fixed-route transit network design, line frequencies, and the ride-sharing fleet size. The original model is an MINLP model with objective function (3.73) and constraints (3.1)-(3.18), (3.22)-(3.35), (3.39)-(3.50), (3.60)-(3.61), (3.63) - (3.71) and (3.74)-(3.76). It is a complicated problem to solve for practical-sized instances for three reasons. First, since we need to apply this to a large network with more than 4,000 OD pairs and 14,000 paths, a large number of integer variables | corresponding to network design and path selection decisions | are involved. Second, the objective function (3.73), as well as the constraints (3.40), (3.41), and (3.42) have bilinear terms involving a product of multiple continuous decision variables. Finally, constraints

(3.40), (3.41), and (3.42) also involve exponential terms. There is no known exact approach to solve this to optimality in a computationally efficient manner. MINLP generalizes both mixed-integer programming problems (MIP) and non-linear programming problems (NLP) which makes the problem difficult to solve in both classes (**?**). Therefore, we strive to separate the complexity of MIP from the NLP and conquer them individually. **?** is one good example of how the separation can be done by repeatedly searching for better integer solutions and solving the NLP subproblem. In this paper, we follow a similar but innovative approach. We apply a new combination of (1) rounding method, and (2) first-order approximation method to solve this complex MINLP model,

3.4.1 Rounding Heuristics

In order to solve a problem with such a large number of integer variables, we adopt the rounding method. The rounding method is often used when solving large-scale integer programs (**???**). **?** designed a heuristic method to solve the robust crew-pairing optimization problem. They first solved the LP relaxation of the original problem, then applied the rounding method to assign the largest fractional variables to 1.

The conventional way of applying rounding heuristics is to run the continuous relaxation of the original model to optimality, then assign the decision variables based on certain criteria. After investigating this approach, we decided that this approach is not feasible. First of all, the runtime for continuous relaxation is extremely long. Due to the existence of exponential constraints and bilinear constraints in the model, Gurobi automatically introduces more integer variables when applying piece-wise linear approximation to these constraints. It usually takes more than 10 days to get a solution with 5% optimality gap. Secondly, Gurobi can provide a much tighter lower bound when running with binary variables than in the continuous relaxation model. We found that Gurobi applies advanced presolving and cutting planes techniques when running the original MINLP model which provides a much better lower bound. In recent years, Gurobi has witnessed remarkable im-

provements in the capabilities of MIP algorithms and four of the biggest contributors have been presolved, cutting planes, heuristics, and parallelism (Gurobi, 2019). When applying rounding heuristics when running the original model, we are able to get a solution in 10 hours. We also observe that many binary variables that should be assigned to 0 are already very close to 0 in the root node. It shows that our approach is superior in both run-time and solution quality. Therefore, rather than applying the rounding method to the continuous relaxation of the original MINLP model, we run the original model directly and apply the rounding method to x_l obtained at the root node. When x_l is determined, y_p and z_p are automatically determined according to constraints (3.6)-(3.18). The significant decrease in the number of integer variables renders the problem feasible to solve.

There are various ways to apply the rounding method. The first one is rounding x_l according to the fractional values of x_l in the root node. Another approach is rounding x_l according to domain-specific metrics like the number of people using the transit line in the root node. We decided to adopt the domain-specific rounding method due to its better performance.

3.4.2 First-order Approximation Method

When solving non-linear, non-convex optimization models, one of the widely used methods is the first-order approximation method (**?**). In **?**'s paper, they encounter similar bilinear and exponential constraints when optimizing transit frequencies and price. Adopting the first-order approximation method, **?** replaced the non-linear multinomial logit choice constraints using first-order local linear constraints.

However, different from **?**'s problem setting, our problem also requires optimizing the transit network design. The MINLP involves a significant number of binary decision variables (e.g., x_l, y_p, z_p, $\xi_{ll'}$, δ_s) that make the optimization model difficult to solve even after the first-order approximation. Therefore, we first apply the rounding heuristics to determine the transit lines (X_l). After the transit lines are determined, we manage to eliminate

x_l, y_p, z_p, and $\xi_{ll'}$. The only remaining integer variables are δ_s, b^R and b^F. In the second step, we apply the first-order approximation method to optimize the model with the objective function (3.77) and constraints (3.1)-(3.18), (3.22)-(3.35), (3.39), (3.43)-(3.61), (3.63)-(3.71), (3.74)-(3.76), and (3.78)-(3.81). The reference points are updated after each iteration until convergence.

$$x_l = X_l \quad \forall l \in \mathcal{L} \tag{3.81}$$

3.4.3 Proposed Method

We propose a two-step solution approach to solve this large-scale non-convex optimization problem. The solution framework is presented in Figure 3.1. The details will be elaborated in the remainder of Section 3.4. We first generate a set of potential transit lines to be selected in the final transit network. Then we generate the parameters needed to run the optimization model. Furthermore, we apply our two-step approach to solve this complicated model. The first step aims to determine the transit network using rounding heuristics by assigning the binary transit line variables (x_{l_m}) to either 0 or 1 based on the number of passengers they serve. The second step aims to decide the frequency and on-demand fleet size given the transit network decided in the first step. First-order Taylor expansion is applied to bilinear and exponential terms and constraints. The MINLP model is transformed into a mixed-integer second-order conic programming (MISOCP) model. Then we iteratively run the MISOCP model with updated Taylor expansion points at each iteration until it converges.

In Step 1 (as described in the third box of Figure 3.1), we run the MINLP model with the objective function (3.73) and constraints (3.1)-(3.18), (3.22)-(3.35), (3.39)-(3.50), (3.60)-(3.61), (3.63) - (3.71) and (3.74)-(3.76). Due to its non-convexity and non-linearity, we cannot run this MINLP model to optimality. Therefore, we apply the rounding heuristics

to assign x_{l_m} to either 1 or 0 based on the number of passengers served by the transit line. After the transit network is decided in the first step, we run the MISOCP model with the objective function (3.77) and constraints (3.1)-(3.18), (3.22)-(3.35), (3.39), (3.43)-(3.61), (3.63)-(3.71), (3.74)-(3.76), and (3.78)-(3.81). Since we applied Taylor expansion in the second step, we need to run this model iteratively and update the Taylor expansion points at each iteration until it converges.

Different from the previous two-step studies that determine the transit network in the first step and frequencies in the second step, this study incorporates the interaction among market share, different travel modes, and passenger choice in both steps while most previous literature doesn't consider them in the first step.

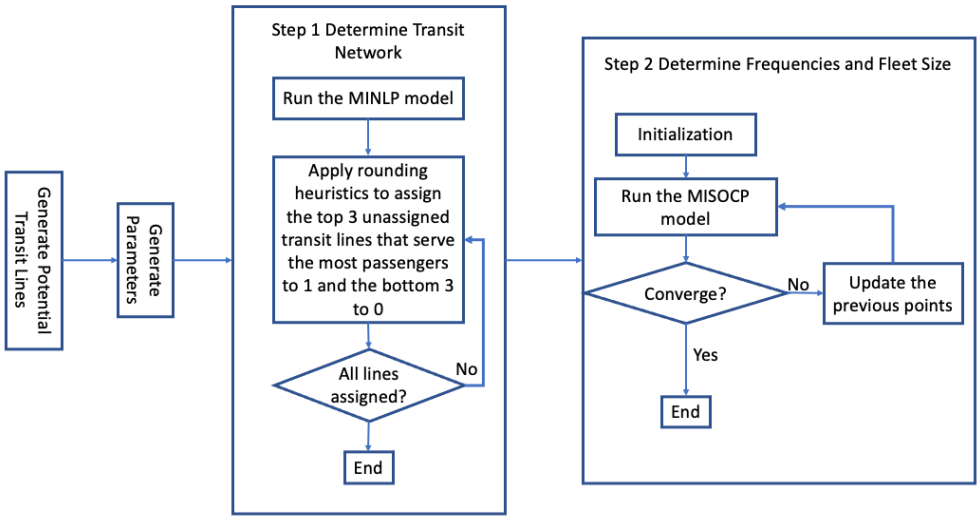


Figure 3.1: Solution Approach Framework

3.4.3.1 First Step: Applying Rounding Heuristics Method to Determine Transit Lines

In the first step, we run the MINLP model with objective function (3.73) and constraints (3.1)-(3.18), (3.22)-(3.35), (3.39)-(3.50), (3.60)-(3.61), (3.63) - (3.71) and (3.74)-(3.76). This model is an MINLP model with 11672 quadratic objective terms, 40897 quadratic constraints, 640724 continuous, and 66485 integers (66458 binary) variables in one of our

case studies. Due to its non-linear, non-convex nature, the model itself takes more than 10 days to get the first feasible solution. Therefore, we decided to apply rounding heuristics to determine the transit networks. One important feature of this problem is that even though there is a large number of integer variables, most of them depend on the transit network. y_p, z_p can be easily obtained once x_{l_m} is determined. Once the transit network is decided, the number of integer variables will decrease to 10. Therefore, when applying rounding heuristics, we applied it only to the binary variables that decide whether to include certain transit lines in the network. We first run the MINLP model using Gurobi 9.1 solver with callback functions displaying x_{l_m}, fr_{l_m} and the number of passengers served by each transit line at each MIP node. Then we round the unassigned binary variables to either 0 or 1 based on the number of commuters they serve. The top three lines that serve the most demand are assigned to 1 and the bottom three lines are assigned to 0. This process is repeated until all lines are assigned.

3.4.3.2 Second Step: Applying an Iterative Linearization Method to Optimize Frequency and Fleet Size

After deciding the transit lines in the network, the second step aims to optimize the frequencies for each transit line and on-demand fleet size. As discussed in Section 3.4.3.1, most of the integer variables come from the transit network decision. Once x_l is fixed, y_p and z_p are both determined. Therefore, we have very few integer variables at this stage (usually fewer than 30). However, the bilinear and exponential constraints will still slow down the process significantly since Gurobi introduces a large number of binary variables when applying piece-wise linear approximation and a special-ordered set. Therefore, we approximate all the bilinear and exponential constraints with first-order linear approximation.

We apply first-order Taylor expansion to the third, fourth, fifth, and sixth terms in the objective function (3.73) as well as constraints (3.40), (3.42), and (3.41). We run the MISOCP model with objective function (3.77) and constraints (3.1)-(3.18), (3.22)-(3.35),

(3.39), (3.43)-(3.61), (3.63)-(3.71), (3.74)-(3.76), and (3.78)-(3.81). It's worth noting that since the transit network is decided, x_{l_m}, y_p, and z_p are constant instead of decision variables. Therefore, the path selection constraints are not included in the second-step optimization model.

First, we obtain an initial solution by fixing the frequencies and on-demand fleet size at the prevalent value. Then, we will have an initial value for s_{od}^F, s_{od}^R, s_{od}^M, s_{od}^O, v_p^F, v_{od}^R, v_p^M, w_p^F, w_p^M in objective function (3.77) and constraints (3.51), (3.52), and (3.53). We run the MISOCP model as described above and update the Taylor expansion points after each iteration until it converges.

3.5 Computational Experiment

We use the greater Boston area as a case study to demonstrate how to apply this framework to address the transit desert problem.

The "focus40" report published by massDOT (2019) identified priority places that need or can support higher-quality transit. In our case study, we focus on the regions defined as Urban Gateways by massDOT (2019). The definition is presented below:

- Urban gateways: Boston is far from the only city within the MBTA service area. Several smaller cities have many of the characteristics that make Boston so amenable to transit services, such as population and employment density and walkable streets. These cities are home to a disproportionate share of the region's immigrant and lower-income workforce and those dependent on public transportation. Housing in these communities is also more affordable than in Boston and its immediate suburbs. In a region where many municipalities can be reluctant to build dense, mixed-use development, these Urban Gateways are often eager to attract new residents and jobs. Collectively, the Urban Gateways represent enormous potential to support the region's economic, environmental, and equity goals. More and better transit could be

the key to unlock the potential of these communities and better connect their residents to the regional economy.

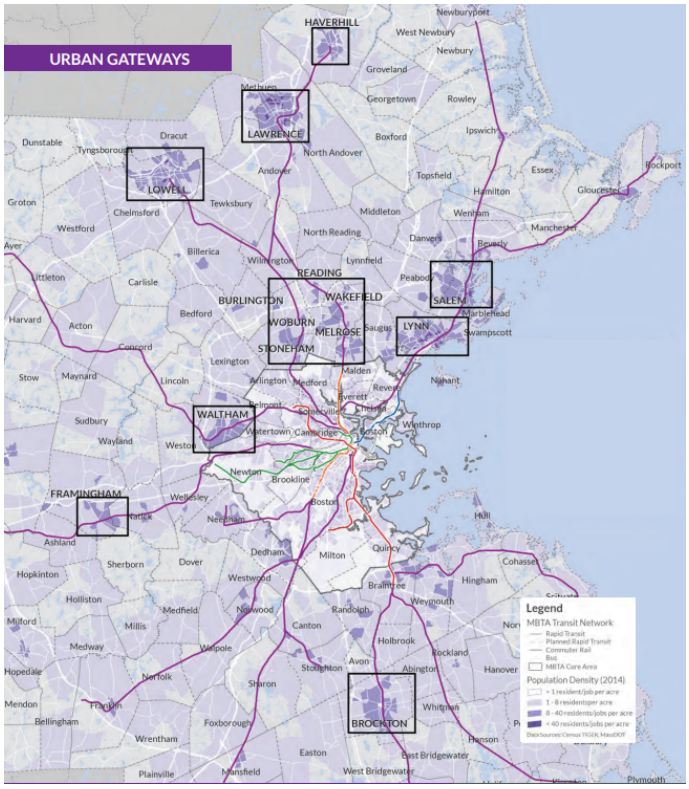


Figure 3.2: Urban Gateways identified by MBTA (Source: Focus40: positioning the MBTA to meet the needs in the regions in 2040)

These regions identified as Urban Gateways are suffered from low transit accessibility. The larger proportion of low-income and household with no private vehicles also exacerbate the situation. massDOT (2019) identified 9 regions in the Urban Gateways categories including Woburn-Melrose region, Salem region, Lynn region, Waltham region, Brockton region, and Framingham region, Lowell region, Lawrence region, and Haverhill region. Woburn-Melrose region, Salem region, Lynn region, and Waltham region are served by MBTA buses. Therefore, we focus on optimizing the bus system for these four regions. Due to the proximity of Salem and Lynn regions with multiple bus lines serving both regions, we combine these two regions and optimize the bus lines within the combined area.

3.5.1 Experimental Setup

Since we consider both on-demand services and fixed-route transit services, we obtain both road networks for the greater Boston area and the current subway and bus networks. The road network consists of 14209 nodes and 71076 road segments. The subway network consists of 7397 nodes and 14876 road segments. We are interested in optimizing the transit schedules during the morning rush hours (6:00 a.m. - 9:00 a.m.).

We get accurate block-level OD pair commuter data from the 2010 Census LODES database. For each OD pair, it provides the Workplace Census Block Code, Residence Census Block Code, the total number of jobs, number of jobs under different age brackets, number of jobs under different income levels, and number of jobs in different industries (United States Census Bureau, 2020).

We obtain all relevant information on subway, buses, and commuter rail from Massachusetts Bay Transportation Authority (MBTA, 2021a) in the format of General Transit Feed Specification (GTFS). It provides the locations of stops, transit lines, lengths and schedule. We also get the transit line geographic information in the format of shapefile for commuter rail, subway, and buses from MassDOT website (MassDOT, 2021). It provides the transit lines which can be easily converted into a graph in Python using networkx package. In our optimization, we enforce N_{min} to be equal to N_{max} which equals the current number of transit lines serving the region.

We categorize four types of commuters in our data. The first type is higher-income commuters with private vehicles, the second type higher-income commuters with no private vehicles, the third type lower-income commuters with private vehicles, and the fourth type lower-income commuters with no private vehicles. The income information is included in the Census LODES data (United States Census Bureau, 2020). For each OD pair, it shows the number of jobs with earnings \$1250/month or less, the number of jobs with earnings \$1251/month to \$3333/month, and the number of jobs with earnings greater than \$3333/month. Furthermore, we obtain the vehicle ownership information from 2019

American Community Survey: 5-Year Data obtained from IPUMS (Manson et al., 2020). Households with an annual income of less than $25,000 are almost nine times as likely to be a zero-vehicle household than households with incomes greater than $25,000. Therefore, we assign 90% of people with no vehicle to low-income with no private vehicle categories and the rest 10% to high-income with no vehicle categories.

We adopted the same utility coefficients as in Wei et al. (2021). However, we assigned different utility coefficients to high-income and low-income commuters with low-income commuters having higher coefficients for cost and fares and high-income commuters having higher coefficients for traveling time, walking time, and waiting time. The coefficients are tuned so that the market shares of different travel modes in the current system fit the market shares in reality. In the objective function, we also apply different cost coefficients for different categories of passengers, the traveling/walking/waiting time cost coefficients for low-income commuters are half of those for high-income commuters since their income is lower.

3.5.2 Computational Results

To prove the efficiency of our method, a few alternative methods were investigated and compared with our method. We listed different scenarios that we want to compare in Table 3.1.

Table 3.1 Different scenarios for computational comparison

				First-order approximation	
				No	Yes
Rounding Heuristics	No			S1	S2
	Yes	LP relaxation		S3	S4
		Root node	Naïve	S5	S6
			Domain-specific	S7	S8

The columns indicate whether we apply the first-order approximation method. The rows indicate whether and how we apply the rounding heuristics. In S1, where we ap-

ply neither rounding heuristics nor first-order approximation, we simply feed the original model to Gurobi and let it handle the MINLP problem directly. In S2, we only apply the first-order approximation method. We optimize the model with objective function (3.77) and constraints (3.1)-(3.18), (3.22)-(3.35), (3.43)-(3.61), (3.63)-(3.71), (3.74)-(3.76), and (3.78)-(3.80). For S3-S8, we apply the rounding heuristics to obtain transit lines in the first stage. In S3 and S4, we first run the LP relaxation of the original model to optimality, then apply the rounding method to obtain the final transit lines. In the second step, we run the original problem with fixed x_l values in S3. In S4, we apply first-order approximation and fix x_l to the values obtained from step 1. In S5-S8, the rounding heuristic is applied to the root node. We first optimize the original model using Gurobi to obtain the root node. Then we either apply rounding heuristics using the naive method (based on the value of x_l) or using domain-specific method (based on the number of passengers using transit line x_l). After determining the transit lines, S5 and S7 will run the original model again with fixed transit lines X_l. S6 and S8 will adopt first-order approximation method in the second step. S8 is the method we are adopting in this paper.

We conducted these different scenarios in the Waltham region. In order to have a fair comparison, we set the runtime to a maximum of six hours. The computational results are listed in Table 3.2.

Table 3.2 Computational Results for Different scenarios

				First-order approximation	
				No	Yes
Rounding Heuristics	No			NaN (bound: 33069)	69653
	Yes	LP relaxation		NaN (bound: 21671)	NaN (bound: 21671)
		Root node	Naïve	NaN (bound: 40807)	53798
			Domain-specific	NaN (bound: 39482)	51565

As shown in Table 3.2, S1, S3, S4, S5, and S7 cannot provide a feasible solution within the runtime limit. It demonstrates the complexity of the problem. In S5 and S7, even after fixing the binary variables, we are not able to solve the NLP successfully without adopting the first-order approximation method. Even though S2 and S6 are able to provide a feasible

solution, the final result is 35% and 4.3% higher than the result obtained using our method (in S8). It demonstrates the necessity of using the rounding method in the first stage.

3.6 Practical Results

We applied our model and solution method to Woburn-Melrose region, Salem-Lynn region, and Waltham region, the total computational times range from 6 to 12 hours which is reasonable for making strategic level decisions.

Table 3.3 summarized the passenger mode share, passenger disutility, social cost for three regions by different commuter types under the current scenarios and optimal scenarios.

Table 3.3 Passenger mode share, disutility breakdown and social cost for Woburn-Melrose, Salem-Lynn, Waltham regions under current scenario and optimal scenario by different commuter types

Region	Scenarios	Commuter Categories	Total Demand	Transit	On-demand	Multi-modal	Outside Option	Disutility per Person	Average Expenditure	Average Travel Time	Social Cost	% Social Cost	% Disutility
Woburn-Melrose	Current Scenario	Overall	6806	8.53%	4.17%	1.20%	86.15%	18.29	10.45	0.27	177038	2.76%	5.73%
		No car	369	49.91%	12.30%	5.88%	32.14%	40.51	7.89	0.54			
		Low-income	1077	26.20%	2.43%	1.60%	69.77%	15.78	6.16	0.30			
		High-income	5728	5.20%	4.49%	1.13%	89.22%	18.76	11.26	0.26			
	Optimal Scenario	Overall	6806	11.67%	2.44%	1.99%	83.90%	17.24	9.80	0.27	172158		
		No car	369	61.64%	13.01%	6.51%	18.84%	30.54	6.36	0.56			
		Low-income	1077	33.81%	2.27%	1.98%	61.94%	13.31	5.63	0.32			
		High-income	5728	7.51%	2.47%	1.99%	88.03%	17.98	10.58	0.26			
Salem-Lynn	Current Scenario	Overall	5222	14.54%	5.63%	1.74%	78.09%	21.38	10.39	0.32	146090	3.84%	6.57%
		No car	881	54.13%	12.28%	5.31%	28.27%	40.34	7.81	0.59			
		Low-income	1010	43.50%	2.95%	2.50%	51.05%	21.48	6.02	0.46			
		High-income	4212	7.59%	6.28%	1.56%	84.58%	21.36	11.44	0.29			
	Optimal Scenario	Overall	5222	17.88%	4.34%	2.28%	75.50%	19.98	9.61	0.32	140487		
		No car	881	60.91%	12.84%	5.55%	20.69%	34.17	6.34	0.60			
		Low-income	1010	53.57%	2.72%	2.73%	40.98%	17.45	4.99	0.47			
		High-income	4212	9.32%	4.74%	2.17%	83.78%	20.59	10.71	0.29			
Waltham	Current Scenario	Overall	2137	10.45%	4.78%	1.20%	83.58%	19.09	9.78	0.25	54980	7.16%	10.09%
		No car	283	48.98%	11.22%	4.83%	34.98%	44.04	7.85	0.56			
		Low-income	290	44.40%	1.96%	1.77%	51.87%	24.67	6.28	0.47			
		High-income	1847	5.13%	5.22%	1.11%	88.55%	18.22	10.33	0.22			
	Optimal Scenario	Overall	2137	14.77%	7.97%	2.09%	75.17%	17.17	9.08	0.25	51565		
		No car	283	54.14%	17.60%	6.95%	21.32%	30.52	6.42	0.45			
		Low-income	290	49.35%	3.25%	2.70%	44.70%	19.52	5.65	0.38			
		High-income	1847	9.35%	8.71%	2.00%	79.95%	16.93	9.61	0.22			

By analyzing the results for the current scenario, it is evident that inequality exists both across and within regions. Comparing these three different regions, we found that Salem-Lynn region has the highest proportion of low-income community (19%) while Waltham has the lowest percentage of low-income community (14%). Salem-Lynn region also has

the highest number of residents without private vehicles (17%). However, under the current scenario, we find that commuters living in Salem-Lynn region have the highest average disutility cost of $21.38 and the highest average commuting time of 0.32 hours. While commuters in Waltham enjoy the shortest average commuting time of 0.25 hours and the lowest average commuting cost of $9.78. With further inspection of different commuter types within the same region, we can also find out that commuters with no cars suffer from much larger travel disutilities compared to high-income commuters. The disutility of people with no cars are almost twice as that of the high-income commuters in all three regions. The disutility for low-income community is lower than the high-income community in Woburn-Melrose region, but it's mainly due to the fact that the unit time cost for low-income commuters is half of the high-income commuters. Even with much longer commuting times, their commuting disutility may still be lower than that of high-income commuters. We will further investigate this in the extension section. These results confirm the existence of inequality both across and within regions among low-income, high-income and no-car commuters.

After applying the model and solution approach mentioned above, we are able to increase the transit ridership in all three regions. Take Woburn-Melrose region for an example, we are able to increase the transit ridership from 8.53% to 11.67%. The transit ridership of no-car and low-income commuters increase considerably by 11.73% and 7.61% respectively. The overall market shares for multi-modal services also increase from 1.20% to 1.99%. The disutility per person decreases from $18.29 to $17.24. When breaking down the benefits across different commuter types, we find that the commuters with no private vehicles enjoy the largest decrease of travel disutility from $40.51 to $30.54. The low-income commuters also get significant disutility decrease from $15.78 to $13.31 while the disutility for high-income people only reduces slightly by $0.78. It shows that our method will disproportionately benefit the underprivileged community. The average expenditures per person decrease the most among no-car commuters because they are able to switch to

lower-cost fixed-route transit instead of taking ride-hailing services. The results are similar for the other regions with under-privileged community benefiting more from the improved transportation systems.

The comparison of environmental cost for three regions under the current and optimal scenarios are presented in Figure 3.3. In all three regions, the environmental costs decrease compared to the current scenario. It shows that the improved transportation network can help mitigate greenhouse emissions. Waltham has the largest percentage reduction of 4% while Woburn has the largest decrease of $1080 per hour in the morning commute.

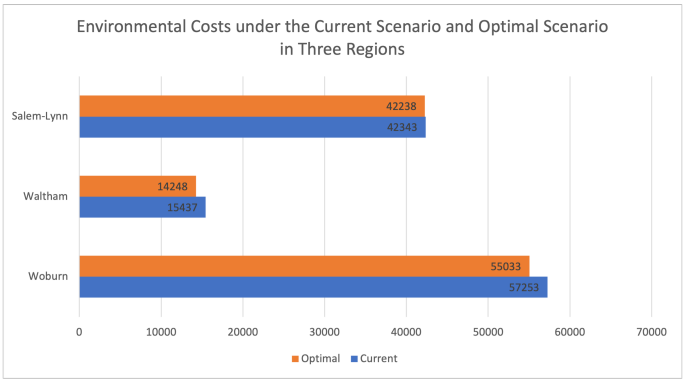


Figure 3.3: Environmental costs for the current scenario and optimal scenario in three regions

The social costs decrease by 2.76%, 3.84%, and 7.16% for Woburn-Melrose region, Salem-Lynn region, and Waltham region respectively. The average travel disutilities also decrease by 5.73%, 6.57%, and 10.09% respectively showing the benefit enjoyed by commuters.

Figure 3.4 shows the disutility decile plot under the current scenario and optimal scenario in Woburn. It shows that for commmuters of all disutility levels have enjoyed decreases in travel disutility. Commuters with higher disutilities generally enjoy higher decreases in travel disutility. However, for commuters with the highest disutilities, the improvement is not significant. It shows that even after the optimization, specific mobility services need to be paid to those commuters with extremely high travel disutilities.

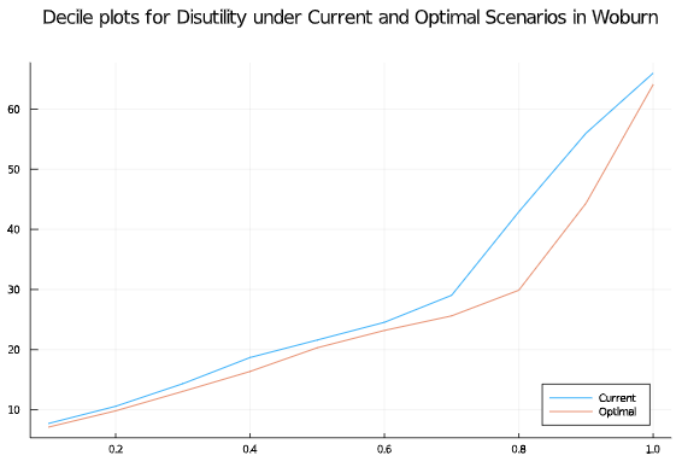


Figure 3.4: Decile plots for Disutility under Current and Optimal Scenarios in Woburn

3.7 Policy Analysis

In this section, we investigate the impact of pricing structure on social welfare and passenger travel disutility. To better integrate the on-demand system with the conventional transit system, multiple cities have provided discounts to passengers if using multi-modal services. In LA, the commuters pay only $3 if they use multi-modal services from selected rail stations during the rush hours (Metro, 2020). We first investigate the impact of different on-demand service pricing levels on commuters' disutility and social welfare. Furthermore, we analyze the impact of discounted multi-modal services on the increase of transit ridership. Here, we evaluate both distance-based pricing scheme and flat pricing scheme for multi-modal services.

3.7.1 On-demand Service Pricing

In the original model, we adopt the prevalent ride-sharing price from Uber's website (Uber, 2021) which has a minimum of price of $7.65, flat price of $2.2, and $1.2 per mile. To analyze the impact of ride-sharing prices on social cost, passenger disutility and profitibility, we apply 0%, 25%, 50%, 75% of the original price.

Figure 3.5 presents the changes of disutilities for all commuter types under different

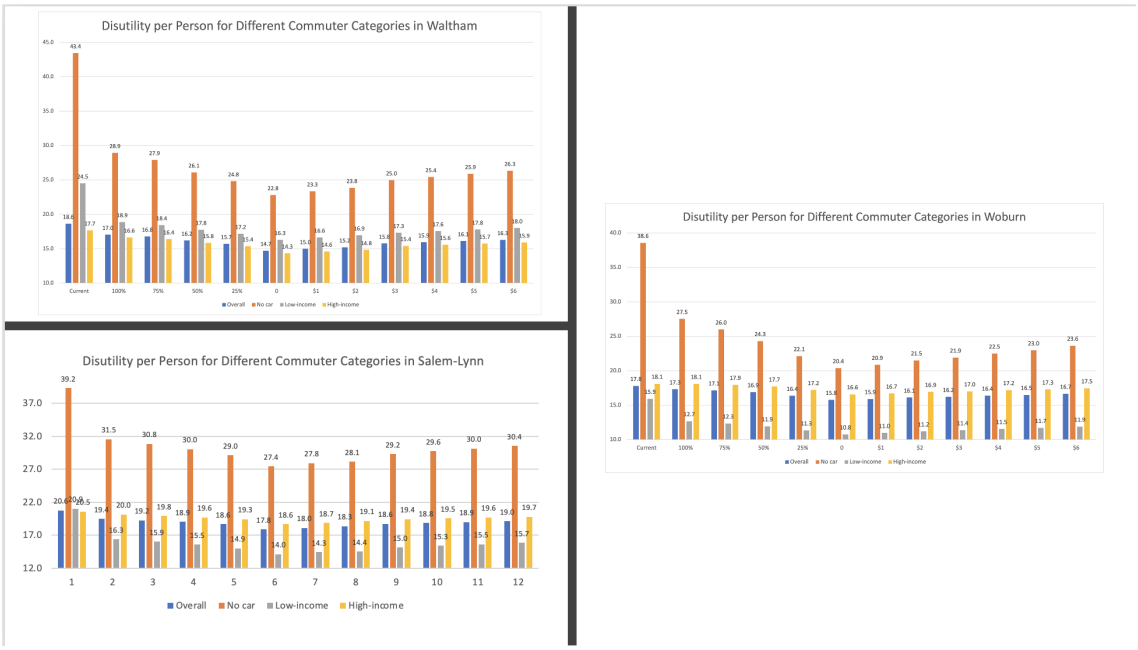


Figure 3.5: Disutility per person for different commuter categories in three regions under different scenarios

ride-sharing prices in all three regions. It is clear that the disutilities decrease when the ride-sharing price decreases for all commuter types. Furthermore, we notice that the on-demand service price has a greater impact on commuters with no private vehicles than the other commuter types. The disutility decreases from \$30.5 to \$22.3 for commuters with no private vehicles in Woburn, while it only decreases by \$2 for low-income commuters and \$1 for high-income commuters. This is mainly due to the fact that on-demand service is the major alternative to public transit for commuters with no private vehicles. When public transit is not easily accessible, a cheaper on-demand service will significantly lessen commuting disutility.

Figure 3.6 further demonstrates the relationship between social welfare and on-demand service price. When the ride-sharing fares go all the way to \$0, the disutility per person decreases more than \$1 per person in all three regions. The social costs are reduced by 1.2%, 1.6%, and 2% in Woburn-Melrose, Waltham, and Salem-Lynn regions respectively. However, it is interesting to note that when reducing the on-demand service price to 75% of its original price, the social welfare increases slightly in Woburn-Melrose region while there

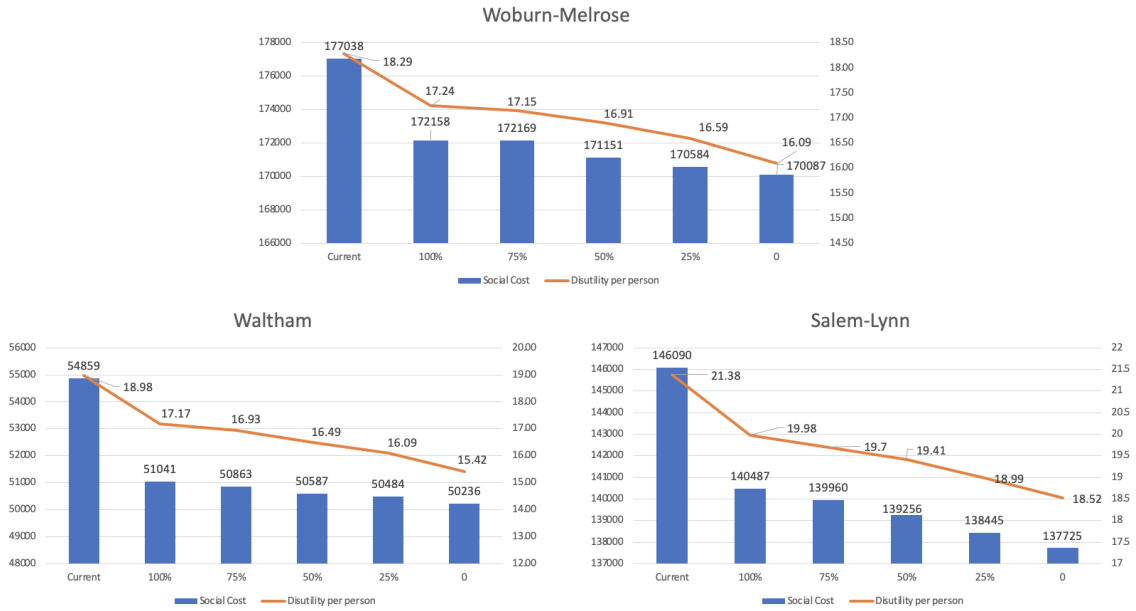


Figure 3.6: Social Costs and Disutilitites in three regions under different scenarios

are no such phenomena in the other two regions. This can be explained by the different compositions of commuters in these three regions. Woburn region has the least percentage of commuters with no private vehicles among the three regions (5.4% compared to 13.2% in Waltham and 16.8% in Salem-Lynn). As shown in Figure 3.5, commuters with no private vehicles benefit the most from the cheaper on-demand services. However, the operational costs for on-demand services will increase due to the increased demand. Due to the small percentage of commuters with no private vehicles in Woburn region, the reduction of travel disutility cannot compensate for the increase in operational costs. Therefore, the overall social cost increases slightly when reducing the on-demand service price by 25% in Woburn. When we further reduce the on-demand service price, the reduction of travel disutility starts to outweigh the increase in operational costs which drives the systemwide cost downwards.

Another interesting finding is that the environmental cost does not follow a monotonic pattern with the price of the on-demand service price. The decrease in on-demand service price can lead to fewer private vehicles on the road but also result in more on-demand vehicles in the system. In Figure 3.7, the environmental costs follow dramatically different pat-

terns in the three different regions. In Woburn-Melrose, the environmental cost decreases with the price except when free on-demand service is offered. In Salem-Lynn region, the environmental cost increases when the on-demand service price is reduced to 75% and decreases afterwards while the environmental cost fluctuates irregularly with the on-demand service price in Waltham. It indicates the complicated impact of on-demand services on the environment. Depending on the commuter types, commuting route patterns, a subsidized on-demand system may lead to more environmental damage than driving private vehicles. The results coincide with the recent literature questioning the positive impact of on-demand services on reducing congestion in the city center. City planners should take demographic information and travel patterns into account when regulating on-demand services.

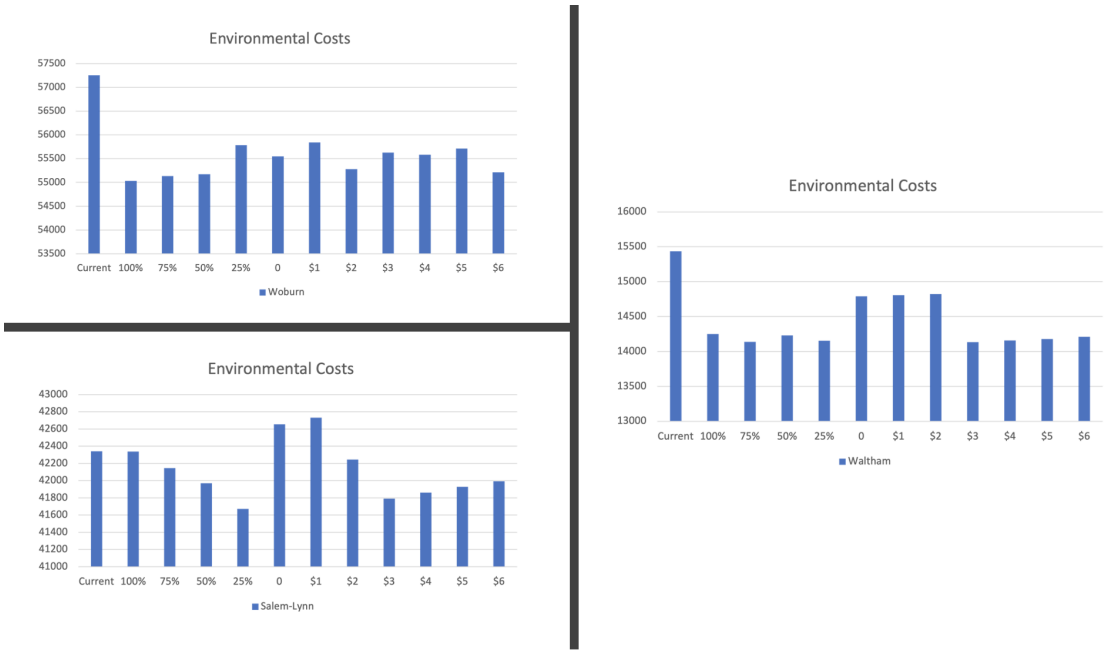


Figure 3.7: Environmental costs in three regions under different scenarios

These results show that even though lower ride-sharing prices can reduce the travel disutility for passengers and overall social cost, the city may suffer from decreasing profits and increasing environmental cost. It is important for the city planner to strike a balance between reducing disutility for commuters and having a healthy financial condition as well as providing a sustainable transportation system.

3.7.2 Multi-model Service Pricing

In order to encourage public transit usage and reduce congestion brought by on-demand services, various cities have carried out policies to promote multi-modal commuting services. One major policy is to provide discounts for passengers adopting multi-modal behavior. Cities like Los Angeles also provide low-price ($3) multi-modal services during rush hours. In this section, we will investigate the impact of different multi-modal service pricing schemes including discounted distance-based pricing and flat pricing. Similar to Section 3.7.1, we analyze scenarios where we charge 0%, 25%, 50%, 75%, and 100% of the original multi-modal price (for the on-demand phase, the fixed-route transit pricing remains the same). We also evaluate the scenarios where flat prices are being charged at $1, $2, $3, $4, $5, and $6.

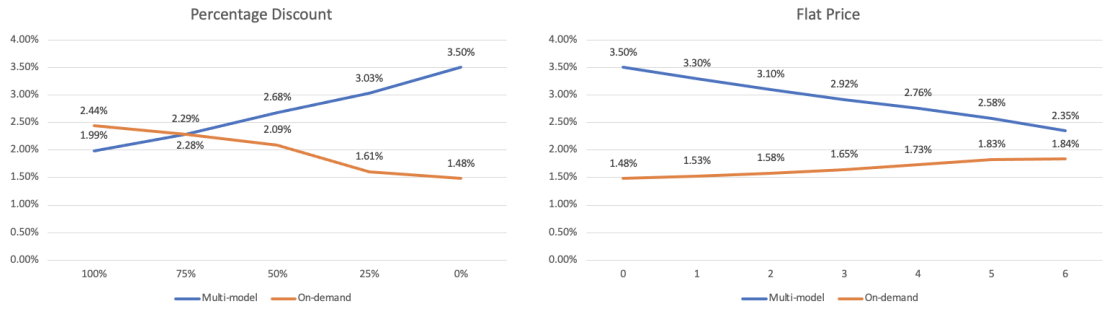


Figure 3.8: Market Shares for On-demand Service and Multi-modal Service in Woburn-Melrose region under Different Pricing Scheme

Figure 3.8 depicts the market share changes for on-demand services and multi-modal services under different pricing schemes. There is a monotonically increasing trend for multi-modal market share when reducing the price to 0. Interestingly, two-thirds of the market share gained by multi-modal came from on-demand service and another one-third came from driving private vehicles. It means commuters are switching from taking on-demand services to using multi-modal services when there is enough financial motivation. It indicates that the policy has a positive effect on increasing transit ridership and reducing congestion caused by on-demand services in the city center.

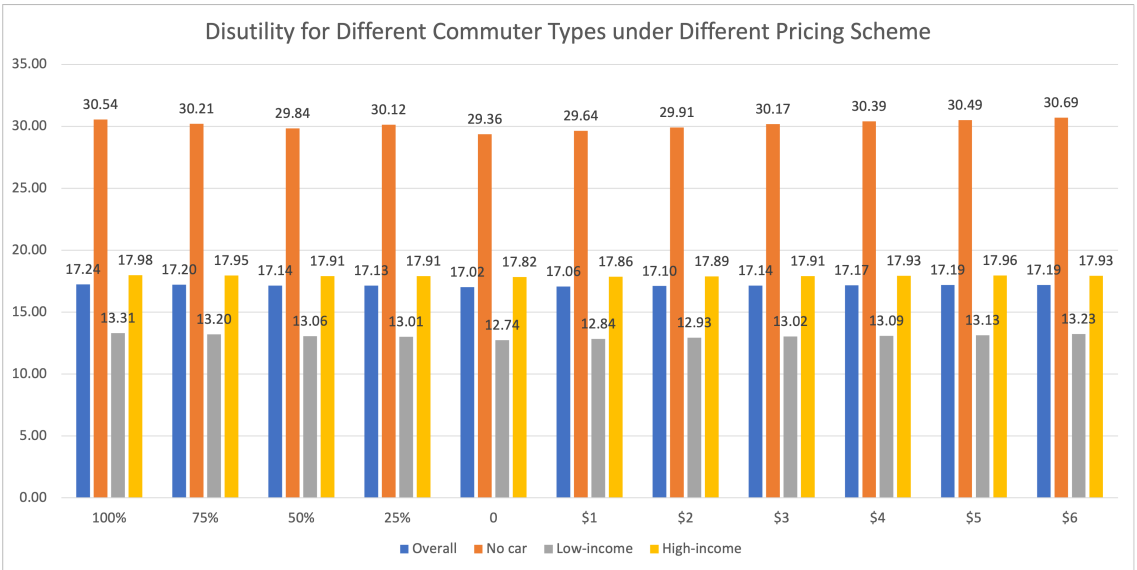


Figure 3.9: Disutility for Different Commuter Types under Different Pricing Schemes

We further investigate which group of commuters benefits the most from cheaper multi-modal service. Figure 3.9 shows the breakdown of disutility for commuters with no private vehicles, low-income commuters, and high-income commuters. The high-income commuters are not affected much by the reduced price since they are not as sensitive to price changes as other groups. Commuters with no private vehicles and low-income commuters benefit the most from the price change. They save \$1.18 and \$0.57 on average respectively when the multi-modal service charges \$0. Furthermore, the environmental cost also reduces when we reduce the multi-modal service prices as shown in Figure 3.10.

We further compare the flat pricing schemes and distance-based pricing schemes. Profitability is an important metric that needs to be considered when deciding multi-modal service price. It would be desirable to increase transit ridership with as little reduction in profitability as possible. Figure 3.11 shows the transit ridership changes and profits under different pricing schemes. We found It is obvious that when more discount is provided, the profits will decrease. However, the overall transit ridership will increase accordingly. In order to balance the trade-off between profitability and transit ridership, we compare the impact of flat pricing schemes and distance-based pricing schemes. Looking at 50%

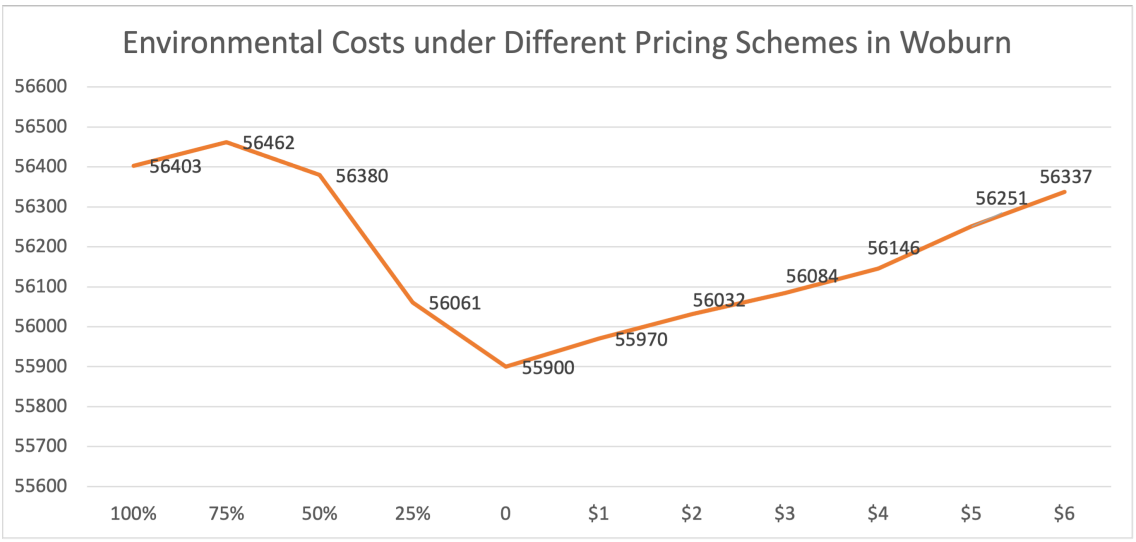


Figure 3.10: Environmental Costs under different Pricing Schemes in Woburn-Melrose

distance-based price and $5 flat price, 50% distance-based price provides both higher profits and higher transit ridership compared to the scenario of $5 flat price. Therefore, in the case of Woburn-Melrose, the distance-based price would be preferable to the flat pricing scheme.

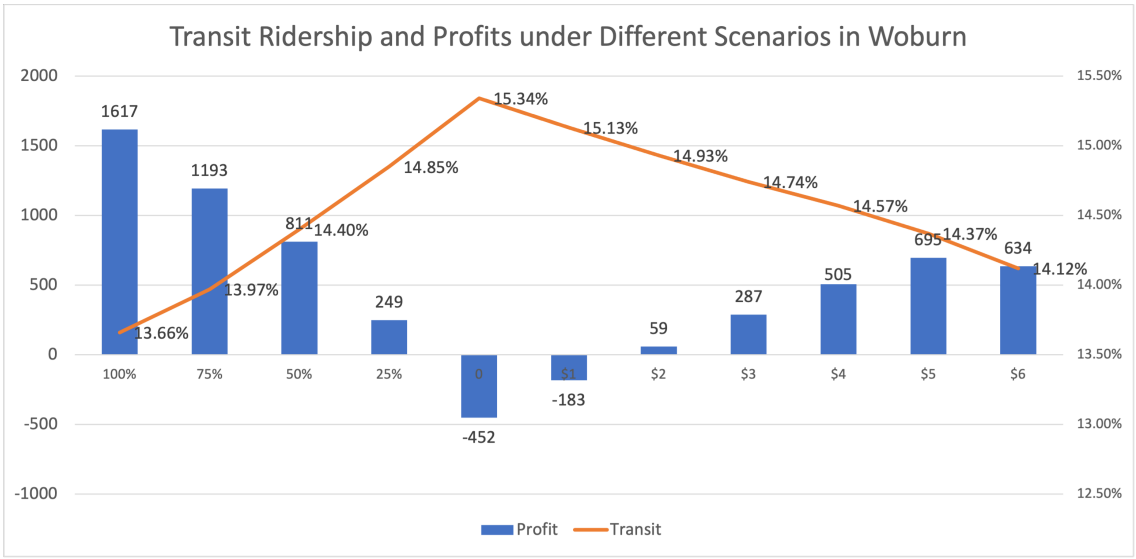


Figure 3.11: Transit Ridership and Profits under Different Pricing Schemes in Woburn-Melrose

3.8 Conclusion

Transit desert has been a severe problem in the US, not only affecting the mobility of daily commuters but also exacerbating the inequality problem for low-income people and marginalized communities. Since this problem cannot be simply addressed by the expansion of the current transit network, we consider developing an integrated transportation system that includes public transit and ride-sharing services to tackle this problem. This paper aims to address the transit desert problem by jointly optimizing the transit network, its corresponding frequencies and on-demand service fleet size considering endogenous passenger demand. We formulate a Mixed-integer non-linear programming model to tackle the problem. A two-step heuristic method is developed to tackle the large-scale MINLP modal. In the first step, we decide on the transit network. In the second step, we determine the transit frequencies and on-demand fleet size.

Case study results in the greater Boston area demonstrate that the optimized transportation system can double the transit ridership compared to the current scenario in Waltham and Woburn regions. The overall social costs are reduced by 2.76%, 3.84%, and 6.96% in Woburn, Salem, and Waltham respectively. The commuter travel disutility also decreases by 5.73%, 6.57%, and 9.56% respectively. We further investigated the impact of ride-sharing pricing strategies on social welfare, disutility, profitability, and environmental cost. We find that social cost and disutility increase with ride-sharing price. A cheaper on-demand service benefits the commuters with no private vehicles the most. However, it has a more complicated effect on environmental costs. The city planner needs more take specific travel patterns into consideration when making policies.

Chapter 4

Bikes and Buses: A Heuristic Adaptive Discretization Scheme for Multimodal Network Design

4.1 Introduction

The recent COVID-19 pandemic posed severe challenges to the current urban transportation system but also demonstrated the importance of building more integrated and sustainable systems. The public transportation system experienced a significant reduction in ridership during the pandemic. When the pandemic struck the US in March 2020, overall transit ridership decreased dramatically. Subways in multiple cities were entirely shut down (Los Angeles Times, 2020; WMATA, 2020). To observe the social distancing regulations, buses suffered a major capacity drop of 60%–90% (Gkiotsalitis and Cats, 2020; Heraldo, 2020; Walawalkar, 2020). That, combined with telework policies, led to a significantly reduced bus ridership for most operators. With social distancing rules and work-from-home arrangements, the overall US transit ridership dropped by nearly 80% in April 2020 and remained more than 60% below the 2019 levels through the rest of the year, as shown in

Figure 4.1 (**?**).

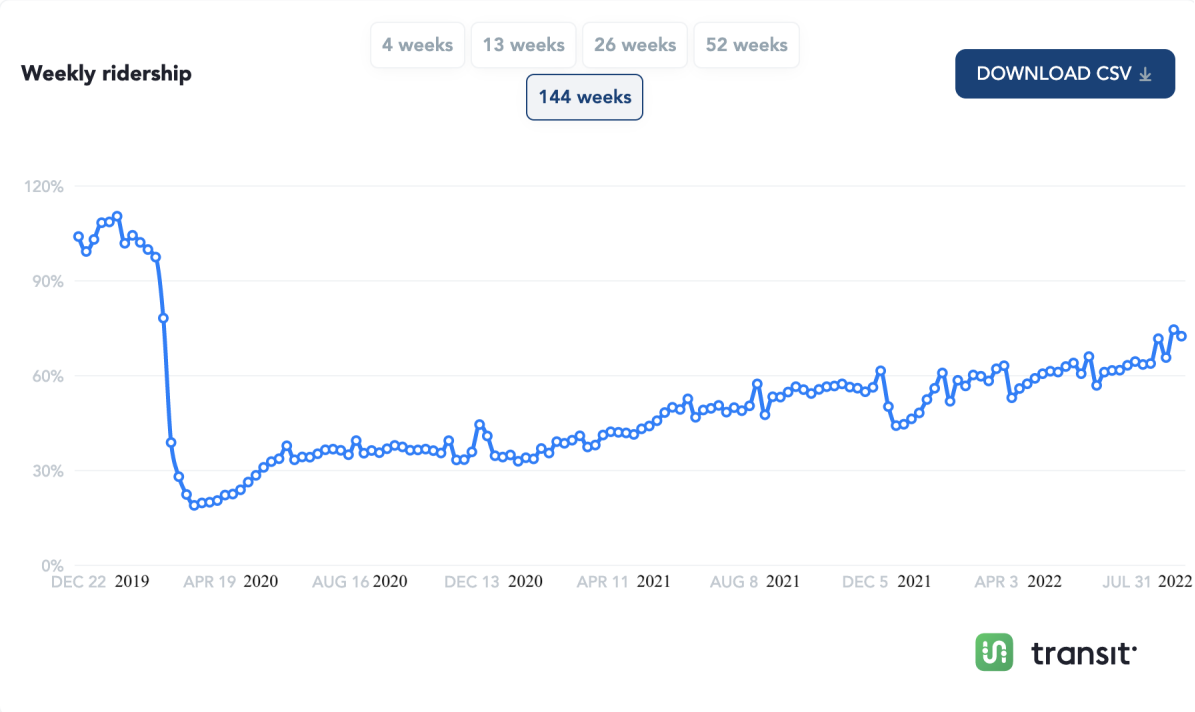


Figure 4.1: Transit Ridership changes from Dec 2019 to Jul 2022. Source: https://transitapp.com/apta

However, the pandemic also led to increased use of micromobility services like bikes and scooters. The percentage of people commuting using bikes and scooters increased from 4.9% to 11.2% according to Association for Commuter Transportation (de Palma et al., 2022). The sale of bikes rocketed in the summer of 2021. The biking activities increased by 28% in the U.S., 45% in Germany, and 82% in the UK from March 2020 to May 2020 (Lindsey et al., 2021). Bluebikes increased their total number of biking stations from 325 to over 400 in the greater Boston area in 2020. Other micromobility also expanded the services significantly during the pandemic. The total scooter fleet in Chicago ballooned from 2,500 total vehicles in 2019, to 10,000 total devices in 2020 (Greenfield and Cobbs, 2020). Lime decided to partner with a commuting benefits provider to prepare for the post-pandemic commuting boom (Holder, 2021).

Shared micromobility has the potential to be an important component of a sustainable

public transportation system to counter climate change. Globally, about a quarter of the total greenhouse gas emissions derives from the transportation sector (Creutzig et al., 2015), and 57% of those are from light-duty vehicles (as compared to only 8% from aircraft). Developing a sustainable transportation system is essential to mitigating climate change, and it is also desirable from a health, economic, and natural resource perspective (Lovelace et al., 2011). For example, a shared e-scooter only generates 75g CO2eq/pkt (CO_2 equivalent per passenger kilometer) while compared to 230g CO2eq/pkt from a gasoline-powered sedan (de Bortoli, 2021). In a pandemic and post-pandemic world, shared micromobility may also allow commuters to adhere to social distancing guidelines. Moreover, it can provide mobility to essential workers without private vehicles when the public transit service level is significantly reduced.

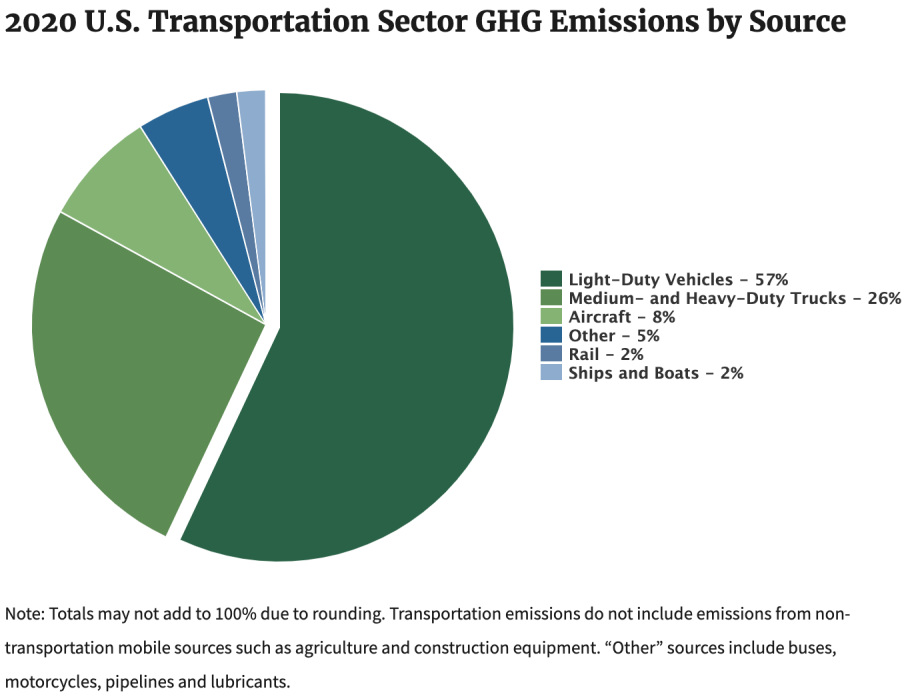


Figure 4.2: 2020 U.S. Transportation Sector GHG Emissions by Source. Source: https://www.epa.gov/greenvehicles/fast-facts-transportation-greenhouse-gas-emissions

Given these potential benefits of micromobility services, local governments have invested enormously in biking infrastructures taking advantage of the empty roads during the

pandemic. Like automobiles, traveling via bikes and scooters also requires a significant infrastructure investment like designated bike lanes and biking racks (Daraei et al., 2021). Philadelphia closed a large 4.4-mile road segment to motor vehicles to create more space for cyclists and pedestrians (LoBasso, 2021). Mexico City proposed plans for 130 kilometers of temporary bike infrastructure providing an alternative to people without private vehicles (COVID Mobility Works, 2020). Berlin created temporary pop-up bike lanes at the height of the pandemic and these pop-up bike lanes are converted to permanent bike lanes with an increasing number of bikers (Kallgren, 2021). Oakland, Minneapolis, Denver, Louisville, Vancouver, and New York City have all implemented similar measures as well (Schwedhelm et al., 2020).

Despite all these benefits, micromobility is not attractive for longer-distance travel, and a significant portion of travelers are still reluctant to adopt micromobility. Dill and McNeil (2013) found that about 30% of adults in Portland, Oregon would never see cycling as an option in any circumstances. With these limitations, it is clear that micromobility is not a viable replacement for commutes during larger transit vehicles like buses and trains. Thus sole dependence on micromobility services won't be sufficient for a more sustainable post-pandemic urban transportation system. However, better integration of micromobility services with the public transit system could encourage commuters to more sustainable alternatives to using private vehicles (Oeschger et al., 2020; Ju, 2022). In many cases, micromobility can serve as the first-mile and last-mile solution by offering flexibility and efficient door-to-door accessibility, while the traditional transit vehicles and routes, such as buses and trains, can take advantage of their higher speeds and greater spatial reach to serve as the mid-mile solutions (Kager et al., 2016).

Both micromobility services, as well as traditional public transit systems, have been studied extensively. There are numerous studies focusing solely on optimizing micromobility systems or solely on optimizing transit systems without considering their integration. The co-design of public transit and micromobility network is the focus of this study, in

which we optimize both transit frequencies and bike-sharing decisions.

The first contribution of this paper is formulating an integrated model to optimize the transit system and the bike-sharing system considering endogenous demand. We optimize the frequency for each transit line, the set of bike-sharing station locations, the number of docks at each station, as well as operational decisions like rebalancing clusters and the number of bikes deployed, loaded and unloaded at each station during rebalancing.

The second contribution is to extend the adaptive discretization method developed by Wang et al. (2022). to multidimensional functions as well as develop heuristics to reduce the runtime and memory usage. Compared to the study of Wang et al. (2022), our problem involves higher dimensional non-linear non-convex constraints. Along with bigger case study sizes, it makes the problem intractable using the original adaptive discretization method. We develop our own heuristic method combining three key ideas: (1) adaptive discretization method, (2) coordinate descent, and (3) anchor point parsimony. For coordinate descent, we separate the decision variables into three categories: A) transit frequencies, B) bike-sharing service levels, and C) all remaining decision variables which include bike station locations, number of docks and bikes at each station, and rebalancing decisions. Our iterative approach first optimizes decision variables in categories B and C by fixing decision variables in category A, and then optimizes those in categories A and C while fixing those in B, until convergence. By applying coordinate descent, we are able to reduce the model size significantly for each iteration. The complexity of the model increases with the number of anchor points added at each iteration when implementing the original adaptive discretization method. By strategically adding and removing anchor points in a parsimonious way, we can significantly reduce the runtime and memory usage while only minimally compromising the optimality of our solution.

The third contribution is that we demonstrate that our algorithm can provide high-quality solutions in a reasonable amount of time by testing it in the city of Boston, Massachusetts. By applying our algorithm to five case studies with increasing sizes within

Boston, we demonstrate the scalability of our algorithm, where all benchmark approaches either have enormous runtimes or fail due to large memory usage. We also compare our algorithm with the original adaptive discretization method which provides an optimality guarantee. We found that the optimality gap between our solution and the true optimal result is within 0.5% but the runtime is 95% smaller. It further demonstrates the superiority of our method when dealing with large-scale, multi-dimensional non-convex relationships within a mixed-integer program. When compared to the current system, our network design solutions reduce the overall travel distance by 4.5% by co-designing the bike-sharing system and public transit system.

4.2 Related Literature

There are three streams of literature related to our work, which we will now discuss in detail. The first stream focuses on bike-sharing system operational level decisions which mainly involve rebalancing clustering and routing decisions. The second stream of literature investigates the optimal bike-sharing strategic level decision like bike station locations and number of docks at each station. In this stream of studies, researchers also have to incorporate operational-level decisions when making strategic-level decisions. The third stream focuses on multi-modal transportation system design. This stream of literature is not limited to the integrated design of the bike-sharing system and transit system. It also includes the co-design of the public transit system with other emerging technologies like ride-hailing services.

4.2.1 Bike-sharing rebalancing

One of the major bike-sharing operational costs is to rebalance the bikes to other bike stations ensuring their availability across the system. The bike rebalancing problem is a $\mathcal{NP}$-hard problem, of which traveling salesman is a special case (Bruck et al., 2019). Zhang

et al. (2020) has developed an adaptive Tabu Search algorithm based on local search to solve this problem. An iterated local search and a route elimination operator were introduced to reduce the number of routes. Shui and Szeto (2018) applied a hybrid rolling horizon artificial bee colony approach to minimize total unmet demand and CO2 emissions. Due to the complexity of the problem, (Schuijbroek et al., 2017) introduced a two-step method where the service level and rebalancing clusters were determined in the first stage, and vehicle routes were decided in the second stage. Similarly, (Forma et al., 2015) developed a three-step algorithm. In the first step, the stations were clustered. In the second step, the route across clusters was determined. In the final step, the stations within the clusters were connected to make one single route for the rebalancing vehicle. Studies have also explored strategies in which the rebalancing vehicle can visit a single station multiple times. There are certain large, central bike stations in which the demand of which cannot be satisfied in a single visit. Bulhões et al. (2018) presented an iterated local search heuristic for the static bicycle relocation problem with multiple vehicles and visits problem. Branch-and-cut algorithm was introduced to tackle the exponential number of constraints. Pal and Zhang (2017) also investigated the possibility of multiple visits but focuses on the free-floating bike-sharing system. A hybrid nested large neighborhood search with a variable neighborhood descent algorithm was presented to solve the mixed integer linear program formulation. Considering the stochastic nature of the demand, (Dell'Amico et al., 2018) developed stochastic programming models and used L-shaped and branch-and-cut methods to solve them. They also developed a heuristic based on the correlation of demands between stations to find good-quality upper bounds.

Our paper incorporates bike-rebalancing operations in our modeling. However, our main focus is co-design the transit-bikesharing system. Instead of solely focusing on bike rebalancing, we also have to optimize the bike station location, number of docks as well as transit frequencies.

4.2.2 Bike station decision

The most commonly used location–allocation modeling approach is maximizing coverage (García-Palomares et al., 2012). Frade and Ribeiro (2015) optimized bike-station locations and the number of bikes at each station to maximize the demand covered within the available budget. Mix et al. (2022) jointly modeled the demand for bike-sharing trips and optimized the location of stations in the system. They first estimated the demand for the bike-sharing system using multiple regressions. Then, maximum demand coverage models are developed to allocate the BSS stations. Similarly, García-Palomares et al. (2012) estimated the spatial distribution of potential demand using GIS data and determine the station locations and their corresponding capacities by maximizing demand coverage. However, Caggiani et al. (2020) found that maximizing coverage without considering equality led to an unequal distribution of accessibility among the population, which would produce discrimination between different groups. Therefore, they proposed a model to minimize inequality in the system while maintaining a certain level of accessibility. The studies mentioned above didn't consider the rebalancing problem when determining the bike station location and capacity, but in reality, the effectiveness of rebalancing has a significant impact on the optimal station capacity and system operation. Fu et al. (2022) investigated the optimal bike-sharing system station location problem with rebalancing vehicle service design. We present a two-stage robust optimization model with a demand-related uncertainty set. In the first stage, the station locations, initial bike inventory, and rebalancing clusters were determined. In the second stage, the rebalancing operation was optimized to maximize profit. In the first stage, an approximate maximum travel distance for each rebalancing cluster using maximum spanning star, which is also applied by Schuijbroek et al. (2017). We adopt this approximation method in our model to simplify the vehicle routing formulation.

Most of the papers in this stream did not consider incorporating rebalancing operations in the planning stage. Besides that, we also consider the stochasticity regarding the orders

of pick-ups and drop-offs, which will be explained in detail in Section 4.3.2.5 and 4.3.2.6. Moreover, we provide an integrated optimization framework to optimize the transit system and bike-sharing system. Therefore, we consider endogenous demand and passenger choice while most of the studies here assumed exogenous demand.

4.2.3 Multi-modal transportation system design

The papers that are most relevant to our study are those investigating multi-modal transportation system design. These papers did not necessarily focus on optimizing the bike-sharing system, but they provided insights into how other emerging transportation modes can be effectively integrated with the public transit system. Yan et al. (2021) conducted a spatiotemporal analysis of the interaction between e-scooters and public transit in Washington DC. They found that e-scooter usage both competes and promotes public transit usage. About 10% of e-scooter rides were taken to get connected to the public transportation system. There is a trade-off between price and time when passengers are deciding whether to take an e-scooter or public transit. Périvier et al. (2021) focused on smart transit systems where a set of cars are used as feeders to high-capacity fixed-route transit lines. They developed an optimization model and applied approximation heuristic to determine the bus routes and frequencies. Shen et al. (2018) and Wei et al. (2021) investigate how to better integrate the ride-hailing/ride-sharing services with public transit system. Shen et al. (2018) proposed to preserve high-demand bus routes and replace low-demand bus routes with shared AVs as an alternative. The agent-based simulation was applied to evaluate the scenarios with different fleet sizes and ride-sharing preferences. The results demonstrated the benefits of the integrated system. However, they didn't provide a framework to obtain the optimal integrated design. Wei et al. (2021) optimize the transit schedules considering road congestions and endogenous passenger choice. They found that the optimized schedule managed to reduce the systemwide cost by 0.4% - 3%. Different from our paper, they didn't consider the possibility of people taking multi-modal travel mode (using ride-

hailing as first-mile and last-mile solution and taking public transit in between). The lack of long-distance travel capability is the key difference between bike-sharing and ride-hailing system. Multi-modal travel mode is an important component in our paper compared to Wei et al. (2021). Moreover, bike-sharing operational decisions involve a different set of problems like bike rebalancing and station location decisions.

4.3 Model

In this section, we present both the original model and the discretized model to co-optimize the transit system and bike-sharing system.

4.3.1 Assumptions

This paper considers morning commutes where four different traveling modes are considered: public transit, bike-sharing services, multi-modal services (commuters taking both fixed-route transit and bike-sharing services), and other traveling options (including private vehicle, biking, not travel etc.). We consider endogenous travel demand where the commuters determine the travel mode based on their corresponding attractiveness.

We co-design the transit system and bike-sharing system from the perspective of a city planner. We are able to decide both transit frequencies and biking decisions. Our objective is to maximize the saved distance not traveled by driving by optimizing bus frequencies, bike-sharing stations, the number of docks at each station, and rebalancing strategies. We don't consider congestion in this study.

The model involves a number of assumptions:

1. Passenger mode choice is captured using a sales-based linear programming model from Gallego et al. (2015) model. The market share of a certain traveling mode increases with its attractiveness which is a function of traveling time, walking time, waiting time, fares, and parking fee.

2. Passengers choose the shortest-distance route when traveling.

3. There is at most one transfer when the commuter is taking fixed-route transit.

4. Commuters are able to enter the transit system with their bikes.

5. The model considers exogenous transit fare and on-demand service pricing. We applied the prevalent flat pricing for both transit and bike-sharing system.

4.3.2 Original Model

4.3.2.1 Objective Function

The objective function of the optimization model is to maximize the total saved traveling distance by switching from driving to other modes of transportation like fixed-route transit, bike-sharing services, and multi-modal services. $\mathcal{T}$ is the set of time periods being considered during the day. $\mathcal{OD}$ is the set of OD pairs. In the Objective Function (4.1), D_{tod} is the total demand for OD pair $od \in \mathcal{OD}$ at time period $t \in \mathcal{T}$. L^D_{od} refers to the travel distance in OD pair od when driving private vehicles. The objective function is equivalent to $\sum_{t \in T, od \in OD}(1 - m^D_{tod}) * D_{tod} * L^D_{od}$. The decision variables are defined below:

m^T_{tod}	the market share for taking fixed-route transit for OD pair $od \in \mathcal{OD}$ during the time period $t \in \mathcal{T}$
m^B_{tod}	the market share for taking bike-sharing services for OD pair $od \in \mathcal{OD}$ during the time period $t \in \mathcal{T}$
m^M_{tod}	the market share for taking multi-modal services for OD pair $od \in \mathcal{OD}$ during the time period $t \in \mathcal{T}$
m^D_{tod}	the market share for driving for OD pair $od \in \mathcal{OD}$ during the time period $t \in \mathcal{T}$

$$Max \quad \sum_{t \in \mathcal{T}} \sum_{od \in \mathcal{OD}} (m^T_{tod} + m^B_{tod} + m^M_{tod}) D_{tod} L^D_{od} \tag{4.1}$$

4.3.2.2 Budget Constraints

Two budget constraints are included to prevent excessive investment in the transit and bike-sharing system. We set the transit budget and biking budget separately since they are usually operated by two departments in most cities.

Constraints (4.2) and (4.3) compute the number of bus vehicles and heavy rail vehicles that are needed to operate under certain frequency levels. Let $\mathcal{L}$ be the set of fixed transit lines. $\mathcal{L}^1 \subseteq \mathcal{L}$ is the set of bus lines whose frequencies are being optimized. $\mathcal{L}^2 \subseteq \mathcal{L}$ is the set of heavy rail lines whose frequencies we are optimizing. By definition, $\mathcal{L}^1 \cup \mathcal{L}^2 = \mathcal{L}$ and $\mathcal{L}^1 \cap \mathcal{L}^2 = \emptyset$. L_l is the transit route length of transit line l. V^{T1} and V^{T2} are the travel speed for buses and heavy rail. f_{tl} is the frequency of transit line $l \in \mathcal{L}$ during time period $t \in \mathcal{T}$.

Constraint (4.4) is the transit budget constraint. The two terms on the left-hand side are the operating cost for buses and heavy rail (T). C_{op}^{F1} and C_{op}^{F2} are the operating cost per unit time for one bus and one heavy rail. B^F and B^B are the operating budgets for the transit system and bike-sharing system respectively.

Constraint (4.5) is the budget constraint for the bike-sharing system. The operating cost involves two major components: (1) the maintenance cost of bike stations and bikes, and (2) the rebalancing cost due to unbalanced demand. $\mathcal{J}$ is the set of candidate regions for bike stations. $\mathcal{G}$ is the set of rebalancing clusters where one rebalancing vehicle traverses all stations picking up and dropping off bikes. C_s^B is the maintenance cost for each bike, and C_c^B is the maintenance cost for each dock at the station. C^{BR} is the cost of picking up each bike during the rebalancing process. C^{BD} is the cost for traveling unit distance per rebalancing vehicle.

$\sum_{t\in\mathcal{T},j\in\mathcal{J}}(C_c^B * c_{tj}^B + C_s^B * s_{tj}^B) + \sum_{t\in\mathcal{T},g\in\mathcal{G}} C_s^B * e_{tg}^B$ represents the maintenance cost. The first term is the maintenance cost for the docks at the bike station. The second and third term represent the maintenance cost for all bikes in the system. s_{tj}^B is the number of bikes that need to be deployed at location j at time t and e_{tg}^B is the number of extra bikes

needed for each rebalancing cluster cl at time t since some clusters may have a deficit of bikes within the rebalancing cluster. Therefore, the total number of bikes needed for the system at time t is $\sum_{j\in\mathcal{J}} s^B_{tj} + \sum_{g\in\mathcal{G}} e^B_{tg}$.

$\sum_{t\in\mathcal{T},j\in\mathcal{J}} C^{BR} * r^B_{tj} + \sum_{t\in\mathcal{T},g\in\mathcal{G}} C^{BR} e^B_{tg} + \sum_{t\in\mathcal{T},g\in\mathcal{G}} C^{BD} d^B_g$ is the rebalancing cost. The first two terms represent the cost of picking up and dropping off bikes when rebalancing. The first term is the cost of picking up and dropping off bikes at the bike stations and the second term is the cost of picking up extra bikes in the case where there is a bike deficit within the cluster. The third term calculates the rebalancing cost based on the vehicle traveling distance.

The decision variables are listed below:

f_{tl}	frequency for transit line $l \in \mathcal{L}$ during the time period $t \in \mathcal{T}$
k^T_{tl}	fleet size for transit line $l \in \mathcal{L}$ during the time period $t \in \mathcal{T}$
s^B_{tj}	number of bikes deployed at location $j \in \mathcal{J}$ at the start of time period $t \in \mathcal{T}$.
c^B_j	number of docks at location $j \in \mathcal{J}$.
r^B_{tj}	number of bikes needed to be picked up or dropped off at location $j \in \mathcal{J}$ at time period $t \in \mathcal{T}$ for rebalancing.
e^B_{tg}	number of extra bikes needed due to demand deficit within cluster $g \in \mathcal{G}$ at time period $t \in \mathcal{T}$ for rebalancing.
d^B_g	distance traveled to traverse all stations within cluster $g \in \mathcal{G}$ for rebalancing.

$$k^T_{tl} \geq \frac{L_l f^T_{tl}}{V^{T1}} \quad \forall t \in T, l \in \mathcal{L}^1 \tag{4.2}$$

$$k^T_{tl} \geq \frac{L_l f^T_{tl}}{V^{T2}} \quad \forall t \in T, l \in \mathcal{L}^2 \tag{4.3}$$

$$\sum_{t\in\mathcal{T},l\in\mathcal{L}^1} C^{T1}_{op} * k^T_{tl} + \sum_{t\in\mathcal{T},l\in\mathcal{L}^2} C^{T2}_{op} * k^T_{tl} \leq B^T \tag{4.4}$$

$$\sum_{t\in\mathcal{T},j\in\mathcal{J}} (C^B_s * s^B_{tj} + C^B_c * c^B_{tj} + C^{BR} * r^B_{tj}) + \sum_{t\in\mathcal{T},g\in\mathcal{G}} (C^B_s * e^B_{tg} + C^{BD} d^B_g + C^{BR} e^B_{tg}) \leq B^B \tag{4.5}$$

4.3.2.3 Path Selection Constraints

Let $\mathcal{P}^M$ and $\mathcal{P}^M_{od}$ represent the set of multi-modal paths and set of multi-modal paths for the OD pair od. $\mathcal{P}^M_{od} \subseteq \mathcal{P}^M$ is a subset of multi-modal path set that serves od pair od. We first generate all possible paths for multi-modal service for each OD pair assuming bike-sharing services are available for all OD pairs. However, only a subset of bike-sharing services will be offered in the final system, we need to identify the available paths and paths taken by the commuters. Therefore, we define the following decision variables:

b_j	=1 if bike-sharing service is available in region $j \in \mathcal{J}$
y^M_p	= 1 if path for multi-modal service $p \in \mathcal{P}^M$ is available for passengers
z^M_p	= 1 if path for multi-modal service $p \in \mathcal{P}^M$ will be taken to serve passengers. $z_{p^M} \leq y_{p^M}$
z^B_{od}	= 1 if $od \in \mathcal{OD}$ can be served using bike-sharing services

One of the major decisions for optimizing bike-sharing system is to decide whether to offer services in one region. The decision variables b_j are created to indicate whether the service is offered in the region j. In the preprocessing stage, we generate all the possible paths to serve each OD pair assuming the bike-sharing service is available in all regions concerned. y^M_p serves as an indicator of whether the path is available. $y^M_p = 1$ when the bike-sharing service is available for both pick-up and drop-off for path p. Similarly, z^B_{od} indicates whether bike-sharing service is available for OD pair $od \in \mathcal{OD}$. j^{BP}_{od} and j^{MP}_p are the bike pickup regions for bike-sharing od-pair $od \in \mathcal{OD}$ and multi-model path $p \in \mathcal{P}^{\mathcal{M}}$. j^{BD}_{od} and j^{MD}_p are the bike drop-off regions for bike-sharing od-pair $od \in \mathcal{OD}$ and multi-model path $p \in \mathcal{P}^{\mathcal{M}}$. Constraints (4.6) and (4.7) ensure that z^B_{od} and y^M_p equal to 1 if the bike-sharing service is available to serve the OD pair (or path).

z^M_p indicates whether the passenger will choose path p. We assume the passenger will choose the shortest path among the available paths for the OD pair. Constraints (4.8)-(4.11) are to ensure that z^M_p is correctly defined. L^M_p is a parameter for the length of multi-modal path p. In constraints (4.8), when $L^M_{p_1} - L^M_{p_2} > 0$ and $y^M_{p2} = 1$ which means the length

of path p_1 is longer than path p_2 and path p_2 is available, $z^M_{p_1}$ on the right-hand side will be forced to be 0. It ensures that the passengers will not choose a longer path when a shorter path is available. Constraints (4.9) guarantee that at most one path will be chosen. Constraints (4.10) ensure that passengers will choose at least 1 path if paths are available for the OD pair. Constraints (4.11) make sure that only available paths will be traveled.

$$z^B_{od} = b_{j^{BP}_{od}} * b_{j^{BD}_{od}} \quad \forall od \in \mathcal{OD} \tag{4.6}$$

$$y^M_p = b_{j^{MP}_p} * b_{j^{MD}_p} \quad \forall p \in \mathcal{P}^M \tag{4.7}$$

$$(L^M_{p_1} - L^M_{p_2}) * y^M_{p_2} \leq (1 - z^M_{p_1}) * M \quad \forall p_1, p_2 \in \mathcal{P}^M_{od} \tag{4.8}$$

$$\sum_{p \in \mathcal{P}^M_{od}} z^M_p \leq 1 \quad \forall od \in \mathcal{OD} \tag{4.9}$$

$$M * \sum_{p \in \mathcal{P}^M_{od}} z^M_p \geq \sum_{p \in \mathcal{P}^M_{od}} z^M_p \quad \forall od \in \mathcal{OD} \tag{4.10}$$

$$z^M_p \leq y^M_p \quad \forall p \in \mathcal{P}^M \tag{4.11}$$

Since there are bilinear terms in constraints (4.6) and (4.7), constraints (4.12)-(4.17) are linear constraints that can replace these two constraints.

$$z^B_{od} \leq b_{j^{BP}_{od}} \quad \forall od \in \mathcal{OD} \tag{4.12}$$

$$z^B_{od} \leq b_{j^{BD}_{od}} \quad \forall od \in \mathcal{OD} \tag{4.13}$$

$$z^B_{od} \geq b_{j^{BP}_{od}} + b_{j^{BD}_{od}} - 1 \quad \forall od \in \mathcal{OD} \tag{4.14}$$

$$y^M_p \leq b_{j^{MP}_{od}} \quad \forall p \in \mathcal{P}^M \tag{4.15}$$

$$y^M_p \leq b_{j^{MD}_{od}} \quad \forall p \in \mathcal{P}^M \tag{4.16}$$

$$y^M_p \geq b_{j^{MP}_{od}} + b_{j^{MD}_{od}} - 1 \quad \forall p \in \mathcal{P}^M \tag{4.17}$$

4.3.2.4 Waiting Time Constraints

Waiting time constraints compute the transit waiting time for passengers taking fixed-route transit and multi-modal services. Assuming that passenger's waiting time for transit is a uniform distribution with lower bound 0 and upper bound of the headway, the average waiting time for the passenger is $\frac{1}{2*frequency}$. Since passengers may make a transfer to another transit line, the total waiting time is the sum of the average waiting times for two segments of the trip. The decision variables are listed below:

w^T_{tod}	waiting time when taking fixed-route transit for OD pair $od \in \mathcal{OD}$ during time period $t \in \mathcal{T}$
w^M_{tod}	waiting time when taking multi-modal services for OD pair $od \in \mathcal{OD}$ during time period $t \in \mathcal{T}$

l^{T1}_{od} and l^{T2}_{od} indicate the transit line taken for the first/second segment of the trip for OD pair od. If no transfer is needed, $l^{T1}_{od} = l^{T2}_{od}$, but F^T_{od} which is a binary parameter indicating whether a transfer is taken will be 0. If a transfer is needed, $F^T_{od} = 1$ and $l^{T1}_{od} \neq l^{T2}_{od}$. Similarly for the multi-modal services, if l^{M1}_p and l^{M2}_p indicate the transit line taken for the first/second segment of the trip for the path p. F^M_p is 1 when a transfer is needed. Constraints (4.18) and (4.19) calculate the waiting time for fixed-route transit and multi-modal service respectively. However, it is worth noticing that both constraints are non-linear. In Section 4.3.3, we will present the algorithms to eliminate the non-linearity.

$$w^T_{tod} = \frac{1}{2 * f_{tl^{T1}_{od}}} + \frac{F^T_{od}}{2 * f_{tl^{T2}_{od}}} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.18}$$

$$w^M_{tp} = \frac{1}{2 * f_{tl^{M1}_p}} + \frac{F^M_p}{2 * f_{tl^{M2}_p}} \quad \forall t \in \mathcal{T}, p \in \mathcal{P}^M \tag{4.19}$$

4.3.2.5 Bike-sharing Availability Constraints

To maintain user satisfaction, it is essential for the bike-sharing system to operate at a reasonable service level. To make it convenient for passengers to pick up and drop off bikes at the station, a certain number of vacant bikes and docks should be available at the station when there is demand. In the following constraints, we will define the relationship between service level, demand, number of bikes deployed at each station, and number of docks at each station. The decision variables are presented below:

h^p_{tj}	number of passengers to pick up bikes at location $j \in \mathcal{J}$ during time period $t \in \mathcal{T}$
h^d_{tj}	number of passengers to drop off bikes at location $j \in \mathcal{J}$ during time period $t \in \mathcal{T}$
s^B_{tj}	number of bikes to be deployed at location $j \in \mathcal{J}$ at the beginning of time period $t \in \mathcal{T}$
l^p_{tj}	$\in [0, 1]$, service level for bike pick-up at bike station location $j \in \mathcal{J}$ which is the probability that a bike is available for pick up at location j
l^d_{tj}	$\in [0, 1]$, service level for bike drop-off at bike station location $j \in \mathcal{J}$ which is the probability that an empty dock is available for drop off at location j
γ^M_{tp}	$= z^M_p * m^M_{tod_p}$

$\mathcal{OD}^p_j$ and $\mathcal{OD}^d_j$ represent the set of OD pairs that pick up (drop off) bikes at location $j \in \mathcal{J}$. Similarly, $\mathcal{P}^p_j$ and $\mathcal{P}^d_j$ represent the set of multi-modal paths that pick up (drop off) bikes at location $j \in \mathcal{J}$. od_p is a parameter denoting the OD pair that is served by path p. Constraints (4.20) and (4.21) calculate the number of pick-ups and drop-offs at location j. Constraints (4.22) and (4.23) ensure that the docks and bikes deployed at the station can sustain service level l^p_{tj} and l^d_{tj}. Constraints (4.24) compute the number of bikes needed to be loaded or unloaded at each location j. Constraints (4.25) compute the number of extra bikes needed for each cluster cl due to the imbalanced demand. To linearize constraints (4.22)-(4.25), we use a linear approximation method which is detailed in Section 4.3.3.

$$h_{tj}^{p} = \sum_{od \in \mathcal{OD}_j^p} m_{tod}^{B} * D_{tod} + \sum_{p \in \mathcal{PJ}_j^p} m_{tod_p}^{M} * D_{tod_p} * z_p^M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.20}$$

$$h_{tj}^{d} = \sum_{od \in \mathcal{OD}_j^d} m_{tod}^{B} * D_{tod} + \sum_{p \in \mathcal{PJ}_j^d} m_{tod_p}^{M} * D_{tod_p} * z_p^M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.21}$$

$$s_{tj}^{B} = f_p(h_{tj}^{p}, h_{tj}^{d}, l_{tj}^{p}) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.22}$$

$$c_j - s_{tj}^{B} = f_d(h_{tj}^{p}, h_{tj}^{d}, l_{tj}^{d}) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.23}$$

$$r_{tj}^{B} = f_r(s_{tj}^{B}, s_{tj+1}, h_{tj}^{p}, h_{tj}^{d}, l_{tj}^{d}) \quad \forall t \in \mathcal{T}/\{T\}, j \in \mathcal{J} \tag{4.24}$$

$$e_{tg}^{B} = f_{ib}(h_{tj}^{p}, h_{tj}^{d}, l_{tj}^{d}) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.25}$$

$$c_j^{B} \leq M * b_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.26}$$

Since both m_{tp}^{M} and z_p^M are both decision variables, we linearize constraints (4.20) and (4.21) by introducing another set of decision variables γ_{tp}^{M}. Constraints (4.20) and (4.21) are replaced by constraints (4.27) and (4.28). Constraints (4.29)-(4.31) are added to the optimization model to define γ_{tp}^{M}.

$$h_{tj}^{p} = \sum_{od \in \mathcal{OD}_j^p} m_{tod}^{B} * D_{tod} + \sum_{p \in \mathcal{PJ}_j^p} \gamma_{tp}^{M} * D_{tod_p} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.27}$$

$$h_{tj}^{d} = \sum_{od \in \mathcal{OD}_j^d} m_{tod}^{B} * D_{tod} + \sum_{p \in \mathcal{PJ}_j^d} \gamma_{tp}^{M} * D_{tod_p} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.28}$$

$$\gamma_{tp}^{M} \leq z_p^M \quad \forall t \in \mathcal{T}, p \in \mathcal{P}^M \tag{4.29}$$

$$\gamma_{tp}^{M} \leq m_{tod_p}^{M} \quad \forall t \in \mathcal{T}, p \in \mathcal{P}^M \tag{4.30}$$

$$\gamma_{tp}^{M} \geq m_{tod_p}^{M} - (1 - z_p^M) \quad \forall t \in \mathcal{T}, p \in \mathcal{P}^M \tag{4.31}$$

4.3.2.6 Service Level Approximation

To linearize constraints (4.22)-(4.25), we first obtain the distribution of demand deficit during the time period. We assume that the pick-ups and drop-offs follow Poisson distribution. The difference between pick-ups and drop-offs follows Skellam distribution. F_{tj} is a Skellam distribution for location j during time period t. Constraints (4.32) and (4.33) ensure that the pick-up and drop-off service level are at least l^p_{tj} and l^d_{tj}. Constraints (4.34) and (4.35) give a conservative estimate for number of bikes to be (un)loaded onto the rebalancing vehicle and number of extra bikes needed for each cluster.

o_{tj}	=1 when $\alpha^p_0 + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha^p_3 * l^p_{tj} \geq 0$, 0 otherwise.
$l^{p'}_{tj}$	= l^p_{tj} if $b_j = 1$, 0 otherwise.
$l^{d'}_{tj}$	= l^d_{tj} if $b_j = 1$, 0 otherwise.
v_{jg}	= 1 when region $j \in \mathcal{J}$ is in cluster $g \in \mathcal{G}$
q^p_{tj}	= $l^{p'}_{tj}$ if $v_{jg} = 1$, 0 otherwise.
q^d_{tj}	= $l^{d'}_{tj}$ if $v_{jg} = 1$, 0 otherwise.
$h^{p'}_{tjg}$	= h^p_{tj} if $v_{jg} = 1$, 0 otherwise.
$h^{d'}_{tjg}$	= h^d_{tj} if $v_{jg} = 1$, 0 otherwise.
v'_{jg}	= v_{jg} if $b_j = 1$, 0 otherwise.
s'_{tjg}	= s^B_{tj} if $v_{jg} = 1$, 0 otherwise.
n^p_{tjg}	= l^p_{tj} if $v'_{jg} = 1$, 0 otherwise.
n^d_{tjg}	= l^d_{tj} if $v'_{jg} = 1$, 0 otherwise.

$$s^B_{tj} = max(0, F^{-1}_{tj}(l^p_{tj})) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.32}$$

$$s^B_{tj} - c^B_j = min(0, F^{-1}_{tj}(1 - l^d_{tj})) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.33}$$

$$r^B_{tj} = max(s^B_{tj} - s^B_{t+1j} - F^{-1}_{tj}(1 - l^d_{tj}), s^B_{t+1j} - s^B_{tj} + F^{-1}_{tj}(l^p_{tj})) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.34}$$

$$e^B_{tg} = max(0, \sum_{j \in \mathcal{J}} (s^B_{t+1j} - s^B_{tj} + F^{-1}_{tj}(l^p_{tj})) * v_{jg} * b_j) \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \tag{4.35}$$

However, we are not able to compute the exact Skellem distributions beforehand since they are affected by variables h^p_{tj} and h^d_{tj}. Therefore, we generate a large pool of different

h^p_{tj}, h^d_{tj}, and service levels to obtain $F^{-1}_{tj}(l^p_{tj})$ and $F^{-1}_{tj}(1-l^d_{tj})$. Then, the results are fitted with two linear regressions, one for available bikes, the other for available docks. $\vec{\alpha^p}$ and $\vec{\alpha^d}$ are the vectors of coefficients from the two linear regressions. Constraints (4.36)-(4.39) are the linearized availability constraints that replace constraints (4.22) and (4.23).

$$s^B_{tj} \geq (\alpha^p_0 + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^p_{tj}) * b_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.36)$$

$$s^B_{tj} \geq 0 \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.37)$$

$$s^B_{tj} - c^B_j \leq (\alpha^d_0 + \alpha^d_1 * h^p_{tj} + \alpha^d_2 * h^d_{tj} + \alpha^d_3 * l^d_{tj}) * b_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.38)$$

$$s^B_{tj} \leq c^B_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.39)$$

$$r^B_{tj} \geq (s^B_{t+1j} - s^B_{tj} + (\alpha^p_0 + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha^p_3 * l^p_{tj})) * b_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.40)$$

$$r^B_{tj} \geq (s^B_{tj} - s^B_{t+1j} - (\alpha^d_0 + \alpha^d_1 * h^d_{tj} + \alpha^d_2 * h^d_{tj} + \alpha^d_3 * l^d_{tj})) * b_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.41)$$

$$e^B_{tg} \geq \sum_{j \in \mathcal{J}} (s^B_{t+1j} - s^B_{tj} + \alpha^d_0 + \alpha^d_1 * h^p_{tj} + \alpha^d_2 * h^d_{tj} + \alpha^d_3 * l^d_{tj}) * v_{jg} * b_j$$

$$\forall t \in \mathcal{T}, g \in \mathcal{G} \quad (4.42)$$

$$e^B_{tg} \geq 0 \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \quad (4.43)$$

Following the same logic, constraints (4.24) and (4.25) will be replaced by constraints (4.40)-(4.43). We adopted a conservative method when calculating r^B_{tj} assuming the maximum number of (un)load bikes given the current service level. Constraints (4.40) refer to the case where bikes need to be unloaded from the rebalancing vehicle while constraints (4.41) refer to the case where bikes need to be loaded onto the vehicle. r^B_{tj} will take the maximum value from the two constraints. Constraints (4.42) and (4.43) also provide a conservative estimation for the number of extra bikes needed for each cluster. By summing up the upper bound of extra bikes needed for each location within the cluster, we obtain an upper bound for e^B_{tg}.

However, since constraints (4.36) are not equality constraints, it is possible that s^B_{tj} may

take an unnecessarily large value. The term $-s^B_{tj}$ and $-s^B_{t+1j}$ in constraints (4.40), (4.42), and (4.41) also provides forces for certain s^B_{tj} extremely large to reduce the value of r^B_{tj} and e^B_{tg}. Therefore, we introduce a binary variable o_{tj} indicating whether $\alpha^p_0 + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha^p_3 * l^p_{tj} \geq 0$. Constraints (4.44) and (4.45) define variable o_{tj}. Constraints (4.46) and (4.47) will be added to the optimization model.

$$(\alpha^p_0 + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^p_{tj}) * b_j \leq o_{tj} * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.44)$$

$$(-(\alpha^p_0 + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^p_{tj})) * b_j \leq (1 - o_{tj}) * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.45)$$

$$s^B_{tj} \leq M * o_{tj} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.46)$$

$$s^B_{tj} \leq (\alpha^p_0 + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^p_{tj}) * b_j + (1 - o_{tj}) * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.47)$$

To eliminate the bilinear terms in constraints (4.36), (4.38), (4.40), (4.41), (4.42), (4.44), (4.45), and (4.47), variables $l^{p'}_{tj}$, $l^{d'}_{tj}$, q^p_{tj}, q^d_{tj}, $h^{p'}_{tjg}$, $h^{d'}_{tjg}$, b'_{jg} and s'_{tjg} are introduced. These new variables are defined in constraints (4.48)-(4.71). Moreover, constraints (4.36), (4.38), (4.40), (4.41), (4.42), (4.44), (4.45), and (4.47) are replaced by constraints (4.72)-(4.79).

$$l_{tj}^{p'} \leq l_{tj}^{p} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.48}$$

$$l_{tj}^{p'} \leq b_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.49}$$

$$l_{tj}^{p'} \geq l_{tj}^{p} - (1 - b_j) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.50}$$

$$l_{tj}^{d'} \leq l_{tj}^{d} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.51}$$

$$l_{tj}^{d'} \leq b_j \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.52}$$

$$l_{tj}^{d'} \geq l_{tj}^{d} - (1 - b_j) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.53}$$

$$n_{tjg}^{p} \leq l_{tj}^{p'} \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \tag{4.54}$$

$$n_{tjg}^{p} \leq v_{jg} \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \tag{4.55}$$

$$n_{tjg}^{p} \geq l_{tj}^{p'} - (1 - v_{jg}) \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \tag{4.56}$$

$$n_{tjg}^{d} \leq l_{tj}^{d'} \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \tag{4.57}$$

$$n^d_{tjg} \leq v_{jg} \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.58)$$

$$n^d_{tjg} \geq l^{d'}_{tj} - (1 - v_{jg}) \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.59)$$

$$h^{p'}_{tjg} \leq h^p_{tj} \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.60)$$

$$h^{p'}_{tjg} \leq v_{jg} * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.61)$$

$$h^{p'}_{tjg} \geq h^p_{tj} - (1 - v_{jg}) * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.62)$$

$$h^{d'}_{tjg} \leq h^d_{tj} \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.63)$$

$$h^{d'}_{tjg} \leq v_{jg} * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.64)$$

$$h^{d'}_{tjg} \geq h^d_{tj} - (1 - v_{jg}) * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.65)$$

$$s'_{tjg} \leq s^B_{tj} \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.66)$$

$$b'_{jg} \leq b_j \quad \forall j \in \mathcal{J}, g \in \mathcal{G} \quad (4.67)$$

$$b'_{jg} \leq v_{jg} \quad \forall j \in \mathcal{J}, g \in \mathcal{G} \quad (4.68)$$

$$b'_{jg} \geq v_{jg} + b_j - 1 \quad \forall j \in \mathcal{J}, g \in \mathcal{G} \quad (4.69)$$

$$s'_{tjg} \leq v_{jg} * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.70)$$

$$s'_{tjg} \geq s^B_{tj} - (1 - v_{jg}) * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J}, g \in \mathcal{G} \quad (4.71)$$

$$s^B_{tj} \geq \alpha^p_0 * b_j + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^{p'}_{tj} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.72)$$

$$s^B_{tj} - c^B_j \leq \alpha^d_0 * b_j + \alpha^d_1 * h^p_{tj} + \alpha^d_2 * h^d_{tj} + \alpha^d_3 * l^{d'}_{tj} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.73)$$

$$r^B_{tj} \geq s^B_{t+1j} - s^B_{tj} + (\alpha^p_0 * b_j + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha^p_3 * l^{p'}_{tj}) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.74)$$

$$r^B_{tj} \geq s^B_{tj} - s^B_{t+1j} - (\alpha^d_0 * b_j + \alpha^d_1 * h^d_{tj} + \alpha^d_2 * h^d_{tj} + \alpha^d_3 * l^{d'}_{tj}) \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.75)$$

$$e^B_{tg} \geq \sum_{j \in \mathcal{J}} (s'_{t+1jg} - s'_{tjg} + \alpha^d_0 * b'_{jg} + \alpha^d_1 * h^{p'}_{tjg} + \alpha^d_2 * h^{d'}_{tjg} + \alpha^d_3 * n^d_{tjg})$$

$$\forall t \in \mathcal{T}, g \in \mathcal{G} \quad (4.76)$$

$$\alpha^p_0 * b_j + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^{p'}_{tj} \leq o_{tj} * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.77)$$

$$-(\alpha^p_0 * b_j + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^{p'}_{tj}) \leq (1 - o_{tj}) * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.78)$$

$$s^B_{tj} \leq \alpha^p_0 * b_j + \alpha^p_1 * h^p_{tj} + \alpha^p_2 * h^d_{tj} + \alpha_3 * l^{p'}_{tj} + (1 - o_{tj}) * M \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \quad (4.79)$$

4.3.2.7 Rebalancing Routing Distance Constraints

A significant cost of operating a bike-sharing system is bike repositioning. Constraints (4.80) define the rebalancing routing distance. The locations are divided into multiple clusters that are served by a single vehicle. The optimization will decide the assignment of locations to rebalancing clusters. We use binary variable v_{jg} to denote whether location j is assigned to cluster cl. After the clusters are defined, the computation of routing distance naturally becomes a traveling salesman problem where a single vehicle traverses all locations and comes back to the original location.

$\mathcal{G}$ is the set of rebalancing clusters. Each cluster is served by one rebalancing vehicle. The vehicle will traverse all locations and come back to the original location. G is a matrix that defines the distance between locations. Constraints (4.80) define the traveling distance as a function of $\vec{b}$ (a binary vector of whether the location offers bike-sharing service), $\vec{v_g}$ (a binary vector indicating whether the locations is served within cluster cl), and G (the distance between locations).

$$d_g^B \geq f_{rb}(\vec{b}, \vec{v_g}, G) \quad \forall g \in \mathcal{G} \tag{4.80}$$

4.3.2.8 Routing Distance Approximation

Due to the complexity of the traveling salesperson problem, constraints (4.80) are simplified to make the optimization model feasible and tractable. Instead of calculating the shortest path for the traveling salesman problem, we use the maximum spanning star to approximate the vehicle traveling distance. This approximation is adopted by Fu et al. (2022) and Schuijbroek et al. (2017). Schuijbroek et al. (2017) proves that Spanning Star is an upper bound on the shortest Hamiltonian path over the cluster. The updated constraints are presented in constraints (4.81) and (4.82). β_{jg} denotes the distance of spanning star with the internal node j and the rest of the nodes within cluster g as leaves. When location j

is served by cluster g and offers bike-sharing services, constraints (4.81) will compute the summation of the distance of every other location within the cluster to location j. However, if location j is not within cluster g or does not offer bike-sharing service, β_{jg} will be 0. Constraints (4.82) return the maximum distance among all the spanning stars within the cluster.

β_{jg}	distance of spanning star with internal node $j \in \mathcal{J}$ in cluster $j \in \mathcal{J}$
β_g^M	the distance of the maximum spanning star in cluster $g \in \mathcal{G}$

$$\beta_{jg} = \sum_k b_k * b_j * v_{kg} * v_{jg} * G_{kj} \quad \forall j \in \mathcal{J}, g \in \mathcal{G} \tag{4.81}$$

$$\beta_g^M = max_{j \in \mathcal{J}}(\beta_{jg}) \quad \forall g \in \mathcal{G} \tag{4.82}$$

However, both constraints (4.81) and (4.82) are non-linear. Therefore, we linearize them as constraints (4.83). When both b'_{jg} equals to 1, constraints (4.83) become equivalent to $d_g^B \geq \beta_{jg} \quad \forall j \in \mathcal{J}$. Therefore, $d_g^B \geq \beta_{jg}^M$.

$$d_g^B \geq \sum_{k \in \mathcal{J}} G_{jk} * (b'_{jg} + b'_{kg} - 1) \quad \forall j \in \mathcal{J}, g \in \mathcal{G} \tag{4.83}$$

4.3.2.9 Capacity Constraints

Constraints (4.84) are capacity constraints for transit. $\mathcal{S}$ is a set of road segments that are traveled by fixed-route transit. We preprocessed the data so that each segment is only traveled by one transit line. Multiple transit lines may share the same segments in reality, but we create distinctive segments for each transit line since different lines don't share one capacity constraint. $\mathcal{OD}_s$ is the set of OD pairs that travels segment s. $\mathcal{P}_s^M$ is the set of multi-modal paths that travels segment s. K_l is the capacity for transit line $l \in \mathcal{L}$. l_s is the

parameter denoting the transit line that travels segment s. Constraints (4.84) make sure that the number of passengers in all segments will not surpass the vehicle capacity.

$$\sum_{od \in \mathcal{OD}_s} m^T_{tod} * D_{tod} + \sum_{p \in \mathcal{P}^M_s} v^M_{tod_p} * D_{tod_p} \leq f^T_{tl_s} * K_{l_s} \quad \forall t \in \mathcal{T}, s \in \mathcal{S} \tag{4.84}$$

4.3.2.10 Attractiveness and Market Share Constraints

Our optimization model considers endogenous demand. The demand for each traveling mode depends on the attractiveness of that mode. Therefore, constraints (4.85)-(4.90) are included to compute disutility and attractiveness for traveling using fixed-route transit, multi-modal service, and bike-sharing service.

u^T_{tod}	the disutility of traveling using fixed-route transit for OD pair $od \in \mathcal{OD}$ during time period $t \in \mathcal{T}$
u^M_{tp}	the disutility of traveling using multi-modal service for path $p \in \mathcal{P}^M$ during time period $t \in \mathcal{T}$
u^B_{tod}	the disutility of traveling using bike-sharing service for OD pair $od \in \mathcal{OD}$ during time period $t \in \mathcal{T}$
a^T_{tod}	the attractiveness of traveling using fixed-route transit for OD pair $od \in \mathcal{OD}$ during time period $t \in \mathcal{T}$
a^M_{tp}	the attractiveness of traveling using multi-modal service for path $p \in \mathcal{P}^M$ during time period $t \in \mathcal{T}$
a^B_{tod}	the attractiveness of traveling using bike-sharing service for OD pair $od \in \mathcal{OD}$ during time period $t \in \mathcal{T}$

Constraints (4.85)-(4.87) calculate the travel disutilities. The disutility of traveling using fixed-route transit includes in-vehicle traveling cost, walking time cost, and waiting time cost. While waiting time is a function of transit frequencies which are decision variables, in-vehicle traveling cost and walking time cost are constant for each OD pair. We use U^T_{od} to denote the in-vehicle traveling cost and walking cost for traveling using fixed-route transit. C^T_w is the cost of waiting for bus per unit time. By adding up U^T_{od} and the waiting cost, constraints (4.85) compute the disutility for taking transit.

Similarly, constraints (4.86) calculate the distuility for using multi-modal service which includes in-vehicle traveling time cost, walking time cost, and waiting time cost for transit as well as traveling time cost for biking and cost for waiting in case bike pick-up or drop-off is unavailable. We assume the cost for waiting for a vacant bike/dock is a decreasing linear function of pick-up/drop-off service level $l^p_{tj^M_p}$ and $l^d_{tj^M_p}$. When the service level is 1, the cost is 0, when the service level is 0, the cost is C^B_{sb}. U^M_p includes in-vehicle traveling cost, walking time cost for transit as well as traveling cost for biking.

Constraints (4.87) follow a similar logic. U^B_{od} is the traveling cost for biking. The total disutility of taking the bike-sharing system is traveling cost and unavailability cost as shown in constraints (4.87).

Attractiveness is the exponential term of negative disutilities. Constraints (4.88)-(4.90) compute the attractiveness for these three travel modes for each OD pair. Since we are taking only one path for multi-modal service, the attractiveness for that OD pair equals to the attractiveness of the path taken. Therefore, in constraints (4.89), we multiply the attractiveness of each path with z^M_p and sum over all paths for OD pair od.

We adopt Sales Based Linear Programming (SBLP) model from Gallego et al. (2015) to formulate the market share for each travel mode in constraints (4.91)-(4.94). The market share of each travel mode is proportional to its corresponding attractiveness unless the capacity constraints are violated. Constraints (4.94) ensure that the total market share is 1.

$$u^T_{tod} = U^T_{od} + C^T_w * w^T_{tod} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.85}$$

$$u^M_{tp} = U^M_p + C^M_w * w^M_{tp} + C^B_{sb} * (1 - l^p_{tj^M_p} + 1 - l^d_{tj^M_{od}}) \quad \forall t \in \mathcal{T}, p \in \mathcal{P}^M \tag{4.86}$$

$$u^B_{tod} = U^B_{od} + C^B_{sb} * (1 - l^p_{tj^B_{od}} + 1 - l^d_{tj^B_{od}}) \quad \forall t \in \mathcal{T}, p \in \mathcal{P}^M \tag{4.87}$$

$$a^T_{tod} = e^{-u^T_{tod}} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.88}$$

$$a^M_{tod} = \sum_{p \in \mathcal{P}^M_{od}} z^M_p e^{-u^M_{tp}} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.89}$$

$$a^B_{tod} = z^B_{od} e^{-u^B_{tod}} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.90}$$

$$m^T_{tod} * A^D_{od} \leq m^D_{tod} * a^T_{tod} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.91}$$

$$m^M_{tod} * A^D_{od} \leq m^D_{tod} * a^M_{tod} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.92}$$

$$m^B_{tod} * A^D_{od} \leq m^D_{tod} * a^B_{tod} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.93}$$

$$m^T_{tod} + m^B_{tod} + m^M_{tod} + m^D_{tod} = 1 \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.94}$$

4.3.2.11 Full model

In the original optimization model, we maximize the saved distance traveled by switching from driving to using transit, bike-sharing, and multi-modal (Objective function (4.1), with constraints (4.2)-(4.5), (4.8)-(4.19), (4.26)-(4.31), (4.37), (4.39), (4.43), (4.43)-(4.79), and (4.83)-(4.94). However, constraints (4.18) and (4.19) are non-linear. The exponential terms in constraints (4.88)-(4.90) and bilinear terms in constraints (4.91)-(4.93) are non-linear which can significantly increase the model complexity. In Section 4.3.3, we introduce the adaptive discretization method to tackle the problem which will be detailed in 4.3.3.

4.3.3 Discretized Model

Due to the non-linearity caused by constraints (4.18), (4.19), (4.90)-(4.93), it is not feasible to solve the optimization model with real-world problem instances in a reasonable amount

of time. Therefore, we decided to discretize the transit frequencies and bike service levels to eliminate the non-linearity in our model.

We use $\mathcal{W}$ to denote the set of candidate frequencies and $\mathcal{Q}$ to denote the set of candidate service levels. We introduce the three sets of main binary decision variables f^b_{tlw}, s^{bp}_{tjq}, and s^{bd}_{tjq} to discretize transit frequency and biking service level.

The extra decision variables are listed below:

f^b_{tlw}	=1 when transit line $l \in \mathcal{L}$ takes frequency $w \in \mathcal{W}$ during time period $t \in \mathcal{T}$, 0 otherwise
l^{bp}_{tjq}	=1 when location $j \in \mathcal{J}$ takes pick-up service level $q \in \mathcal{Q}$ during time period $t \in \mathcal{T}$, 0 otherwise
l^{bd}_{tjq}	=1 when location $j \in \mathcal{J}$ takes drop-off service level $q \in \mathcal{Q}$ during time period $t \in \mathcal{T}$, 0 otherwise
$x^T_{todw_1w_2}$	=1 when the first transit segment takes frequency $w_1 \in \mathcal{W}$ and the second transit segment takes frequency $w_2 \in \mathcal{W}$ for OD pair $od \in \mathcal{OD}$ during the time period $t \in \mathcal{T}$, 0 otherwise
$x^B_{todq_1q_2}$	=1 when the pick-up service level is $q_1 \in \mathcal{Q}$ and the drop-off service level is $q_2 \in \mathcal{Q}$ for OD pair $od \in \mathcal{OD}$ during the time period $t \in \mathcal{T}$, 0 otherwise
$x^M_{tpw_1w_2q_1q_2}$	=1 when the first transit segment takes frequency $w_1 \in \mathcal{W}$, the second transit segment takes frequency $w_2 \in \mathcal{W}$, the pick-up service level is $q_1 \in \mathcal{Q}$, and the drop-off service level is $q_2 \in \mathcal{Q}$ for OD pair $p \in \mathcal{P}^M$ during the time period $t \in \mathcal{T}$, 0 otherwise

Constraints (4.95) define the relationship between $b^T_{todw_1w_2}$ and f^b_{tlw}. If the transit line of the first segment takes frequency index w_1 and the transit line of the second segment takes frequency index w_2, then $b^T_{todw_1w_2}$ will be 1. Constraints (4.96) and (4.97) follow similar logic. $j^{BP}_{od}(j^{BD}_{od})$ returns the pick-up (drop-off) location for OD pair od for bike-sharing service. $j^{MP}_p(j^{MD}_p)$ returns the pick-up (drop-off) location for path p for multi-modal service. Constraints (4.98)-(4.100) ensure that only 1 frequency or service level will be chosen for each transit line or bike location.

Constraints (4.101)-(4.103) can replace constraints (4.91)-(4.93) as market share constraints. In the pre-processing stage, we calculate the attractiveness under all combinations of frequencies and service levels. Therefore, the attractiveness values in constraints

(4.101)-(4.103) become parameters instead of decision variables. Constraints (4.104)-(4.106) ensure that the values of frequencies and service levels are consistent with the binary variables f^b_{tlw}, s^{bp}_{tjq} and s^{bd}_{tjq}.

$$x^T_{todw_1w_2} = f^b_{tl^{T1}_{od}w_1} * f^b_{tl^{T2}_{od}w_2} \quad \forall t \in T, od \in \mathcal{OD} \tag{4.95}$$

$$x^B_{todq_1q_2} = l^{bp}_{tj^{BP}_{od}q_1} * l^{bd}_{tj^{BD}_{od}q_2} \quad \forall t \in T, od \in \mathcal{OD} \tag{4.96}$$

$$x^M_{tpw_1w_2q_1q_2} = f^b_{tl^{M1}_{p}w_1} * f^b_{tl^{M2}_{p}w_2} * l^{bp}_{tj^{MP}_{p}q_1} * l^{bd}_{tj^{MD}_{p}q_2} \quad \forall t \in T, p \in \mathcal{P} \tag{4.97}$$

$$\sum_{q \in \mathcal{Q}} s^{bp}_{tjq} = 1 \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.98}$$

$$\sum_{q \in \mathcal{Q}} s^{bd}_{tjq} = 1 \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.99}$$

$$\sum_{w \in \mathcal{W}} f^b_{tlw} = 1 \quad \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{4.100}$$

$$m^T_{tod} * A^D_{od} \leq m^D_{tod} * \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} A^T_{todw_1w_2} x^T_{todw_1w_2} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.101}$$

$$m^M_{tod} * A^D_{od} \leq m^D_{tod} * \sum_{p \in \mathcal{P}^M_{od}} z^M_p \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^M_{tpw_1w_2q_1q_2} x^M_{tpw_1w_2q_1q_2}$$

$$\forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.102}$$

$$m^B_{tod} * A^D_{od} \leq m^D_{tod} * \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^B_{todq_1q_2} x^B_{todq_1q_2} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.103}$$

$$f^T_{tl} = \sum_{w \in \mathcal{W}} F_{tlw} * f^b_{tlw} \quad \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{4.104}$$

$$l^p_{tj} = \sum_{q \in \mathcal{Q}} l_{tjq} * l^{bp}_{tlq} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.105}$$

$$l^d_{tj} = \sum_{q \in \mathcal{Q}} S_{tjq} * l^{bd}_{tlq} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.106}$$

However, constraints (4.95)-(4.97) and constraints (4.101)-(4.103) are still bilinear. We introduce several auxiliary variables below to linearize them.

Constraints (4.95)-(4.97) are replaced by (4.107)-(4.117). Constraints (4.101)-(4.103)

$\theta^{TD}_{todw_1w_2}$	$=m^D_{tod}$ when $x^T_{todw_1w_2}=1$, 0 otherwise
$\theta^{BD}_{todq_1q_2}$	$=m^D_{tod}$ when $x^B_{todq_1q_2}=1$, 0 otherwise
γ^{MD}_{tp}	$=m^D_{tod_p}$ when $z^M_p=1$, 0 otherwise
$\gamma^{MD'}_{tpw_1w_2q_1q_2}$	$=\gamma^{MD}_{tp}$ when $x^M_{tpw_1w_2q_1q_2}=1$, 0 otherwise

are replaced by (4.118)-(4.120). $\theta^{TD}_{todw_1w_2}$, $\theta^{BD}_{todq_1q_2}$, γ^{MD}_{tp}, and $\gamma^{MD'}_{tpw_1w_2q_1q_2}$ are defined by constraints (4.124)-(4.132).

$$x^T_{todw_1w_2} \leq f^b_{tl^{T1}_{od}w_1} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W} \tag{4.107}$$

$$x^T_{todw_1w_2} \leq f^b_{tl^{T2}_{od}w_2} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W} \tag{4.108}$$

$$x^T_{todw_1w_2} \geq f^b_{tl^{T1}_{od}w_1} + f^b_{tl^{T2}_{od}w_2} - 1 \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W} \tag{4.109}$$

$$x^B_{todq_1q_2} \leq l^{bp}_{tj^{BP}_{od}q_1} \quad \forall t \in T, od \in \mathcal{OD}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.110}$$

$$x^B_{todq_1q_2} \leq l^{bd}_{tj^{BD}_{od}q_2} \quad \forall t \in T, od \in \mathcal{OD}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.111}$$

$$x^B_{todq_1q_2} \geq l^{bp}_{tj^{BP}_{od}q_1} + l^{bd}_{tj^{BD}_{od}q_2} - 1 \quad \forall t \in T, od \in \mathcal{OD}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.112}$$

$$x^M_{tpw_1w_2q_1q_2} \leq f^b_{tl^{M1}_{p}w_1} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.113}$$

$$x^M_{tpw_1w_2q_1q_2} \leq f^b_{tl^{M2}_{p}w_2} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.114}$$

$$x^M_{tpw_1w_2q_1q_2} \leq l^{bp}_{tj^{MP}_{p}q_1} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.115}$$

$$x^M_{tpw_1w_2q_1q_2} \leq l^{bd}_{tj^{MD}_{od}q_2} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.116}$$

$$x^M_{tpw_1w_2q_1q_2} \geq f^b_{tl^{M1}_{p}w_1} + f^b_{tl^{M2}_{p}w_2} + l^{bp}_{tj^{MP}_{p}q_1} + l^{bd}_{tj^{MD}_{od}q_2} - 4$$
$$\forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \tag{4.117}$$

$$m^T_{tod} * A^D_{od} \leq \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} A^T_{todw_1w_2}\theta^{TD}_{todw_1w_2} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.118}$$

$$m^M_{tod} * A^D_{od} \leq \sum_{p \in \mathcal{P}^M_{od}} \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^M_{tpw_1w_2q_1q_2}\gamma^{MD'}_{tpw_1w_2q_1q_2}$$
$$\forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.119}$$

$$m^B_{tod} * A^D_{od} \leq \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^B_{todq_1q_2} \theta^{BD}_{todq_1q_2} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \quad (4.120)$$

$$\theta^{TD}_{todw_1w_2} \leq x^T_{todw_1w_2} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W} \quad (4.121)$$

$$\theta^{TD}_{todw_1w_2} \leq m^D_{tod} \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W} \quad (4.122)$$

$$\theta^{TD}_{todw_1w_2} \geq m^D_{tod} - (1 - x^T_{todw_1w_2}) \quad \forall t \in T, od \in \mathcal{OD}, w_1 \in \mathcal{W}, w_2 \in \mathcal{W} \quad (4.123)$$

$$\theta^{BD}_{todq_1q_2} \leq x^B_{todq_1q_2} \quad \forall t \in T, od \in \mathcal{OD}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \quad (4.124)$$

$$\theta^{BD}_{todq_1q_2} \leq m^D_{tod} \quad \forall t \in T, od \in \mathcal{OD}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \quad (4.125)$$

$$\theta^{BD}_{todq_1q_2} \geq m^D_{tod} - (1 - x^B_{todq_1q_2}) \quad \forall t \in T, od \in \mathcal{OD}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \quad (4.126)$$

$$\gamma^{MD}_{tp} \leq z^M_p \quad \forall t \in T, p \in \mathcal{P}^M \quad (4.127)$$

$$\gamma^{MD}_{tp} \leq m^D_{tod_p} \quad \forall t \in T, p \in \mathcal{P}^M \quad (4.128)$$

$$\gamma^{MD}_{tp} \geq m^D_{tod_p} - (1 - z^M_p) \quad \forall t \in T, p \in \mathcal{P}^M \quad (4.129)$$

$$\gamma^{MD'}_{tpw_1w_2q_1q_2} \leq \gamma^{MD}_{tp} \quad \forall t \in T, p \in \mathcal{P}^M, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \quad (4.130)$$

$$\gamma^{MD'}_{tpw_1w_2q_1q_2} \leq x^M_{tpw_1w_2q_1q_2} \quad \forall t \in T, p \in \mathcal{P}^M, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \quad (4.131)$$

$$\gamma^{MD'}_{tpw_1w_2q_1q_2} \geq \gamma^{MD}_{tp} - 1 + x^M_{tpw_1w_2q_1q_2}$$

$$\forall t \in T, p \in \mathcal{P}^M, w_1 \in \mathcal{W}, w_2 \in \mathcal{W}, q_1 \in \mathcal{Q}, q_2 \in \mathcal{Q} \quad (4.132)$$

The discretized model has the objective function (4.1) and constraints (4.2)-(4.5), (4.8)-(4.17), (4.26)-(4.31), (4.37), (4.39), (4.43), (4.43)-(4.79), and (4.83)-(4.84), (4.94), (4.98)-(4.100), and (4.107)-(4.132).

4.4 Solution method

In Section 4.3, we introduced the original model and discretized model. The original model is a non-linear and non-convex model which cannot be solved in a reasonable amount of time for a real-life size city. The discretized model eliminates the non-linearity and non-

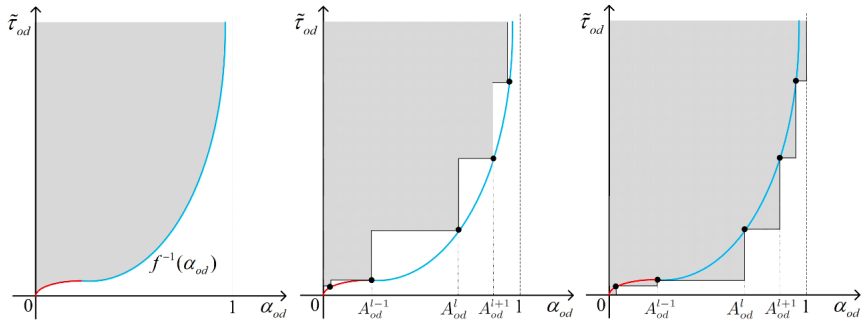
convexity in the model, but it still poses two significant issues. (1) We are not able to evaluate the performance of our solution to the true optimal solution, and (2) we still need an algorithm for the selection of the anchor points (the selected frequencies and service levels) in the discretization model. Therefore, we adopted the adaptive discretization method in the study of Wang et al. (2022). This method develops a conservative model (which provides a feasible solution) and a relaxed model (which provides an upper bound) for the original non-linear optimization problem. Convergence to a global optimum is secured by alternatively solving the conservative model and relaxed model. We will elaborate on this method in Section (4.4.1).

4.4.1 Adaptive Discretization Method

The adaptive discretization method is developed by Wang et al. (2022). as an exact algorithm to solve the non-convex optimization problems converging to the global optimum. The non-convex constraints were approximated using piece-wise constant segments. Figure 4.3 shows how Wang et al. (2022) discretized the non-convex constraints and the corresponding feasible regions for the original model, the conservative model, and the relaxed model. In Figure 4.3(b), the feasible region is a subset of the feasible regions of the original model (Figure 4.3(a)). The feasible region of the relaxed model (Figure 4.3(c)) contains the feasible region of the original model. Therefore, the conservative model will provide a feasible solution to the original model, while the relaxed model will provide an upper bound.

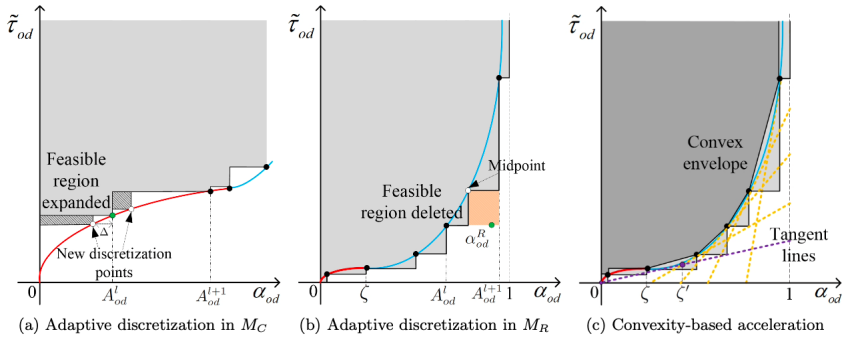
The adaptive method is guaranteed to converge to a global optimum by iteratively solving the conservative model and the relaxed model. After each iteration, a new anchor point will be added based on the solution from the current model as shown in Figure 4.4. The algorithm can also be accelerated by adding tangent lines in the convex region as shown in 4.4(c).

The adaptive discretization method fits perfectly with our problem which is also a non-



(a) M^* feasible region (b) M_C feasible region (c) M_R feasible region

Figure 4.3: Example of S-curve function and feasible regions of the original model, conservative model, and relaxed model from Wang et al. (2022)



(a) Adaptive discretization in M_C (b) Adaptive discretization in M_R (c) Convexity-based acceleration

Figure 4.4: Adaptive discretization scheme, and acceleration based on local convexity from Wang et al. (2022)

convex optimization problem. The major non-convexity in our problem comes from the market share constraints (4.91)-(4.93). Market share is a monotonically increasing function of the attractiveness of that travel mode. Therefore, if we discretize the attractiveness, we can apply the adaptive discretization method mentioned above. With closer scrutiny of the disutility constraints (4.85)-(4.87), we found that disutilities and attractiveness are functions of transit frequencies and bike-sharing service levels. Therefore, instead of discretizing the attractiveness, we discretize transit frequencies and bike-sharing service level as shown in Section 4.3.3. Some adjustment needs to be made to the discretization model. We replace constraints (4.104)-(4.106) with (4.133)-(4.138). Constraints (4.118)-(4.120) will be replaced by constraints (4.139)-(4.141) in the conservative model and constraints (4.142)-(4.144) in the relaxed model.

$$f_{tl} \geq \sum_{w \in \mathcal{W}} F_{tlw} * f^b_{tlw} \quad \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{4.133}$$

$$f_{tl} \leq \sum_{w \in \mathcal{W} \setminus \{0\}} F_{tlw} * f^b_{tlw-1} \quad \forall t \in \mathcal{T}, l \in \mathcal{L} \tag{4.134}$$

$$l^p_{tj} \geq \sum_{q \in \mathcal{Q}} L_{tjq} * l^{bp}_{tlq} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.135}$$

$$l^p_{tj} \leq \sum_{q \in \mathcal{Q} \setminus \{0\}} L_{tjq} * l^b_{tlq-1} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.136}$$

$$l^d_{tj} \geq \sum_{q \in \mathcal{Q}} L_{tjq} * l^{bd}_{tlq} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.137}$$

$$l^d_{tj} \leq \sum_{q \in \mathcal{Q} \setminus \{0\}} L_{tjq} * l^{bd}_{tlq-1} \quad \forall t \in \mathcal{T}, j \in \mathcal{J} \tag{4.138}$$

Conservative model:

$$m^T_{tod} * A^D_{od} \leq \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} A^T_{todw_1w_2} \theta^{TD}_{todw_1w_2} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.139}$$

$$m^M_{tod} * A^D_{od} \leq \sum_{p \in \mathcal{P}^M_{od}} \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^M_{tpw_1w_2q_1q_2} \gamma^{MD'}_{tpw_1w_2q_1q_2}$$

$$\forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.140}$$

$$m^B_{tod} * A^D_{od} \leq \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^B_{todq_1q_2} \theta^{BD}_{todq_1q_2} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.141}$$

Relaxed model:

$$m^T_{tod} * A^D_{od} \leq \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} A^T_{todw_1w_2} \theta^{TD}_{todw_1-1w_2-1} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.142}$$

$$m^M_{tod} * A^D_{od} \leq \sum_{p \in \mathcal{P}^M_{od}} \sum_{w_1 \in \mathcal{W}} \sum_{w_2 \in \mathcal{W}} \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^M_{tpw_1w_2q_1q_2} \gamma^{MD'}_{tpw_1-1w_2-1q_1-1q_2-1}$$

$$\forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.143}$$

$$m^B_{tod} * A^D_{od} \leq \sum_{q_1 \in \mathcal{Q}} \sum_{q_2 \in \mathcal{Q}} A^B_{todq_1q_2} \theta^{BD}_{todq_1-1q_2-1} \quad \forall t \in \mathcal{T}, od \in \mathcal{OD} \tag{4.144}$$

The major difference from the study of Wang et al. (2022) is that our discretization is multi-dimensional instead of one-dimensional. There are at most 4 combinations of frequencies and service levels. For example, when we increase the number of anchor points from 2 to 3, the number of possible attractiveness values and its corresponding binary decision variables ($x_{tpw_1w_2q_1q_2}$) increase from 16 to 81 for each multi-modal path. Therefore, the complexity of our model will grow significantly with more anchor points. Besides the model complexity, our case study is also larger than that in the study of Wang et al. (2022). They focused on long-distance aerial trips, but we investigate transit and bike-sharing users in urban areas with much more OD pairs and data points. In order to reduce runtime, we further apply the coordinate descent method and anchor point parsimony method which will be explained in Section 4.4.2 and 4.4.3.

4.4.2 Coordinate Descent Method

Instead of focusing on one system as in the study of Wang et al. (2022), we co-optimize two systems: the public transit system and the bike-sharing system. In order to reduce the complexity of the original problem, we introduce the coordinate descent method where we iteratively optimize decision variables in one system while fixing certain decision variables in another system. We separate the decision variables into three categories:

A Bike-sharing service level decision variables (l^p_{tj} and l^d_{tj})

B Transit frequencies decision variables (f_{tl}^T)

C All the other decision variables (including c_j^B, s_{tj}^B, r_{tj}^B, e_{tg}^B etc.)

We first fix the decision variables B (transit frequencies) and optimize the decision variables A (bike-sharing service level) and C (all the other decision variables), then we fix the decision variables A (bike-sharing service level) and optimize decision variables B (the transit frequencies) and C (all the other decision variables). We iteratively run these two optimizations until it converges.

4.4.3 Anchor Point Parsimony

As discussed in Section 4.4.1, model complexity increases significantly with the number of anchor points. When the number of anchor points increase to 7, the runtime is already more than 30 hours for a single iteration. Furthermore, with a closer observation of the added anchor points, we find that many of them are redundant. For example, we start the algorithm with 3 service levels with anchor points [0.75, 0.85, 0.95]. After 4 iterations, the anchor points become [0.75, 0.85, 0.9, 0.9125, 0.91875, 0.925, 0.95]. As you can see, the algorithm is mostly exploring and adding points to the range of [0.9, 0.95]. [0.75, 0.85] is redundant only adding complexity to the model.

Therefore, we developed an algorithm where we only keep three anchor points at all iterations. We start with $[a, b, c]$, we presented how we update the anchors below:

1. If the optimal solution x^* is not a, b, or c, then the anchors becomes $[\frac{a+x^*}{2}, x^*, \frac{x^*+c}{2}]$
2. If the optimal solution x^* is a, then the anchors becomes $[a, \frac{a+b}{2}, \frac{a+c}{2}]$
3. If the optimal solution x^* is b, then the anchors becomes $[\frac{a+b}{2}, b, \frac{b+c}{2}]$
4. If the optimal solution x^* is c, then the anchors becomes $[\frac{a+c}{2}, \frac{b+c}{2}, c]$

4.5 Case Studies and Results

We applied our algorithm to Boston city in Massachusetts. We obtain the origin-destination data from Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES) (United States Census Bureau, 2020) and the transit network and schedule from MBTA (2021a). The major bike-sharing system in Boston is Bluebikes. We are able to get access to data like station location, number of docks at each location, and the number of trips from their website (Bluebikes, 2022).

Figure 4.5: Map of Boston City, Massachusetts

4.5.1 Case Studies

We develop 5 different case studies with increasing size within the city of Boston (Figure 4.5). There are several reasons we are doing this. (1) The scale for running the optimization model on the entire Boston city is very big, we are not able to apply the method of Wang et al. (2022) for Boston proper directly. However, we still want to evaluate our results against the optimal solution to get an optimality gap. Therefore, we run our method and the method developed by Wang et al. (2022) in smaller case study to show obtain the optimality gap of our method. (2) Other than obtaining the optimality gap, the increasing size of our

case studies provides strong evidence of the scalability of our method. The runtime of our method is significantly smaller than that when using the method of Wang et al. (2022) in the smallest case study. And the method of Wang et al. (2022) will fail to provide a feasible solution in the following case studies since it runs out of memory, while our method can still provide results in a reasonable amount of time. (3) With different case study sizes, we are able to evaluate the impact of the three different solution methods mentioned above. Our final solution algorithm applies all three methods mentioned in Section 4.4 which are the adaptive discretization method, coordinate descent method, and anchor point parsimony method. We want to evaluate the contribution of each of them regarding the reduction of runtime and optimality gap.

The case instances and their included regions are listed in Table 4.1.

Table 4.1 The regions included for 5 case instances

Case Instances	Regions included
Case 1	Downtown
Case 2	Downtown and Beacon Hill
Case 3	Downtown, Beacon Hill and South End neighborhood
Case 4	Downtown, Beacon Hill, South End, Fenway, and South Boston neighborhood
Case 5	Downtown, Beacon Hill, South End, Fenway, South Boston, Charlestown, East Boston, and Allston-Brighton neighborhood

4.5.2 Computational Results

In this section, we further investigate how the three solution methods contribute to the reduction of runtime and how accurate these methods are compared to the true optimal solution. The method of Wang et al. (2022) is used as an optimal solution benchmark. The result for the current scenario is evaluated to be compared with our solution. We evaluate the scenarios where we apply all three solution methods as well as the scenarios we only apply two of them. We set a runtime limit of 24 hours and a memory limit of 50 GB.

Table 4.2 presents the objective values and the runtime under the current scenario, when applying all three solution methods, applying two solution methods as well as the method

of Wang et al. (2022) (with and without time limit). Since the method of Wang et al. (2022) guarantees optimality, it provides an optimal benchmark for our algorithm. The percentage improvement of applying all three methods compared to the current scenario is presented in the second last column. The optimality gap between applying all three methods and the method of Wang et al. (2022) is presented in the last column. In Table 4.2, AD, CD, and AP denotes The runtime is presented in the parentheses. Column AD+CD+AP presents the results when we apply all three solution methods mentioned in Section 4.4. Column AD+CD presents the results where we only apply the adaptive discretization method and coordinate descent method. Comparing columns AD+CD+AP and AD+CD, we find that the results are within 1% difference for all cases. However, the runtime for AD+CD is much longer than AD+CD+AP in Instances 2,3, and 4. In Instance 5, AD+CD reached the time limit of 24 hours and we cannot obtain a feasible solution since the memory limit has been reached.

Table 4.2 The objectives and the runtime for 5 case instances under the current scenario and different combinations of 3 methods

	Current scenario	AD+CD+AP	AD+CD	AD+AP	CD + AP	AD without time limit	AD	% increase	Optimality gap
1	1994 (7s)	2000 (2m)	1997 (34m)	1987 (24m)	1995 (1m)	2002 (256m)	2002 (256m)	0.31%	0.09%
2	8929 (9s)	9102 (25m)	9086 (89m)	9063 (6h)	9057 (10m)	NA (50h)	9081 (24h)	1.94%	NA
3	11670 (9s)	11957 (2h)	11975 (6.5h)	11973 (24h)	11973 (72m)	NA	NA	2.46%	NA
4	47838 (15s)	48436 (3.7h)	48410 (12h)	NA	48442 (2h)	NA	NA	1.25%	NA
5	149192 (42s)	154182 (20h)	NA	NA	154006 (15h)	NA	NA	3.34%	NA

*AD denotes adaptive discretization; CD denotes coordinate descent; AP denotes anchor point parsimony. The second last column indicates the percentage difference between Column AD+CD+AP and the current scenario. The last column indicates the percentage difference between Column AD+CD+AP and Column AD without a time limit. If not specified, all solutions are evaluated with a time limit of 24 hours.

Column AD+AP applies adaptive discretization and anchor point parsimony methods. Due to the large decision space, it does not manage to obtain a feasible solution for Instances 4 and 5. The runtime is also much larger in Instances 1, 2, and 3. It shows that coordinate descent is indispensable in co-optimize the transit system and bike-sharing system in large real-world instances.

Column CD+AP applies coordinate descent and anchor point parsimony methods. Instead of applying the adaptive discretization method stated in the study of Wang et al.

(2022), we made a modification to the method. Rather than iteratively running the relaxed model and the conservative model, we only run the conservative model and update the anchor points solely based on the results of the conservative models. Comparing the results with Column AD+CD+AP, we find that it offers similar quality results (less than 0.5%) within a similar, if not shorter, amount of time. It shows that iteratively running a conservative model and a relaxed model may not provide much value in reducing runtime or improving the quality of our solution in our case studies.

Column AD without time limit is the method of Wang et al. (2022) which uses only the adaptive discretization method. It guarantees optimality and provides a benchmark for our own solution. Comparing this column with Column AD+CD+AP, we find that the optimality gap for the first instance is within 0.1%. It proves that our method provides a near-optimal solution. In Instances 2 - 5, we are not able to obtain a solution by applying the method of Wang et al. (2022) within a reasonable amount of time. In the next column AD, we set a time limit of 24 hours for the method of Wang et al. (2022). We find that the runtimes reach the limit for instances 2 and 3. We are not able to obtain a feasible solution for Instances 3, 4 and 5. This explains why we cannot simply apply the method of Wang et al. (2022) to our instances. It again demonstrates the importance of coordinate descent and anchor point parsimony in obtaining a near-optimal solution in a shorter runtime.

Comparing the current scenarios and our solution in columns 2 and 3, it is clear that our solution provides a more efficient integrated design. The percentages of objective gain range from 0.31% (case study 1) to 3.34% (case study 5). The percentage improvement increases with the size of the instances in general. It shows that our tool is more powerful in larger instances where multi-modal services can have a bigger impact.

Table 4.3 compares the market share under the current scenario and the proposed solution. From Case 1 to Case 5, there is a clear trend of decreasing bike-sharing market shares. This is due to the fact that smaller case instances are within the city center while larger instances involve more longe-distance commuting. Biking is not suitable for long-

Table 4.3 The market share for the current scenario and proposed solution of case 1-5

Instances	Current Scenario				Proposed Solution			
	Transit	Bike-sharing	Multi-modal	Outside option	Transit	Bike-sharing	Multi-modal	Outside option
Case 1	53%	8%	0%	39%	52%	12%	0%	36%
Case 2	53%	3%	0%	43%	51%	9%	1%	40%
Case 3	52%	2%	1%	46%	49%	7%	2%	42%
Case 4	58%	2%	2%	38%	55%	8%	1%	36%
Case 5	57%	0%	0%	42%	56%	2%	2%	40%

distance traveling. Therefore, the market share for bike-sharing generally decreases with the size of the case instance. However, the market share for multi-modal increases slightly in the proposed solution with the size of the instance. It demonstrates the potential for an integrated bike-sharing and transit system in longer-distance commuting.

Comparing the market shares between the current scenario and the proposed solution, there is an obvious increase in bike-sharing and multi-modal market share in the proposed solution. There is also a slight decline in transit market share. However, in Case 5, the total transit ridership increases if we count commuters taking only fixed-route transit and multi-modal services. In all instances, we are able to reduce the outside option market share by 3%. It demonstrates the value of our proposed solution in discouraging private vehicle driving.

4.6 Conclusion

The micromobility industry has been recovering rapidly after the hit of the pandemic. More innovative micro-mobility platforms and services have provided an alternative for a more sustainable urban transportation system. However, due to the limited travel range of micro-mobility services, the integration of micro-mobility and public transit is the key to its prevalence. This paper aims to co-optimize the public transit and bike-sharing system. It provides an integrated framework for future urban transportation planning.

We develop an integrated model framework to optimize transit frequencies, bike-sharing station locations, the number of docks at each location as well as the rebalancing operations. A mixed-integer non-linear programming model is formulated. In order to solve

this complicated model, we applied three important solution algorithms: (1) adaptive discretization method, (2) coordinate descent method, and (3) anchor point parsimony method. We compare our algorithm to the method of Wang et al. (2022) which guarantees optimality. We found that our solution algorithm can provide a near-optimal solution (within 0.1% optimality gap) with only 5% of runtime. We also evaluate the impact of each individual algorithm and find that coordinate descent and anchor point parsimony contribute significantly to reducing runtime as well as reducing memory usage.

Our model and solution are evaluated using data from Boston city. We are able to increase the bike-sharing market share as well as the multi-modal service market share.

Regarding future research, this model framework can be extended to e-scooter and e-bike services with more consideration in charging facilities. Furthermore, our paper only considers dock-based bike-sharing services. Future studies can also investigate the integration of dockless micro-mobility services with public transit systems. Lastly, our model does not optimize the transit network. By further optimizing the transit network, it can potentially provide more mobility and accessibility inside the city.

Chapter 5

Conclusions and Future Directions

In this thesis, we focus on optimizing the public transit system considering the interactions with both traditional operators (parking operators) and emerging technologies (ride-hailing services and micro-mobility services). In this section, we summarize our contributions and present future research directions.

5.1 Chapter 2: Comprehensive Public Transit Design Considering Parking Operator's Response Using a Tractable Two-stage Framework

This chapter aims to optimize transit frequencies and transit fare considering the interactions with the parking operator as well as endogenous passenger choice. There was extensive literature on transit network design and frequency setting. However, few studies considered the response from the parking operator and how this interaction exerted an impact on the optimal transit frequencies and transit fare. In this chapter, we capture the interrelationship between the transit operator and the parking operator as well as obtain the optimal transit frequencies and transit fare under different strategies and assumptions.

Methodologically, this is the first paper presenting a two-stage game-theoretic model

and a heuristic solution to optimize transit frequencies and transit fare considering the parking operator's response, road congestion, and endogenous passenger demand. A mixed-integer non-linear model is developed to optimize transit frequencies and transit fare. The coordinate descent method and acceleration method is applied to the base models to solve the complicated non-linear optimization model. Regression method is introduced to solve incorporate the Stage II response to Stage I decision making. The sequential Optimization Heuristic Method is applied to obtain Nash Equilibrium in the second stage. From the methodological standpoint, it provides a prototype for solving two-stage sequential games.

From a practical standpoint, it demonstrates that it is necessary for researchers to incorporate the responses from other major players in the system when optimizing the transit system. The results demonstrate that the optimal frequencies and transit fare is different when considering the parking operator's response. The positive impact on systemwide passenger travel cost under the optimal solution is also overestimated when assuming the parking fees and parking capacities will remain unchanged. Two different transit pricing strategies (upfront pricing and flexible second-stage pricing strategy) are evaluated. We find that the flexible second-stage pricing strategy provides a solution with a lower systemwide passenger travel cost compared to the upfront pricing strategy. Interestingly, it shows that pricing flexibility will actually put the transit operator at disadvantage. This argument is also proved theoretically. Our model framework enables the evaluation of different government policies like subsidies, taxation, road pricing, etc.

This research leads to numerous future directions of research. It provides a model framework to optimize transit operations considering the response from the parking operator. The framework can be easily extended to other transportation systems where competition happens between other operators. For example, the interactions between transit and ride-hailing system, and between private vehicle driving and autonomous vehicle ride-hailing systems can be potentially analyzed using this framework. Furthermore, more government regulations and policies can be evaluated using this framework. For example, the

impacts of road space rationing in Beijing, and congestion pricing in NYC can be evaluated using this framework.

5.2 Chapter 3: Optimizing Transit Network Planning and Local On-demand Services in Transit Desert Regions

This chapter aims to optimize the transit network, frequencies as well as fleet size of on-demand services to provide a solution to the transit desert problem. Commuters can choose from four travel modes, fixed-route transit, on-demand services, multi-modal services, and outside options, depending on the attractiveness of each travel mode.

From a methodological standpoint, we develop a mixed-integer non-linear programming model to optimize the transit network, frequencies, and fleet size of on-demand services. We manage to incorporate the complicated ride-sharing service process into our model in a tractable manner. In order to solve the non-linear optimization model, a two-step heuristic solution approach is developed. In the first step, we solve the transit network design problem and determine the transit lines. In the second step, we determine the transit frequencies and fleet size. In the first step, we apply rounding heuristics using callback functions in Gurobi. We also compare our approach with the other rounding heuristic which runs the linear relaxation model to optimality and then applies rounding heuristic. We argue that our approach is faster and more effective. This research provides an alternative way to tackle complicated mixed-integer non-linear programming models by using callback functions.

Practically, we apply our framework to the greater Boston city. We first identify the inequality of commuting costs among different regions. Our proposed solution managed to reduce the overall social costs and travel disutility. The overall social costs are reduced by 2.76%, 3.84%, and 6.96% in Woburn, Salem, and Waltham respectively. The commuter travel disutility also decreases by 5.73%, 6.57%, and 9.56% respectively. We also find that

our solution disproportionately benefits low-income commuters and commuters with no cars. A cheaper on-demand service benefits commuters with no private vehicles the most. However, it has a more complicated effect on environmental costs. The city planner needs more take specific travel patterns into consideration when making policies.

There are several future research directions. First, we do not consider road congestion in this paper. Congestion can be incorporated into the framework in future studies. Secondly, we significantly simplify the on-demand service operations in the current model. Future studies can also focus on how different dispatching, matching, and routing rules can influence the systemwide cost. Thirdly, we only consider low-capacity on-demand service vehicles to serve as feeders to transit lines. There are other forms of integration that are worth investigating. For example, high-capacity on-demand vehicles can be used as feeders to the system. Flexible and hybrid transit should also be investigated as a solution to transit desert problems.

5.3 Chapter 4: Bikes and Buses: A Heuristic Adaptive Discretization Scheme for Multimodal Network Design

This chapter aims to co-optimize the transit system as well as the bike-sharing system deciding the optimal transit frequencies, bike-sharing station location, number of bike-sharing docks and rebalancing operations to maximize the saved vehicle miles traveled by private vehicles. The passengers can choose from four different travel modes: fixed-route transit, bike-sharing services, multi-modal services, and outside options based on the attractiveness of each travel mode.

From the methodological perspective, we are the first paper to formulate a mixed-integer non-linear optimization model to optimize both transit frequencies and bike-sharing

operations decisions like bike station location and rebalancing operations. We also manage to solve this complicated model by adopting the adaptive discretization method, coordinate descent method, and anchor point parsimony method. By comparing our results to the optimal solution, we are able to achieve an optimality gap of less than 0.5% while reducing the runtime by more than 95%. The framework is tested using data from the city of Boston. We develop 6 study instances with increasing sizes. Our method is able to finish within 24 hours. It demonstrates our ability to provide a near-optimal solution in a reasonable time for large-scale realistic cases.

From a practical standpoint, we provide the optimal transit frequencies and bike-sharing operations in Boston city. We are able to reduce the total vehicle miles traveled by private vehicles by more than 5%. We also further demonstrate how the optimal solution changes under different budget levels. The integration of bike-sharing and transit systems plays a greater role in longer-distance travel. It encourages more people to switch from a private vehicle driving.

This research generates several future directions. First of all, we only consider bike-sharing system in this paper, but this model framework can be easily extended to other micro-mobility services like scooters and e-bikes. Secondly, we have only considered dock-based bike-sharing system in this paper, future studies can further investigate the co-optimization of dockless micro-mobility services. Thirdly, our model does not optimize the transit network. Future research can further integrate the transit system and the bike-sharing system by adjusting the current transit lines. It has the potential to provide more mobility in the transit desert regions inside the city. Lastly, there are other micro-mobility infrastructures decisions that need to be optimized in the integrated system. For example, what's the optimal bike lane network to maximize biking experience? Where to locate the charging facilities if e-bikes and e-scooters are adopted? These are some questions needing to be addressed in future works.

References

A. S. Abdelfatah and M. A. Taha. Parking capacity optimization using linear programming. *Journal of Traffic and Logistics Engineering*, 2.3:176–181, 2014.

T. Achterberg and E. Towle. Non-convex quadratic optimization, May 2020. URL https://www.gurobi.com/resource/non-convex-quadratic-optimization/.

N. Adler. Competition in a deregulated air transportation market. *European Journal of Operational Research*, 129(2):337–345, 2001.

N. Adler. Hub-spoke network choice under competition with an application to Western Europe. *Transportation Science*, 39(1):58–72, 2005.

R. O. Arbex and C. B. da Cunha. Efficient transit network design and frequencies setting multi-objective optimization by alternating objective genetic algorithm. *Transportation Research Part B: Methodological*, 81:355–376, 2015. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2015.06.014. URL https://www.sciencedirect.com/science/article/pii/S0191261515001435. Optimization of Urban Transportation Service Networks.

R. Arnott. Spatial competition between parking garages and downtown parking policy. *Transport Policy*, 13(6):458–469, 2006.

R. Arnott and E. Inci. An integrated model of downtown parking and traffic congestion.

Journal of Urban Economics, 60(3):418 – 442, 2006. ISSN 0094-1190. doi: https://doi.org/10.1016/j.jue.2006.04.004. URL http://www.sciencedirect.com/science/article/pii/S0094119006000386.

R. Arnott and J. Rowse. Modeling parking. *Journal of Urban Economics*, 45(1):97 – 124, 1999. ISSN 0094-1190. doi: https://doi.org/10.1006/juec.1998.2084. URL http://www.sciencedirect.com/science/article/pii/S0094119098920848.

S. Asadi Bagloee and A. A. Ceder. Transit-network design methodology for actual-size road networks. *Transportation Research Part B: Methodological*, 45(10):1787–1804, 2011. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2011.07.005. URL https://www.sciencedirect.com/science/article/pii/S0191261511001056.

K. A. Azari, S. Arintono, H. Hamid, and S. R. Davoodi. Evaluation of demand for different trip purposes under various congestion pricing scenarios. *Journal of Transport Geography*, 29:43 – 51, 2013.

L. J. Basso and S. R. Jara-Díaz. Integrating congestion pricing, transit subsidies and mode choice. *Transportation Research Part A: Policy and Practice*, 46(6):890–900, 2012.

D. Bertsimas, Y. Sian Ng, and J. Yan. Joint frequency-setting and pricing optimization on multimodal transit networks at scale. *Transportation Science*, 54(3):839–853, 2020. doi: 10.1287/trsc.2019.0959. URL https://doi.org/10.1287/trsc.2019.0959.

O. Besbes, F. Castro, and I. Lobel. Surge pricing and its spatial supply response. *Management Science*, 67(3):1350–1367, 2021. doi: 10.1287/mnsc.2020.3622.

BigBlueBus. Big blue bus partners with lyft to transform dial-a-ride program to on-demand service for older adults living in santa monica, Jun 2018. URL https://www.bigbluebus.com/Newsroom/Press/Big-Blue-Bus-Partners-

with-Lyft-to-Transform-Dial-A-Ride-Program-to--On-Demand-Service-for-Older-Adults-Living-in-Santa-Monica.aspx?type=Press.

L. Bliss. How much traffic do uber and lyft cause?, Aug 2019. URL https://www.bloomberg.com/news/articles/2019-08-05/uber-and-lyft-admit-they-re-making-traffic-worse.

M. Blodgett, A. Khani, D. Negoescu, and S. Benjaafar, Aug 2017. URL https://conservancy.umn.edu/bitstream/handle/11299/192846/MCOTA%2017-04.pdf?sequence=1&isAllowed=y.

Bluebikes. Bluebikes system data, 2022. URL https://www.bluebikes.com/system-data.

Boston Transportation Department. Parking in boston, Sep 2017. URL https://www.boston.gov/sites/default/files/document-file-09-2017/eviction$_g$uide$_2$017.pdf.

L. A. Bowman and M. A. Turnquist. Service frequency, schedule reliability and passenger wait times at transit stops. *Transportation Research Part A: General*, 15(6):465–471, 1981. doi: 10.1016/0191-2607(81)90114-x.

B. P. Bruck, F. Cruz, M. Iori, and A. Subramanian. The static bike sharing rebalancing problem with forbidden temporary operations. *Transportation Science*, 53(3): 882–896, 2019. doi: 10.1287/trsc.2018.0859. URL https://doi.org/10.1287/trsc.2018.0859.

A. T. Buba and L. S. Lee. A differential evolution for simultaneous transit network design and frequency setting problem. *Expert Systems with Applications*, 106:277–289, 2018. doi: 10.1016/j.eswa.2018.04.011.

T. Bulhões, A. Subramanian, G. Erdoğan, and G. Laporte. The static bike relocation problem with multiple vehicles and visits. *European Journal of Operational Research*, 264(2):508–523, 2018. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2017.06.028. URL https://www.sciencedirect.com/science/article/pii/S0377221717305672.

. Bureau of Public Roads. Traffic assignment manual for application with a large, high speed computer. 1964.

S. Burer and A. N. Letchford. Non-convex mixed-integer nonlinear programming: A survey. *Surveys in Operations Research and Management Science*, 17(2):97–106, 2012.

A. Burke. World ghg emission flowchart, Dec 2010. URL https://climateinsight.wordpress.com/science/greenhouse-gases/world-flowchart/.

L. Caggiani, A. Colovic, and M. Ottomanelli. An equality-based model for bike-sharing stations location in bicycle-public transport multimodal mobility. *Transportation Research Part A: Policy and Practice*, 140:251–265, 2020. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2020.08.015. URL https://www.sciencedirect.com/science/article/pii/S0965856420306972.

D. Canca, A. De-Los-Santos, G. Laporte, and J. A. Mesa. Integrated railway rapid transit network design and line planning problem with maximum profit. *Transportation Research Part E: Logistics and Transportation Review*, 127:1–30, 2019. ISSN 1366-5545. doi: https://doi.org/10.1016/j.tre.2019.04.007. URL https://www.sciencedirect.com/science/article/pii/S1366554518309979.

J. Castillo, D. T. Knoepfle, and E. G. Weyl. Surge pricing solves the wild goose chase. *SSRN Electronic Journal*, 2017. doi: 10.2139/ssrn.2890666.

J. Cavadas and A. Antunes. An optimization model for integrated transit-parking policy planning. *Transportation*, pages 1–25, 2018.

CBS. Mta restoring bus routes cut in 2010, Jan 2013. URL https://www.cbsnews.com/newyork/news/mta-restoring-bus-routes-cut-in-2010/.

CBS. First cta fare increases in 9 years takes effect sunday, Jan 2018. URL https://www.cbsnews.com/chicago/news/first-cta-fare-increases-in-9-years-takes-effect-sunday/.

S. Chandra and L. Quadrifoglio. A model for estimating the optimal cycle length of demand responsive feeder transit services. *Transportation Research Part B: Methodological*, 51:1–16, 2013. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2013.01.008. URL https://www.sciencedirect.com/science/article/pii/S0191261513000192.

P. W. Chen and Y. M. Nie. Analysis of an idealized system of demand adaptive paired-line hybrid transit. *Transportation Research Part B: Methodological*, 102:38–54, 2017. doi: 10.1016/j.trb.2017.05.004.

P. W. Chen and Y. M. Nie. Optimal design of demand adaptive paired-line hybrid transit: Case of radial route structure. *Transportation Research Part E: Logistics and Transportation Review*, 110:71–89, 2018. doi: 10.1016/j.tre.2017.12.006.

S. Chen. Who's afraid of a transit desert? *The New York Times*, 2019. URL https://www.nytimes.com/2019/10/11/realestate/whos-afraid-of-a-transit-desert.html.Accessed30Aug2021.

E. Cipriani, S. Gori, and M. Petrelli. Transit network design: A procedure and an application to a large urban area. *Transportation Research Part C:*

Emerging Technologies, 20(1):3–14, 2012a. ISSN 0968-090X. doi: https://doi.org/10.1016/j.trc.2010.09.003. URL https://www.sciencedirect.com/science/article/pii/S0968090X10001397. Special issue on Optimization in Public Transport+ISTT2011.

E. Cipriani, S. Gori, and M. Petrelli. A bus network design procedure with elastic demand for large urban areas. *Public Transport*, 4(1):57–76, July 2012b. ISSN 1613-7159. doi: 10.1007/s12469-012-0051-7. URL https://doi.org/10.1007/s12469-012-0051-7.

City of Boston. Parking freezes, Apr 2017. URL https://www.boston.gov/departments/environment/air-pollution-control-commission/parking-freezes#.

E. Codina and F. Rosell. A heuristic method for a congested capacitated transit assignment model with strategies. *Transportation Research Part B: Methodological*, 106: 293–320, 2017. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2017.07.008. URL https://www.sciencedirect.com/science/article/pii/S0191261517303211.

COVID Mobility Works. 130 km of temporary bicycle lanes to support safe travel in mexico city, Sep 2020. URL https://www.covidmobilityworks.org/responses/130-km-of-temporary-bicycle-lanes-to-support-safe-travel-in-mexico-city-2ea581c6cb.

F. Creutzig, P. Jochem, O. Y. Edelenbosch, L. Mattauch, D. P. Vuuren, D. McCollum, and J. Minx. Transport: A roadblock to climate change mitigation? *Science*, 350(6263): 911–912, 2015. doi: 10.1126/science.aac8033.

J. Cuba. State bill would eliminate parking minimums in the citynbsp;, Jan

2022. URL https://nyc.streetsblog.org/2022/01/31/state-bill-would-eliminate-parking-minimums-in-the-city/#.

C. F. Daganzo and Y. Ouyang. A general model of demand-responsive transportation services: From taxi to ridesharing to dial-a-ride. *Transportation Research Part B: Methodological*, 126:213–224, 2019. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2019.06.001. URL https://www.sciencedirect.com/science/article/pii/S0191261518307793.

S. Daraei, K. Pelechrinis, and D. Quercia. A data-driven approach for assessing biking safety in cities. *EPJ Data Science*, 10(1), 2021. doi: 10.1140/epjds/s13688-021-00265-y.

A. de Bortoli. Environmental performance of shared micromobility and personal alternatives using integrated modal lca. *Transportation Research Part D: Transport and Environment*, 93:102743, 2021. doi: 10.1016/j.trd.2021.102743.

A. de Palma, S. Vosough, and F. Liao. An overview of effects of covid-19 on mobility and lifestyle: 18 months since the outbreak. *Transportation Research Part A: Policy and Practice*, 159:372–397, 2022. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2022.03.024. URL https://www.sciencedirect.com/science/article/pii/S0965856422000714.

G. de Rus and M. P. Socorro. Pricing and investment in alternative transport infrastructures. *Transportation Research Part A: Policy and Practice*, 119:96 – 107, 2019. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2018.10.040. URL http://www.sciencedirect.com/science/article/pii/S0965856417308315.

N. DeCosta-Klipa. Mbta fares have gone up. here's what that means for you., Jun 2016. URL https://www.boston.com/news/local-news/2016/06/30/mbta-fares-up-heres-means/.

M. Dell'Amico, M. Iori, S. Novellani, and A. Subramanian. The bike sharing rebalancing problem with stochastic demands. *Transportation Research Part B: Methodological*, 118:362–380, 2018. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2018.10.015. URL https://www.sciencedirect.com/science/article/pii/S0191261518301711.

M. Diao, H. Kong, and J. Zhao. Impacts of transportation network companies on urban mobility. *Nature Sustainability*, 4(6):494–500, 2021. doi: 10.1038/s41893-020-00678-z.

J. Dill and N. McNeil. Four types of cyclists?: Examination of typology for better understanding of bicycling behavior and potential. *Transportation Research Record*, 2387(1): 129–138, 2013. doi: 10.3141/2387-15. URL https://doi.org/10.3141/2387-15.

EBP. The impact of the covid-19 pandemic on public transit funding needs in the u.s., Jan 2021. URL https://www.apta.com/research-technical-resources/research-reports/the-impact-of-the-covid-19-pandemic-on-public-transit-funding-needs-in-the-u-s/.

H. R. Eftekhari and M. Ghatee. The lower bound for dynamic parking prices to decrease congestion through CBD. *Operational Research*, 17(3):761–787, Oct. 2017. ISSN 1866-1505. doi: 10.1007/s12351-016-0238-9. URL https://doi.org/10.1007/s12351-016-0238-9.

G. D. Erhardt, S. Roy, D. Cooper, B. Sana, M. Chen, and J. Castiglione. Do transportation network companies decrease or increase congestion? *Science Advances*, 5(5), 2019. doi: 10.1126/sciadv.aau2670.

W. Fan and R. B. Machemehl. Tabu search strategies for the public transportation network

optimizations with variable transit demand. *Computer-Aided Civil and Infrastructure Engineering*, 23(7):502–520, 2008. doi: 10.1111/j.1467-8667.2008.00556.x.

L. Feng, J. Xie, Y. M. Nie, and X. Liu. Efficient algorithm for the traffic assignment problem with side constraints. *Transportation Research Record*, 2674(4):129–139, 2020. doi: 10.1177/0361198120912234. URL https://doi.org/10.1177/0361198120912234.

X. Feng, X. Zhu, X. Qian, Y. Jie, F. Ma, and X. Niu. A new transit network design study in consideration of transfer time composition. *Transportation Research Part D: Transport and Environment*, 66:85–94, 2019. ISSN 1361-9209. doi: https://doi.org/10.1016/j.trd.2018.03.019. URL https://www.sciencedirect.com/science/article/pii/S1361920917306107. Special Issue on Electromobility for Green Transportation Systems and Sustainable Environment.

A. D. Fernandez. Dart partners with uber to provide free and low cost rides to commuters, Mar 2019. URL https://www.wfaa.com/article/news/dart-partners-with-uber-to-provide-free-and-low-cost-rides-to-commuters/287-02481cce-3c6b-4eb6-99c4-3d56f3a5996a.

B. Ford. To boost ridership, chicago plans to slash transit fares, Oct 2021. URL https://www.bloomberg.com/news/articles/2021-10-25/why-chicago-plans-to-cut-fares-on-public-transportation.

I. A. Forma, T. Raviv, and M. Tzur. A 3-step math heuristic for the static repositioning problem in bike-sharing systems. *Transportation Research Part B: Methodological*, 71:230–247, 2015. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2014.10.003. URL https://www.sciencedirect.com/science/article/pii/S0191261514001726.

I. Frade and A. Ribeiro. Bike-sharing stations: A maximal covering location ap-

proach. *Transportation Research Part A: Policy and Practice*, 82:216–227, 2015. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2015.09.014. URL https://www.sciencedirect.com/science/article/pii/S0965856415002487.

S. F. Franco. Downtown parking supply, work-trip mode choice and urban spatial structure. *Transportation Research Part B: Methodological*, 101:107–122, 2017. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2017.03.012. URL https://www.sciencedirect.com/science/article/pii/S0191261516306750.

Y. Freemark. Us public transit has struggled to retain riders over the past half century. reversing this trend could advance equity and sustainability., Jun 2021. URL https://www.urban.org/urban-wire/us-public-transit-has-struggled-retain-riders-over-past-half-century-reversing-trend-could-advance-equity-and-sustainability.

C. Fu, N. Zhu, S. Ma, and R. Liu. A two-stage robust approach to integrated station location and rebalancing vehicle service design in bike-sharing systems. *European Journal of Operational Research*, 298(3):915–938, 2022. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2021.06.014. URL https://www.sciencedirect.com/science/article/pii/S0377221721005294.

G. Gallego, R. Ratliff, and S. Shebalov. A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research*, 63(1):212–232, 2015. doi: 10.1287/opre.2014.1328. URL https://doi.org/10.1287/opre.2014.1328.

R. García and A. Marín. Parking capacity and pricing in park'n ride trips: A continuous equilibrium network design problem. *Annals of Operations Research*, 116(1):153–178, 2002. doi: 10.1023/A:1021332414941. URL https://doi.org/10.1023/A:1021332414941.

J. C. García-Palomares, J. Gutiérrez, and M. Latorre. Optimizing the location of stations in bike-sharing programs: A gis approach. *Applied Geography*, 35(1):235–246, 2012. ISSN 0143-6228. doi: https://doi.org/10.1016/j.apgeog.2012.07.002. URL https://www.sciencedirect.com/science/article/pii/S0143622812000744.

K. Gkiotsalitis and O. Cats. Public transport planning adaption under the covid-19 pandemic crisis: Literature review of research needs and directions. *Transport Reviews*, 41 (3):374–392, 2020. doi: 10.1080/01441647.2020.1857886.

A. Gordon. Uber and lyft say they reduce car ownership, but that might not be true, Jan 2021. URL https://www.vice.com/en/article/n7v4kd/uber-and-lyft-say-they-reduce-car-ownership-but-that-might-not-be-true.

J. Greenfield and C. Cobbs. Despite a 400% fleet increase, total scooter use dropped in this year's pilot, Dec 2020. URL https://chi.streetsblog.org/2020/12/11/despite-a-400-fleet-increase-total-scooter-use-dropped-in-this-years-pilot/.

Q. Guo, Y. Sun, P. Schonfeld, and Z. Li. Time-dependent transit fare optimization with elastic and spatially distributed demand. *Transportation Research Part A: Policy and Practice*, 148:353–378, 2021. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2021.04.002. URL https://www.sciencedirect.com/science/article/pii/S096585642100094X.

Gurobi. Mixed-integer programming (mip) - a primer on the basics, Jun 2019. URL https://www.gurobi.com/resource/mip-basics/.

Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2022a. URL https://www.gurobi.com.

Gurobi Optimization, LLC. Non-Convex Quadratic Optimization, 2022b. URL https://www.gurobi.com/resource/non-convex-quadratic-optimization/.

Gurobi Optimization, LLC. Solving Bilinear Programming Problems, 2022c. URL https://www.gurobi.com/resource/solving-bilinear-programming-problems/.

C. Guse. Reduce nyc congestion by charging uber, lyft when their cars cruise empty in manhattan, says city council report, Jan 2021. URL https://www.nydailynews.com/new-york/ny-council-report-fhv-uber-surcharge-per-minute-20210112-npeh54zshjedfkh2v5eh2ohhfq-story.html.

C. GUSE. Mta leaders mull big train, bus, subway service cuts for 2023; expect transit ridership won't hit pre-pandemic levels. *New York Daily News*, 2021. URL https://www.nydailynews.com/new-york/ny-mta-mulls-big-service-cuts-despite-fed-bailout-20210721-afkxnz5gjnggzn3e2zcpl4dasy-story.html.

M. Hansen. Airline competition in a hub-dominated environment: An application of non-cooperative game theory. *Transportation Research Part B: Methodological*, 24(1):27–43, 1990.

R. Harder and V. Vaze. Two-stage game theoretic modelling of airline frequency and fare competition. *ACM SIGMETRICS Performance Evaluation Review*, 44(3):21–21, 2017.

R. Hemmecke, M. Köppe, J. Lee, and R. Weismantel. Nonlinear integer programming. In *50 Years of Integer Programming 1958-2008*, pages 561–618. Springer, 2010.

E. Heraldo. Transporte público debe operar a un 35% de su capacidad: Duque, Apr 2020. URL https://www.elheraldo.co/colombia/transporte-publico-debe-operar-un-35-de-su-capacidad-duque-719138.

D. B. Hess. Effect of free parking on commuter mode choice: Evidence from travel diary data. *Transportation Research Record*, 1753(1):35–42, 2001. doi: 10.3141/1753-05. URL https://doi.org/10.3141/1753-05.

S. Holder. When commuting comes back, the e-scooters will be ready, Jan 2021. URL https://www.bloomberg.com/news/articles/2021-01-22/e-scooters-get-ready-for-the-commuting-comeback.

D. Huang, Z. Liu, P. Liu, and J. Chen. Optimal transit fare and service frequency of a non-linear origin-destination based fare structure. *Transportation Research Part E: Logistics and Transportation Review*, 96:1 – 19, 2016. ISSN 1366-5545. doi: https://doi.org/10.1016/j.tre.2016.10.004. URL http://www.sciencedirect.com/science/article/pii/S1366554516304501.

ITS. A dynamic time-of-day parking meter pricing system in los angeles increased revenues by 2.5 percent and lowered the average parking meter rate by $0.19 per hour., Aug 2017. URL https://www.itskrs.its.dot.gov/its/benecost.nsf/ID/a756406f684aa9588525817600521371.

Y. Jiang, T. K. Rasmussen, and O. A. Nielsen. Integrated optimization of transit networks with schedule- and frequency-based services subject to the bounded stochastic user equilibrium. *Transportation Science*, 0(0):null, 2022. doi: 10.1287/trsc.2022.1148. URL https://doi.org/10.1287/trsc.2022.1148.

J. Jiao and C. Bischak. People are stranded in 'transit deserts' in dozens of us cities. 2018. URL https://theconversation.com/people-are-stranded-in-transit-deserts-in-dozens-of-us-cities-92722.

J. Jiao and M. Dillivan. Transit deserts: The gap between demand and supply. *Journal of Public Transportation*, 16(3):23–39, 2013. doi: 10.5038/2375-0901.16.3.2.

M. Ju. Micromobility toward the future – an opportunity arose in the pandemic, Mar 2022. URL https://www.enotrans.org/article/guest-op-ed-micromobility-toward-the-future-an-opportunity-arose-in-the-pandemic/.

S. Jäppinen, T. Toivonen, and M. Salonen. Modelling the potential effect of shared bicycles on public transport travel times in greater helsinki: An open data approach. *Applied Geography*, 43:13–24, 2013. ISSN 0143-6228. doi: https://doi.org/10.1016/j.apgeog.2013.05.010. URL https://www.sciencedirect.com/science/article/pii/S014362281300132X.

R. Kager, L. Bertolini, and M. Te Brömmelstroet. Characterisation of and reflections on the synergy of bicycles and public transport. *Transportation Research Part A: Policy and Practice*, 85:208–219, 2016. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2016.01.015. URL https://www.sciencedirect.com/science/article/pii/S0965856416000240.

J. Kallgren. Pop-up bike lanes in berlin are here to stay, Sep 2021. URL https://www.euronews.com/next/2021/09/02/berlin-s-pop-up-bike-lanes-made-permanent.

P. N. Kechagiopoulos and G. N. Beligiannis. Solving the urban transit routing problem using a particle swarm optimization based algorithm. *Applied Soft Computing*, 21: 654–676, 2014. doi: 10.1016/j.asoc.2014.04.005.

S. Kimes. Revenue management: A retrospective. *Cornell Hotel and Restaurant Administration Quarterly*, 44(5):131–138, 2003.

N. Korolko, D. Woodard, C. Yan, and H. Zhu. Dynamic pricing and matching in ride-hailing platforms. *SSRN Electronic Journal*, 2018. doi: 10.2139/ssrn.3258234.

A. Lee and M. Savelsbergh. An extended demand responsive connector. *EURO Journal on Transportation and Logistics*, 6(1):25–50, 2017. ISSN 2192-4376. doi: https://doi.org/10.1007/s13676-014-0060-6. URL https://www.sciencedirect.com/science/article/pii/S2192437620300881.

C. Li, J. Ma, T. H. Luan, X. Zhou, and L. Xiong. An incentive-based optimizing strategy of service frequency for an urban rail transit system. *Transportation Research Part E: Logistics and Transportation Review*, 118:106–122, 2018. ISSN 1366-5545. doi: https://doi.org/10.1016/j.tre.2018.07.005. URL https://www.sciencedirect.com/science/article/pii/S1366554517307433.

Z.-C. Li, W. H. K. Lam, and S. C. Wong. The Optimal Transit Fare Structure under Different Market Regimes with Uncertainty in the Network. *Networks and Spatial Economics*, 9(2):191–216, June 2009. ISSN 1572-9427. doi: 10.1007/s11067-007-9058-z. URL https://doi.org/10.1007/s11067-007-9058-z.

Z.-C. Li, W. H. Lam, and S. Wong. Modeling intermodal equilibrium for bimodal transportation system design problems in a linear monocentric city. *Transportation Research Part B: Methodological*, 46(1):30 – 49, 2012. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2011.08.002. URL http://www.sciencedirect.com/science/article/pii/S0191261511001238.

J. Liang, J. Wu, Z. Gao, H. Sun, X. Yang, and H. K. Lo. Bus transit network design with uncertainties on the basis of a metro network: A two-step model framework. *Transportation Research Part B: Methodological*, 126:115–138, 2019. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2019.05.011. URL https://www.sciencedirect.com/science/article/pii/S0191261518306969.

Z. Liao. Real-time taxi dispatching using global positioning systems. *Communications of the ACM*, 46(5):81–83, 2003. doi: 10.1145/769800.769806.

J. Lindsey, T. Engel, M. Pskowski, and T. Heming. The pandemic bike boom is here to stay, Jun 2021. URL https://www.outsideonline.com/outdoor-gear/bikes-and-biking/pandemic-bike-boom-here-stay/.

R. Lindsey. Road pricing and investment. *Economics of Transportation*, 1(1):49 – 63, 2012. ISSN 2212-0122. doi: https://doi.org/10.1016/j.ecotra.2012.07.001. URL http://www.sciencedirect.com/science/article/pii/S2212012212000044.

H. Liu and X. Yang. Bus transit route network design using genetic algorithm. *International Conference on Transportation Engineering 2007*, 2007. doi: 10.1061/40932(246)187.

W. Liu and N. Geroliminis. Modeling the morning commute for urban networks with cruising-for-parking: An mfd approach. *Transportation Research Part B: Methodological*, 93:470 – 494, 2016. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2016.08.004. URL http://www.sciencedirect.com/science/article/pii/S0191261516305677.

Y. Liu and Y. Ouyang. Mobility service design via joint optimization of transit networks and demand-responsive services. *Transportation Research Part B: Methodological*, 151:22–41, 2021. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2021.06.005. URL https://www.sciencedirect.com/science/article/pii/S0191261521001168.

Y. Liu, X. Guo, and H. Yang. Pareto-improving and revenue-neutral congestion pricing schemes in two-mode traffic networks. *NETNOMICS: Economic Research and Electronic Networking*, 10(1):123–140, Apr. 2009. ISSN 1573-7071. doi: 10.1007/s11066-008-9018-x. URL https://doi.org/10.1007/s11066-008-9018-x.

R. LoBasso. Mlk drive: Closed to motor vehicles, open to (socially-distanced) people, Mar 2021. URL https://bicyclecoalition.org/mlk-drive-closed-to-motor-vehicles-open-to-socially-distanced-people/.

Los Angeles Times. 'last train': New york subway shutdown because of coronavirus is another historic first, May 2020. URL https://www.latimes.com/world-nation/story/2020-05-11/in-new-yorks-subway-shutdown-an-unthinkable-departure.

R. Lovelace, S. Beck, M. Watson, and A. Wild. Assessing the energy implications of replacing car trips with bicycle trips in sheffield, uk. *Energy Policy*, 39(4):2075–2087, 2011. doi: 10.1016/j.enpol.2011.01.051.

S. Luo and Y. Nie. Impact of ride-pooling on the nature of transit network design. *Transportation Research Part B: Methodological*, 129:175–192, 2019. doi: 10.1016/j.trb.2019.09.007.

S. Manson, J. Schroeder, D. V. Riper, T. Kugler, and S. Ruggles. Ipums national historical geographic information system: Version 15.0 [dataset]. 2020. URL http://doi.org/10.18128/D050.V15.0.

J. C. Martín and C. Román. Hub location in the south-atlantic airline market: A spatial competition game. *Transportation Research Part A: Policy and Practice*, 37(10):865–888, 2003.

massDOT. Focus 40 The 2040 Investment Plan for the MBTA — Focus40: The Plan, Mar. 2019. URL https://www.mbtafocus40.com/focus40theplan.

MassDOT. MassDOT Open Data Portal, 2021. URL https://geo-massdot.opendata.arcgis.com/.

G. Mazzuoccolo and L. Mella. On the ratio between the maximum weight of a perfect matching and the maximum weight of a matching. *Discrete Applied Mathematics*, 301:19–25, 2021. doi: 10.1016/j.dam.2021.05.009.

MBTA. GTFS | Developers | MBTA, 2021a. URL https://www.mbta.com/developers/gtfs.

MBTA. The ride flex, 2021b. URL https://www.mbta.com/accessibility/the-ride/on-demand-pilot.

J. McGill and G. van Ryzin. Revenue management: Research overview and prospects. *Transportation Science*, 33(2):233–256, 1999.

L. Metro. Can public transit and tncs get along? expanding the reach of transit with lyft, Mar 2020. URL https://thesource.metro.net/2020/03/09/can-public-transit-and-tncs-get-along-expanding-the-reach-of-transit-with-lyft/.

D. Meyer. Mta service cuts still on the table even after billions in federal aid. *New York Post*, 2021. URL https://nypost.com/2021/07/21/mta-service-cuts-still-on-the-table-even-after-billions-in-federal-aid/.

P. Miller. *Sustainability and Public Transportation Theory and Analysis*. PhD thesis, University of Calgary, 2014.

B. Milligan. Tc eyeing higher, fluctuating parking rates based on demand, Sep 2020. URL https://www.traverseticker.com/news/tc-eyeing-higher-fluctuating-parking-rates-based-on-demand/.

R. Mix, R. Hurtubia, and S. Raveau. Optimal location of bike-sharing stations: A built environment and accessibility approach. *Transportation Research Part A: Policy and Practice*, 160:126–142, 2022. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2022.03.022. URL https://www.sciencedirect.com/science/article/pii/S0965856422000696.

MTA. 2010 service reduction implementation plan, Mar 2010. URL http://web.mta.info/mta/news/books/docs/service%20implementation%20presentation%205-24-10%20%28CAB1172%29.pdf.

National Geographic. Greendex 2009: Consumer choice and the environment – a worldwide ..., May 2009. URL https://www.nationalgeographic.com/greendex/assets/Greendex$_{H}$ighlights$_{R}$eport$_{M}$ay09.pdf.

N. Nebojsa. page 621. Cambridge University Press, published for the Intergovernmental Panel on Climate Change, 2000.

NeighborhoodX Boston. Boston proper, NA. URL http://boston.neighborhoodx.com/lists/index?g=144.

J. Nicas. Now prices can change from minute to minute. *The Wall Street Journal*, 2015.

M. Nourinejad and J. M. Roorda. Impact of hourly parking pricing on travel demand. *Transportation Research Part A: Policy and Practice*, 98:28 – 45, 2017. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2017.01.023. URL http://www.sciencedirect.com/science/article/pii/S0965856415303396.

G. Oeschger, P. Carroll, and B. Caulfield. Micromobility and public transport integration: The current state of knowledge. *Transportation Research Part D: Transport and Environment*, 89:102628, 2020. ISSN 1361-9209. doi: https://doi.org/10.1016/j.trd.2020.102628. URL https://www.sciencedirect.com/science/article/pii/S1361920920308130.

N. Oliker and S. Bekhor. A frequency based transit assignment model that considers online information at the boarding stop. *2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC)*, 2016. doi: 10.1109/itsc.2016.7795552.

R. O'Toole. Transit: The urban parasite, Apr 2020. URL https://www.cato.org/policy-analysis/transit-urban-parasite.

A. Pal and Y. Zhang. Free-floating bike sharing: Solving real-life large-scale static rebalancing problems. *Transportation Research Part C: Emerging Technologies*, 80:92–116, 2017. ISSN 0968-090X. doi: https://doi.org/10.1016/j.trc.2017.03.016. URL https://www.sciencedirect.com/science/article/pii/S0968090X17300992.

M. Patriksson. *The traffic assignment problem: Models and methods*. Dover Publications, Inc., 2015.

N. Périvier, C. Hssaine, S. Samaranayake, and S. Banerjee. Real-time approximate routing for smart transit systems. *Proc. ACM Meas. Anal. Comput. Syst.*, 5(2), jun 2021. doi: 10.1145/3460091. URL https://doi.org/10.1145/3460091.

F. R. Proulx, B. Cavagnolo, and M. Torres-Montoya. Impact of parking prices and transit fares on mode choice at the university of california, berkeley. *Transportation Research Record*, 2469(1):41–48, 2014. doi: 10.3141/2469-05. URL https://doi.org/10.3141/2469-05.

L. Quadrifoglio and X. Li. A methodology to derive the critical demand density for designing and operating feeder transit services. *Transportation Research Part B: Methodological*, 43(10):922–935, 2009. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2009.04.003. URL https://www.sciencedirect.com/science/article/pii/S0191261509000502.

RTD. Rtd, uber transit debut on-demand service for access-a-ride customers, Apr 2021. URL https://www.rtd-denver.com/news-stop/news/rtd-uber-transit-debut-demand-service-access-ride-customers.

I. Savage. The dynamics of fare and frequency choice in urban transit. *Transportation Research Part A: Policy and Practice*, 44(10):815 – 829, 2010. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2010.08.002. URL http://www.sciencedirect.com/science/article/pii/S0965856410001230.

C. Schaller, Jul 2018. URL http://www.schallerconsult.com/rideservices/automobility.pdf.

S. Schechner. Why do gas station prices constantly change? blame the algorithm. *The Wall Street Journal*, 2017.

P. L. Schiller, E. C. Bruun, and J. R. Kenworthy. *An introduction to sustainable transportation: Policy, planning and implementation*. Earthscan, 2010.

J. Schuijbroek, R. Hampshire, and W.-J. van Hoeve. Inventory rebalancing and vehicle routing in bike sharing systems. *European Journal of Operational Research*, 257(3):992–1004, 2017. ISSN 0377-2217. doi: https://doi.org/10.1016/j.ejor.2016.08.029. URL https://www.sciencedirect.com/science/article/pii/S0377221716306658.

A. Schwedhelm, W. Li, L. Harms, and C. Adriazola-Steil. Biking provides a critical lifeline during the coronavirus crisis, Apr 2020. URL https://www.wri.org/insights/biking-provides-critical-lifeline-during-coronavirus-crisis.

Y. Sheffi. *Urban Transportation Networks: Equilibrium Analysis with mathematical programming methods*. Prentice-Hall, 1985.

Y. Shen, H. Zhang, and J. Zhao. Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in singapore. *Transportation Research Part A: Policy and Practice*, 113:125–136, 2018.

ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2018.04.004. URL https://www.sciencedirect.com/science/article/pii/S096585641730681X.

C. Shui and W. Szeto. Dynamic green bike repositioning problem – a hybrid rolling horizon artificial bee colony algorithm approach. *Transportation Research Part D: Transport and Environment*, 60:119–136, 2018. ISSN 1361-9209. doi: https://doi.org/10.1016/j.trd.2017.06.023. URL https://www.sciencedirect.com/science/article/pii/S136192091730500X. Special Issue on Traffic Modeling for Low-Emission Transport.

M. Staff. St. louis metro, via team for new on-demand service, Jun 2020. URL https://www.metro-magazine.com/10119937/st-louis-metro-via-team-for-new-on-demand-service?utm$_s$ource=email&utm$_m$edium=enewsletter&utm$_c$ampaign=20200702-NL-MET-Express-BOBCD200626002&omdt=NL-MET-Express&omid=1000479443&oly$_e$nc$_i$d=2561D4106245H1W.

statista. Ride-hailing taxi - worldwide: Statista market forecast, 2021. URL https://www.statista.com/outlook/mmo/mobility-services/ride-hailing-taxi/worldwide.

R. Stern. Passenger transfer system review. *Transit Cooperative Research Program (TCRP) Synthesis 19*, 1996. URL https://www.worldtransitresearch.info/research/2821/.

M. Stiglic, N. Agatz, M. Savelsbergh, and M. Gradisar. Enhancing urban mobility: Integrating ride-sharing and public transit. *Computers Operations Research*, 90: 12–21, 2018. ISSN 0305-0548. doi: https://doi.org/10.1016/j.cor.2017.08.016. URL https://www.sciencedirect.com/science/article/pii/S0305054817302228.

S. Sun and W. Szeto. Logit-based transit assignment: Approach-based formulation and paradox revisit. *Transportation Research Part B: Methodological*, 112:191–215, 2018. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2018.03.018. URL https://www.sciencedirect.com/science/article/pii/S0191261517306276.

S. Sun and W. Szeto. Optimal sectional fare and frequency settings for transit networks with elastic demand. *Transportation Research Part B: Methodological*, 127:147–177, 2019. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2019.06.011. URL https://www.sciencedirect.com/science/article/pii/S0191261518306532.

S. T. Syed, B. S. Gerber, and L. K. Sharp. Traveling towards disease: Transportation barriers to health care access. *Journal of Community Health*, 38(5):976–993, 2013. doi: 10.1007/s10900-013-9681-1.

W. Szeto and Y. Jiang. Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. *Transportation Research Part B: Methodological*, 67:235–263, 2014a. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2014.05.008. URL https://www.sciencedirect.com/science/article/pii/S0191261514000812.

W. Szeto and Y. Jiang. Transit route and frequency design: Bi-level modeling and hybrid artificial bee colony algorithm approach. *Transportation Research Part B: Methodological*, 67:235–263, 2014b. doi: 10.1016/j.trb.2014.05.008.

M. Tarduno. The congestion costs of uber and lyft. *Journal of Urban Economics*, 122: 103318, 2021. doi: 10.1016/j.jue.2020.103318.

Uber, Apr 2021. URL https://www.uber.com/global/en/price-estimate/.

United States Census Bureau. LEHD Origin-Destination Employment Statistics (LODES) Dataset Structure, Jan. 2020. URL https://lehd.ces.census.gov/data/lodes/LODES7/LODESTechDoc7.5.pdf. publisher: United States Census Bureau.

U.S. Census Bureau. Urban areas facts. 2010. URL https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/ua-facts.html.

U.S. Department of Transportation. Commute mode share: 2015, October 2016. URL https://www.bts.gov/content/commute-mode-share-2015.

V. Vaze and R. Harder. A game-theoretic modeling approach to air traffic forecasting. In *12th USA/Europe Air Traffic Management Research and Development Seminar, Seattle, WA*, 2017.

A. Walawalkar. Back to work: 'capacity of transport network will be down by 90%', May 2020. URL https://www.theguardian.com/world/2020/may/09/back-to-work-capacity-of-transport-network-will-be-down-by-90.

Y. Wan and A. Zhang. Urban road congestion and seaport competition. *Journal of Transport Economics and Policy*, 47(1):55–70, 2013. ISSN 00225258. URL http://www.jstor.org/stable/24396352.

M. Wanek-Libman. Miami-dade transit partnering with uber, lyft for late night essential ride options, Apr 2020. URL https://www.masstransitmag.com/alt-mobility/shared-mobility/article/21133166/miamidade-transit-partnering-with-uber-lyft-for-late-night-options?utm_source=MASS+NewsViews+Newsletter&utm_medium=email&utm_campaign=MASS200408003&oeid=0541B0916423H4E&rdx.ident%5Bpull%5D=omeda%7C0541B0916423H4E.

K. Wang, A. Jacquillat, and V. Vaze. Vertiport planning for urban aerial mobility: An adaptive discretization approach. *Manufacturing & Service Operations Management*, 24 (6):3215–3235, 2022.

W. W. Wang, D. Z. W. Wang, H. Sun, and J. Wu. Public transit service operation strategy under indifference thresholds-based bi-modal equilibrium. *Journal of Advanced Transportation*, 50(6):1124–1138, 2016. doi: 10.1002/atr.1393. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/atr.1393.

Z. Wang, J. Yu, W. Hao, T. Chen, and Y. Wang. Designing high-freedom responsive feeder transit system with multitype vehicles. *Journal of Advanced Transportation*, 2020:1–20, 2020. doi: 10.1155/2020/8365194.

K. Wei, V. Vaze, and A. Jacquillat. Transit planning optimization under ride-hailing competition and traffic congestion. *Transportation Science*, 0(0):null, 2021. doi: 10.1287/trsc.2021.1068. URL https://doi.org/10.1287/trsc.2021.1068.

Wikipedia. New york city transit fares, May 2022. URL https://en.wikipedia.org/wiki/New$_{\text{Y}}$ork$_{\text{C}}$ity$_{\text{t}}$ransit$_{\text{f}}$ares#Fare$_{\text{h}}$istory.

J. P. Williams. Stranded without transit. *U.S. News*, 2018. URL https://www.usnews.com/news/healthiest-communities/articles/2018-04-24/transit-deserts-a-growing-problem-in-the-us.

WMATA. Metro to close additional 17 stations and selected entrances amid low-ridership, need to conserve critical cleaning supplies during covid-19 response, Mar 2020. URL https://www.wmata.com/about/news/COVID-19-Service-Update-4-Station-Closures.cfm.

W. Wu, R. Liu, W. Jin, and C. Ma. Stochastic bus schedule coordination considering demand assignment and rerouting of passengers. *Transportation Research Part B: Methodological*, 121:275–303, 2019. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2019.01.010. URL https://www.sciencedirect.com/science/article/pii/S019126151830688X.

J. Xie and C. Xie. New insights and improvements of using paired alternative segments for traffic assignment. *Transportation Research Part B: Methodological*, 93:406–424, 2016. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2016.08.009. URL https://www.sciencedirect.com/science/article/pii/S0191261516305902.

J. Xie, Y. M. Nie, and X. Liu. A greedy path-based algorithm for traffic assignment. *Transportation Research Record*, 2672(48):36–44, 2018. doi: 10.1177/0361198118774236. URL https://doi.org/10.1177/0361198118774236.

J. Xie, S. Wong, S. Zhan, S. Lo, and A. Chen. Train schedule optimization based on schedule-based stochastic passenger assignment. *Transportation Research Part E: Logistics and Transportation Review*, 136:101882, 2020. ISSN 1366-5545. doi: https://doi.org/10.1016/j.tre.2020.101882. URL https://www.sciencedirect.com/science/article/pii/S1366554518311086.

Z. Xu, J. Xie, X. Liu, and Y. M. Nie. Hyperpath-based algorithms for the transit equilibrium assignment problem. *Transportation Research Part E: Logistics and Transportation Review*, 143:102102, 2020. ISSN 1366-5545. doi: https://doi.org/10.1016/j.tre.2020.102102. URL https://www.sciencedirect.com/science/article/pii/S136655452030750X.

X. Yan, J. Levine, and R. Marans. The effectiveness of parking policies to reduce parking demand pressure and car use. *Transport Policy*, 73:41–50, 2019. ISSN 0967-070X. doi: https://doi.org/10.1016/j.tranpol.2018.10.009. URL https://www.sciencedirect.com/science/article/pii/S0967070X18304402.

X. Yan, W. Yang, X. Zhang, Y. Xu, I. Bejleri, and X. Zhao. A spatiotemporal analysis of e-scooters' relationships with transit and station-based bikeshare. *Transportation Research Part D: Transport and Environment*, 101:103088, 2021. ISSN 1361-9209. doi: https:

//doi.org/10.1016/j.trd.2021.103088. URL https://www.sciencedirect.com/science/article/pii/S1361920921003849.

Y. Yan, Z. Liu, Q. Meng, and Y. Jiang. Robust optimization model of bus transit network design with stochastic travel time. *Journal of Transportation Engineering*, 139(6):625–634, 2013. doi: 10.1061/(ASCE)TE.1943-5436.0000536. URL https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29TE.1943-5436.0000536.

H. Yang and Y. Tang. Managing rail transit peak-hour congestion with a fare-reward scheme. *Transportation Research Part B: Methodological*, 110:122–136, 2018. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2018.02.005. URL https://www.sciencedirect.com/science/article/pii/S0191261517300541.

D. Zhang, W. Xu, B. Ji, S. Li, and Y. Liu. An adaptive tabu search algorithm embedded with iterated local search and route elimination for the bike repositioning and recycling problem. *Computers Operations Research*, 123:105035, 2020. ISSN 0305-0548. doi: https://doi.org/10.1016/j.cor.2020.105035. URL https://www.sciencedirect.com/science/article/pii/S0305054820301520.

J. Zhang, R. Lindsey, and H. Yang. Public transit service frequency and fares with heterogeneous users under monopoly and alternative regulatory policies. *Transportation Research Part B: Methodological*, 117:190–208, 2018. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2018.08.020. URL https://www.sciencedirect.com/science/article/pii/S0191261518302054.

R. Zhang and M. Pavone. Control of robotic mobility-on-demand systems: A queueing-theoretical perspective. *Robotics: Science and Systems X*, 2014. doi: 10.15607/rss.2014.x.026.

N. Zheng and N. Geroliminis. Modeling and optimization of multimodal ur-

ban networks with limited parking and dynamic pricing. *Transportation Research Part B: Methodological*, 83:36–58, 2016. ISSN 0191-2615. doi: https://doi.org/10.1016/j.trb.2015.10.008. URL https://www.sciencedirect.com/science/article/pii/S0191261515002234.

J. Zhou, M. Zhang, and P. Zhu. The equity and spatial implications of transit fare. *Transportation Research Part A: Policy and Practice*, 121:309–324, 2019. ISSN 0965-8564. doi: https://doi.org/10.1016/j.tra.2019.01.015. URL https://www.sciencedirect.com/science/article/pii/S0965856417304937.