Supply-Demand Equilibrium of Private and Shared Mobility in a Mixed Autonomous/Human Driving Environment

By

MOHAMADHOSSEIN NORUZOLIAEE

M.S., Tarbiat Modares University, Iran, 2012
B.S., Iran University of Science and Technology, Iran, 2009

Dissertation

Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Civil Engineering in the Graduate College of the University of Illinois at Chicago, 2018

Chicago, Illinois

Defense Committee:

Dr. Bo Zou, Chair and Advisor
Dr. Abolfazl (Kouros) Mohammadian
Dr. Jie (Jane) Lin
Dr. Kazuya Kawamura, College of Urban Planning and Public Affairs
Dr. Theja Tulabandhula, Department of Information and Decision Sciences
Dedicated to

My parents

and

My wife
ACKNOWLEDGEMENTS

The achievement of my academic goals so far would have not been possible without the encouragement, company, help, and guidance of many others, a complete list of whom requires a few pages. I would hereby like to deliver my gratitude to some of them.

Foremost, I would like to express my deepest appreciation to my PhD advisor, Professor Bo Zou, for his generous support and skillful guidance throughout this journey. I am very impressed by his academic and personal attitudes, conscientious, and diligence. Bo, your encouraging words were always inspirational to me at the hopelessness moments of research. I am indebted to you for teaching me not only to stand fearlessly against the hardship of research, but to move forward firmly with all my strengths. Now, I know how to step into the unknown realms of research.

I am grateful to my advisory committee members Professors Abolfazl Mohammadian, Jane Lin, Kazuya Kawamura, and Theja Tulabandhula for their invaluable comments and ideas. My special thanks goes to Professor Abolfazl Mohammadian for all his moral and professional support throughout these years.

Next, I would like to extend my sincere thanks to my Master’s supervisor and advisor, Professors AmirReza Mamdoohi and Mahmoud Saffarzadeh, at Tarbiat Modares University in Iran. I am deeply indebted to them for their unconditional support and guidance that steered my research life into good directions.

I have been fortunate to have good friends who supported and helped me in different ways during these years, especially Mohsen Hajikazemi, Hossein Taghizadeh Khamesi, Hooman Motahari, and Hossein Joshaghani. My next appreciation goes to my lab mates for their company in happy and sad moments, especially Ahmadreza Talebian, Nabin Kafle, Amirhassan Kermanshah, Ashkan Mahdavi, Thanakorn Siriakorn, Tanvir Ahamed, and Limon Barua.
I am very grateful to God for all his blessings, especially for my first and best teachers of life: my beloved father and mother. During the past five years, geographical distance and travel restrictions separated us, but I have been always having you in my heart. I owe my accomplishments to your non-stop prayers, pure love, dedication, and support that are not anything new in my life. How could I perfectly thank you? Proudly dedicating this dissertation to you, I hope it can partly compensate for your support and for being away from you. Words cannot express my appreciation to my dear sisters, Zeinab, Zahedeh, and Faezeh, and my brother-in-law, Reza.

Last but not the least, I could never imagine this achievement without the constant support, company, and encouragement of my wonderful wife. Noosheen, you have been my strength by kindly standing next to me in good and bad times over the years of PhD studies. Thanks to you I have my every dream come true. I would like to also thank my parents-in-law for their encouragement and prayers.
**Contribution of Authors**

Chapter 1 is an introduction to the dissertation which presents the overall framework and contributions of the research. Chapter 2 represents a published manuscript (“Noruzoliaee, M., Zou, B., Liu, Y., 2018. Roads in transition: Integrated modeling of a manufacturer-traveler-infrastructure system in a mixed autonomous/human driving environment. Transportation Research Part C: Emerging Technologies 90, 307-333.”), for which I was the primary author and major driver of the research. My advisor, Dr. Bo Zou, and Dr. Yang Liu contributed to the writing of the manuscript. Chapter 3 represents an unpublished work, for which I was the primary author and major driver of the research. My advisor, Dr. Bo Zou, contributed to the writing of the manuscript. Chapter 3 addresses the transportation system performance in the era of shared autonomous vehicles.
# TABLE OF CONTENTS

1. Introduction ................................................................................................................................. 1


   2.1. Introduction .......................................................................................................................... 5
   2.2. Review of relevant literature ............................................................................................... 9
      2.2.1. Previous research on AVs ............................................................................................ 9
      2.2.2. Transportation network optimization problems .......................................................... 10
   2.3. Problem formulation ............................................................................................................ 11
      2.3.1. Road link capacity with presence of AVs ................................................................. 12
      2.3.2. AV market penetration .................................................................................. 14
      2.3.3. AV Manufacturer pricing .................................................................................. 16
   2.4. Linearization to approximate the MPCC .............................................................................. 18
      2.4.1. Linearization of the UE complementarity condition .............................................. 18
      2.4.2. Linearization of the link travel time function ....................................................... 19
      2.4.3. Linearization of the market penetration function .............................................. 22
      2.4.4. Linearization of the profit function .................................................................. 24
      2.4.5. Final formulation of the MILP .................................................................................. 26
   2.5. Strategies to reduce the approximation error ........................................................................ 27
      2.5.1. Designing non-uniform distribution of breakpoints .................................................. 28
      2.5.1.1. Choosing |I| + 1 breakpoints from a large set of candidate breakpoints .......... 28
      2.5.1.2. Predetermining the set of candidate breakpoints ........................................... 29
      2.5.2. Feasibility-based domain reduction technique ....................................................... 30
   2.6. Overall solution algorithm ................................................................................................... 31
   2.7. Numerical experiments ........................................................................................................ 32
      2.7.1. The simplified Singapore network ........................................................................... 32
      2.7.1.1. The base scenario ............................................................................................... 33
      2.7.1.2. Sensitivity analyses ............................................................................................ 36
      2.7.1.2.1. VOTT savings and AV headway .................................................................... 36
      2.7.1.2.2. Additional AV technology cost ..................................................................... 37
      2.7.1.2.3. Cost perception variation of travelers ......................................................... 40
      2.7.1.2.4. Market size .................................................................................................... 42
      2.7.1.3. Computational effectiveness and efficiency ....................................................... 43
      2.7.2. The Sioux Falls network ............................................................................................. 47
   2.8. Conclusions ............................................................................................................................ 49
   Appendix 2.1. Algorithm to identify the breakpoints .............................................................. 51

3. One-to-many matching and section-based formulation of autonomous ridesharing equilibrium ........ 52

   3.1. Introduction .......................................................................................................................... 52
   3.2. Basic setup ........................................................................................................................... 57
      3.2.1. Representation of a road network and notion of sections ....................................... 57
      3.2.2. Workflows of vehicles ............................................................................................. 60
   3.3. Formulation of the multimodal autonomous ridesharing and parking user equilibrium (MARPUE) ....................... 62
      3.3.1. Autonomous ridesharing user equilibrium (ARUE) for SAV travelers .............. 62
      3.3.1.1. Waiting time due to matching and meeting ....................................................... 63
      3.3.1.2. Waiting time due to limited SAV seat capacity .............................................. 69
      3.3.1.3. In-vehicle time and fare .................................................................................. 71
      3.3.2. Autonomous parking user equilibrium (APUE) ................................................... 72
      3.3.3. User equilibrium (UE) for occupied HVs/AVs ..................................................... 75
      3.3.4. Endogenous HV/AV/SAV market share ............................................................... 77
      3.3.5. Overall formulation of MARPUE and its existence ............................................ 78
LIST OF TABLES

Table 2.1. Network and demand data ........................................................................................................... 33
Table 2.2. Main modeling parameters in the base scenario .............................................................................. 34
Table 2.3. AV manufacturer strategy and transportation system performance under base scenario .......... 35
Table 2.4. Equilibrium flows and AV market penetration under base scenario ............................................. 35
Table 2.5. Summary descriptive statistics of computational effectiveness and efficiency .......................... 44
Table 2.6. AV manufacturer strategy and transportation system performance under base scenario (Sioux Falls) .... 49

Table 3.1. Synthesis of existing network equilibrium studies of ridesharing and the present study ............... 54
Table 3.2. Main modeling parameters in the baseline scenario ...................................................................... 88
Table 3.3. Baseline results ............................................................................................................................ 89
LIST OF FIGURES

Figure 2.1. The overarching framework of the leader (car manufacturer)-follower (travelers) game in this chapter...6
Figure 2.2. Ratio of link capacity with mixed AV and HV traffic and with only HVs as a function of the proportion of AVs on the link (HV headway: 1.6 sec) ........................................................................................................14
Figure 2.3. Outer approximation of the univariate link travel time function ..................................................................................19
Figure 2.4. Illustration of modeling λ’s as SOS2 variables by mapping intervals onto binary vectors and value branching of the newly introduced binary variables η’s ........................................................................22
Figure 2.5. Outer approximation of a univariate logarithmic function .....................................................................................................23
Figure 2.6. Construction of triangles ..........................................................................................................................................................24
Figure 2.7. Two strategies to reduce the approximation error of the travel time function (top) and the logarithmic function (bottom)...27
Figure 2.8. Approximation error for the interval between h = 3 and h = 4 for a travel time function ..................................................29
Figure 2.9. The study area in Singapore and the abstract network representation ..................................................................................33
Figure 2.10. Effect of VOTT savings and AV headway on: (a) AV price; (b) AV manufacturer profit; (c)/(d) AV market penetration among high-/low-VOTT travelers; (e) total travel time and generalized travel cost; and (f) network capacity ........................................................................................................38
Figure 2.11. Effect of additional AV technology cost on: (a) AV price and AV manufacturer profit; (b) AV market penetration; (c) total travel time and generalized travel cost; and (d) network capacity ...............................................39
Figure 2.12. Effect of cost perception variation of travelers on: (a) AV price and AV manufacturer profit; (b) AV market penetration; (c) total travel time and generalized travel cost; and (d) network capacity ................................................41
Figure 2.13. Effect of market size on: (a) AV price and AV manufacturer profit; (b) AV market penetration; (c) total travel time and generalized travel cost; and (d) network capacity ................................................42
Figure 2.14. Solution time and linearization errors using the solution algorithm in section 6 and alternative algorithms ..................45
Figure 2.15. Sensitivity analysis of the tradeoff among the maximum number of iterations for domain reduction, the linearization error, and the computation time ..............................................................................46
Figure 2.16. The Sioux Falls network ......................................................................................................................................................48

Figure 3.1. The overarching framework of this study...........................................................................................................................56
Figure 3.2. Representations of road network, routes for SAVs, and sections for SAV travelers .............................................................59
Figure 3.3. SAV workflow ...........................................................................................................................................................................60
Figure 3.4. Changes in system performance compared to the baseline scenario, without ridesharing and without AV self-parking ......................................................................................................................92
Figure 3.5. Changes in system performance compared to the scenario with only HVs, without ridesharing and without AV self-parking ....................................................................................................................................92
Figure 3.6. Changes in system performance compared to the baseline scenario, with different SAV sizes .................................................................93
Figure 3.7. Changes in system performance compared to the scenario with only HVs, with different SAV sizes .........................93
Figure 3.8. Changes in system performance compared to the baseline scenario, with different vehicle headways ................94
Figure 3.9. Changes in system performance compared to the scenario with only HVs, with different vehicle headways .................................................................................................................................................95

Figure A.1. Impacts of matching elasticities on system performance compared to the baseline scenario .........................103
Figure A.2. Impacts of matching elasticities on system performance compared to the scenario with only HVs .........................103
Figure A.3. Impact of (S)AV price on system performance compared to the baseline scenario ..................................................103
Figure A.4. Impact of (S)AV price on system performance compared to the scenario with only HVs ............................................104
Figure A.5. Impact of in-vehicle VOT saving on system performance compared to the baseline scenario .................................105
Figure A.6. Impact of in-vehicle VOT saving on system performance compared to the scenario with only HVs ............................106
SUMMARY

This dissertation research presents a methodological investigation of the transportation system performance with coexistence of privately owned human-driven and autonomous vehicles (HVs and AVs) and shared AVs (SAVs) that provide ridesharing services. Built upon the network optimization and game theory concepts, two models are proposed that holistically account for the complex interplay among three players, i.e., transportation supply, demand, and infrastructure, which contribute to the system performance.

The first model explicitly accounts for the interplay among the AV manufacturer, travelers with heterogeneous in-vehicle values of time (VOT), and road infrastructure capacity. By making in-vehicle time use more leisurely or productive, AVs reduce travelers’ VOT. In addition, AVs can move closer together than HVs because of shorter safe reaction time, which leads to increased road capacity. On the other hand, the use of AV technologies means added manufacturing cost and higher price. Thus, traveler adoption of AVs will trade VOT savings with additional out-of-pocket cost. The model is structured as a leader (AV manufacturer)-follower (traveler) game. Given the cost of producing AVs, the AV manufacturer sets AV price to maximize profit while anticipating AV market penetration. Given an AV price, the vehicle and routing choice of heterogeneous travelers are modeled by combining a multinomial logit model with multi-modal multi-class user equilibrium (UE). The overall problem is formulated as a mathematical program with complementarity constraints (MPCC), which is challenging to solve. A solution approach is proposed which is based on piecewise linearization of the MPCC as a mixed-integer linear program (MILP) and solving the MILP to global optimality. Non-uniform distribution of breakpoints that delimit piecewise intervals and feasibility-based domain reduction are further employed to reduce the approximation error brought by linearization. The model is implemented in a simplified Singapore network with extensive sensitivity analyses and the Sioux Falls network. Computational results demonstrate the effectiveness and efficiency of the solution approach and yield valuable insights about transportation system performance in a mixed autonomous/human driving environment.
The second model characterizes the market shares of both private and shared AVs in mixed traffic with HVs. The model accounts for ridesharing with SAVs, which allows multiple travelers to share one autonomous vehicle at the same time. In doing so, two major contributions are made. First and foremost, a novel one (SAV)-to-many (travelers) matching is proposed to characterize the waiting times of an SAV and multiple travelers who share rides in the SAV while being matched, which generalizes the one-to-one matching absent ridesharing in previous taxi/ridesourcing studies. The proposed matching characterization explicitly considers the possibilities of a traveler matched with an SAV which starts from the same origin and previously moved to the origin either for pickup or as a result of relocation; or with an en-route SAV that goes through the traveler’s origin and does not incur extra stopping other than picking up the traveler. Theoretical insights are derived by comparing the waiting times of an SAV/traveler due to matching with and without ridesharing. The second contribution is the introduction of a section-based formulation for SAV ridesharing equilibrium. The notion of sections is introduced to represent traveler flows, which prevents undesired traveler en-route transfer and frees up the seat after a traveler leaves an SAV, while respecting the SAV seat capacity constraint. The utilization of SAV seat capacity is characterized at the section level, which takes into account travelers of the section under study as well as travelers of competing sections that can use the same SAV. Besides these two major contributions, the proposed model further considers inter- and intra-zone self-parking of privately owned AVs after dropping off the users. A new multimodal autonomous ridesharing and parking user equilibrium (MARPUE) is then put forward with a proof of its existence. Anticipating traveler reactions as characterized by MARPUE, the problem of optimally determining the SAV fleet size, fare rates, and vehicle allocation/relocation of a transportation network company (TNC) is formulated and solved. Several original insights are obtained from numerical implementation of the model, which advance the understanding of how transportation system performance can be reshaped by SAV ridesharing in the future.
1. Introduction

It is envisaged that the emerging autonomous vehicles (AVs) will revolutionize private and shared urban mobility. Privately owned AVs can drop off owners at desired locations and then self-park in cheaper areas, thereby transforming urban parking demand pattern (Fagnant and Kockelman, 2015). Shared AVs (SAVs) can provide improved shared mobility services compared to existing taxi/ridesourcing systems through reduced fares in the absence of human drivers and shorter waiting times by proactively relocating unoccupied to balance vehicle supply and travel demand (Fagnant et al., 2015). Using SAVs, ridesharing (i.e., sharing rides with other passengers in an SAV) further allows for efficient vehicle capacity utilization and lower fares (Levin et al., 2017). Besides, (S)AV users will enjoy reduced in-vehicle value of time (VOT) by being able to perform other more productive or leisurely activities than driving (Steck et al., 2018), and experience more efficient traffic operations as (S)AVs increase road capacity by safely moving closer together compared to human-driven vehicles (HVs) (Talebpour and Mahmassani, 2016). On the other hand, (S)AVs could adversely impact transportation system due to the deadhead miles generated by unoccupied (S)AVs in self-parking, relocation, and pickup trips (Fagnant and Kockelman, 2018). The net effect of (S)AVs on transportation system performance has yet to be rigorously studied. To this end, it is necessary to adopt a holistic approach that accounts for the decisions and interactions of different players who contribute to the system performance, as elaborated below.

On the demand side, traveler adoption behavior of (S)AVs will govern the penetration of (S)AVs on the roads, which will affect the extent of system-wide impacts of (S)AVs. On the supply side, car manufacturers will sell (S)AVs at possibly higher prices than that of human-driven vehicles (HVs) due to the new technologies used in (S)AVs such as sensors, navigation and communication systems, software, and Light Detection and Ranging systems (LIDAR) (Fagnant and Kockelman, 2015). Moreover, the level of ridesharing service will depend directly on the supply attributes of SAVs (e.g., fare, fleet size/allocation/relocation), which are decided by transportation network companies (TNCs) such as Uber and Lyft, and indirectly on the efficiency of the matching technology used for establishing SAV-traveler
contacts which manifests through service waiting times. The demand for and supply of (S)AVs further interact with transportation infrastructure through changes in road network capacity and urban parking demand. The former is caused by shorter car following headways of automated/connected driving technologies used in (S)AVs compared to HVs. The latter is attributed to the self-parking capability of AVs in distant/cheaper areas and the reduced number of private cars in the presence of ridesharing. Such supply-demand-infrastructure interactions will be more complicated in mixed traffic of (S)AVs and HVs, which is expected to dominate at least for a few decades until human driving becomes obsolete. In view of these, this dissertation presents two methodological studies to explore the intertwined relationship of supply, demand, and infrastructure in the era of (S)AVs.

The first study (chapter 2) proposes an integrated model of a manufacturer-traveler-infrastructure system with coexistence of HVs and AVs that encompasses AV manufacturing cost and pricing, combined vehicle and route choice of travelers who have heterogeneous in-vehicle VOTs, endogenous road capacity with mixed AV and HV traffic, and transportation system performance. The integrated model is developed under a Stackelberg leader-follower game structure, in which an AV manufacturer acts as the leader. Given the cost of producing AVs, the profit-maximizing AV manufacturer determines AV price anticipating the consequent AV market penetration. With the AV price, the vehicle and routing choice of travelers are modeled by combining a multinomial logit model with multi-modal multi-class user equilibrium (UE). The overall problem is formulated as a mathematical program with complementarity constraints (MPCC), which is challenging to solve mainly due to its non-convex feasible region caused by the UE complementarity constraints. To overcome this challenge, a new solution approach is proposed based on linearizing the MPCC as a mixed-integer linear program (MILP) and solving the MILP to global optimality. Two contributions are made in developing the solution approach. The first contribution lies in linearizing nonlinear functions by constructing piecewise linear functions in a way that requires the number of additional binary variables and constraints to be only logarithmic – rather than linear – in the number of the piecewise intervals, which significantly reduces problem size. The second contribution is made by devising
two strategies to reduce approximation error brought by the linearization. The first strategy is to non-uniformly locate breakpoints that delimit the piecewise intervals by formulating a constrained shortest path problem using dynamic programming. The second strategy is to employ a feasibility-based domain reduction technique that shrinks the domains of the linearized functions by discarding infeasible solutions to the MILP.

The second study (chapter 3) proposes an integrated model of a TNC-traveler-infrastructure system with coexistence of private (HV/AV) and shared (SAV) mobility under a leader (TNC)-follower (travelers) game structure. Given the SAV purchase price, the profit-maximizing TNC determines the autonomous ridesharing supply attributes including fleet size, fare rates, and vehicle allocation/relocation in the road network, while anticipating travelers’ reactions which are reflected in aggregate as the market share for each vehicle type. By observing the TNC’s strategies and the AV/HV purchase price, travelers’ vehicle type decisions are formulated as what we designate as the multimodal autonomous ridesharing and parking user equilibrium (MARPUE) comprising four components. First, an autonomous ridesharing user equilibrium (ARUE) is proposed, which assigns travelers to SAVs while respecting the vehicle seat capacity using a section-based representation of SAV travelers’ network, and captures the one (SAV)-to-many (travelers) matching involved in ridesharing via the notion of search intensity. Second, an autonomous parking user equilibrium (APUE) is proposed to characterize self-parking of vacant AVs right after dropping off AV users at their destination, given the fee and capacity of an urban parking infrastructure that is shared with HVs and SAVs. Third is a link-node based user equilibrium (UE) formulation for loading HVs and occupied AVs to road network, which accounts for SAV flow as centrally determined by the TNC and considers increase in road capacity as proportional to the (S)AV flow on each road link. These three modules are coupled by endogenously determining the market penetrations of HV/AV/SAV based on the corresponding travel costs obtained from the three modules using a multinomial logit model. Specifically, travelers’ costs encompass in-vehicle time cost (for HV/AV/SAV users who have different in-vehicle VOT), waiting time cost (during parking search for HV users, and due to matching and limited ridesharing
capacity for SAV users), and out-of-pocket costs such as vehicle operating costs while en-route and during parking search (for HV/AV users), parking fee (for HV/AV users), and fare (for SAV users). The overall problem is formulated as an MPCC, which is solved using the active-set method (Lawphongpanich and Yin, 2010).
2. Roads in transition: Integrated modeling of a manufacturer-traveler-infrastructure system in a mixed autonomous/human driving environment

The materials of the current chapter are partially published with the following citation:

2.1. Introduction

Recent advances in autonomous vehicle (AV) technologies and legislations show significant prospect of AV use in the future (Fagnant and Kockelman, 2015; Litman, 2017). Among the benefits that AVs can bring to urban transportation, this chapter focuses on two of them. First, by making in-vehicle time use more leisurely or productive, AVs reduce travelers’ value of travel time (VOTT) (Le Vine et al., 2015; van den Berg and Verhoef, 2016). Second, AVs can move with reduced headways compared to human-driven vehicles (HVs) because of shorter reaction time of in-vehicle computers than human brains. This will lead to higher road network capacity (Le Vine et al., 2015; Wang et al., 2015). On the other hand, due to the new technologies used in AVs such as sensors, navigation and communication systems, software, and Light Detection and Ranging systems (LIDAR), AVs will be more expensive (Fagnant and Kockelman, 2015). The price of AVs is further affected by the market supply-demand interactions, in which the willingness of travelers to use AVs will depend on the changes in the travelers’ VOTTs using AVs vs. HVs. To understand the impacts of the above-mentioned factors brought by AVs on urban transportation, this chapter contributes to the literature by developing an integrated model that encompasses AV manufacturing cost and pricing,
combined vehicle and route choice of travelers who have heterogeneous VOTTs, endogenous road capacity with mixed AV and HV traffic, and transportation network performance.

The integrated model is developed under a leader-follower game structure (Figure 2.1), in which an AV manufacturer acts as the leader. Given the cost of producing AVs, the profit-maximizing AV manufacturer determines AV price anticipating the consequent AV market penetration. With the AV price, the vehicle and routing choice of travelers are modeled by combining a multinomial logit model with multi-modal multi-class user equilibrium (UE). The overall problem is formulated as a mathematical program with complementarity constraints (MPCC), which is challenging to solve due to: the non-convex feasible region caused by the UE complementarity constraints; the nonlinear multivariate travel time and multinomial logit functions in the constraints; and the non-concave objective function. To overcome this challenge, we propose a new solution approach based on linearizing the MPCC as a mixed-integer linear program (MILP) and solving the MILP to global optimality.

![Figure 2.1. The overarching framework of the leader (car manufacturer)-follower (travelers) game in this chapter](image)

Two contributions are made in developing the solution approach. The first contribution lies in linearization techniques. In the MPCC formulation, the UE complementarity constraints are linearized using disjunctive constraints and additional binary variables. Other nonlinear functions in the constraints and the objective function are approximated by piecewise linear functions. We construct the piecewise
linear functions in a way that requires the number of additional binary variables and constraints to be only logarithmic – rather than linear as in existing methods (Farvaresh and Sepehri, 2011; Liu and Wang, 2015; Wang et al., 2015) – in the number of the piecewise intervals, based on a recent development in the disjunctive programming literature (Vielma and Nemhauser, 2011). To our knowledge, this logarithmic-sized linearization has never been considered in the transportation network modeling literature. Compared to the existing methods, the logarithmic-sized linearization considerably reduces the size of the MILP obtained from linearization, yet without compromising the approximation error.

Besides linearization techniques, we make a second contribution by devising two strategies to reduce approximation error brought by the linearization. The first strategy is to non-uniformly locate breakpoints that delimit the piecewise intervals in linearizing a constraint. The basic idea is to locate the breakpoints such that the total approximation error across the piecewise intervals is minimized, given a set of candidate breakpoints and the desired number of breakpoints to select from the candidate breakpoints. The idea is materialized by formulating a constrained shortest path problem using dynamic programming, built on the work of Dahl and Realfsen (2000) who cast the general problem of constrained shortest path into a dynamic programming formulation. The second strategy is to employ a feasibility-based domain reduction technique (Caprara and Locatelli, 2010) that shrinks the domains of the linearized functions by discarding infeasible solutions to the MILP. Same as the logarithmic-sized linearization, we are not aware of its use in transportation network modeling. This strategy allows all breakpoints to be placed in the feasible domains, which also contributes to reducing approximation error.

The model and the solution approach are implemented in simplified Singapore and Sioux Falls networks. The computational experience demonstrates the effectiveness and efficiency of the solution approach. In addition, several important insights are obtained. Specifically, in the scenario that producing an AV costs $10,000 more than an HV and AV travelers save 20% VOTT, the AV market share will be 2.4% (Singapore) and 6.1% (Sioux Falls). The change in aggregate system performance such as total travel time and generalized travel cost will be marginal. The smaller AV market share in Singapore is due to the
significantly high overhead cost of buying a car and smaller VOTTs assumed in Singapore than in the US (Sioux Falls). Extensive sensitivity analyses are performed on the Singapore network to investigate the impacts on system performance of VOTT savings, AV manufacturing cost, cost perception variation of travelers, and market size, with the following major findings:

- As AV travelers enjoy a greater extent of VOTT savings, AV price will increase (up to 90% higher than HV). Most AV users will be with high VOTT and their choice between AVs and HVs is more sensitive to VOTT savings than low-VOTT travelers. The AV market share among high-VOTT travelers will increase to over 90%. Because of the AV use, network capacity will increase by up to 40%. Total travel time and generalized travel cost will decrease by up to 11% and 9%.

- As the AV technology cost falls, AV manufacturer profit will increase. So will AV market share (up to 55%), total travel time saving (up to 5%), total generalized travel cost saving (up to 4%), and network capacity (up to 19%). High-VOTT travelers are more sensitive to AV technology cost than low-VOTT travelers.

- As the cost perception variation of travelers decreases (travelers perceive more “accurately” the benefit and cost of using AVs), AV price will decrease. Travelers become more rational in using AVs: low-VOTT travelers will use AVs less, whereas high-VOTT travelers will first decrease then increase AV use. System travel time and network capacity follow the trend of AV market share. System generalized travel cost will decrease.

- As the market size increases, the AV manufacturer will earn more profit with fluctuating price. AV market share will first increase and then decrease, the latter due to more severe congestion which offsets travel time savings after switching to AVs. Change in network capacity follows a similar trend as AV market share.

These results are expected to help researchers and policy makers to gain further understanding about the impact of AVs on urban transportation and inform infrastructure investment decisions in the advent of vehicle automation.
The rest of this chapter is organized as follows. Section 2.2 reviews the literature on AVs and transportation network optimization, which relates to the model and the solution approach proposed in this chapter. The model formulation and its linear approximation are presented in sequence in sections 2.3 and 2.4. Section 2.5 details the two strategies to reduce the approximation error brought by the linearization. The overall solution algorithm is summarized in section 2.6. Section 2.7 reports model implementation in the Singapore and Sioux Falls networks, which demonstrates the computational effectiveness and efficiency of the solution algorithm. Section 2.8 concludes and suggests directions for future research.

2.2. Review of relevant literature

2.2.1. Previous research on AVs

AVs are expected to exhibit multi-faceted impacts on transportation systems including improved road safety (Fagnant and Kockelman, 2015; Kalra and Paddock, 2016), enhanced mobility (Bansal et al., 2016; Harper et al., 2016; Krueger et al., 2016), increased road capacity (Talebpour and Mahmassani, 2016; Levin and Boyles, 2016b; Chen et al., 2017), more efficient traffic operations (Gong et al., 2016; Le Vine et al., 2016; Levin and Boyles, 2016a; Li et al., 2014; Wang et al., 2015), and new patterns for urban parking (Correia and van Arem, 2016; Zhang et al., 2015). AVs will also affect travel demand because travelers can use their in-vehicle time more productively (Jamson et al., 2013; Fagnant and Kockelman, 2015; van den Berg and Verhoef, 2016), which results in reduced generalized cost of travel. In addition, AVs will decrease energy consumption and emissions (Fagnant and Kockelman, 2014; Mersky and Samaras, 2016; Wadud et al., 2016) and shape land use in the long run (Bansal et al., 2016). To keep the literature review most relevant, in what follows we focus on studies that investigate the capacity impact and market penetration of AVs. Interested readers are referred to Anderson et al. (2014) and Fagnant and Kockelman (2015) for more comprehensive reviews of AVs and the policy implications.

The capacity impact of AVs comes from changes in macroscopic traffic flows (Wang et al., 2015) – more specifically, the reduction in vehicle headways. van den Berg and Verhoef (2016) reviewed existing
predictions of road capacity increase with AVs and found a large variation (from 1% to 400%) in the literature. In the case of mixed AV and HV traffic, the capacity increase is generally moderate and considered a function of the proportion of AVs in the total traffic. From the methodological perspective, Mahmassani (2016) and Talebpour and Mahmassani (2016) used microsimulation to investigate road capacity with mixed traffic. The authors found that as AV market share increases, AVs will have a greater influence on capacity. Levin and Boyles (2016b) incorporated driver reaction time in a car following model to predict road link capacity, also with mixed traffic. van den Berg and Verhoef (2016) computed shared road capacity based on a dynamic bottleneck model which assumes that AVs and HVs travel separately over time and derived road capacity as the weighted mean of AV capacity and HV capacity.

Research on predicting the market penetration of AVs has been based on empirical evidence of adoption of earlier vehicle technologies (Lavasani et al., 2016; Litman, 2017), stated-preference surveys (Shin et al., 2015; Kyriakidis et al., 2015; Yap et al., 2016; Bansal et al., 2016; Nazari et al., 2018), and simulation techniques (Bansal and Kockelman, 2017). Fagnant and Kockelman (2015) and Bansal and Kockelman (2017) highlighted that a comprehensive market penetration analysis should consider the interactions between traveler willingness-to-pay and AV price. However, because AVs are not yet commercially available, the actual market price of AVs is unknown. In van den Berg and Verhoef (2016), an analytical model was developed to derive closed-form solutions for AV price under three AV supply regimes: public supply, private monopoly, and perfect competition. For AV market penetration, existing studies mostly assumed it to be exogenous (Chen et al., 2016, 2017; Levin and Boyles, 2016b). One exception is Chen et al. (2016) who presented an endogenous evolution path of AV market penetration, although the initial AV market penetration was assumed known.

2.2.2. Transportation network optimization problems

Given the presence of a leader (the AV manufacturer) who anticipates decisions of the followers (the travelers), which depend on transportation network performance, the problem considered in this chapter
belongs to the broad class of transportation network optimization problems. Popular examples in the context of urban networks include but are not limited to network design (Farahani et al., 2013; Yang and Bell, 1998), toll design (Chen et al., 2015), and OD matrix estimation (Yang et al., 1992; Yousefikia et al., 2016). Interested readers are referred to Farahani et al. (2013) for a recent review. Due to the presence of equilibrium constraints and nonlinear functions such as the travel time, transportation network optimization problems are typically nonconvex and many existing algorithms yield local optimal solutions. Several efforts were made towards devising nonlinear programming based algorithms to generate global optimal solutions (Li et al., 2012; Wang et al., 2013). In addition, approximation by linearization has received growing attention. Relevant work includes Wang and Lo (2010), Luathep et al. (2011), Farvaresh and Sepehri (2011), Liu and Wang (2015), and Wang et al., 2015) for transportation network design, Ekström et al. (2012) and Ekström et al. (2014) for tolling, and Faturechi and Miller-Hooks (2014) for transportation system resilience.

2.3. Problem formulation

In this section, we present in sequence the formulations of road capacity with AV presence (section 2.3.1), AV market penetration (section 2.3.2), and AV manufacturer pricing decision (section 2.3.3). An urban transportation network is modeled as a directed graph comprising a set of directional links $a \in A$. Each traveler, who belongs to user class $n \in N$ based on VOTT, makes his/her trip between OD pair $w \in W$ by taking route $r \in R_w$ using vehicle type $m \in M$. The following notations are employed in the subsequent model formulations.

<table>
<thead>
<tr>
<th>Nomenclature</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Acronyms</strong></td>
</tr>
<tr>
<td>AV</td>
</tr>
<tr>
<td>HV</td>
</tr>
<tr>
<td>VOTT</td>
</tr>
<tr>
<td>MPCC</td>
</tr>
<tr>
<td>MILP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sets</th>
</tr>
</thead>
</table>

11
### Variables
- $\rho_m$: price of vehicle type $m$ ($)
- $\pi$: AV manufacturer’s profit ($/year)
- $q_w^m$: demand of vehicle type $m$ by user class $n$ for travel between OD pair $w$ (passengers/hr)
- $f_r^{nm}$: flow on route $r$ associated with vehicle type $m$ and user class $n$ (vehicles/hr)
- $p_w^m$: market penetration of vehicle type $m$ for user class $n$ in OD pair $w$
- $c_w^m$: minimum generalized cost of travel by vehicle type $m$ incurred by user class $n$ in OD pair $w$ ($)
- $C_r^{nm}$: generalized cost of travel by vehicle type $m$ on route $r$ incurred by user class $n$ ($)
- $c_a^m$: out-of-pocket cost of travel by vehicle type $m$ on link $a$ ($)
- $t_a$: flow-dependent travel time on link $a$ (hours)
- $x_a$: total vehicle flow on link $a$ (vehicles/hr)
- $x_a^m$: flow of vehicle type $m$ on link $a$ (vehicles/hr)
- $p_a^m$: proportion of flow of vehicle type $m$ on link $a$
- $K_a$: capacity of link $a$ (vehicles/hr)

### Parameters
- $q_w^n$: travel demand between OD pair $w$ by user class $n$ (passengers/hr)
- $\gamma^{nm}$: VOTT of user class $n$ when traveling with vehicle type $m$ ($/hr$)
- $\tau_m$: average headway between vehicles of type $m$ (hours)
- $K_a^{HV}$: base capacity of link $a$ when only HVs are on road (vehicles/hr)
- $l_a$: length of link $a$ (miles)
- $t_{0,a}$: free-flow travel time of link $a$ (hours)
- $\mu_m$: marginal manufacturing cost of vehicle type $m$ ($)
- $V C_m$: variable cost of using vehicle type $m$ ($/mile$)

#### 2.3.1. Road link capacity with presence of AVs

We establish a function that relates steady-state capacity of a road link to the proportion of AVs on the link. Specifically, we assume that link capacity monotonically increases with the proportion of AVs in the total traffic flow. As link capacity is a linear function of critical density at maximum flow based on the Greenshields relationship, we start by considering critical density of an arbitrary link $a$ with mixed traffic ($k_a^{C_{mix}}$) and with only HVs ($k_a^{CHV}$). It is known that critical density of a link is reciprocal of average vehicle spacing (Cassidy, 1999). Following Bose and Ioannou (2003), the average vehicle spacing in mixed flow...
is considered as a weighted average of vehicle spacing with only AVs and only HVs, weighted by the proportions of AVs and HVs in the total flow. Under the assumption of identical speeds for HVs and AVs in mixed traffic, vehicle spacing is proportional to vehicle headway. The assumption is reasonable given that AVs may adapt to the speed of surrounding vehicles (Levin and Boyles, 2016b). With the above reasoning, the following relationship between critical density and vehicle headway on a link can be derived:

\[
\frac{k_{a}^{c,\text{mix}}}{k_{a}^{c,\text{HV}}} = \frac{\tau_{HV}}{\sum_{m \in M} \tau_{m} p_{a}^{m}} \quad \forall a \in A
\]  

(2.1)

where

\[
p_{a}^{m} = \frac{x_{a}^{m}}{x_{a}} \quad \forall a \in A, x_{a} \neq 0
\]  

(2.2)

\(\tau_{m}\) is the average headway of vehicle type \(m\) at maximum flow (i.e., capacity), which relates to the perception-reaction time of the vehicle type. \(p_{a}^{m}\) is the proportion of vehicle type \(m\) traffic flow \((x_{a}^{m})\) in the total link flow \(x_{a} = \sum_{m \in M} x_{a}^{m}\).

Using Eq. (2.1) and the linear relationship between link capacity and critical density, the capacity of link \(a\) with mixed traffic \((K_{a})\) can be expressed in Eq. (2.3), which is computed by adjusting the base capacity with only HVs, \(K_{a}^{HV}\), by the ratio of average vehicle headway with only HVs and mixed traffic.

\[
K_{a} = \frac{\tau_{HV}}{\sum_{m \in M} \tau_{m} p_{a}^{m}} K_{a}^{HV} \quad \forall a \in A, x_{a} \neq 0
\]  

(2.3)

Given that automation reduces headways, i.e., \(\tau_{HV} \geq \tau_{AV}\), Eq. (2.3) suggests that the capacity of a link monotonically increases with the proportion of AVs on the link. Figure 2.2 illustrates the link capacity increase considering two headways for AVs. The headway for HVs is assumed 1.6 sec (Nowakowski et al., 2010). It can be seen that link capacity will increase by up to 45% and 167% when vehicles are all AVs compared to all HVs.
Figure 2.2. Ratio of link capacity with mixed AV and HV traffic and with only HVs as a function of the proportion of AVs on the link (HV headway: 1.6 sec)

2.3.2. AV market penetration

Travelers choose between AVs and HVs based on generalized travel cost, which comprises two components: out-of-pocket cost and travel time cost. For a traveler on a road link, the out-of-pocket cost is modeled as the sum of distance-based capital and variable costs, as follows:

\[ c_{am} = \frac{b_{m,1} b_{m,2}}{b_{m,3} b_{m,4} b_{m,5}} \rho_m l_a + \frac{V C_m}{b_{m,5}} l_a \quad \forall a \in A; m \in M \]  \hfill (2.4)

For vehicle type \( m \), \( c_{am} \) is the out-of-pocket travel cost on link \( a \); \( \rho_m \) is the vehicle price; \( l_a \) is the length of link \( a \); \( VC_m \) represents the variable cost (fuel cost, maintenance cost, insurance premium, etc.) per mile; \( b_{m,1} \) scales the vehicle price by capturing other overhead costs (e.g., tax, registration fee); \( b_{m,2} \) is the portion of lost value in vehicle price due to depreciation at the end of the vehicle’s lifetime; \( b_{m,3} \) is the average vehicle lifetime (in years); \( b_{m,4} \) is the average travel distance per year; \( b_{m,5} \) is the average vehicle occupancy rate. Thus, the first term on the right-hand side corresponds to the distance-based capital cost; the second term the distance-based variable cost.

Travel time cost is determined by the VOTT of travelers and travel time, the latter affected by the level of congestion. The flow-dependent travel time on a link \( a \in A \) is computed using the Bureau of Public Roads (1964) function:
\[ t_a = t_{0,a} \left[ 1 + \alpha \left( \frac{x_a}{k_a} \right)^\beta \right] \quad \forall a \in A \quad (2.5) \]

Plugging the link capacity function (Eq. (2.2)-(2.3)) into Eq. (2.5) and performing simple algebra yield the following travel time function:

\[ t_a = t_{0,a} \left[ 1 + \alpha \left( \sum_{m \in M} \tau_m x_a^m \right)^\beta \right] \quad \forall a \in A, x_a \neq 0 \quad (2.6) \]

where

\[ x_a^m = \sum_{n \in N} \sum_{w \in W} \sum_{r \in R_w} \delta_{ar} f_{rnm} \quad \forall a \in A; m \in M \quad (2.7) \]

\[ t_{0,a} \text{ and } t_a \text{ are free-flow travel time and flow-dependent travel time on link } a; f_{rnm} \text{ is the flow of vehicle type } m \text{ and user class } n \text{ on route } r; \delta_{ar} \text{ is the 0-1 link-route incidence indicator; } \alpha, \beta \text{ are the parameters.} \]

The travel time is monetized by the user class- and vehicle type-specific VOTTs (\( \gamma_{nm} \) in Eq. (2.8)). The travel time cost is then added to the out-of-pocket cost to obtain the generalized travel cost on a link (\( C_{anm} \)). The generalized travel cost on a route \( C_{rnm} \) is the sum of the travel costs on all links used by the route (Eq. (2.9)).

\[ C_{anm} = c_{anm}^m + \gamma_{nm} t_a \quad \forall a \in A; n \in N; m \in M \quad (2.8) \]

\[ C_{rnm} = \sum_{a \in A} \delta_{ar} C_{anm} \quad \forall r \in R_w; w \in W; n \in N; m \in M \quad (2.9) \]

Traveler route choice is determined using Wardrop’s UE principle (Wardrop, 1952). The multi-modal multi-class UE condition can be mathematically formulated as follows (Yang and Huang, 2004):

\[ f_{rnm} (C_{rnm} - C_{wnm}) = 0 \quad \forall r \in R_w; w \in W; n \in N; m \in M \quad (2.10) \]

\[ C_{rnm} - C_{wnm} \geq 0 \quad \forall r \in R_w; w \in W; n \in N; m \in M \quad (2.11) \]

\[ \sum_{r \in R_w} f_{rnm} = q_w^m b_{ms} \quad \forall w \in W; n \in N; m \in M \quad (2.12) \]
where $C_{wnm}$ is the minimum generalized travel cost by vehicle type $m$ in OD pair $w$ incurred by user class $n$. Expressions (2.10)-(2.11) state that user class $n$ with vehicle type $m$ will use route $r$ only if the associated generalized travel cost on the route equals the minimum generalized travel cost of that OD pair, for the same user class and vehicle type. Constraints (2.12)-(2.13) are user class- and vehicle type-specific OD demand conservation and non-negativity constraints.

The market share of each vehicle type is modeled endogenously using a multinomial logit form, inspired by the pioneering works of Yang (1998) and Yin and Yang (2003) who investigated market penetration of advanced traveler information systems. In Eq. (2.14), the share of trips using vehicle type $m$ for user class $n$ and OD pair $w$, $P_{wnm}$, depends on the generalized travel cost of all vehicle types for the same user class and OD pair:

$$P_{wnm} = \frac{\exp(-\varphi C_{wnm})}{\sum_{m' \in M} \exp(-\varphi C_{wnm'})} \quad \forall w \in W; n \in N; m \in M$$

(2.14)

where $\varphi > 0$ is the scale factor (Train, 2003) capturing cost perception variation of travelers.

Under the assumption of constant total travel demand for each user class and OD pair, and with the above market share function (Eq. (2.14)), the travel demand for using vehicle type $m$ from user class $n$ of OD pair $w$ is:

$$q_{wnm} = q_{wn} \cdot P_{wnm} \quad \forall w \in W; n \in N; m \in M$$

(2.15)

2.3.3. AV Manufacturer pricing

In this chapter, we consider an AV manufacturer as the leader which maximizes profit by choosing an AV price while anticipating the response of the followers, i.e., travelers in terms of vehicle type choice.
The price of HVs is assumed known and constant. The profit maximization problem is formulated in (2.16)-(2.17), as an MPCC.

\[
\max_{\rho_m | m = AV} \pi = \frac{\theta}{b_{m,5}}(\rho_m - \mu_m) \sum_{n \in N} \sum_{w \in W} q_{wn}^{nm}
\]

s.t.

\[
\rho_m \leq \rho_m \leq \bar{\rho}_m \quad m = AV
\]

(a) travel costs: (2.4), (2.6)-(2.9)

(b) travelers’ route choice: (2.10)-(2.13)

(c) travelers’ vehicle choice: (2.14)-(2.15)

where \(\mu_m\) is the unit manufacturing cost of vehicle type \(m\); \(\rho_m\) and \(\bar{\rho}_m\) are the lower and upper bounds of AV price. The bounds, which can be set very wide, are necessary to develop the equivalent linear approximation of MPCC in the next section. \(\theta\) is a conversion factor for the AV manufacturer to estimate the number of AVs sold in a year based on AV trips per hour in the urban transportation network which is

\[
\frac{1}{b_{m,5}} \sum_{n \in N} \sum_{w \in W} q_{wn}^{nm}, \quad m = AV
\]

The underlying assumption in the conversion is that each AV makes a certain average number of trips in a year.\(^1\) The value of the conversion factor \(\theta\) may depend on the nature of AV ownership. For example, if all AVs in the network are privately owned and used, \(\theta\) will take a larger value (i.e., a smaller number of vehicle trips per hour per AV) than if a portion of the AVs are owned and operated by ride sharing services. On the other hand, as \(\theta\) only appears in the objective function (2.16) and is considered a pre-specified parameter in this chapter, the specific values for \(\theta\) do not affect the insights obtained from solving the model.

\(^1\) As an illustration, suppose that the lifetime of an AV \((b_{m,3}, m = AV)\) is 10 years and that \(Z\) number of AVs are sold each year. Then at equilibrium the city has 10\(Z\) AVs. We use \(\omega_y\) to denote the average number of trips that an AV makes in a year, and \(\zeta\) the traffic k-factor that converts the number of hourly AV trips, \(\frac{1}{b_{m,5}} \sum_{n \in N} \sum_{w \in W} q_{wn}^{nm}\), to the number of daily AV trips, \(\frac{1}{\omega_y} \left(\frac{1}{b_{m,5}} \sum_{n \in N} \sum_{w \in W} q_{wn}^{nm}\right), m = AV\). Consequently \(\frac{365.1}{\omega_y} \zeta \left(\frac{1}{b_{m,5}} \sum_{n \in N} \sum_{w \in W} q_{wn}^{nm}\right), m = AV\) will represent the number of AVs needed to fulfill the trips in a year. Equating the expression to 10\(Z\) yields \(Z = \frac{1}{10 \omega_y} \left(\frac{365.1}{\omega_y} \zeta \left(\frac{1}{b_{m,5}} \sum_{n \in N} \sum_{w \in W} q_{wn}^{nm}\right), m = AV\). By comparing the two expressions for the AV manufacturer’s annual profit, i.e., \(Z(\rho_m - \mu_m)\) and Eq. (2.16), we obtain \(\theta = \frac{365.1}{10 \omega_y} \zeta\).
2.4. Linearization to approximate the MPCC

Solving the above MPCC is challenging due to: the non-convex feasible region caused by the UE complementarity constraints (2.10)-(2.11); the nonlinear multivariate travel time and multinomial logit functions in the constraints; and the non-concave objective function. We consider approximating the MPCC as an MILP. Specifically, the non-convexity of the feasible region is removed using disjunctive constraints based linear transformation of the complementarity constraints (section 2.4.1). The nonlinear multivariate travel time and multinomial logit functions (Eq. (2.6) and (2.14)) are first reformulated into univariate functions, and then linearized using an outer approximation technique (sections 2.4.2 and 2.4.3). The nonlinear objective function is linearized by a triangulation of its two-dimensional domain (section 2.4.4).

2.4.1. Linearization of the UE complementarity condition

The only source of nonlinearity in the UE condition is the nonlinear complementarity constraints (2.10)-(2.11). We transform these constraints into equivalent linear constraints using additional binary variables as follows (Siddiqui and Gabriel, 2013).

\[ f_r^{nm} \leq U_f (1 - \psi_r^{nm}) \quad \forall r \in R, w \in W; n \in N; m \in M \]  
\[ C_r^{nm} - C_w^{nm} \leq U_C \psi_r^{nm} \quad \forall r \in R, w \in W; n \in N; m \in M \]  
\[ C_r^{nm} - C_w^{nm} \geq 0 \quad \forall r \in R, w \in W; n \in N; m \in M \]  
\[ f_r^{nm} \geq 0 \quad \forall r \in R, w \in W; n \in N; m \in M \]  
\[ \psi_r^{nm} \in \{0,1\} \quad \forall r \in R, w \in W; n \in N; m \in M \]

where \( U_f \) and \( U_C \) are sufficiently large numbers. It can be verified that if \( \psi_r^{nm} = 0 \), then \( C_r^{nm} - C_w^{nm} = 0 \) which means that the generalized travel cost on route \( r \) equals the minimum generalized travel cost of the corresponding OD pair \( w \). Thus, the route will take flow \( (f_r^{nm} \leq U_f) \). Similarly, \( \psi_r^{nm} = 1 \) implies that the route will not take flow.
2.4.2. Linearization of the link travel time function

The multivariate polynomial travel time function of a link \( a \in A \) in Eq. (2.6) can be transformed into a univariate nonlinear function by introducing a new variable \( h_a \):

\[
\begin{align*}
    t_a(h_a) &= t_{0,a} \left[ 1 + \alpha \left( \frac{h_a}{\tau_{HV} K_{HV}} \right)^\beta \right] \quad \forall a \in A \tag{2.23} \\
    h_a &= \sum_{m \in M} \tau_m x_a^m \quad \forall a \in A \tag{2.24}
\end{align*}
\]

We linearize the univariate function (2.23) using an outer approximation technique following Liu and Wang (2015) and Wang et al. (2015). Specifically, the domain of travel time function, i.e., \( h_a \in [\underline{h}_a, \overline{h}_a] \), is partitioned into a set of mutually exclusive intervals with indices \( i \in I = \{1, 2, ..., |I|\} \) which are associated with a set of predetermined breakpoints indexed by \( j \in J = \{1, 2, ..., |I| + 1\} \). At a breakpoint \( j \in J \), the value for \( h_a \) is denoted by \( h_a^j \). The first and the last breakpoints are the lower and upper bounds \( \underline{h}_a \) and \( \overline{h}_a \). The convex travel time curve is approximated using tangent lines at each breakpoint and chord lines that connect two adjacent breakpoints (Figure 2.3).

![Figure 2.3. Outer approximation of the univariate link travel time function](image)

The tangent lines at each breakpoint \( j \in J \) provide a lower bound of convex travel time function. These tangent lines are expressed using the first-order Taylor expansion around each breakpoint:

\[
\begin{align*}
    t_a(h_a) &= t_{0,a} + \alpha \frac{h_a}{\tau_{HV} K_{HV}} \beta \quad \forall a \in A
\end{align*}
\]
\[ t_a(h_a) \geq t_{0,a} \left[ 1 + \alpha \left( \frac{h_a^j}{h_a} \right)^\beta + \beta \left( \frac{h_a^j}{h_a} \right)^{\beta-1} \left( h_a - h_a^j \right) \right] \quad \forall a \in A; j \in J \] (2.25)

The chord lines provide an upper bound of the travel time function. We express these chord lines with the introduction of additional variables and constraints. Specifically, travel time at a point \( h_a \in [h_a, h_a] \) is uniquely represented as a convex combination of travel time values associated at the two breakpoints of the active interval containing \( h_a \) (constraints (2.26)-(2.27)). Constraints (2.28)-(2.29) ensure non-negative combination factors \( \lambda_{h_a}^j \) which sum up to one.

\[ t_a(h_a) \leq \sum_{j \in J} \lambda_{h_a}^j \cdot t_a(h_a^j) \quad \forall a \in A \] (2.26)

\[ h_a = \sum_{j \in J} \lambda_{h_a}^j \cdot h_a^j \quad \forall a \in A \] (2.27)

\[ \sum_{j \in J} \lambda_{h_a}^j = 1 \quad \forall a \in A \] (2.28)

\[ \lambda_{h_a}^j \geq 0 \quad \forall a \in A; j \in J \] (2.29)

How can we have at most two \( \lambda_{h_a}^j \)'s \((j \in J)\), which should be adjacent and associated with the active interval, non-zero? This is done by considering \( \lambda_{h_a}^j \)'s \((j \in J)\) as Special Ordered Sets of type 2 (SOS2) variables (Beale and Forrest, 1976). We model the SOS2 variables using a recently developed logarithmic-sized method (Vielma and Nemhauser, 2011). The salient feature of the method is a significant reduction in the size of the linearized MILP, with the number of binary variables and additional constraints only logarithmic (rather than linear) in the number of intervals.

Specifically, each interval \( i \in I \) is mapped onto a binary vector with \( \log_2 |I| \) elements using a bijective function: \( B: \{1,2,\ldots,|I|\} \rightarrow \{0,1\}^{\log_2 |I|} \). We index each element in a binary vector by \( \ell = 1,2,\ldots,\log_2 |I| \).

To reflect the adjacency of two intervals, we require that vectors \( B(i) \) for intervals \( i \) and \( B(i + 1) \) for \( i + 1 \), \( \forall i \in I \setminus \{ |I| \} \), differ only in one element. Lemma 1 in Vielma and Nemhauser (2011) presents an inductive
method to construct such bijective functions. Using the bijective functions, the SOS2 conditions for \( \lambda^j_{ha} \)'s, i.e., at most two adjacent \( \lambda^j_{ha} \) are non-zero, is satisfied by introducing \( \log_2 |I| \) binary variables \( \eta^\ell_{ha} (\ell = 1, 2, ..., \log_2 |I|) \) and \( 2 \log_2 |I| \) constraints as follows:

\[
\sum_{j \in J^+(\ell, B)} \lambda^j_{ha} \leq \eta^\ell_{ha} \quad \forall a \in A; \ell = 1, 2, ..., \log_2 |I| \tag{2.30}
\]

\[
\sum_{j \in J^-(\ell, B)} \lambda^j_{ha} \leq 1 - \eta^\ell_{ha} \quad \forall a \in A; \ell = 1, 2, ..., \log_2 |I| \tag{2.31}
\]

\[
\eta^\ell_{ha} \in \{0, 1\} \quad \forall a \in A; \ell = 1, 2, ..., \log_2 |I| \tag{2.32}
\]

where \( J^+(\ell, B) = \{ j \in J: \forall i \in I_j, B(i)_{\ell} = 1 \} \); \( J^0(\ell, B) = \{ j \in J: \forall i \in I_j, B(i)_{\ell} = 0 \} \). \( I_j \) is a set whose elements are the two adjacent intervals to breakpoint \( j \), i.e., \( I_j = \{ j - 1, j \} \). Note that if \( j = 1, I_j = 1 \). If \( j = |I| + 1, I_j = |I| \). \( B(i)_{\ell} \) is the \( \ell \)th element of the binary vector for interval \( i \). Put in plain words, for a given \( \ell \), \( J^+(\ell, B) \) represents the set of breakpoints whose adjacent intervals have the \( \ell \)th element in the corresponding binary vectors equal to 1. Similarly, for a given \( \ell \), \( J^0(\ell, B) \) represents the set of breakpoints whose adjacent intervals have the \( \ell \)th element in the corresponding binary vectors equal to 0.

To illustrate how the SOS2 conditions for \( \lambda^j \)'s are satisfied, let us consider a variable whose domain is partitioned into four intervals \( j = 1, 2, 3, 4, 5 \) (Figure 2.4(a)). Each interval is mapped onto a binary vector with \( \log_2 4 = 2 \) elements. Following Lemma 1 in Vielma and Nemhauser (2011), the intervals can be mapped onto binary vectors as: \( B(1) = (0, 1), B(2) = (1, 1), B(3) = (1, 0), B(4) = (0, 0) \). These \( B(i) \)'s result in \( J^+(1, B) = \{ 3 \} \), \( J^0(1, B) = \{ 1, 5 \} \), \( J^+(2, B) = \{ 1, 2 \} \), \( J^0(2, B) = \{ 4, 5 \} \). For any combination of \( \eta^\ell (\ell = 1, 2) \) values, constraints (2.30)-(2.32) ensure that at most two adjacent \( \lambda^j \)'s are non-zero. The combinations of \( \eta^\ell \) values are represented in Figure 2.4(b) using binary branching of two (\( \log_2 4 \)) levels. Each branching level \( \ell \) is associated with an \( \eta^\ell \). For example, the combination of \( \eta^1 = 0 \) and \( \eta^2 = \)

---

1 For brevity, subscript \( h_a \) is omitted here and in the rest of section 2.4.2.
1 will result in $\lambda^3 = \lambda^4 = \lambda^5 = 0$ (the indices 3, 4, and 5 are in the parentheses along the branches). The reasoning is as follows. When $\eta^1 = 0$, $\lambda^3$ must be 0 because $J^+(1, B) = \{3\}$ which implies that $\lambda^3 \leq 0$ according to constraint (2.30) but $\lambda^3$ must be non-negative (constraint (2.29)). When $\eta^2 = 1$, $\lambda^4$ and $\lambda^5$ must be 0 because $J^0(2, B) = \{4, 5\}$ which implies that $\lambda^4 + \lambda^5 \leq 0$ according to constraint (2.31) but $\lambda^4$ and $\lambda^5$ again must be non-negative (constraint (2.29)). Thus, at most $\lambda^1$ and $\lambda^2$ are non-zero and interval 1 is the active interval.

![Diagram](image)

Figure 2.4. Illustration of modeling $\lambda^j$'s as SOS2 variables by mapping intervals onto binary vectors and value branching of the newly introduced binary variables $\eta^j$'s

### 2.4.3. Linearization of the market penetration function

The logit function for AV market penetration of Eq. (2.14) can be rewritten as follows:

$$
\ln(p_{wn}^{nm}) - \ln(p_{wn}^{nm'}) = \phi(c_{wn}^{nm'} - c_{wn}^{nm}) \quad \forall m, m' \in M, m \neq m'; w \in W; n \in N
$$

$$
\sum_{m \in M} p_{wn}^{nm} = 1 \quad \forall w \in W; n \in N
$$
Similar to the linearization of the travel time function in section 2.4.2, the concave logarithmic function is outer approximated by chord lines connecting consecutive breakpoints to provide lower bounds, and tangent lines at breakpoints which provide upper bounds (Figure 2.5). A new variable $L_{P_{nm}}$ is introduced to represent the logarithmic function $L_{P_{nm}} = \ln(P_{nm})$, which is linearized using convex combination.

The tangent lines at breakpoints $j \in J$ are again expressed by the first-order Taylor expansion of the logarithmic function around each breakpoint (constraint (2.35)). Based on the method presented in section 2.4.2, the chord lines are expressed using linear constraints (2.36)-(2.42).

\[
L_{P_{nm}} \leq \ln(P_{nm,j}) + \frac{P_{nm,j}}{P_{nm}} - 1 \quad \forall w \in W; n \in N; m \in M; j \in J
\]  

(2.35)

\[
L_{P_{nm}} \geq \sum_{j \in J} \lambda_{P_{nm}}^{j} \ln(P_{nm,j}) \quad \forall w \in W; n \in N; m \in M
\]  

(2.36)

\[
P_{nm} = \sum_{j \in J} \lambda_{P_{nm}}^{j} \cdot P_{w}^{nm,j} \quad \forall w \in W; n \in N; m \in M
\]  

(2.37)

\[
\sum_{j \in J} \lambda_{P_{nm}}^{j} = 1 \quad \forall w \in W; n \in N; m \in M
\]  

(2.38)

\[
\sum_{j \in J^{\ell}(\ell, B)} \lambda_{P_{nm}}^{j} \leq \eta_{P_{nm}}^{\ell} \quad \forall w \in W; n \in N; m \in M; \ell = 1, 2, \ldots, \log_{2} |I|
\]  

(2.39)

\[
\sum_{j \in J^{N}(\ell, B)} \lambda_{P_{nm}}^{j} \leq 1 - \eta_{P_{nm}}^{\ell} \quad \forall w \in W; n \in N; m \in M; \ell = 1, 2, \ldots, \log_{2} |I|
\]  

(2.40)
\[ \lambda_{pw}^j \geq 0 \quad \forall w \in W; n \in N; m \in M; j \in J \quad (2.41) \]

\[ \eta_{pw}^j \in [0,1] \quad \forall w \in W; n \in N; m \in M; \ell = 1,2,\ldots,\log_2 |I| \quad (2.42) \]

2.4.4. Linearization of the profit function

The objective function (2.16) involves a bivariate nonlinear revenue function \( \rho_m q_m \) \((m = AV)^1\), where \( q_m = \sum_{n \in N} \sum_{w \in W} q_{wnm} \). We linearly approximate \( \rho_m q_m \) by a new variable \( \tilde{\rho}_m \) using convex combination. The approximation builds on Carathéodory’s theorem, which says that any point on a \( d \)-dimension domain can be uniquely represented by the convex combination of \( d + 1 \) points (Cook et al., 1986; Lee and Leyffer, 2011). Given that the domain of \( \tilde{\rho}_m \) has two dimensions \((\rho_m, q_m)\): \( \rho_m \in [\underline{\rho}_m, \bar{\rho}_m] \) and \( q_m \in [\underline{q}_m, \bar{q}_m] \), any \( \tilde{\rho}_m \) can be uniquely represented by a convex combination of three points. To this end, we partition the domain \((\rho_m, q_m)\): \( \rho_m \in [\underline{\rho}_m, \bar{\rho}_m] \) and \( q_m \in [\underline{q}_m, \bar{q}_m] \) into mutually exclusive triangles.

![Diagram](image_url)

(a) partitioning the revenue function domain into rectangles  
(b) Union Jack triangulation of the revenue function domain

**Figure 2.6. Construction of triangles**

The construction of the triangles is as follows. We first divide the plane into mutually exclusive rectangles based on the breakpoints for \( \rho_m \) and \( q_m \), i.e., \( j \in J = \{1,2,\ldots,|I| + 1\} \) and \( j' \in J = \ldots \)

---

1 In the rest of section 2.4.4, we always assume \( m = AV \).
\{1, 2, ..., |I| + 1\}, as shown in Figure 2.6(a) in the case of |I| = 4. Each rectangle is then divided into two triangles using Union Jack triangulation (Todd, 1977), which is known to yield smaller problem size among all triangulation classes (e.g., the K-triangulation in Luathep et al. (2011)) (Vasudeva, 2015). This is shown in Figure 2.6(b).

We express \( \tilde{\rho}_m \) as a linear combination of the \( \rho_m, q_m \) values at the three vertices of the triangle that contains \( (\rho_m, q_m) \), i.e., the active triangle. This is shown in constraints (2.43)-(2.45). Constraints (2.46)-(2.47) ensure that non-negative convex combination factors \( \lambda_{ij'}^{jj'} \) sum up to one.

\[
\tilde{\rho}_m = \sum_{j' \in J} \sum_{j \in J} \lambda_{ij'}^{jj'} \rho_m^j \cdot q_m^{j'}
\]

(2.43)

\[
\rho_m = \sum_{j' \in J} \sum_{j \in J} \lambda_{ij'}^{jj'} \rho_m^j
\]

(2.44)

\[
q_m = \sum_{j' \in J} \sum_{j \in J} \lambda_{ij'}^{jj'} \cdot q_m^{j'}
\]

(2.45)

\[
\sum_{j' \in J} \sum_{j \in J} \lambda_{ij'}^{jj'} = 1
\]

(2.46)

\[
\lambda_{ij'}^{jj'} \geq 0
\]

(2.47)

Only three \( \lambda_{ij'}^{jj'} \)'s that are associated with the active triangle should be non-zero. These \( \lambda_{ij'}^{jj'} \)'s are identified in two steps. First, the active rectangle is identified by applying the same binary branching as in section 2.4.2 to each of the \( \rho_m \) and \( q_m \). This is mathematically shown by constraints (2.48)-(2.52). After this step, only four non-zero \( \lambda_{ij'}^{jj'} \)'s will remain.

\[
\sum_{j' \in J} \sum_{j \in J} \lambda_{ij'}^{jj'} \leq \eta_{\rho_m}^{\ell} \quad \forall \ell = 1, 2, ..., \log_2 |I| 
\]

(2.48)

\[
\sum_{j' \in J} \sum_{j \in J} \lambda_{ij'}^{jj'} \leq 1 - \eta_{\rho_m}^{\ell} \quad \forall \ell = 1, 2, ..., \log_2 |I| 
\]

(2.49)

\[
\sum_{j \in J} \sum_{j' \in J} \lambda_{ij'}^{jj'} \leq \eta_{q_m}^{\ell} \quad \forall \ell = 1, 2, ..., \log_2 |I| 
\]

(2.50)
\[
\sum_{j \in J_0} \sum_{j' \in J_0} \lambda_{\rho_m}^{jj'} \leq 1 - \eta_{q_m}^\ell \quad \forall \ell = 1, 2, \ldots, \log_2 |I| 
\]
(2.51)

\[
\eta_{\rho_m}^\ell, \eta_{q_m}^\ell \in \{0, 1\} \quad \forall \ell = 1, 2, \ldots, \log_2 |I| 
\]
(2.52)

where the number of \(\eta_{\rho_m}^\ell\)'s is logarithmic in the number of intervals specified for \(\rho_m\). Similarly, for \(\eta_{q_m}^\ell\).

In the second step, the active triangle inside the active rectangle is chosen based on further branching, in which one branch selects a white triangle and the other branch picks a gray triangle (Figure 2.6(b)). This is formulated as constraints (2.53)-(2.55) with the new binary variable \(\chi_{\rho_m}\). If \(\chi_{\rho_m} = 0\), constraint (2.53) will prevent any triangle that involves a vertex with odd \(j\) and even \(j'\) (square vertices in Figure 2.6(b)) from being chosen. Such triangles are gray triangles. Similarly, if \(\chi_{\rho_m} = 1\), constraint (2.54) will prevent any triangle that involves a vertex with even \(j\) and odd \(j'\) (diamond vertices in Figure 2.6(b)) from being chosen. Such triangles are white triangles. Since an active rectangle consists of only one white triangle and one gray triangle, only one triangle will be ultimately chosen as the active triangle. In other words, one non-zero \(\lambda_{\rho_m}^{jj'}\) in the active rectangle of step one will be removed, and three non-zero \(\lambda_{\rho_m}^{jj'}\)'s will remain.

\[
\sum_{j' \in \text{even}} \sum_{j \in \text{odd}} \lambda_{\rho_m}^{jj'} \leq \chi_{\rho_m} 
\]
(2.53)

\[
\sum_{j' \in \text{odd}} \sum_{j \in \text{even}} \lambda_{\rho_m}^{jj'} \leq 1 - \chi_{\rho_m} 
\]
(2.54)

\[
\chi_{\rho_m} \in \{0, 1\} 
\]
(2.55)

### 2.4.5. Final formulation of the MILP

With the above linearization, MPCC (2.16)-(2.17) can be approximated as the following MILP:

\[
\max_{\rho_m | m = \text{AV}} \pi = \frac{\theta}{b_{m, 5}} (\bar{\rho}_m - \mu_m q_m) 
\]
(2.56)

s.t.

(a) bounds of the partitioned variables: 
\(h_a \in [\underline{h}_a, \bar{h}_a]\); 
\(P_w^{nm} \in [P_w^{nm}, \bar{P}_w^{nm}]\); \(\rho_m \in [\underline{\rho}_m, \bar{\rho}_m]\);

and \(q_m \in [\underline{q}_m, \bar{q}_m]\)
(b) travel costs: (2.4), (2.8)-(2.9)
(c) traffic flows: (2.7), (2.12)-(2.13)
(d) linearized UE conditions: (2.18)-(2.22)
(e) linearized link travel time function: (2.24)-(2.32)
(f) linearized market penetration of vehicle types: (2.33)-(2.42)
(g) linearized profit function: (2.43)-(2.55)

2.5. Strategies to reduce the approximation error

Given that MILP (2.56) is indeed a linear approximation of MPCC (2.16)-(2.17), reducing the approximation error is desirable. However, a smaller approximation error can entail more breakpoints/intervals and consequently a larger problem size (recall that the number of added binary variables and constraints in linearization is logarithmic in the number of intervals). An increase in the MILP size can lead to longer computation time to solve the relaxed LP at each node in the branch-and-bound tree.

![Figure 2.7](image_url)

(a) uniform distribution of breakpoints without domain reduction
(b) non-uniform distribution of breakpoints
(c) domain reduction

Figure 2.7. Two strategies to reduce the approximation error of the travel time function (top) and the logarithmic function (bottom)
We iteratively implement two strategies to reduce the approximation error while solving MILP (2.56), without increasing the number of breakpoints. The first strategy is to design non-uniform distribution of breakpoints through a dynamic program (Section 2.5.1). The second strategy tightens the variable bounds using a feasibility-based domain reduction technique (Section 2.5.2). Figure 2.7 presents the concepts of the two strategies.

2.5.1. Designing non-uniform distribution of breakpoints

2.5.1.1. Choosing |I| + 1 breakpoints from a large set of candidate breakpoints

As shown in Figure 2.7(b), the approximation error from piecewise linearization can be reduced by locating more breakpoints in parts of the domain where the nonlinear function has greater curvature. To do this, a large set of candidate breakpoints indexed by \( v \in V = \{1, 2, ..., |V|\} \) is first given. Any two candidate breakpoints (no matter they are adjacent or not) can form an interval. As we want to divide the domain of the nonlinear function into |I| intervals, a piecewise linear representation is constructed by choosing |I| + 1 out of the |V| candidate breakpoints (including the lower and upper bounds) to form |I| non-overlapping intervals that together cover the whole domain. The choice of the |I| + 1 breakpoints is to minimize the total approximation error over the constructed intervals. For each interval, the approximation error is defined as the difference between the middle point values on the nonlinear function and the approximate linear function (Figure 2.8).

This problem of choosing |I| + 1 out of the |V| candidate breakpoints to minimize total approximation error can be formulated as a constrained shortest path problem and solved in polynomial time using dynamic programming (Dahl and Realfsen, 2000). A path corresponds to a sequence of non-overlapping intervals starting from breakpoint 1 (lower bound of the domain) to a specified breakpoint. We recursively solve dynamic program (2.57) to find the shortest paths, i.e., the paths that have the minimum total approximation error, from breakpoint 1 to \( v \) using \( s \) intervals (\( s = 1, ..., |I| \)). Details about the recursive algorithm is presented in Appendix 2.1.
\[ E(v, s) = \min_{s \leq v' \leq v - 1} \{ E(v', s - 1) + e_{v'v} \} \]  

(2.57)

where \( E(v, s) \) is the approximation error of the shortest path from breakpoint 1 to \( v \in V \setminus \{1\} \) using \( s \) intervals; \( e_{v'v} \) is the approximation error for interval \( v'v \) where \( v' \) is an intermediate breakpoint between 1 and \( v \).

\[ \begin{align*}
0.5(t_a(h^3_a) + t_a(h^4_a)) & \\
t_a(h^*_a) & \\
0.5(t_a(h^3_a) + t_a(h^4_a)) & \\
\end{align*} \]

**Figure 2.8.** Approximation error for the interval between \( h^3_a \) and \( h^4_a \) for a travel time function

### 2.5.1.2. Predetermining the set of candidate breakpoints

Clearly, the set of candidate breakpoints is important in affecting the approximation error. On the one hand, too few candidate breakpoints cannot approximate a nonlinear function well no matter what breakpoints are chosen. On the other hand, too many candidate breakpoints are not desirable given that the number of possible intervals is \( O((|V| - 1)^2) \) and the computational complexity of the recursive algorithm for solving (2.57) is \( O((|V| - 1)^2 |I|) \). It is necessary to consider both approximation quality and computational cost while predetermining the set of candidate breakpoints.
Because the solution value is unknown \textit{a priori}, it is desirable to place candidate breakpoints such that the approximation errors of each interval formed by two adjacent candidate breakpoints are close to each other. In this way, more candidate breakpoints will be located where the corresponding nonlinear function has greater curvature, which is similar to the way we choose $|I| + 1$ breakpoints from the set of candidate breakpoints in section 2.5.1.1.

For the logarithmic function $L_{p_{wnm}}$, this is done by first specifying a desired approximation error for any interval formed by two adjacent candidate breakpoints. Then, starting from the lower bound $B_{wnm}$ (which is the first candidate breakpoint), the next candidate breakpoint is identified based on the previous candidate breakpoint and the desired approximation error. This process reiterates. Note that doing so does not require a predetermined number of candidate breakpoints $|V|$. In case that we end up having $|V| < |I| + 1$, we set $|I| = |V| - 1$.

For the travel time function $t_a(h_a)$, using the same method as for $L_{p_{wnm}}$ can be computationally difficult due to the higher order involved (note that $\beta$ in Eq. (2.23) typically takes value 4). An alternative approach is adopted. We set a predetermined number of candidate breakpoints $|V|$. These candidate breakpoints are located in the domain $h_a$ according to $h_a^v = h_a + [h_a - h_a][(v - 1)/(|V| - 1)]^\sigma$ with $\sigma$ being a positive parameter. If $\sigma < 1$, candidate breakpoints will be accumulated toward $h_a$. If $\sigma > 1$, more candidate breakpoints will be placed close to $h_a$. $\sigma = 1$ is a special case that the candidate breakpoints will be evenly located. Because the curvature of the travel time function increases as $h_a \to h_a^*$, a $\sigma < 1$ will be desired. After performing some numerical experiments, we choose $\sigma = 0.5$ as it yields very close approximation errors among the formed intervals.

2.5.2. Feasibility-based domain reduction technique

Besides designing non-uniform distribution of breakpoints, the approximation error of linearization can be further reduced by shrinking the domain without eliminating feasible solutions. A feasibility-based
domain reduction technique (Caprara and Locatelli, 2010) is employed to update the bounds of the partitioned variables \( h_a, p_{w}^{nm}, \rho_m, \) and \( q_m \). For a given variable, this entails minimizing/maximizing the variable value over the feasible region of MILP (2.56) to obtain updated lower/upper bounds. For example, MILP (2.58)-(2.59) is used to find an updated lower bound for \( \rho_m \), i.e., \( \underline{\rho}_m^{new} \):

\[
\rho_m^{new} = \min \rho_m \quad m = AV
\]

s.t. \( \underline{\rho}_m^{old} \leq \rho_m \leq \bar{\rho}_m^{old} \)

constraints of the relaxed MILP (2.56)

An updated upper bound for \( \rho_m \), i.e., \( \bar{\rho}_m^{new} \), is obtained by solving MILP (2.60)-(2.61):

\[
\bar{\rho}_m^{new} = \max \rho_m \quad m = AV
\]

s.t. \( \underline{\rho}_m^{old} \leq \rho_m \leq \bar{\rho}_m^{old} \)

constraints of the relaxed MILP (2.56)

Similar MILPs can be formulated for \( h_a \)'s, \( p_{w}^{nm} \)'s, and \( q_m \). To achieve even tighter variable bounds, the above domain reduction is implemented iteratively over all \( \rho_m, h_a \)'s, \( p_{w}^{nm} \)'s, and \( q_m \) until either a convergence threshold is met or after a desired number of iterations. It has been shown that the sequence of updated bounds converges to some limit which is independent of the order in which the MILPs of variables are solved (Caprara and Locatelli, 2010; Caprara et al., 2016).

2.6. Overall solution algorithm

Summarizing sections 2.4 and 2.5, we solve MPCC (2.16)-(2.17) using the following algorithm.

**Step 1.** Initialization. For each variable \( \rho_m, h_a, p_{w}^{nm}, \) and \( q_m \) (\( m = AV \)):

**Step 1.1.** Set the initial upper and lower bounds.

**Step 1.2.** Set the number of desired breakpoints.
Step 2. Pre-processing. For each variable $\rho_m$, $h_a$, $P_w^{nm}$, and $q_m$ ($m = AV$):

Step 2.1. Choose breakpoints. Solve dynamic program (2.57) for each $h_a$ and $P_w^{nm}$. For $\rho_m$ and $q_m$, because they are parts of the multivariate revenue function $\tilde{\rho}_m$, we still divide their domains into equal intervals. Set the iteration counter $\vartheta$ to 1.

Step 2.2. Domain reduction. Solve MILP (2.58)-(2.59) and MILP (2.60)-(2.61) to update the lower and upper bounds for $\rho_m$. Update the bounds for $h_a$, $P_w^{nm}$, and $q_m$ by solving similar MILPs. If $\vartheta$ is less than the maximum number of iterations, update $\vartheta = \vartheta + 1$ and go to Step 2.1. Otherwise, go to Step 3.

Step 3. Using branch-and-cut algorithm to solve the relaxed MILP (2.56) to global optimality.

2.7. Numerical experiments

2.7.1. The simplified Singapore network

We implement the model in a simplified Singapore network (Figure 2.9) to investigate transportation system performance with AVs and with only HVs. In addition, extensive sensitivity analyses on the impacts of VOTT savings, AV headway, additional AV technology cost, cost perception variation of travelers, and market size are performed. We also discuss computational effectiveness and efficiency in model implementation.

The study area consists of five traffic analysis zones: Woodlands (1), Jurong West (2), Bukit Timah (3), Ang Mo Kio (4), and CBD (5), in total 20 OD pairs connected by 16 links (Figure 2.9). Link characteristics including length ($l_a$), free-flow travel time ($t_{0,a}$), and base capacity ($K_a^{HV}$) are presented in Table 2.1. We consider two user classes with VOTTs being $5.6/hr and $16.8/hr when driving HVs\(^1\). The OD demand data for the two user classes is also shown in Table 2.1.

---

\(^1\) VOTTs are computed based on the monthly household incomes of $1,068 and $3,195, which are associated with the 25\(^{th}\) and 75\(^{th}\) percentiles of Singapore population (Singstat, 2016). The monthly incomes are converted to hourly values assuming 44 working hours per week (Singapore ministry of manpower, 2017).
Table 2.1. Network and demand data

<table>
<thead>
<tr>
<th>Link parameters</th>
<th>OD demands</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a \in A$</td>
<td>$l_a$ (mi)</td>
</tr>
<tr>
<td>1, 2</td>
<td>13.2</td>
</tr>
<tr>
<td>3, 4</td>
<td>13.3</td>
</tr>
<tr>
<td>5, 6</td>
<td>8.0</td>
</tr>
<tr>
<td>7, 8</td>
<td>7.3</td>
</tr>
<tr>
<td>9, 10</td>
<td>10.5</td>
</tr>
<tr>
<td>11, 12</td>
<td>6.3</td>
</tr>
<tr>
<td>13, 14</td>
<td>6.6</td>
</tr>
<tr>
<td>15, 16</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*: 1↔2 means for both OD pairs 1→2 and 2→1. In other words, the demand is 1745 pax/hr for each OD pair. Likewise for other OD pairs in the table.

2.7.1.1. The base scenario

We first establish the base scenario with AVs. For comparison, a corresponding scenario with only HVs is constructed as well. The main parameters used are listed in Table 2.2. We follow Nowakowski et al. (2010) and consider the average vehicle headways ($\tau_m$) to be 1.6 and 1.1 seconds for HVs and AVs. We assume the price and the manufacturing cost for an HV to be $\rho_m |_{m=HV} = $20,000 (Sgcarmart, 2017) and $\mu_m |_{m=HV} = $18,000, based on the assumption of 10% profit margin of a car manufacturer. For an AV, we
assume its manufacturing cost to be $\mu_{m|m=AV} = 28,000$ to account for the additional cost associated with new technologies (Fagnant and Kockelman, 2015). Parameter $\theta$ in the objective (2.16) is set to 0.5\(^1\).

On the traveler side, the base scenario assumes 20% saving in VOTT when travelers use AVs. Recall from Eq. (2.4) that the capital cost consists of depreciation and overhead costs. We follow Sgcarmart (2017) and assume depreciation cost to be 90% of the vehicle price ($b_{m,2} = 0.9$) over an average ownership period of 10 years ($b_{m,3} = 10$), and overhead cost to be $82,000 per vehicle, which translates to $b_{m,1} = 5.1$ (($82,000 + 20,000)/20,000$). We assume that the average travel distance of a vehicle ($b_{m,4}$) is 12,427 mi/year (20,000 km/year). $b_{m,5}$ is set to 1.75 following Fwa and Chua (2007). In this chapter, we consider identical $b_{m,1}$ values for HVs and AVs. $VC_{m|m=HV}$ and $VC_{m|m=AV}$ are $0.442$/mi and $0.455$/mi respectively, with the difference reflecting the additional maintenance cost associated with new technologies, and reduced insurance and fuel cost (Mersky and Samaras, 2016; Litman, 2017). Finally, the scale factor $\varphi$ in Eq. (2.14) is set to 4.

<table>
<thead>
<tr>
<th>Vehicle type</th>
<th>VOTT ($/hr)</th>
<th>Headway (sec)</th>
<th>Price ($)</th>
<th>Manufacturing cost ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-VOTT travelers</td>
<td>High-VOTT travelers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HV</td>
<td>5.60</td>
<td>16.80</td>
<td>1.6</td>
<td>20,000</td>
</tr>
<tr>
<td>AV</td>
<td>4.48 (20% saving)</td>
<td>13.44 (20% saving)</td>
<td>1.1</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.2. Main modeling parameters in the base scenario

Table 2.3 presents the optimal AV price and some aggregate results of system performance. AV will be sold at $29,700, which generates a total profit of $0.47 million per year for the AV manufacturer. We note that the AV manufacturer charges only a small price markup ($1,700) and ends up with a 5.7% profit margin as opposed to a 10% profit margin for HVs. Given the significant AV technology cost ($10,000), part of which will be passed onto travelers, the out-of-pocket cost of travelers will increase by 0.48% compared to the HVs-only scenario. Total travel time in the network is reduced by 0.03% with AVs due to reduced vehicle headway and increased road capacity. The overall generalized travel cost, however, will

\(^1\) This is obtained by assuming $\omega_y = 730$ (each AV makes on average 2 trips per day) and $\zeta = 0.1$. 

34
experience a slight increase (0.09%), suggesting that the effect of a higher AV price dominates the benefits of travel time reduction.

Table 2.3. AV manufacturer strategy and transportation system performance under base scenario

<table>
<thead>
<tr>
<th>AV manufacturer</th>
<th>System performance</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Performance measure</td>
<td>With AVs</td>
<td>Only HVs</td>
</tr>
<tr>
<td>AV price ($)</td>
<td>Total generalized cost ($)</td>
<td>272,076</td>
<td>271,833</td>
</tr>
<tr>
<td>Profit margin (%)</td>
<td>Total out of pocket cost ($)</td>
<td>171,843</td>
<td>171,013</td>
</tr>
<tr>
<td>Profit ($million)</td>
<td>Total time cost ($)</td>
<td>100,233</td>
<td>100,820</td>
</tr>
<tr>
<td></td>
<td>Total travel time (hr)</td>
<td>9,020</td>
<td>9,023</td>
</tr>
</tbody>
</table>

Table 2.4 lists the equilibrium vehicle flow and travel time on each link with AVs and with only HVs, and link capacity increase after AVs are introduced. The share of AV use for each OD pair is also presented. Not surprisingly, travelers with a high VOTT are much more inclined to use AVs than travelers with a low VOTT – almost 4.6% of high-VOTT travelers will use AVs compared to only 0.3% among low-VOTT travelers. This translates into an overall 2.4% AV market share and a network capacity increase by 0.51%.

Table 2.4. Equilibrium flows and AV market penetration under base scenario

<table>
<thead>
<tr>
<th>$a \in A$</th>
<th>Link flow, travel time, and capacity improvement</th>
<th>AV market penetration (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>With AVs $x_a$ (veh/hr) $t_a$ (min)</td>
<td>Only HVs $x_a$ (veh/hr) $t_a$ (min)</td>
<td>Capacity increase (%)</td>
</tr>
<tr>
<td>1, 2</td>
<td>1994</td>
<td>22.3</td>
<td>1994</td>
</tr>
<tr>
<td>3, 4</td>
<td>949</td>
<td>20.0</td>
<td>949</td>
</tr>
<tr>
<td>5, 6</td>
<td>2755</td>
<td>18.2</td>
<td>2720</td>
</tr>
<tr>
<td>7, 8</td>
<td>3226</td>
<td>19.4</td>
<td>3192</td>
</tr>
<tr>
<td>9, 10</td>
<td>657</td>
<td>17.0</td>
<td>692</td>
</tr>
<tr>
<td>11, 12</td>
<td>540</td>
<td>17.0</td>
<td>574</td>
</tr>
<tr>
<td>13, 14</td>
<td>1757</td>
<td>14.1</td>
<td>1757</td>
</tr>
<tr>
<td>15, 16</td>
<td>2047</td>
<td>24.0</td>
<td>2047</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.7.1.2. Sensitivity analyses

The results in the base scenario obviously depend on model parameter values. To gain further insights into the impacts of model parameters on system performance, extensive sensitivity analysis is conducted. The parameters under investigation include VOTT savings, AV headway, additional AV technology cost, cost perception variation of travelers, and market size.

2.7.1.2.1. VOTT savings and AV headway

Even though the literature admits traveler VOTT savings as a fundamental benefit brought by AVs, there is no evidence of the exact percentage (van den Berg and Verhoef, 2016). The percentage saving represents the portion of in-vehicle time that could be used for more leisurely or productive activities. It reflects the impact of the automation level: a higher level of automation requires less human interference (Yap et al., 2016). For AV headway, we test an alternative headway of 0.6 seconds which may result from connected vehicle technologies in addition to automation (Nowakowski et al., 2010). We also experiment with a larger headway of 2.0 seconds, which might occur at an early stage when AVs in mixed traffic have to rely on relatively imperfect vision/radar-based vehicle recognition instead of connected vehicle technology (Seo and Asakura, 2017). However, the results turn out to be very similar to the results under the base scenario (the AV market share, for example, would be 2.3% compared to 2.4% in the base scenario), and thus are not presented in detail here.

Figure 2.10 presents the results. The VOTT saving of AV users is varied from 0% to 90% with two AV headways (1.1 and 0.6 seconds). First, the AV manufacturer will charge a greater mark-up and enjoys larger profits as traveler VOTT saving increases (Figure 2.10(a)-(b)). If travelers receive no VOTT saving with AVs, benefits of using AVs will only come from capacity improvement, which is very small, as shown in Figure 2.10(f). In this case, AVs will be priced at $29,200, or a profit margin of 4.1%. As VOTT saving increases from 0% to 70%, AV price roughly follows an S-curve shape where the price first increases marginally in the range of 0-30% of VOTT saving and then more significantly. At the extreme of 90%
VOTT saving, the manufacturer will sell an AV at $37,600, which is approximately 90% more expensive than an HV. If AV headway is reduced to 0.6 seconds, AV price will slightly decrease (by up to $900).

We also note that low- and high-VOTT travelers behave very differently in vehicle choice. The AV adoption rate of high-VOTT travelers is very sensitive to VOTT saving. As shown in Figure 2.10(c), over 90% high-VOTT travelers will use AVs when VOTT saving reaches 90%. In contrast, less than 1% of low-VOTT travelers will choose AVs regardless of VOTT saving (Figure 2.10(d)). The AV adoption curve among low-VOTT travelers is non-monotonic with respect to VOTT saving and peaks when VOTT saving is 30%. This is because, as VOTT saving grows, AV price keeps increasing and will exceed the travel time cost savings for some low-VOTT travelers, which reduces the appeal of AVs among low-VOTT travelers.

Finally, Figure 2.10(e)-(f) show the reduction in system total travel time and generalized travel cost, and increase in network capacity with AVs and with only HVs. More visible changes occur after VOTT saving is greater than 20%. The changes are also more significant when AVs operate with a smaller headway. For example, with 90% VOTT saving, the network capacity will increase by 15% and 40% when AV headway is 1.1 seconds and 0.6 seconds, respectively.

2.7.1.2.2. Additional AV technology cost

At present, AVs are produced with significantly higher costs than HVs due to the use of new technologies such as sensors, navigation and communication systems, software, and LIDAR. Fagnant and Kockelman (2015) reviewed some of the AV manufacturing cost estimates and concluded that AVs are currently not affordable for many people. However, the authors noted that the additional AV technology cost (i.e., $\mu_{m|m=AV} - \mu_{m|m=HV}$) can decrease drastically to around $10,000 per vehicle, and even to as low as $1000 by the time when AVs are produced at a large scale. To understand the impact of AV technology cost on system performance, we vary the additional AV technology cost from $0 to $14,000 (Figure 2.11).
Figure 2.10. Effect of VOTT savings and AV headway on: (a) AV price; (b) AV manufacturer profit; (c)/(d) AV market penetration among high-/low-VOTT travelers; (e) total travel time and generalized travel cost; and (f) network capacity.
In general, the benefits to the AV manufacturer, travelers, and network capacity increase as AV technology costs less in manufacturing. Figure 2.11(a) shows that the AV manufacturer will charge a small mark-up ($610) when the additional AV technology cost is $14,000. As AV technology cost decreases, AV price will go down. However, the ability of the AV manufacturer to mark up will go up, reaching $4,820 if no cost difference exists between manufacturing an AV and an HV. The AV manufacturer profit also keeps growing, reaching a maximum of $29.1 million.

Figure 2.11. Effect of additional AV technology cost on: (a) AV price and AV manufacturer profit; (b) AV market penetration; (c) total travel time and generalized travel cost; and (d) network capacity
Figure 2.11(b) shows that the AV market penetration will vary widely – from 0.4% with $14,000 additional AV technology cost to 54.8% when there is no additional cost. Again, AV adoption among high- and low-VOTT travelers is very different. Moreover, the rates of reduction in system travel time and generalized travel cost, and increase in network capacity all diminish as AVs become more expensive to produce (Figure 2.11(c)-(d)).

2.7.1.2.3. Cost perception variation of travelers

Travelers’ perception of generalized travel costs is subject to error. This error is captured in the AV market penetration (logit) model of Eq. (2.14) by the scale parameter $\varphi$. A larger $\varphi$ suggests smaller errors in travelers’ cost perception, which result in greater probability of choosing a lower cost alternative. To test the effect of cost perception variation of travelers, we consider four possible $\varphi$ values: 2, 4 (base scenario), 6, and 8.

Figure 2.12(a) shows that, as the cost perception variation decreases ($\varphi$ increases), the AV manufacturer prices with smaller mark-ups and earns less profit. The changing trend of high-VOTT travelers in using AVs is not monotonic. For low-VOTT travelers, they will use AVs less as their cost perception variation decreases (Figure 2.12(b)). An explanation is provided as follows. Imagine when $\varphi = 0$, i.e., the largest perception variation, exactly 50% of low- and high-VOTT travelers would choose AV. As $\varphi$ increases (i.e., perception variation becomes smaller), travelers can distinguish more “clearly” between the generalized costs of the two vehicle types. For low-VOTT travelers, the generalized cost of travel by AV will always be larger than by HV because AVs are more expensive than HVs, yet the travel time saving benefits with AVs are limited due to their low VOTT. The larger generalized cost of travel by AV is invariant to the AV price drop and the change in system total travel time as $\varphi$ increases from 2 to 8. On the other hand, as $\varphi$ increases, the low-VOTT travelers realize more clearly the larger generalized cost of AV. As a consequence, the AV market penetration for these travelers continues to decrease.
For high-VOTT travelers, the story is a little different due to their higher VOTT. When \( \phi \) increases from 2 to 4, the AV market penetration follows the same decreasing trend as for low-VOTT travelers. But when \( \phi \) increases from 4 to 8, the continuous AV price drop plus the decline in system total travel time – which is now associated with higher VOTT – will reduce the generalized cost of travel by AV for high-VOTT travelers. This cost reduction will also be more clearly perceived as \( \phi \) continues to increase, leading to an increased AV market penetration among the high-VOTT travelers. Finally, for all travelers, the changing trend of AV market penetration is similar to that of high-VOTT travelers because of the dominance of high-VOTT travelers among AV users.

Figure 2.12. Effect of cost perception variation of travelers on: (a) AV price and AV manufacturer profit; (b) AV market penetration; (c) total travel time and generalized travel cost; and (d) network capacity
The changes in system travel time and increase in network capacity have consistent trends with AV market penetration (Figure 2.12(c)-(d)): greater AV market penetration is associated with lower system travel time and higher network capacity. On the other hand, we find that system total generalized cost monotonically decreases as the perception variation becomes smaller (Figure 2.12(c)).

2.7.1.2.4. Market size

The system performance can be different if we change market size. An increase in market size may be possible due to growing population. We rerun the model by increasing demand for each OD pair by 5-25%.

Figure 2.13. Effect of market size on: (a) AV price and AV manufacturer profit; (b) AV market penetration; (c) total travel time and generalized travel cost; and (d) network capacity
Figure 2.13(a) indicates that the AV price will change somewhat, fluctuating between $29,550 and $29,900. With greater market size, it is not surprising that the AV manufacturer profit will always increase. In Figure 2.13(b), the market penetration of AVs first increases and then decreases. A possible explanation for the decrease is that as the network becomes more severely congested, for some OD trips the travel time saving benefit cannot offset the added out-of-pocket cost of using AVs. Thus, the market penetration of AVs will drop. Figure 2.13(c) shows that traveler experience will deteriorate (increased time and generalized cost) compared to the base scenario as more travelers lead to greater congestion and generalized travel cost. Finally, the change in network capacity follows the same trend as AV market penetration (Figure 2.13(d)).

2.7.1.3. Computational effectiveness and efficiency

To report the computational effectiveness and efficiency of the overall solution algorithm, we first specify the initial bounds of the partitioned variables $h_\alpha, P_{wn}^{nm}, \rho_m|_{m=AV}$, and $q_m|_{m=AV}$. The lower and upper bounds of $h_\alpha$ are set to zero and the total OD demand that can use link $\alpha$. Each $P_{wn}^{nm}$ is bounded between $10^{-6}$ and 1. $\rho_m|_{m=AV}$ is bounded between one time and 10 times the unit AV manufacturing cost. $q_m|_{m=AV}$ is bounded between zero and the total OD demand over all ODs. We partition each $h_\alpha, P_{wn}^{nm}, \rho_m|_{m=AV}$, and $q_m|_{m=AV}$ into $|I|=8$ intervals. After model linearization and implementing the two strategies in section 2.5, we come with MILP (2.56) which has 471 binary variables, 1604 continuous variables, and 2755 constraints. The entire problem is coded in GAMS 24.7 and solved using CPLEX 12.6 on a personal computer with Intel(R) Core(TM) i7 CPU @ 3.40 GHz with 12 GB RAM. The optimality gap of the domain reduction MILPs (2.58)-(2.61) and the main MILP (2.56) are set to $10^{-10}$ and $10^{-6}$ respectively (IBM, 2015).

Table 2.5 presents three computation statistics: 1) approximation error (i.e., the absolute value difference for $t_\alpha, P_{wn}^{nm}, \tilde{P}_m|_{m=AV}$ obtained from solving MILP (2.56) and from plugging the price and flow solutions of MILP (2.56) into constraints (2.6) and (2.14), and $\rho_m q_m$); 2) percentage reduction of the initial domains of the variables; and 3) solution time. The first two statistics measure the computational
effectiveness and the third one the computational efficiency of the solution algorithm. The statistics are based on the computation of the base scenario in section 2.7.1.1 and all sensitivity analysis scenarios in section 0 (in total, 36 scenarios). We find that link travel times $t_a$’s, market shares $P_w^{nm}$’s, and revenue $\tilde{\rho}_m | m = AV$ can all be very closely approximated, with the average approximation errors at $2.0\times10^{-5}$ hours, $1.1\times10^{-3}$, and $1.7\times10^4$. The table shows that the domain reduction technique is very effective as well: the initial domains for $h_a$, $P_w^{nm}$, $\rho_m | m = AV$, and $q_m | m = AV$ are, on average, reduced by 78.3%, 59.7%, 74.6%, and 16.7%. The average solution time is only 13.4 minutes, in which almost 92% is attributed to solving MILPs (2.58)-(2.61) iteratively for $h_a$, $P_w^{nm}$, $\rho_m$, and $q_m$ domain reduction as the pre-processing step (we set the maximum number of iterations (step 2.2 of the algorithm) to 3).

<table>
<thead>
<tr>
<th>Linearization error</th>
<th>Domain reduction (%)</th>
<th>CPU time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t_a$ (hr)</td>
<td>$P_w^{nm}$</td>
<td>$\tilde{\rho}_m</td>
</tr>
<tr>
<td>Ave.</td>
<td>$2.0\times10^{-5}$</td>
<td>$1.1\times10^{-3}$</td>
</tr>
<tr>
<td>Std. dev.</td>
<td>$3.7\times10^{-5}$</td>
<td>$1.7\times10^{-3}$</td>
</tr>
</tbody>
</table>

Recall from sections 2.4 and 2.5 that the computational effectiveness and efficiency of the solution algorithm are attributed to three elements: logarithmic-sized linearization (sections 2.4.2-2.4.4), non-uniform distribution of breakpoints (section 2.5.1), and feasibility-based domain reduction (section 2.5.2). To further investigate the computational benefits of employing the three elements, we solve the model using alternative algorithms each lacking at least one of the three elements. For simplicity, we use “Log”, “NU”, and “DR” to denote logarithmic-sized linearization, non-uniform distribution of breakpoints, and feasibility-based domain reduction. Lacking “Log” would modify step 3 in the overall solution algorithm (section 2.6) by formulating a linear-sized MILP that uses the number of additional binary variables and constraints that is linear in the number of piecewise intervals, as in previous studies (Luathep et al., 2011; Farvaresh and Sepehri, 2011; Wang et al., 2015). Lacking “NU” would modify step 2.1 in the algorithm by dividing domains of each $h_a$ and $P_w^{nm}$ into equal intervals. Lacking “DR” means removing step 2.2 in the algorithm. Simple permutation results in a total of seven such algorithms: None (none of the three elements
is used), Log, NU, DR, NU-Log, DR-NU, and DR-Log. DR-NU-Log represents the solution algorithm in section 6 that employs all three elements.

![Solution time and linearization errors](image)

**Figure 2.14. Solution time and linearization errors using the solution algorithm in section 6 and alternative algorithms (Note: “Log” = logarithmic-sized linearization; “NU” = Non-uniform distribution of breakpoints; “DR” = Feasibility-based domain reduction)**

Figure 2.14 shows the solution time and linearization errors for link travel time, AV market share, and AV manufacturer revenue using all eight algorithms, for the base scenario in section 2.7.1.1. Three points are worth noting. First, the overall computational time will be primarily consumed by domain reduction if it is employed. But the benefits are also evident. Compared to the four algorithms without domain reduction
(Log, None, NU-Log, and NU), linearization errors are drastically reduced with domain reduction. Second, using non-uniform distribution of breakpoints results in substantial reduction in linearization error for market shares. Third, the use of logarithmic-sized linearization also contributes to reduction in solution time (e.g., comparing DR-NU-Log with DR-NU). It should be noted that the time savings from using logarithmic-sized linearization could be greater if considering a larger number of intervals to partition variable domains and larger network sizes.

![Figure 2.15](a) Sensitivity analysis of the tradeoff among the maximum number of iterations for domain reduction, the linearization error, and the computation time

Finally, we investigate the trade-offs among the maximum number of iterations for domain reduction, the linearization error, and the computation time for the base scenario. As shown in Figure 2.15(a), the linearization errors of travel time, market share, and revenue almost disappear after only two iterations. Thus we could reduce the computation time by considering two instead of three as the maximum number of iterations at the price of slightly increased linearization error. Figure 2.15(b) shows the tradeoff in computation time between domain reduction and solving the main MILP. As we spend more time on domain reduction (with greater number of iterations), the time needed for solving the main MILP will be generally reduced. This is because more iterations for domain reduction lead to a smaller feasible region for the main MILP. We also note that the reduction in solution time for the main MILP is quite small after three iterations, while the time for domain reduction continues to increase. Considering additionally the
miniscule improvement in linearization error when the maximum number of iterations is greater than three (Figure 2.15(a)), having more than three iterations is probably not worthwhile for domain reduction.

2.7.2. The Sioux Falls network

To further investigate the system performance with AVs and the efficiency of the proposed solution algorithm, we implement the model in the Sioux Falls network (Figure 2.16). The network has 76 links, 24 nodes, 12 of which are considered as trip generation / attraction zones (i.e., 132 OD pairs). The corresponding data are obtained from Transportation Test Networks (2017).

As in Section 2.7.1, we establish a base scenario with AVs and a scenario with only HVs. Model parameter values are also similar to the Singapore example with the following exceptions. The VOTTs of the two user classes are $7.2/hr and $23.3/hr when driving HVs (the same 20% VOTT saving is assumed for AV users). The price and the manufacturing cost for an HV are assumed to be $33,500 (which is the average market price of HVs in the US (USA Today, 2015)) and $30,150, the latter based on the assumption of 10% profit margin of a car manufacturer. An additional $10,000 in manufacturing cost is considered for an AV ($40,150). $0.247/mi and $0.280/mi respectively. The depreciation ratio in Eq. (2.4) is assumed 55% of the vehicle price ($0.55) over an average ownership period of 5 years ($5), and the average travel distance of a vehicle ($15,000 mi/year (AAA, 2015)). $1.1 accounts for overhead costs pertaining to 10% sales tax. $5 is set to 1.

The initial bounds and the number of intervals of the partitioned variables $h_a$, $P^m_w$, $\rho_m|_{m=AV}$, and $q_m|_{m=AV}$ are set the same as in the Singapore example. With the above setting, MILP (2.56) has 7099 binary variables, 17968 continuous variables, and 33751 constraints, which is solved to an optimality gap

---

1. Nodes 1, 2, 4, 5, 10, 11, 13, 14, 15, 19, 20, and 21 in Figure 2.16.
2. VOTTs are computed based on the annual household incomes of $30,090 and $96,870, which are associated with the 25th and 75th percentiles of the US population (US Census Bureau, 2015). The annual incomes are converted to hourly values assuming 2080 working hours per year and a VOTT equal to 50% of hourly income (US DOT, 2016).
of $10^2$ within 34 minutes, plus 103 minutes for solving the domain reduction MILPs to an optimality gaps of $10^{-10}$ (thus in total about 137 minutes). The average approximation errors of the nonlinear functions $t_a$, $P_w^{nm}$, $\tilde{\rho}_{m|m=AV}$ are respectively $4.9\times10^{-5}$ hours, $2.1\times10^{-4}$, and $2.5\times10^4$.

![Figure 2.16. The Sioux Falls network (Transportation Test Networks, 2017)](image)

Table 2.6 reports the optimal AV price and some aggregate results of system performance. An AV will be sold at $45,990, or a 12.7% profit margin, which is larger than 10% as for HVs. This translates to an annual profit of $17.24 million to the AV manufacturer. The total out-of-pocket cost of travelers will increase by 0.92% compared to the HVs-only scenario owing to the significant AV technology cost ($10,000). The total travel time in the network is reduced by 1.01% with AVs due to reduced vehicle headway and increased road capacity by 0.70%. The generalized travel cost will experience a slight increase (0.46%), suggesting that the effect of a higher AV price dominates the benefits of travel time reduction. For AV market penetration, travelers with a high VOTT are more inclined to use AVs than low VOTT travelers.
– about 8.3% of high-VOTT travelers will use AVs compared to 3.9% among low-VOTT travelers. Overall, AVs hold a 6.1% market share, which is larger than 2.4% as in the base scenario of the Singapore network. This can be explained by the larger VOTTs used in the Sioux Falls example than in the Singapore case, and the significantly larger overhead cost in Singapore (i.e., $b_{m,1} = 5.1$ in Singapore versus $b_{m,1} = 1.1$ in Sioux Falls).

Table 2.6. AV manufacturer strategy and transportation system performance under base scenario (Sioux Falls)

<table>
<thead>
<tr>
<th>AV manufacturer</th>
<th>System performance</th>
<th>Performance measure</th>
<th>With AVs</th>
<th>Only HVs</th>
<th>Change (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV price ($)</td>
<td></td>
<td>Total generalized cost ($)</td>
<td>611,209</td>
<td>608,391</td>
<td>0.46</td>
</tr>
<tr>
<td>Profit margin (%)</td>
<td></td>
<td>Total out of pocket cost ($)</td>
<td>465,186</td>
<td>460,890</td>
<td>0.92</td>
</tr>
<tr>
<td>Profit ($million)</td>
<td></td>
<td>Total time cost ($)</td>
<td>146,023</td>
<td>147,501</td>
<td>-0.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total travel time (hr)</td>
<td>9,670</td>
<td>9,678</td>
<td>-1.01</td>
</tr>
</tbody>
</table>

2.8. Conclusions

The use of autonomous vehicles (AVs) allows for more leisurely or productive use of in-vehicle time, which reduces travelers’ value of travel time (VOTT). In addition, AVs can move closer together than human-driven vehicles because of shorter safe reaction times of computers than human brains, which lead to increased road capacity. However, the use of new technologies in AVs means added manufacturing cost and higher price. Consequently, traveler adoption of AVs will trade VOTT savings with the additional out-of-pocket cost. This chapter develops an integrated model to characterize AV market penetration by accounting for the interplay among the AV manufacturer, travelers with heterogeneous VOTTs, and road infrastructure capacity. The overall problem is formulated as a mathematical program with complementarity constraints (MPCC). A linearization-based solution approach is developed which approximates the MPCC by a logarithmic-sized MILP. Non-uniform distribution of breakpoints and feasibility-based domain reduction are further employed to reduce the approximation error.
We implement the model and the solution approach in both simplified Singapore and Sioux Falls networks, with the computation showing high efficiency and effectiveness. In addition, several important findings are generated. In the base scenario with low VOTT savings (20%) and high AV technology cost ($10,000), the AV market share reaches 2.4% (Singapore) and 6.1% (Sioux Falls). The share is greater among high-VOTT travelers. However, the reduction in system travel time is marginal. Sensitivity analyses on the Singapore network further reveals that: (1) as traveler VOTT savings increases, AV price and AV manufacturer profit will increase. High-VOTT travelers are very sensitive to VOTT savings in choosing AVs, whereas low-VOTT travelers are much less so; (2) as AV technology becomes cheaper, benefits to the AV manufacturer, travelers, and network capacity will all increase; (3) as traveler cost perception variation decreases, AV price will drop. Low-VOTT travelers will use less AVs; (4) as market size increases, AV market share and network capacity will first increase and then decrease. A possible explanation is that larger market size implies greater congestion. As network congestion reaches a certain level, for some trips travel time savings cannot offset the added out-of-pocket cost of using AVs.

This chapter presents a start towards understanding the impact of AVs on urban transportation. Future research can be directed in several ways. First, the assumption of fixed total demand can be relaxed to capture potential induced demand after AVs are introduced, as may occur to the previously travel-restrictive population (seniors, youth, etc.). Second, competition of multiple car manufacturers producing both AVs and HVs could be considered. Third, benefits of AVs, including parking cost avoidance and shared-use mobility which reduces the travel expenses per rider, could be further and explicitly modeled. Finally, as this chapter uses a route based UE formulation, future research may consider alternative formulations, for example, link-node based formulation, which may result in a smaller number of complementary constraints and binary variables after linearization. This would be helpful when route enumeration is difficult for large networks. In this case, algorithms other than branch-and-cut, for example metaheuristics and column generation, may also be explored to solve the MILP. Overall, we hope that the model, solution approach,
and results presented in this chapter will stimulate further research in this important yet still wide-open area in the era of vast mobility transformation.

Appendix 2.1. Algorithm to identify the breakpoints

Following Vasudeva (2015), the recursive algorithm below is used to solve dynamic program (2.57).

All notations are already defined in section 2.5.1, except for Pred(v, s) which represents the predecessor of breakpoint \( v \) in the shortest s-interval path from breakpoint 1 to breakpoint \( v \).

---

**Algorithm**

1. **begin procedure**
2. \( E(v, 1) := e_{1v} \quad \forall v = \{2, \ldots, |\mathcal{V}|\} \)
3. \( \text{Pred}(v, 1) := 1 \quad \forall v = \{2, \ldots, |\mathcal{V}|\} \)
4. **for** \((s = 2, \ldots, |\mathcal{I}|)\)
5. \( \quad \text{for } (v = 2, \ldots, |\mathcal{V}|)\)
6. \( \quad \quad E(v, s) := E(v, s - 1) \)
7. \( \quad \quad \text{Pred}(v, s) := \text{Pred}(v, s - 1) \)
8. \( \quad \quad \text{if } (v - 1 \geq s) \) then
9. \( \quad \quad \quad \text{for } (v' = s \text{ to } v - 1)\)
10. \( \quad \quad \quad \quad \text{if } E(v', s - 1) + e_{v'v} < E(v, s) \) then
11. \( \quad \quad \quad \quad \quad E(v, s) := E(v, s - 1) + e_{v'v} \)
12. \( \quad \quad \quad \quad \quad \text{Pred}(v, s) := v' \)
13. \( \quad \quad \quad \text{End if} \)
14. \( \quad \quad \text{next } v' \)
15. \( \quad \text{next } v \)
16. **next** \( s \)
17. \( v := |\mathcal{V}| \)
18. \( s := |\mathcal{I}| \)
19. **while** \((v \neq 1)\)
20. \( \quad \text{save } h^w_v \text{ (or } h^{nm,v} \text{)} \)
21. \( \quad v := \text{Pred}(v, s) \)
22. \( \quad s := s - 1 \)
23. **end while**
24. **end procedure**
3. One-to-many matching and section-based formulation of autonomous ridesharing equilibrium

3.1. Introduction

Shared autonomous vehicles (SAVs) enable ridesharing — *multiple travelers share one vehicle at the same time* — which reduces VMT while serving the same level of travel demand. In addition, SAV travelers may enjoy lower fare than using current taxi/ridesourcing services due to the elimination of driver cost, and shorter waiting time as empty SAVs can be centrally controlled and proactively relocated before demand arises. This chapter makes a first attempt in the literature to model ridesharing with SAVs in a network equilibrium setting. Two major contributions are made. First and foremost, we propose a novel *one (SAV)-to-many (travelers) matching* characterization of the waiting times of an SAV and multiple travelers who share rides in the SAV while being matched via an online platform. In the most relevant literature of ridesharing network equilibrium with human-driven vehicles (HVs), however, waiting times of riders and vehicles due to matching are overlooked (Table 3.1). By introducing the notion of search intensity, the one-to-many matching characterization proposed in this chapter generalizes the one-to-one matching in previous taxi/ridesourcing studies (e.g., Yang et al. 2010; Zha et al. 2016) which ends when a vehicle is matched with one traveler. The generalization is nontrivial due to the nonlinearity of SAV/traveler waiting times in the number of travelers sharing rides in an SAV, the latter of which needs to respect the number of available seats in an SAV and be endogenously determined. Theoretical insights are derived by comparing the waiting times of an SAV/traveler due to matching with and without ridesharing.

In characterizing one-to-many matching, our contribution also includes comprehensive considerations of traveler-SAV matching possibilities. Specifically, a traveler can be matched with an SAV which has the same origin as the traveler, or with an en-route SAV that goes through the traveler’s origin and does not incur extra stopping other than picking up the traveler. For travelers matched with an SAV which has the
same origin, we derive the total waiting time of these travelers due to matching and subsequently meeting the matched SAV. In doing so, our derivation recognizes the fact that SAVs starting from an origin consist of those moving to the origin either for pickup, or as a result of proactive relocation. These two types of SAVs incur different total waiting times to travelers. For travelers matched with “going-through” SAVs, the total waiting time of travelers due to matching and waiting for the matched SAV is also derived.

The second major contribution made in this chapter is the introduction of a section-based formulation for SAV ridesharing equilibrium. Given that SAV ridesharing considered in this chapter allows: 1) an SAV and travelers in the SAV to have different ODs, and 2) different travelers in an SAV to have different ODs, modeling the utilization of SAV seat capacity by travelers on and off an SAV is critical yet challenging. This issue similarly arises in ridesharing network equilibrium with HVs, but in our view, has not been satisfactorily addressed (Table 3.1). Specifically, the literature of ridesharing network equilibrium with HVs has link- and route-based vehicle capacity constraints. Under the link-based constraints (Xu et al. 2015b; Di et al. 2018), it is possible to have a rider dropped off from a vehicle on the way and picked up by another vehicle to complete the trip. Such en-route transfer would be undesirable for riders and is not commonly seen in practice. The route-based constraints (Li et al. 2018) prevent en-route transfer, but do not free up the seat after a rider leaves the vehicle, leading to an underutilization of vehicle seat capacity along a vehicle route.

The proposed section-based formulation both prevents en-route transfer and frees up the seat after a traveler leaves the SAV, while respecting the SAV seat capacity constraint. Specifically, the notion of sections is introduced to represent traveler flows, in addition to routes and links on which SAV flows are modeled. We define a section as a sequence of zones connecting a traveler OD and traversed by at least one SAV route. Following this definition, multiple sections can be traversed by a common SAV route and thus compete for available seats of SAVs on the route. The utilization of SAV seat capacity is characterized at the section level, considering travelers of the section under study and of its competing sections. By doing
so, the section-based formulation further allows for capturing additional waiting time of travelers of a section due to limited available SAV seats on the section.

**Table 3.1. Synthesis of existing network equilibrium studies of ridesharing and the present study**

<table>
<thead>
<tr>
<th>Model capability</th>
<th>Distributed ridesharing with HVs</th>
<th>Centralized ridesharing with SAVs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Considers vehicle/rider wait time in one-to-many matching</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Considers vehicle relocation</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Prevents rider en-route transfer</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Frees up vehicle seat after a rider leaves the vehicle</td>
<td>NR*</td>
<td>✓</td>
</tr>
</tbody>
</table>

*NR means “not relevant”, as the indicated studies considered the same OD for a vehicle and riders in the vehicle.

Besides the above two major contributions, this chapter also contributes by modeling inter- and intra-zone *self-parking* with privately owned autonomous vehicles (AVs). AVs can drop off users (who own the AVs) at desired locations and then self-park in cheaper areas, thereby transforming urban parking demand pattern. To our knowledge, only Levin and Boyles (2015) and Correia and van Arem (2016) studied AV self-parking. Levin and Boyles (2015) incorporated a nested logit model of AV parking location (either at the traveler destination or home, which is more restricted than our case) into a modified four-step model. The main difference of our work from Correia and van Arem (2016) is that we consider AV self-parking as a separate stage of a trip, thus having its own equilibrium. In characterizing the equilibrium for AV self-parking, we explicitly consider parking search time and limited capacity of parking infrastructure that is shared by SAV/AV/HV.

Building on the proposed SAV ridesharing and AV self-parking formulations, we put forward a new *multimodal autonomous ridesharing and parking user equilibrium (MARPUE)* and provide a proof for its existence. MARPUE models the coexistence of (S)AVs and HVs in a mixed traffic environment with the market share of each vehicle type endogenously determined. Previous research only distinguished between HV and AV without considering ridesharing (Chen et al. 2016; van den Berg and Verhoef 2016;
Noruzoliaee et al. 2018). In our study, MARPUE further accounts for increased road capacity as (S)AVs can move closer together than HVs (Levin and Boyles 2015; Talebpour and Mahmassani 2016), and in-vehicle value of time (VOT) saving of (S)AV traveler(s) as they can perform more productive or leisurely activities than driving (Steck et al. 2018; Nazari et al. 2018). Anticipating traveler reactions as characterized by MARPUE, the supply decisions of SAVs, assumed to be made by a centralized transportation network company (TNC) and including SAV fleet size, fare rates, and vehicle allocation/relocation, are optimally determined and thoroughly investigated.

This chapter makes a helpful contribution to the growing efforts of modeling SAV operations, which so far has relied only on optimization methods (Liang et al. 2016; Ma et al. 2017; Levin 2017; Hyland and Mahmassani 2018) or simulation techniques (Fagnant and Kockelman 2014; Fagnant et al. 2015; Chen et al. 2016; Chen and Kockelman 2016; Azevedo et al. 2016; Boesch et al. 2016). To our knowledge, ridesharing was considered in only a few simulation studies (Farhan and Chen 2018; Levin et al. 2017; Fagnant and Kockelman 2018). While valuable insights were offered, existing simulation studies have not provided answers to two key questions: 1) the optimal supply (fleet size, fare rates, and vehicle allocation/relocation) of a centralized SAV service while competing with AV/HV, and 2) the equilibrium state of the overall transportation system that simultaneously allows for autonomous ridesharing, autonomous single-occupancy riding, conventional driving, and (S)AV self-parking.

Figure 3.1 outlines the overall model. Given the SAV purchase price, the profit-maximizing TNC determines its SAV fleet size, fare rates, and vehicle allocation/relocation in the road network, anticipating traveler reactions, SAV level of service, and road congestion. On the traveler side, MARPUE characterizes vehicle type choice, composed of four parts. First is an autonomous ridesharing user equilibrium (ARUE), which assigns travelers to SAVs using a section-based representation of the SAV traveler network. ARUE explicitly accounts for one-to-many matching and the SAV seat capacity constraint on each section. Second, an autonomous parking user equilibrium (APUE) is established to capture AV self-parking after dropping off users, given parking fee and capacity of parking infrastructure that is shared by HV/AV/SAV. Third, a
link-node based user equilibrium (UE) for loading HVs and occupied AVs to a road network is formulated which considers the presence of SAV flow and road capacity increase brought by smaller headway of (S)AVs. Lastly, ARUE, APUE, and UE for HVs and occupied AVs are integrated with endogenous HV/AV/SAV market shares. We formulate MARPUE as nonlinear complementarity and variational inequality problems, and develop its equivalent nonlinear program using the primal-dual gap function. The overall TNC-traveler-infrastructure problem is formulated as a mathematical program with complementarity constraints (MPCC) and solved using the active-set algorithm (Lawphongpanich and Yin 2010).

Figure 3.1. The overarching framework of this study

Several original insights are obtained from numerical implementation of the model. First, we find that SAV ridesharing brings much reduced traveler waiting time while being matched, increased TNC profit, and large SAV market share compared to a system in which ridesharing is not allowed in using SAVs. Second, the overall transportation system will benefit from increasing SAV size, but the marginal effect is diminishing. Third, vacant SAV miles will incur, in which miles due to picking up travelers substantially outnumber miles due to SAV relocation. Fourth, AV self-parking significantly degrades the overall system.
performance by increasing traveler en-route time and generalized cost, total and occupied VMT, parking demand, and reduced TNC profit. In fact, the vacant VMT from AV self-parking turns out to be much larger than the vacant VMT from SAVs. This suggests the negative externalities imposed by selfish AVs to reduce their own parking cost at the expense of greater system cost. These insights, which are first-of-its-kind in a mixed environment with SAV/AV/HV, advance the understanding of how transportation system performance can be reshaped by SAV ridesharing.

The rest of the chapter is organized as follows. Section 3.2 presents the basic setup. Formulations of the MARPUE and the TNC problem are detailed respectively in sections 3.3 and 3.4. Section 3.5 presents the solution method. Numerical analyses are discussed in section 3.6. The chapter concludes in section 3.7.

3.2. Basic setup

This section presents the basic setup for model formulation. Section 3.2.1 delineates the representation of a road network and introduces the notion of sections which are associated with SAV travelers. The workflow of each vehicle type is outlined in section 3.2.2. The notations used herein and throughout this chapter are detailed in Appendix 3.1.

3.2.1. Representation of a road network and notion of sections

We characterize an urban road network as a directed graph $G(N, A)$ comprising nodes $n \in N$ and directional links $a \in A$ representing intersections and streets respectively. Some of the nodes are considered as zones that generate/attract travel demand. Traveler demand from zone $i$ to zone $j$, referred to as OD $(i, j)$, is assigned to a set of routes $r \in R$. In Figure 3.2(a), $i_1, i_2,$ and $i_3$ are zones and $n$ is a non-zone node. We consider three ODs in the network: $(i_3, i_2)$, $(i_1, i_3)$, and $(i_2, i_3)$. Figure 3.2(b) shows the routes for the three ODs. For instance, three routes are possible for OD $(i_1, i_3)$: $r_1(a_1)$, $r_2(a_2, a_3, a_4)$, and $r_3(a_2, a_3, a_5)$. $r_1(a_1)$ means route 1 that traverses link $a_1$. Likewise for $r_2(a_2, a_3, a_4)$ and $r_3(a_2, a_3, a_5)$. Such a link-route based representation of road networks works well when ridesharing is not involved. In this case,
travelers and vehicles have the same ODs. Traveler flow and vehicle flow on each link and each route only differ by a vehicle occupancy rate.

The use of SAVs involves ridesharing. Along its route, an SAV can have multiple travelers whose ODs may be different: SAV travelers can be picked up (from their origin) and dropped off (in their destination) in the middle of an SAV’s route. Thus, the origin and destination of an SAV traveler are not necessarily the same as the start and end of the SAV route the traveler takes. SAV traveler flow can no longer be considered proportional to SAV flow. To represent SAV traveler flows, below we introduce the notion of sections which is inspired by the transit assignment literature (De Cea and Fernández, 1993; Szeto and Jiang, 2014).

**Definition 1: Section.** A section \( l \in L \) is associated with an SAV traveler OD and defined by a sequence of zones. The sequence starts from the origin zone of the SAV traveler OD, ends at the destination zone of the SAV traveler OD, and must be traversed by at least one SAV route.

The requirement that a section must be traversed by at least one SAV route ensures the usability of the section. If no SAV route traverses a section, then SAV travelers of the section cannot be picked up and dropped off by SAVs. We use Figure 3.2(c) to illustrate how sections are constructed given SAV traveler ODs and SAV routes. Two sections, \( l_1 \) and \( l_2 \), are associated with SAV traveler OD (\( i_1, i_3 \)). \( l_1 \) is defined by zone sequence \( i_1 \rightarrow i_3 \) and traversed by route \( r_1(a_1) \). \( l_2 \) is defined by zone sequence \( i_1 \rightarrow i_2 \rightarrow i_3 \) and traversed by routes \( r_2(a_2, a_3, a_4) \) and \( r_3(a_2, a_3, a_5) \). One section, \( l_3 \), is associated with SAV traveler OD (\( i_1, i_2 \)). \( l_3 \) is defined by zone sequence \( i_1 \rightarrow i_2 \) and traversed by routes \( r_2(a_2, a_3, a_4) \) and \( r_3(a_2, a_3, a_5) \). One section, \( l_4 \), is associated with SAV traveler OD (\( i_2, i_3 \)). \( l_4 \) is defined by zone sequence \( i_2 \rightarrow i_3 \) and traversed by routes \( r_2(a_2, a_3, a_4), r_3(a_2, a_3, a_5), r_4(a_3, a_4), \) and \( r_5(a_3, a_5) \).
Figure 3.2. Representations of road network, routes for SAVs, and sections for SAV travelers

From the above illustration, it is clear that an SAV route may traverse multiple sections. For example, route \( r_2(a_2, a_3, a_4) \) traverses sections \( l_2, l_3, \) and \( l_4 \). This is important as SAV travelers of different sections will interact by competing for seats in the same SAVs whose route traverses these sections. In view of this, below we introduce the notion of a competing section.

**Definition 2: Competing section.** Section \( l' \) is a competing section of section \( l \) if \( l' \) and \( l \) are traversed by a common route, and either of the following conditions is satisfied. We use binary indicator \( \chi_{l'l'} \) to show if \( l' \) is a competing section of \( l \).

(a) \( l' \) and \( l \) have the same origin but different destinations. SAV travelers of the two sections start from the same zone and compete for seats of SAVs whose route traverses both sections. In this case, \( l \) and \( l' \) are mutually competing: \( \chi_{l'l'} = \chi_{l'l} = 1 \). For example, \( l_2 \) and \( l_3 \) in Figure 3.2(c) are mutually competing.

(b) On the commonly traversed route, the origin of \( l \) is between the origin and the destination of \( l' \). Travelers of section \( l' \) in an SAV thus limit the available seats for travelers of section \( l \) boarding
the same SAV. In this case, \( \chi_{l l'} = 1 \) but \( \chi_{l' l} = 0 \). For example, \( l_2 \) in Figure 3.2(c) is a competing section of \( l_4 \) but not vice versa.

3.2.2. Workflows of vehicles

Each HV/AV trip consists of two stages: en-route and parking. At the en-route stage, an HV/AV moves its user from his/her origin zone to destination zone. Upon arriving at the destination zone, an HV user will search in the same zone for a parking spot, park the HV, and walk to the desired location in the zone. In contrast, an AV user will get off at the desired location in the destination zone, and have the AV self-park in any zone by taking the shortest route.

The workflow of an SAV is more complicated. The movement of an SAV can be in one of three states: in-service, pickup, and relocation, denoted as \( s = I, II, III \). The in-service state \( (s = I) \) is shown in the three colored boxes in Figure 3.3. An SAV starts its in-service state right after picking up the first traveler. Then onboard travelers may be dropped off (at their destinations) and new travelers get onboard (at their origins). After dropping off all travelers onboard, the SAV exits the in-service state, and faces two choices:

![Diagram of SAV workflow](image)

**Figure 3.3. SAV workflow**

(a) Remaining in the current zone, waiting to be matched with future traveler(s) from any zone and then moving unoccupied to that zone to pick up the matched traveler(s). The movement is referred
to as the *pickup* state \((s = II)\). While waiting to be matched, we assume that the SAV will search for a parking spot in the current zone. The parking search terminates when a parking spot is found, or the SAV is matched with new traveler(s), whichever comes sooner.

(b) Relocating from the current zone to a new zone based on anticipated demand. The movement is referred to as the *relocation* state \((s = III)\). Note that the SAV may be matched with new traveler(s) from the zone it relocates to before arriving at the zone. If the SAV is unmatched when arriving at the zone, the SAV will search for parking in the zone. The parking search terminates when a parking spot is found, or the SAV is matched with new traveler(s) in the zone, whichever comes sooner.

The TNC’s choices with respective to SAV movement states in the road network are characterized by variables \(y^s_r (\forall s \in S, r \in R)\), which signifies the SAV flow rate on route \(r\) in movement state \(s\). While it is possible to observe intra-zone SAV trips in the pickup state \((y^{II}_r \geq 0, O(r) = D(r))\), where \(O(r)\) and \(D(r)\) denote the origin and destination zones of route \(r\), we assume that in-service and relocating SAVs do not contribute to intra-zone flow (i.e., \(y^I_r = y^{III}_r = 0, O(r) = D(r)\)).

Note that an SAV traveler waiting at zone \(i \in Z\) may be picked up by an in-service SAV that originates from \(i\) or by an in-service SAV going through \(i\). In this study, “an SAV route going through a zone” means that the zone is an *intermediate* zone of the route. A “going-through” SAV can pick up a traveler when the SAV is matched with the traveler no later than the time the SAV arrives at \(i\). We introduce variable \(g_{ri}\) which describes the possibility that travelers waiting in zone \(i\) can be picked up by in-service SAVs on route \(r\) which goes through \(i\).

\[
    g_{ri} = 0.5 \zeta_{ri} \left[ 1 + \text{sgn}(t_{ri} - \omega^Y_r) \right] \quad \forall r \in R; i \in Z
\]

(3.1)

where \(\zeta_{ri}\) is a route-zone correspondence indicator equaling 1 if zone \(i\) is an *intermediate* zone on route \(r\). \(t_{ri}\) is the SAV travel time from the origin of route \(r\) to zone \(i\). As an SAV which can pick up travelers in zone \(i\) may be already in the middle of route \(r\), \(t_{ri}\) represents a conservative estimate of the SAV travel
time (which probably is the best we can do in modeling a static equilibrium). $\omega_t^i$ is the matching time of such an SAV. There are three possibilities for $g_{ri}$: 1) if $t_{ri} > \omega_t^i$, $g_{ri} = 1$, i.e., SAVs on route $r$ can pick up traveler(s) waiting at zone $i$; 2) if $t_{ri} < \omega_t^i$, $g_{ri} = 0$, i.e., SAVs on route $r$ cannot pick up traveler(s) waiting at zone $i$; 3) if $t_{ri} = \omega_t^i$, $g_{ri} = 0.5$, i.e., SAV-traveler matching occurs exactly when the SAV arrives at $i$. In this case, a 50% chance for pickup is considered. The derivation of $\omega_t^i$ and $t_{ri}$ is presented in the next section (Eqs. (3.14) and (3.19)).

3.3. Formulation of the multimodal autonomous ridesharing and parking user equilibrium (MARPUE)

With the basic setup in section 3.2, this section formulates MARPUE which consists of four components: ARUE (section 3.3.1), APUE (section 3.3.2), UE for occupied HVs/AVs (section 3.3.3), and endogenous market share of each vehicle type (section 3.3.4). Section 3.3.5 formalizes the overall MARPUE and provides proof of its existence.

3.3.1. Autonomous ridesharing user equilibrium (ARUE) for SAV travelers

As defined in section 3.2, SAV travelers for each OD are assigned to sections associated with the OD. It is assumed that each SAV traveler makes choice of a section such that under ARUE, no SAV traveler can be better off by unilaterally switching to a different section. This assumption seems reasonable given that the static ARUE intends to capture the long-term behavior of SAV travelers. Constraints (3.2)-(3.3) formalize the definition of ARUE.

$$0 \leq q_l \perp h_l - H_{ij} \geq 0 \quad \forall l \in L_{ij}; (i,j) \in \Omega \quad (3.2)$$

$$\sum_{l \in L_{ij}} q_l = Q_{ij}^m \quad \forall (i,j) \in \Omega; \ m = SAV \quad (3.3)$$

where $q_l$ is SAV traveler flow of section $l$, i.e., travelers who have the same origin and destination as section $l$ and traverse in sequence all the zones on section $l$. $h_l$ is the generalized travel cost of SAV travelers of
section $l$. $L_{ij} \subseteq L$ denotes the set of sections associated with OD $(i,j)$. The complementarity constraint (3.2) means that at equilibrium no SAV traveler takes a section that has a larger generalized travel cost than the minimum generalized travel cost of the OD ($H_{ij}$). Eq. (3.3) expresses SAV traveler flow conservation: the sum of SAV traveler flow on all sections connecting an OD equals the total SAV traveler demand for the OD.

Eq. (3.4) expresses the generalized travel cost of SAV travelers of section $l$, which is the sum of four components. $\sigma_{O(l)}$ represents the average waiting time due to matching and meeting at the origin of $l$ (see section 3.3.1.1), $(\eta_l + \tilde{\eta}_l)$ is the total extra waiting time due to limited SAV seat capacity for an SAV traveler on $l$ (see section 3.3.1.2). $t_l$ is the in-vehicle travel time on $l$. $f_l$ is the fare charged by the TNC on $l$ (see section 3.3.1.3). $\gamma_1$ and $\gamma_2^m$ ($m = SAV$) are SAV traveler VOTs for waiting and in-vehicle time.

$$h_l = \gamma_1 \sigma_{O(l)} + \gamma_1 (\eta_l + \tilde{\eta}_l) + \gamma_2^m t_l + f_l \quad \forall l \in L; \ m = SAV$$

(3.4)

3.3.1.1. Waiting time due to matching and meeting

An SAV traveler first waits to be matched with an SAV, and then waits to meet the matched SAV. In this section, we derive in sequence the waiting time due to matching by introducing a one-to-many matching function, and the waiting time due to meeting based on the relative locations of the SAV and the matched traveler(s).

To derive the SAV traveler waiting time due to matching, we start by expressing the SAV traveler flow originating from zone $i \in Z$ ($n^q_i$) and the SAV flow that can serve those travelers ($\bar{n}^y_i$), as Eqs. (3.5)-(3.6).

$$n^q_i = \sum_{l \in L: O(l) = i} q_l \quad \forall i \in Z$$

(3.5)

$$\bar{n}^y_i = \sum_{r \in R: O(r) = i} y^r_l + \sum_{r \in R: O(r) \neq i} g_{rl} \left( y^r_l - \frac{1}{K} \sum_{l \in L} \epsilon_{il} q^r_{il} \right) \quad \forall i \in Z$$

(3.6)
where $\kappa$ is the number of seats in an SAV. $\epsilon_{it}$ is a section-zone correspondence indicator equaling 1 if zone $i$ is an intermediate zone on section $l$. $q_{lt}$ denotes SAV traveler flow of section $l$ and using SAV route $r$. The calculation of $q_{lt}$ is discussed in detail in section 3.3.1.2.

The right-hand side (RHS) of Eq. (3.5) is the sum of SAV travelers on all sections originating from zone $i$. In Eq. (3.6), the RHS decomposes the SAV flow that can serve travelers originating from zone $i$ into two parts. The first part is the in-service SAV flow originating from $i$. The second part is the in-service SAV flow which goes through zone $i$, without incurring SAV waiting (as governed by $g_{r_{-i}}$). The subtracted term in the parentheses uses $\kappa$ to convert the number of occupied seats when the going-through SAVs arrive at zone $i$ to the equivalent occupied SAV flow.

With $n_{i}^{q}$ and $\tilde{n}_{i}^{y}$, we now specify the matching function. It is important to note that an SAV can be matched with multiple travelers as long as the SAV has available seats. Hence, one-to-many matching is appropriate. In one-to-many matching, we introduce the notion of a traveler group, which refers to a group of travelers originating from the same zone and sharing the same SAV. Thus, on the one hand, each traveler group is matched with one SAV. Each traveler in a traveler group should wait until everyone in the group is matched with an SAV. On the other hand, each SAV is also matched with one traveler group. An SAV should wait until matched with all travelers in a traveler group.

It is intuitive that the rate of SAV-traveler matching depends on the size of traveler groups. To understand this further, we define search intensity following Petrongolo and Pissarides (2001).

**Definition 3: Search intensities of a traveler group and an SAV.** In ridesharing matching, the search intensity of a traveler group is the group size, because each traveler is considered to perform individual search for an SAV without coordinating with others in the group. The search intensity of an SAV is the maximum number of traveler groups that can be accommodated by the SAV, which depends on the traveler group size and the SAV seat capacity.
Following Definition 3, Eq. (3.7) expresses the search intensity of a traveler group in zone $i$ ($\sigma^q_i$). The min and max operators in Eq. (3.7) impose the constraint that the size of a traveler group should be between 1 and $\kappa$.

$$\sigma^q_i = \min\left\{\kappa, \max\left\{1, \frac{n^q_i}{n^q_i}\right\}\right\} \quad \forall i \in Z \tag{3.7}$$

Based on Eq. (3.7), the SAV flow that is matched with SAV traveler groups in zone $i$, $n^y_i$, can be derived as in Eq. (3.8).

$$n^y_i = \frac{n^q_i}{\sigma^q_i} \quad \forall i \in Z \tag{3.8}$$

Eq. (3.9) expresses the search intensity of an SAV in zone $i$, $\sigma^y_i$, which is the number of traveler groups an SAV can accommodate in the zone:

$$\sigma^y_i = \frac{\kappa}{\sigma^q_i} \quad \forall i \in Z \tag{3.9}$$

In one-to-many matching, the matching rate in zone $i$, $m_i$, can be described as the number of traveler groups matched per hour or the number of SAVs matched per hour, which are the same under equilibrium. $m_i$ is specified as a function of the total search intensities of all traveler groups and of all SAVs (Eq. (3.10)). The total search intensity of all traveler groups is equal to the search intensity of one traveler group ($\sigma^q_i$), multiplied by the number of traveler groups to be matched ($N^q_i$). Similarly for the total intensity of all SAVs ($\sigma^y_i N^y_i$).

$$m_i = M_i\left(\sigma^q_i N^q_i, \sigma^y_i N^y_i\right) \quad \forall i \in Z \tag{3.10}$$

In Eq. (3.10), $N^q_i$ is derived using Little’s law: $N^q_i = \left(n^q_i / \sigma^q_i\right) \omega^q_i$, where $n^q_i / \sigma^q_i$ gives the number of traveler groups to be matched per hour and $\omega^q_i$ is the waiting time of an SAV traveler group in matching.
By the matching rate definition, \( m_i = n_i^q / \sigma_i^q \). Likewise, for SAVs: \( N_i^y = n_i^y \omega_i^y \) where \( \omega_i^y \) is the waiting time of an SAV in matching, and \( m_i = n_i^y \). Replacing \( N_i^q \) and \( N_i^y \) by \( (n_i^q / \sigma_i^q) \omega_i^q \) and \( n_i^y \omega_i^y \) in Eq. (3.10), we obtain:

\[
m_i = M_i(n_i^q \omega_i^q, \sigma_i^y n_i^y \omega_i^y) \quad \forall i \in Z
\]

In this study, we consider the Cobb-Douglas form for \( m_i \), which is used in existing taxi/ridesourcing literature (Yang et al. 2010; Yang and Yang 2011; He and Shen 2015; Zha et al. 2016; Wang et al. 2016; Xu et al. 2017).

\[
m_i = B_i(n_i^q \omega_i^q)^{b_1}(\sigma_i^y n_i^y \omega_i^y)^{b_2} \quad \forall i \in Z
\]

where \( b_1 \) and \( b_2 \) are matching rate elasticities. It is generally expected that \( b_1, b_2 \in (0,1] \). If \( b_1 + b_2 > 1 \) (\( = 1 \), or \(< 1 \)), the matching function exhibits increasing (constant, or decreasing) returns-to-scale. \( B_i > 0 \) is a zone-specific constant reflecting the meeting location properties, e.g., point meeting as for taxi stands (larger \( B_i \)) versus zonal meeting as for ride-hailing taxis which cruise on streets to find riders (smaller \( B_i \)).

The waiting times of travelers and SAVs can then be obtained as \( \omega_i^q = N_i^q / m_i \) and \( \omega_i^y = N_i^y / m_i \). Simple algebra on Eq. (3.12) and \( m_i = n_i^q / \sigma_i^q = n_i^y \) yields the waiting time functions of travelers and SAVs. As shown in Eq. (3.13), traveler waiting time in a zone is associated with the traveler group size \( \sigma_i^q \), the SAV traveler flow originating the zone \( n_i^q \), and the number of SAVs to be matched in the zone \( n_i^y \omega_i^y \). Similarly for the SAV waiting time in Eq. (3.14). In the special case of \( b_1 = b_2 = 1 \) as in some taxi/ridesourcing studies (Yang et al., 2002; Wong et al., 2008; Qian and Ukkusuri, 2017), Eqs. (3.13)-(3.14) collapse to \( \omega_i^q = 1/(\kappa B_i N_i^q) \) and \( \omega_i^y = 1/(\kappa B_i N_i^y) \).

\[
\omega_i^q = (\kappa)^{-b_2/b_1}(\sigma_i^q)^{-b_2/b_1} \left[ (B_i)^{-1/b_1}(n_i^q)^{1-b_1/b_2}(n_i^y \omega_i^y)^{-b_2/b_1} \right] \quad \forall i \in Z
\]

(3.13)
\[
\omega_i^y = \left(\sigma_i^y\right)^{-1} \left[ \beta_i \frac{1}{b_2} \left( n_i^q \right)^{1-b_2} \left( n_i^q \omega_i^q \right)^{\frac{b_1}{b_2}} \right] \quad \forall i \in Z
\] (3.14)

The terms outside the brackets in Eqs. (3.13) and (3.14), i.e., \(\frac{b_2}{b_1} \left( \sigma_i^q \right)^{\frac{b_2-1}{b_1}} \) and \(\left( \sigma_i^y \right)^{-1}\), are less than or equal to one given that \(\sigma_i^q, \sigma_i^y \in [1, \kappa]\) and \(b_1, b_2 \in (0,1]\). This suggests that holding the bracket terms constant, the waiting times of a traveler group and an SAV in matching will be reduced (at least not increase) with ridesharing.

The following two propositions further explore the relationship between traveler group and SAV waiting times in matching, and parameters \(b_1\) and \(b_2\) in the Cobb-Douglas matching function.

**Proposition 1.** Given the SAV traveler flow originating from zone \(i\) \(\left(n_i^q\right)\) and SAV waiting time in matching \(\omega_i^q\), the waiting time of an SAV traveler group in matching \(\omega_i^q\) is non-decreasing with ridesharing vs. without ridesharing if \(b_2 > 0.5\). In addition, \(\omega_i^q\) becomes larger with larger \(\sigma_i^q\), i.e., more intense ridesharing. Similarly for \(\omega_i^y\) given \(n_i^q\) and \(\omega_i^q\).

**Proof.** Let us first look at \(\omega_i^q\). Recall that \(\frac{n_i^q}{\sigma_i^q} = \omega_i^y\). We replace \(n_i^q\) by \(\frac{n_i^q}{\sigma_i^q}\) in Eq. (3.13) and obtain

\[
\omega_i^q = \left(\kappa\right)^{-\frac{b_2}{b_1}} \left(\sigma_i^q\right)^{\frac{2b_2-1}{b_1}} \left[ \beta_i \frac{1}{b_2} \left( n_i^q \right)^{1-b_2-b_1} \left( \omega_i^q \right)^{\frac{b_1}{b_2}} \right].
\]

If \(b_2 > 0.5\), \(\left(\sigma_i^q\right)^{\frac{2b_2-1}{b_1}} \geq 1\) as \(\sigma_i^q \geq 1\) and \(b_1 > 0\). Thus holding \(n_i^q\) and \(\omega_i^q\) constant, \(\omega_i^q\) is non-decreasing with ridesharing vs. without ridesharing. As \(\frac{2b_2-1}{b_1} > 0\), it is evident that \(\omega_i^q\) becomes larger with larger \(\sigma_i^q\), i.e., more intense ridesharing.

For \(\omega_i^y\), we again replace \(n_i^q\) by \(\frac{n_i^q}{\sigma_i^q}\) in Eq. (3.14) and obtain

\[
\omega_i^y = \left(\kappa\right)^{-1} \left(\sigma_i^y\right)^{\frac{2b_2-1}{b_2}} \left[ \beta_i \frac{1}{b_2} \left( n_i^q \right)^{1-b_2-b_1} \left( \omega_i^q \right)^{\frac{b_1}{b_2}} \right].
\]

Holding \(n_i^q\) and \(\omega_i^q\) constant, it is easy to see that similar conclusions can be reached for \(\omega_i^y\) with respect to \(b_2\) and \(\sigma_i^y\). \(\square\)
**Proposition 2.** Given the traveler group search intensity \((\sigma_i^q)\) and SAV waiting time in matching \((\omega_i^y)\) in a zone \(i\), the waiting time of an SAV traveler group in matching \((\omega_i^q)\) increases with the SAV traveler flow originating from zone \(i\) \((n_i^q)\) if the Cobb-Douglas matching function exhibits decreasing returns-to-scale and \(n_i^q > 1\). Similarly for \(\omega_i^y\) given \(\sigma_i^q\) and \(\omega_i^q\).

**Proof.** As shown in the proof for Proposition 1, \(\omega_i^q = (\kappa)\frac{b_2}{b_1}\left(\sigma_i^q\right)^{\frac{b_2 - 1}{b_1}}\left(B_i\right)^{\frac{1}{b_1}}\left(n_i^q\right)^{\frac{1 - b_2}{b_1}}\left(n_i^q\right)^{- \frac{b_2}{b_1}}\). Holding \(\sigma_i^q\) and \(\omega_i^y\) constant, decreasing returns-to-scale of the Cobb-Douglas matching function means \(b_1 + b_2 < 1\). Thus \(\frac{1 - b_2}{b_1} > 0\). If the SAV traveler flow is greater than 1 \((n_i^q > 1)\), it is clear that \(\omega_i^q\) increases with \(n_i^q\). The same can be said for \(\omega_i^y\) holding \(\sigma_i^q\) and \(\omega_i^q\) constant. □

Once the matching ends, travelers originating from zone \(i\) may need to wait further for meeting the matched SAVs. As shown in Eq. (3.6), such SAVs either originate from or go through \(i\). We derive \(\omega_i\), the average waiting time of an SAV traveler originating from zone \(i\) due to matching and meeting. \(\omega_i\) is calculated as the sum of total traveler waiting time for SAVs both originating from and going through zone \(i\) (termed \(TTW_i^+\) and \(TTW_i^-\)) divided by the SAV traveler flow originating from the zone:

\[
\omega_i = \frac{TTW_i^+ + TTW_i^-}{\sum_{l \in L: O(l) = i} q_l} \quad \forall i \in Z
\]  

(3.15)

We need to further express \(TTW_i^+\) and \(TTW_i^-\). For \(TTW_i^+\), it is the SAV traveler flow originating from zone \(i\) and taking SAVs that also originate from \(i\) (i.e., \(\sum_{r \in R: O(r) = i} \sum_{l \in L: O(l) = i} q_{lr}\)), multiplied by the average waiting time incurred by each of these travelers. Since SAVs originating from \(i\) consist of those moving to \(i\) in pickup state \((s = II)\) and in relocation state \((s = III)\), the average waiting time of a traveler is the weighted average of waiting time for taking SAVs moving to \(i\) in the two states, weighted by the SAV flows in the two states. The waiting time for SAVs in pickup state is the sum of matching time and
pickup time, i.e., $\omega_i^q + t_r$. The waiting time for SAVs in relocation state is the maximum of matching time and relocation time, i.e., $\max\{\omega_i^q, t_r\}$. Eq. (3.16) gives the final expression for $TTW_i^+$. A small positive number $\varepsilon$ is added to avoid zero in the denominator.

$$TTW_i^+ = \sum_{r \in R : O(r) = i} \left[ y_r^{II} (\omega_i^q + t_r) + y_r^{III} \max\{\omega_i^q, t_r\} \right] \sum_{r \in R : O(r) = i} \sum_{i \in L : O(i) = i} q_{ir} \quad \forall i \in Z \quad (3.16)$$

$TTW_i^-$ is computed in Eq. (3.17) as the SAV traveler flow originating from zone $i$ and taking SAVs that go through $i$, multiplied by the average waiting time incurred by each of these travelers. We use $\zeta_{ri}$ because zone $i$ must be an intermediate zone on the SAV route. $\max\{\omega_i^q, t_{ri}\}$ reflects the fact that an SAV traveler can only be picked up when the traveler is matched and the SAV has arrived at zone $i$ from its origin.

$$TTW_i^- = \sum_{r \in R} \zeta_{ri} \max\{\omega_i^q, t_{ri}\} \sum_{i \in L : O(i) = i} q_{ir} \quad \forall i \in Z \quad (3.17)$$

Eqs. (3.18)-(3.19) shows how $t_r$ and $t_{ri}$ are calculated, where $\delta_{ar}$ is the link-route correspondence indicator equaling 1 if link $a$ lies on route $r$. $\delta_{ar}$ is a link-route-zone correspondence indicator equating 1 if: 1) link $a$ lies on route $r$; 2) zone $i$ lies on route $r$; 3) the end of link $a$ is no later than $i$ on route $r$.

$$t_r = \sum_{a \in A} \delta_{ar} t_a \quad \forall r \in R \quad (3.18)$$

$$t_{ri} = \sum_{a \in A} \delta_{ar}^i t_a \quad \forall r \in R; \ i \in Z \quad (3.19)$$

3.3.1.2. Waiting time due to limited SAV seat capacity

In addition to waiting for matching and meeting, another type of waiting may occur when there are not enough seats provided by SAVs. To capture this, we establish the following SAV seat capacity constraint on each section.
\[ q_l + \tilde{q}_l \leq \kappa \left( \sum_{r \in R : d(r) \neq 0(l)} \lambda_{lr} g_{r, O(l)} y_r^l l + \sum_{r \in R : d(r) = 0(l)} \lambda_{lr} y_r^l \right) \quad \forall l \in L \]  

(3.20)

where \( \tilde{q}_l \) is the portion of SAV traveler flow of competing sections of section \( l \) and using routes that traverse \( l \). \( \lambda_{lr} \) is a section-route correspondence indicator equaling 1 if route \( r \) traverses section \( l \), and 0 otherwise.

The RHS of Eq. (3.20) denotes total SAV seat flow on section \( l \), which equals the number of seats per SAV (\( \kappa \)) times the total SAV flow on the section. Inside the parentheses, the first term refers to the in-service SAV flow going through section \( l \). \( g_{r, O(l)} \) is used to ensure that only those SAVs that can pick up travelers at the origin of section \( l \) are counted. The second term represents the in-service SAV flow whose routes start at zone \( i \).

The left-hand side (LHS) of Eq. (3.20) denotes total traveler flow competing for SAV seats on section \( l \), which is SAV traveler flow of section \( l \) (\( q_l \)), plus SAV traveler flow of competing sections of section \( l \) and using routes that traverse \( l \) (\( \tilde{q}_l \)). \( \tilde{q}_l \) is calculated as:

\[ \tilde{q}_l = \sum_{l' \in L \setminus l} \chi_{l l'} \sum_{r \in R} \lambda_{lr} q_{l' r} \quad \forall l \in L \]  

(3.21)

where \( q_{l' r} \) is the SAV traveler flow of section \( l' \) and using SAV route \( r \), which is computed in Eq. (3.22).

The basic idea of Eq. (3.22) is to assign SAV travelers of a section to different routes based on SAV flows of these routes. In Eq. (3.22), \( \pi_{lr} \) is the SAV flow of route \( r \) that traverses section \( l \). \( \pi_{lr} \) is calculated by Eq. (3.23), with the two terms already explained in Eq. (3.20). A small positive number \( \varepsilon \) in the denominator is added to avoid zero when \( \pi_l = 0 \). In Eq. (3.24), \( \pi_l \) is the total SAV flows on all routes traversing section \( l \). It can be easily seen that \( \pi_l \) is equal to the parenthesized term on the RHS of Eq. (3.20).

\[ q_{lr} = \frac{\pi_{lr}}{\pi_l + \varepsilon} q_l \quad \forall l \in L; r \in R \]  

(3.22)
\[ \pi_{l'r} = \begin{cases} \lambda_{l'r} y_{r}^{l} & \text{if } O(r) \neq O(l) \\ \lambda_{l'r} y_{r}^{l} & \text{if } O(r) = O(l) \end{cases} \quad \forall l \in L; r \in R \quad (3.23) \]

\[ \pi_{l} = \sum_{r \in R} \pi_{l'r} \quad \forall l \in L \quad (3.24) \]

Following Larsson and Patriksson (1999), Eq. (3.20) represents a side constraint of ARUE. Therefore, its Lagrange multiplier, \( \eta_{l} \), is used in the complementarity constraint (3.25). \( \eta_{l} \) can be interpreted as the extra waiting time for SAV travelers of section \( l \) due to limited SAV seat capacity on the section. Obviously, \( \eta_{l} = 0 \) when \( \kappa \pi_{l} - q_{l} - \bar{q}_{l} > 0 \).

\[ 0 \leq \eta_{l} \perp \kappa \pi_{l} - q_{l} - \bar{q}_{l} \geq 0 \quad \forall l \in L \quad (3.25) \]

An SAV route may traverse multiple sections. Therefore, the waiting time for SAV travelers of section \( l \) due to limited SAV seat capacity arises not only from \( \eta_{l} \), but from similar waiting time for SAV travelers of competing sections (\( \bar{\eta}_{l} \)). The calculation of \( \bar{\eta}_{l} \) is shown in Eq. (3.26). For a competing section \( l' \), its contribution to the extra waiting time for SAV travelers of section \( l \) is \( \eta_{l'} \) times the proportion of SAV flow on routes that traverse sections \( l \) and \( l' \), in the total SAV flows over all routes traversing \( l \), i.e., \( \lambda_{l'r} \frac{\pi_{l'r}}{\pi_{l} + \varepsilon} \).

A small number \( \varepsilon \) is added to both the denominator and the numerator. The addition of \( \varepsilon \) to the denominator is to avoid zero in the denominator when \( \pi_{l} = 0 \). The addition of \( \varepsilon \) to the numerator is to prevent \( \bar{\eta}_{l} = 0 \), which would occur when no SAV traverses section \( l \) (\( \pi_{l} = 0 \), and subsequently \( \pi_{l'r} = 0 \)). But this would be unrealistic if for some competing section \( l' \) of \( l \), \( \eta_{l'} \neq 0 \).

\[ \bar{\eta}_{l} = \sum_{l' \in L \setminus l} \chi_{l'l'} \eta_{l'} \sum_{r \in R} \lambda_{l'r} \lambda_{l'r} \frac{\pi_{l'r} + \varepsilon}{\pi_{l} + \varepsilon} \quad \forall l \in L \quad (3.26) \]

3.3.1.3. In-vehicle time and fare

Besides waiting time, in-vehicle time and fare need to be quantified. Eqs. (3.27)-(3.28) compute the average in-vehicle travel time (\( t_{l} \)) and the average fare (\( f_{l} \)) incurred by SAV travelers of section \( l \). Same as
the treatment in Eq. (3.26), $t_l$ and $f_l$ come from weighted averages over all routes traversing the section, weighted by SAV flows on the routes. The addition of a small number $\varepsilon$ to the denominator and to the numerator is similarly justified as that in Eq. (3.26).

$$t_l = \sum_{r \in R} \frac{\pi_{lr} + \varepsilon}{\pi_l + \varepsilon} \lambda_{lr} t_{lr} \quad \forall l \in L$$ (3.27)

$$f_l = \sum_{r \in R} \frac{\pi_{lr} + \varepsilon}{\pi_l + \varepsilon} \lambda_{lr} f_{lr} \quad \forall l \in L$$ (3.28)

where $t_{lr}$ and $f_{lr}$ represent the in-vehicle travel time and fare on a portion of route $r$ that is section $l$:

$$t_{lr} = \lambda_{lr} \sum_{a \in A} \left( \delta^{D(l)}_{ar} - \delta^{O(l)}_{ar} \right) t_a \quad \forall r \in R, l \in L$$ (3.29)

$$f_{lr} = \lambda_{lr} (f_0 + f_d \sum_{a \in A} \left( \delta^{D(l)}_{ar} - \delta^{O(l)}_{ar} \right) d_a) \quad \forall l \in L, r \in R$$ (3.30)

where $f_0$ and $f_d$ are the constant part ($$/ride$$) and the distance-based part ($$/mile$$) of fare determined by the TNC.

### 3.3.2. Autonomous parking user equilibrium (APUE)

Let us now turn to the user equilibrium for autonomous parking. As mentioned in section 3.2.2, an AV faces a self-parking location problem after dropping off its user at the destination zone. We model this problem by defining the APUE condition, under which no AV can be better off by unilaterally switching to a different parking zone. Constraints (3.31)-(3.32) establish the APUE conditions.

$$0 \leq V_{ij} \perp e_{ij} - E_i \geq 0 \quad \forall i, j \in Z$$ (3.31)

$$\sum_{j \in Z} V_{ij} = \sum_{(j,i) \in \Omega} Q^m_{ji} \quad \forall i \in Z; m = AV$$ (3.32)
where $V_{ij}$ is the flow of vacant self-parking AVs from zone $i$ where they drop off users to parking zone $j$ (including $i = j$). $e_{ij}$ is the associated generalized cost of self-parking. $E_i$ represents the minimum generalized cost of self-parking for an AV whose user destination is zone $i$. The complementarity constraint (3.31) ensures that the chosen parking zone should have the minimum cost. Eq. (3.32) holds the flow conservation, i.e., the total flow of vacant self-parking AVs from $i$ equals the total flow of occupied AVs destined to zone $i$.

The generalized cost of self-parking $e_{ij}$ comprises three components as shown in Eq. (3.33). The first component, $\theta^m(d_{ij} + d_{ji}), m = AV$, is the distance-based AV operating cost for both parking and return bounds of the self-parking trip. $d_{ij}$ is the length of the shortest route from zone $i$ to zone $j$. $\theta^m (m = AV)$ is the unit distance-based AV operating cost. The other two components are AV operating cost during parking search and parking fee in zone $j$ ($PC_j^m, m = AV, and PF_j$).

$$e_{ij} = \theta^m(d_{ij} + d_{ji}) + PC_j^m + PF_j \quad \forall i, j \in Z; \ m = AV$$ (3.33)

The AV operating cost during parking search in a zone $i$ is considered a linear function of the parking search time $PT_i$ and a (constant) vehicle speed $s_i$ during search. For HVs, we consider the parking search cost in a similar fashion. These are shown in Eq. (3.34).

$$PC_i^m = \theta^m s_i PT_i \quad \forall i \in Z; \ m = HV, AV$$ (3.34)

Following Lam et al. (2006) and Jiang et al. (2014), we derive the parking search time in a zone as a BPR-like function of total parking demand and parking infrastructure capacity in the zone.

$$PT_i = PT_{0,i} \left[ 1 + \alpha_i \left( \frac{\sum_{m \in M} PD_i^m}{PK_i} \right)^\beta_i \right] \quad \forall i \in Z$$ (3.35)
where $PT_{0i}$ is the free-flow parking search time, $\alpha_i$ and $\beta_i$ are parameters, $PD_i^m$ is the parking demand of vehicle type $m \in M$, and $PK_i$ is the parking infrastructure capacity, all for zone $i$.

Eqs. (3.36)-(3.38) show how vehicle type-specific parking demand in a zone is computed. For HVs, parking demand in zone $i$ equals the total HV flow destined to $i$. AV parking demand in zone $i$ comes from the AVs which self-park in zone $i$ from any zone $j$.

The parking demand of vacant SAVs in zone $i$ is more complicated as it depends on movement states. For vacant SAVs in zone $i$ waiting to be matched with travelers in any zone $j$, and then moving to zone $j$ in pickup state ($s = II$), demand for parking in zone $i$ occurs only when the matching time is longer than the time to find a parking spot in zone $i$, i.e., $\omega_j^r > PT_i$. This is characterized by $\max\{0, \omega_j^r - PT_i\}$ in Eq. (3.38). For vacant SAVs moving to $i$ from another zone $j$ in relocation state ($s = III$), demand for parking in zone $i$ occurs only when it takes longer for the SAVs to be matched with travelers in zone $i$ than to travel from $j$ to $i$ plus the parking search time in zone $i$, i.e., $\omega_i^r > t_r + PT_i$. This is reflected by $\max\{0, \omega_i^r - t_r - PT_i\}$ in Eq. (3.38). Note further that parking demand is defined on an hourly basis, to be consistent with the hourly parking capacity $PK_i$ in Eq. (3.35). Thus, if an SAV parks for more than an hour, only the first hour contributes to the hourly parking demand. This is realized by using the min functions in Eq. (3.38).

\[
P_D^m = \sum_{(j,i) \in N} Q^m_{ji} \quad \forall i \in Z; \ m = HV \tag{3.36}
\]

\[
P_D^m = \sum_{j \in Z} V_{ji} \quad \forall i \in Z; \ m = AV \tag{3.37}
\]

\[
P_D^m = \sum_{j \in Z} \left( \sum_{r \in R_{ij}} y_{jr}^{II} \min\{1, \max\{0, \omega_j^r - PT_i\}\} + \sum_{r \in R_{ji}} y_{ri}^{III} \min\{1, \max\{0, \omega_i^r - t_r - PT_i\}\} \right) \tag{3.38}
\]

\[\forall i \in Z; \ m = SAV\]
3.3.3. User equilibrium (UE) for occupied HVs/AVs

We model the routing behavior of occupied HVs/AVs based on Wardrop’s first principle (Wardrop, 1952), whereby no one can reduce his/her en-route travel cost at equilibrium by unilaterally changing routes. To prevent route enumeration and keep the problem size tractable, we employ the link-node formulation (Ban et al., 2008).

\[
0 \leq x_{ai}^m \cdot \sum_{a \in A} C_a^m + C_{D(a),i}^m - C_{O(a),i}^m \geq 0 \quad \forall a \in A; \; i \in Z; \; m = HV, AV
\]  

\[
\sum_{a \in A, O(a) = n} x_{ai}^m - \sum_{a \in A, D(a) = n} x_{ai}^m = Q_{ni}^m \quad \forall n \in N; \; i \in Z; \; m = HV, AV
\]

where

\[
Q_{ni}^m = \begin{cases} 
0 & \text{if } n \in N - Z \\
Q_{ni}^m & \text{if } n \in Z \setminus i \\
- \sum_{(j,i) \in \Omega} Q_{ji}^m & \text{if } n = i 
\end{cases} \quad \forall n \in N; \; i \in Z; \; m = HV, AV
\]

\(x_{ai}^m\) is the occupied vehicle flow of type \(m\) on link \(a\) and destined to zone \(i\). \(c_a^m\) is the generalized travel cost using vehicle type \(m\) on link \(a\). \(C_{D(a),i}^m\) is the minimum generalized travel cost using vehicle type \(m\) from the end of link \(a\) to zone \(i\). We conventionally set \(C_{D(a),i}^m = 0\) if \(D(a) = i\). Similarly for \(C_{O(a),i}^m\). \(Q_{ni}^m\) is the travel demand of vehicle type \(m\) from zone \(n\) to zone \(i\).

The complementarity constraint (3.39) establishes the UE conditions for each vehicle type \(m = HV, AV\). In plain words, if link \(a\) carries flow from any zone to zone \(i\) \((x_{ai}^m > 0)\), then travel cost on the link plus the minimum generalized travel cost from the end of the link \(D(a)\) to zone \(i\) should equal the minimum generalized travel cost from the beginning of the link \(O(a)\) to zone \(i\). Flow conservation condition of Eq. (3.40) ensures that, for each vehicle type \(m = HV, AV\), the \(i\)-bound flow leaving node \(n\) minus the \(i\)-bound flow entering node \(n\) should equal the \(i\)-bound demand generated from node \(n\), unless \(n = i\).
Eq. (3.42) computes the generalized travel cost with HV/AV on a link as the sum of distance-based vehicle operating cost (first term) and in-vehicle time cost (second term).

\[
c_a^m = \theta^m d_a + \gamma_a^m t_a \quad \forall a \in A; \ m = HV, AV
\] (3.42)

where \(\gamma_a^m\), \(d_a\), and \(t_a\) denote in-vehicle VOT using vehicle type \(m\), length of link \(a\), and travel time on link \(a\).

Following Noruzoliaee et al. (2018), \(t_a\) is derived using a modified BPR function (3.43) which endogenously models link capacity as a function of vehicle flows of different types (\(x_a^m\)) on the link and vehicle-type specific headways at maximum flow (i.e., capacity) (\(\tau_m\)).

\[
t_a = t_{0,a} \left[ 1 + \alpha_a \left( \frac{\sum_{m \in M} \tau_m x_a^m}{\tau_{HV} K_a^{HV}} \right)^{\beta_a} \right] \quad \forall a \in A
\] (3.43)

where parameters \(t_{0,a}\) and \(K_a^{HV}\) represent free-flow travel time and base capacity with only HVs on link \(a\). \(\alpha_a\) and \(\beta_a\) are the link’s BPR function parameters. \(x_a^m\) for different vehicle types is derived in (3.44)-(3.46).

\[
x_a^m = \sum_{i \in Z} x_{ai}^m \quad \forall a \in A; \ m = HV
\] (3.44)

\[
x_a^m = \sum_{i \in Z} x_{ai}^m + \sum_{r \in R'} \delta_{ar} v_r \quad \forall a \in A; \ m = AV
\] (3.45)

\[
x_a^m = \sum_{s \in S} \sum_{r \in R} \delta_{ar} y_r^s \quad \forall a \in A; \ m = SAV
\] (3.46)

where

\[
v_r = V_{O(r),D(r)} + V_{D(r),O(r)} \quad \forall r \in R'
\] (3.47)

The RHS of Eq. (3.44) aggregates HV flows destined to all zones using link \(a\). The RHS of Eq. (3.45) sums up occupied AV flows destined to any zone \(i\) using link \(a\) (first term) and vacant AV flows using link
\( a \) (second term). In the second term, \( v_r \) is the vacant AV flow on the shortest route of an OD pair, i.e., \( r \in R' \), where \( R' \) denotes the set of shortest (distance) routes for all ODs. Because vacant AVs contribute to network congestion in both parking and return bounds of self-parking trips, \( v_r \) accounts for vacant AV flows in both directions (Eq. (3.47)). Finally, SAV flows using link \( a \) in any movement state \( s \in S \) is captured in RHS of Eq. (3.46).

### 3.3.4. Endogenous HV/AV/SAV market share

The ARUE, APUE, and UE conditions are tied together using a logit-based market share function (3.48), which splits the total demand by vehicle type for each OD.

\[
\frac{q_{ij}^m}{q_{ij}} = \frac{\exp(-\varphi U_{ij}^m)}{\sum_{m' \in M} \exp(-\varphi U_{ij}^{m'})} \quad \forall (i,j) \in \Omega; \ m \in M
\]  

(3.48)

where

\[
U_{ij}^m = C_{ij}^m + \gamma_1 P_{Tj} + P_{Cj}^m + P_{Fj} \quad \forall (i,j) \in \Omega; \ m = HV
\]  

(3.49)

\[
U_{ij}^m = C_{ij}^m + E_j \quad \forall (i,j) \in \Omega; \ m = AV
\]  

(3.50)

\[
U_{ij}^m = H_{ij} \quad \forall (i,j) \in \Omega; \ m = SAV
\]  

(3.51)

In the above equations, \( \varphi > 0 \) is the logit scale factor (Train, 2003). \( U_{ij}^m \) is the perceived generalized travel cost using vehicle type \( m \) for OD \((i,j)\). HV users incur in-vehicle time and vehicle operating cost \((C_{ij}^m)\), parking time cost \((P_{Tj})\), parking search cost \((P_{Cj}^m)\), and parking fee \((P_{Fj})\). We assume that HV users during parking search have the same VOT as travelers when waiting for SAVs. Thus \( \gamma_1 \) is placed before \( P_{Tj} \). AV users incur \( C_{ij}^m \) and AV self-parking cost \((E_j)\). SAV travelers incur ridesharing cost \((H_{ij})\).
3.3.5. Overall formulation of MARPUE and its existence

This section presents the overall formulation of MARPUE and proves its existence. We note that sgn, max, and min functions, which are not continuously differentiable, are used in ARUE and APUE. For computational tractability, we first approximate these functions using Eqs. (3.52)-(3.54), where $\varepsilon$ is a small positive number.

\[\text{sgn}(a) = a/|a| \approx a/\sqrt{a^2 + \varepsilon}\]  \hspace{1cm} (3.52)

\[\max\{a, b\} = 0.5(a + b + |b - a|) \approx 0.5 \left( a + b + \sqrt{(b - a)^2 + \varepsilon} \right)\]  \hspace{1cm} (3.53)

\[\min\{a, b\} = 0.5(a + b - |a - b|) \approx 0.5 \left( a + b - \sqrt{(a - b)^2 + \varepsilon} \right)\]  \hspace{1cm} (3.54)

Lemma 1 below presents the nonlinear complementarity problem (NCP) formulation of MARPUE. Complementarity constraints (3.55)-(3.57) replace the flow conservation constraints of ARUE (Eq. (3.3)), APUE (Eq. (3.32)), and UE (Eq. (3.40)). The equivalence can be verified similar to that for traditional UE (e.g., in Facchinei and Pang (2003)). Complementarity constraint (3.58) satisfies the logit-based market share function (3.48) due to the strict positivity of $Q_{ij} \frac{\exp(-\phi U_{ij}^m)}{\sum_{m' \in M} \exp(-\phi U_{ij}^{m'})}$, which forces $Q_{ij}^m$ to be strictly positive.

**Lemma 1.** NCP (P1) yields MARPUE.

**(P1): NCP-MARPUE**

\[0 \leq H_{ij} \perp \sum_{l \in L_{ij}} q_l - Q_{ij}^m \geq 0 \quad \forall (i, j) \in \Omega; \ m = SAV\]  \hspace{1cm} (3.55)

\[0 \leq E_i \perp \sum_{j \in Z} V_{ij} - \sum_{(j, i) \in D} Q_{ji}^m \geq 0 \quad \forall i \in Z; \ m = AV\]  \hspace{1cm} (3.56)

\[0 \leq C_{ni}^m \perp \sum_{a \in A: O(a) = n} x_{ai}^m - \sum_{a \in A: D(a) = n} x_{ai}^m - Q_{ni}^m \geq 0 \quad \forall n \in N; \ i \in Z; \ m = HV, AV\]  \hspace{1cm} (3.57)

\[0 \leq Q_{ij}^m \perp Q_{ij}^m - Q_{ij} \frac{\exp(-\phi U_{ij}^m)}{\sum_{m' \in M} \exp(-\phi U_{ij}^{m'})} \geq 0 \quad \forall (i, j) \in \Omega; \ m \in M\]  \hspace{1cm} (3.58)
Complementarity constraints (3.2), (3.25), (3.31), and (3.39)
and definitional constraints (3.1), (3.4)-(3.9), (3.13)-(3.19), (3.21)-(3.24), (3.26)-(3.30), (3.33)-(3.38),
(3.41)-(3.47), (3.49)-(3.54)

Proposition 3 establishes an equivalent variational inequality (VI) problem to MARPUE, which is used
later in section 3.5 for solving the TNC’s problem since our numerical experiments find that it is more
efficient to solve the VI than the NCP (P1).

**Proposition 3.** VI (P2) is equivalent to NCP (P1).

(P2): VI-MARPUE

\[
\sum_{l \in L} [h_l(q_l - \tilde{q}_l) + (\kappa \pi_l - q_l - \tilde{q}_l)(\eta_l - \tilde{\eta}_l)] + \sum_{j \in Z} \sum_{l \in Z} e_{ij} (v_{ij} - \tilde{v}_{ij})
\]

\[
+ \sum_{m=HV, AV} \sum_{i \in Z} \sum_{a \in A} c^m_a (x^m_{ai} - \hat{x}^m_{ai})
\]

\[
+ \sum_{(i,j) \in \Omega} \left[ \sum_{m=HV} \left( \gamma_1 PT_j + PC^i_j + PF_j + \frac{1}{\varphi} \ln (\hat{Q}_{ij}^m) \right) (Q_{ij}^m - \hat{Q}_{ij}^m) \right]
\]

\[
+ \frac{1}{\varphi} \sum_{m \in M \setminus HV} \ln (\hat{Q}_{ij}^m) (Q_{ij}^m - \hat{Q}_{ij}^m) \right] \geq 0 \quad \forall (q, \eta, V, x, Q) \in \Phi
\] (3.59)

where \((q, \eta, V, x, Q)\) is the vector form of the corresponding variables. The variables with a hat denote the
solution of the VI. \(\Phi\) is the feasible region of the VI determined by the following constraints.

\[q_l \geq 0 \quad \forall l \in L\] (3.60)

\[\eta_l \geq 0 \quad \forall l \in L\] (3.61)

\[V_{ij} \geq 0 \quad \forall i, j \in Z\] (3.62)

\[x^m_{ai} \geq 0 \quad \forall a \in A; \ i \in Z; \ m = HV, AV\] (3.63)

\[\sum_{m \in M} Q_{ij}^m = Q_{ij} \quad \forall (i, j) \in \Omega\] (3.64)

Flow conservation constraints (3.3), (3.32), (3.40)

and definitional constraints of problem (P1)
Proposition 4 below proves the existence of an equilibrium solution to MARPUE under a specific condition.

**Proposition 4.** Assume that a set of Lagrange multipliers \( \eta \) for the SAV seat capacity constraint (3.20) exists. Then there is at least one solution to NCP-MARPUE (P1) and VI-MARPUE (P2).

**Proof.** All terms used in P1 are continuous given the continuous approximations in Eqs. (3.52)-(3.54). Therefore, there exists a solution to P1, according to Facchinei and Pang (2003). P2 is continuous as well. In addition, \( \Phi \) is convex as it is composed of linear and non-negativity constraints. Under the above assumption about \( \eta \), \( \Phi \) is also compact. Thus, P2 also has at least one solution, according to Facchinei and Pang (2003).

### 3.4. Formulation of the TNC’s decisions

The profit-maximizing TNC centrally decides on the SAV fleet size \( (F) \), fixed and distance-based fare rates \( (f = (f_0, f_d)) \), and vehicle allocation/relocation \( (y = (..., y_r^s, ...), r \in R, s \in S) \). The TNC’s profit is written in Eq. (3.65) as the total after-tax revenue raised (first term) minus parking cost (second term), en-route operating cost (third term), and fleet ownership cost (fourth term).

\[
\text{Profit} = (1 - \text{income tax rate}) \sum_{t \in L} q_t f_t - \sum_{m=\text{SAV}} \sum_{t \in Z} (PC_t^m + PF_t \cdot PD_t^m) - \mu_d \sum_{s \in S} \sum_{r \in R} d_r y_r^s - \mu_F F \quad (3.65)
\]

where \( \mu_d \) is the SAV operating cost per unit distance, \( \mu_F \) is the capital cost of an SAV per unit time, and \( d_r \) is the length of route \( r \in R \).

The parking cost in a zone \( t \) consists of parking search cost \( (PC_t^m) \) and parking fee \( (PF_t \cdot PD_t^m) \). As shown in Eq. (3.66), \( PC_t^m \) depends on the state of vacant SAVs. For vacant SAVs in pickup state waiting...
in zone \( i \), parking search lasts until finding a parking spot or being matched with new traveler(s) from any zone \( j \), whichever occurs first (as captured by \( \min\{PT_i, \omega_j^y\} \)). For vacant SAVs relocating to zone \( i \) from another zone \( j \), parking search starts only when an SAV is not matched with new traveler(s) when arriving in zone \( i \) (as captured by \( \max\{0, \omega_i^y - t_r\}, r \in R_{ji}\)), and lasts until either the parking search ends or the SAV picks up the matched traveler(s) (as captured by \( \min\{PT_i, \max\{0, \omega_i^y - t_r\}\}\)).

\[
P_{CM}^i = \mu_d s_i \sum_{j \in Z} \left( \sum_{r \in R_{ij}} y_r^{II} \min\{PT_i, \omega_j^y\} + \sum_{r \in R_{ji}} y_r^{III} \min\{PT_i, \max\{0, \omega_i^y - t_r\}\} \right) \quad \forall i \in Z; \; m = SAV
\]

Constraints (3.67)-(3.69) ensure the non-negativity of the supply attributes.

\[
f_0 \geq 0 \quad (3.67)
\]

\[
f_d \geq 0 \quad (3.68)
\]

\[
y_r^s \geq 0 \quad \forall s \in S; \; r \in R \quad (3.69)
\]

The TNC’s allocation/relocation decisions should further satisfy SAV flow conservation constraints and service time constraints, as detailed in sections 3.4.1 and 3.4.2.

**3.4.1. SAV flow conservation**

The total in-service SAV flow destined to a zone \( i \) should be equal to the total vacant SAV flow originating from \( i \) (Eq. (3.70)). In addition, the total vacant SAV flow destined to zone \( i \) must be balanced with the total in-service SAV flow originating from \( i \) (Eq. (3.71)). Note that the pickup SAV flow can be both intra- and inter-zone.

\[
\sum_{r \in R: D(r) = i} y_r^I = \sum_{r \in R: O(r) = i} (y_r^{II} + y_r^{III}) \quad \forall i \in Z
\]

\[
\sum_{r \in R: D(r) = i} \left( \sum_{r \in R: O(r) = i} (y_r^{II} + y_r^{III}) \right) \quad \forall i \in Z
\]
\[ \sum_{r \in R: o(r) = i} (y_r^{II} + y_r^{III}) = \sum_{r \in R: d(r) = i} y_r^{I} \quad \forall i \in Z \quad (3.71) \]

### 3.4.2. SAV time conservation

Assuming a one-hour analysis period, constraint (3.72) establishes the relationship between SAV flow and SAV fleet size \( F \). In an hour, \( F \) equals the total en-route SAV hours (first term) plus the total waiting SAV hours (second and third terms). En-route SAV hours include occupied, pickup, and relocation SAV hours. Waiting SAV hours are associated with vacant SAVs. Specifically, for vacant SAVs in zone \( i \) waiting to be matched with travelers from any zone \( j \) and then picking up the matched travelers, the SAV waiting time is \( \omega_j^y \). For vacant SAVs relocating from zone \( j \) to another zone \( i \), the SAV waiting time in zone \( i \) is non-zero only when matching with travelers ends after arriving in \( i \) (as captured by \( \max\{0, \omega_i^y - t_r\}, r \in R_{ji}\)).

\[
F = \sum_{s \in S} \sum_{r \in R} y_r^s t_r + \sum_{j \in Z} \sum_{i \in Z} \left( \sum_{r \in R_{ij}} y_r^{II} \omega_j^y + \sum_{r \in R_{ji}} y_r^{III} \max\{0, \omega_i^y - t_r\} \right) \quad (3.72)
\]

### 3.4.3. Overall formulation of the TNC’s problem

The overall problem encompassing MARPUE (section 3.3) and the TNC’s decisions (section 3.4) is integrated into an MPCC shown in P3. For computational convenience, we transform the logit function (3.48) to an equivalent one-dimensional function using Eqs. (3.64) and (3.73). All sgn, max, and min functions are approximated by continuous functions following Eqs. (3.52)-(3.54).

(P3): MPCC-TNC

\[
\begin{bmatrix}
\max_{f,f',s_y, d, q, \eta V, x, q, H, E, C} & (3.65) \\
\end{bmatrix}
\]

s.t.

TNC’s constraints (3.66)-(3.72)

Complementarity constraints (3.2), (3.25), (3.31), and (3.39)
Flow and demand conservation constraints (3.3), (3.32), (3.40), (3.64)

\[ \ln Q_{ij}^m - \ln Q_{ij}^{m'} = \varphi \left( U_{ij}^{m'} - U_{ij}^m \right) \quad \forall (i,j) \in \Omega; \ m, m' \in M \vert m \neq m' \]  

(3.73)

and definitional constraints of problem (P1)

3.5. Solution method

We use the active-set algorithm (Lawphongpanich and Yin, 2010) to solve P3. The basic idea is to sequentially solve a restricted problem of P3 where its complementarity constraints are replaced by regular (in)equality constraints. For notational simplicity, we hereafter use \( \rho_{q,t}, \rho_{\eta,l}, \rho_{\nu,ij}, \) and \( \rho_{x,aim} \) to represent the terms paired with variables \( q_t, \eta_l, V_{ij}, \) and \( x_{ai}^m \) in the complementarity constraints (3.2), (3.25), (3.31), and (3.39). We start by defining a pair of active sets \( \Gamma_q \) and \( \Gamma_{\rho_q} \) for each group of the above complementarity constraints, as in Eqs. (3.74)-(3.77). Note that the intersection of each active set pair (e.g., \( \Gamma_q \cap \Gamma_{\rho_q} \)) is not necessarily empty since strict complementarity is not required in (3.2), (3.25), (3.31), and (3.39).

\[
\begin{align*}
\Gamma_q &= \{ l \in L : q_l = 0 \} , & \Gamma_{\rho_q} &= \{ l \in L : \rho_{q,l} = 0 \} \\
\Gamma_\eta &= \{ l \in L : \eta_l = 0 \} , & \Gamma_{\rho_\eta} &= \{ l \in L : \rho_{\eta,l} = 0 \} \\
\Gamma_\nu &= \{ (i,j) \in Z \times Z : V_{ij} = 0 \} , & \Gamma_{\rho_\nu} &= \{ (i,j) \in Z \times Z : \rho_{\nu,ij} = 0 \} \\
\Gamma_x &= \{ (m,a,i) \in M \times A \times Z : x_{ai}^m = 0 \} , & \Gamma_{\rho_x} &= \{ (m,a,i) \in M \times A \times Z : \rho_{x,aim} = 0 \}
\end{align*}
\]

(3.74) (3.75) (3.76) (3.77)

Given the active sets, we consider the following nonlinear program P4 termed NLP-TNC. Compared to P3, the complementarity constraints (3.2), (3.25), (3.31), and (3.39) are replaced by regular (in)equality constraints (3.78)-(3.85) using the defined active sets.

\[ \textbf{(P4): NLP-TNC} \]

\[
\begin{align*}
\max_{\substack{\mathbf{f}, \mathbf{g}, \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}, \mathbf{u}, \mathbf{v}, \mathbf{h}, \mathbf{e}, \mathbf{c}}} & \quad (3.65) \\
\text{s.t.} & \quad \text{Constraints (3.78)-(3.85)}
\end{align*}
\]
TNC’s constraints (3.66)-(3.72)

Flow and demand conservation constraints (3.3), (3.32), (3.40), (3.64)

Market share constraint (3.73)

Definitional constraints in P1

\[
\begin{align*}
\{q_l = 0 & \quad \forall l \in I_q^r \\
q_l \geq 0 & \quad \forall l \notin I_q^r \\
\rho_{q,l} = 0 & \quad \forall l \in I_{\rho_q}^r \\
\rho_{q,l} \geq 0 & \quad \forall l \notin I_{\rho_q}^r \\
\eta_l = 0 & \quad \forall l \in I_\eta^r \\
\eta_l \geq 0 & \quad \forall l \notin I_\eta^r \\
\rho_{\eta,l} = 0 & \quad \forall l \in I_{\rho_\eta}^r \\
\rho_{\eta,l} \geq 0 & \quad \forall l \notin I_{\rho_\eta}^r \\
V_{ij} = 0 & \quad \forall (i,j) \in I_V^n \\
V_{ij} \geq 0 & \quad \forall (i,j) \notin I_V^n \\
\rho_{v,ij} = 0 & \quad \forall (i,j) \in I_{\rho_v}^n \\
\rho_{v,ij} \geq 0 & \quad \forall (i,j) \notin I_{\rho_v}^n \\
x_{ai}^m = 0 & \quad \forall (m,a,i) \in I_X \\
x_{ai}^m \geq 0 & \quad \forall (m,a,i) \notin I_X \\
\rho_{x,aim} = 0 & \quad \forall (m,a,i) \in I_{\rho_x}^n \\
\rho_{x,aim} \geq 0 & \quad \forall (m,a,i) \notin I_{\rho_x}^n
\end{align*}
\]

The active-set algorithm proceeds as follows.

**Step 1: Initialization.** Find a feasible solution \((F, f, y, q, \eta, V, x, Q, H, E, C)\) to the original P3. Set iteration counter \(k = 1\) and construct the active sets \(I_{\rho_x}^k\) and \(I_{\rho_\eta}^k\) as defined in Eqs. (3.74)-(3.77) based on the feasible solution.

**Step 2: Solving P4.** Set \(I_{\rho_x}^k = I_{\rho_{x,aim}}^k\) and \(I_{\rho_\eta}^k = I_{\rho_{\eta,i}}^k\) in Eqs. (3.74)-(3.77). Solve P4 to obtain solution \((F^k, f^k, y^k, q^k, \eta^k, V^k, x^k, Q^k, H^k, E^k, C^k)\). Because the active sets are given which satisfy the complementarity constraints (3.2), (3.25), (3.31), and (3.39), P4 is a restricted problem of the original problem P3.
Step 3: Convergence check. Denote the Lagrange multipliers associated with the equality constraints in (3.78)-(3.85) by $\psi_{q,l}, \psi_{\rho q,l}, \psi_{\eta,l}, \psi_{\rho \eta,l}, \psi_{\nu,l}, \psi_{\rho \nu,l}, \psi_{x,aim}, \psi_{p,aim}$. Construct new pairs of sets $\psi_{(\cdot)}$ and $\psi_{\rho(\cdot)}$ as follows.

$$
\psi^k_q = \left\{ l \in (\Gamma^k_q \cap \Gamma^k_{\rho q}): \psi_{q,l} > 0 \right\}, \quad \psi^k_{\rho q} = \left\{ l \in (\Gamma^k_q \cap \Gamma^k_{\rho q}): \psi_{\rho q,l} > 0 \right\};
$$

$$
\psi^k_{\eta} = \left\{ l \in (\Gamma^k_{\eta} \cap \Gamma^k_{\rho \eta}): \psi_{\eta,l} > 0 \right\}, \quad \psi^k_{\rho \eta} = \left\{ l \in (\Gamma^k_{\eta} \cap \Gamma^k_{\rho \eta}): \psi_{\rho \eta,l} > 0 \right\};
$$

$$
\psi^k_{\nu} = \left\{ (i, j) \in (\Gamma^k_{\nu} \cap \Gamma^k_{\rho \nu}): \psi_{\nu,i,j} > 0 \right\}, \quad \psi^k_{\rho \nu} = \left\{ (i, j) \in (\Gamma^k_{\nu} \cap \Gamma^k_{\rho \nu}): \psi_{\rho \nu,i,j} > 0 \right\};
$$

$$
\psi^k_x = \left\{ (m, a, i) \in (\Gamma^k_x \cap \Gamma^k_{\rho x}): \psi_{x,aim,i} > 0 \right\}, \quad \psi^k_{\rho x} = \left\{ (m, a, i) \in (\Gamma^k_x \cap \Gamma^k_{\rho x}): \psi_{p,aim} > 0 \right\}.
$$

If $\psi^k_q = \psi^k_{\rho q} = \psi^k_{\eta} = \psi^k_{\rho \eta} = \psi^k_{\nu} = \psi^k_{\rho \nu} = \psi^k_x = \psi^k_{\rho x} = \emptyset$, stop and $(F^k, f^k, y^k, q^k, \eta^k, \nu^k, x^k, Q^k, H^k, E^k, C^k)$ is the strongly stationary solution of P3. Otherwise, go to Step 4.

Step 4: Updating the active sets. For each pair of active sets $\Gamma_{(\cdot)}$ and $\Gamma_{\rho(\cdot)}$ in Eqs. (3.74)-(3.77), do the following:

If $\psi^k_{(\cdot)} \cap \psi^k_{\rho(\cdot)} = \emptyset$

Set $\Gamma^{k+1}_{(\cdot)} = \Gamma^k_{(\cdot)} - \psi^k_{(\cdot)}$ and $\Gamma^{k+1}_{\rho(\cdot)} = \Gamma^k_{\rho(\cdot)} - \psi^k_{\rho(\cdot)}$

End If

If $\psi^k_{(\cdot)} \cap \psi^k_{\rho(\cdot)} \neq \emptyset$

For each $l \in \psi^k_{(\cdot)} \cap \psi^k_{\rho(\cdot)}$

If $\psi_{(\cdot),l} > \psi_{\rho(\cdot),l}$

$$
\psi^k_{\rho(\cdot)} = \psi^k_{\rho(\cdot)} - l
$$

Else

$$
\psi^k_{(\cdot)} = \psi^k_{(\cdot)} - l
$$

End If

85
End For

\[ \psi_{(\cdot)}^{k+1} = \psi_{(\cdot)}^{k} \]  and  \[ \psi_{\rho(\cdot)}^{k+1} = \psi_{\rho(\cdot)}^{k} \]

End If

Set \( k = k + 1 \) and go back to Step 2.

In the algorithm, step 1 requires an initial feasible solution to P3. The initial solution is found by solving an equivalent NLP of P2 using a primal-dual gap function (see Appendix 3.3), provided a feasible decision (including \( f_0, f_d, \) and \( y_r^s, r \in R, s \in S \)) for the TNC. For the feasible decision for the TNC, arbitrary values are given to \( f_0 \) and \( f_d \). For \( y_r^s \), we fix the SAV relocation flows to zero, and equally assign all in-service and pickup SAVs to the shortest route for each OD, so that a given SAV traveler demand (as an arbitrary portion of total demand) for each OD is satisfied assuming that each SAV takes one traveler.

### 3.6. Numerical analysis

We implement the proposed model in the Sioux Falls network (Figure 2.16) to investigate the system performance with coexistence of HV/AV/SAV. The road network has 76 links and 24 nodes, of which 12 (i.e., nodes 1, 2, 4, 5, 10, 11, 13, 14, 15, 19, 20, and 21 in Figure 2.16) represent zones (in total 132 OD pairs). Link characteristics such as free-flow travel time \( (t_{0,a}) \) and base capacity \( (K_{a}^{HV}) \) are borrowed from Transportation Test Networks (2017). Link length \( (d_{a}) \) is assumed as \( t_{0,a} \) times a constant link speed of 40 mi/hr. The demand matrix \( (Q_{ij}) \) in Transportation Test Networks (2017) is scaled up such that the total trip generation from each of the 12 zones becomes similar to that of the original demand matrix. We create 1,992 routes (including 12 virtual intra-zone routes with zero costs) and 634 sections using Yen’s algorithm (Yen, 1971).

Several interesting results are presented including the baseline (section 3.6.1) and scenario analysis of the effects of ridesharing, AV self-parking, SAV size, and (S)AV headway (section 3.6.2). Sensitivity
analysis with respect to key model parameters such as matching elasticity, (S)AV price, and in-vehicle VOT saving is further conducted, with results shown in Appendix 3.4. In total, 27 scenarios are computed.

Before reporting the results, we provide below some summary statistics of the problem size and the solution algorithm performance. The MPCC (P3) has 25,337 variables and 25,120 constraints, of which 3,236 are complementarity constraints. The entire problem is coded in GAMS and solved using the generalized reduced gradient algorithm of CONOPT on a personal computer with Intel(R) Core(TM) i7 CPU @ 3.40 GHz with 12 GB RAM. The average (respectively, standard deviation) of the number of iterations and the solution time (in minutes) of the active-set algorithm over the 27 scenarios are respectively 8.4 (2.9) and 17.4 (3.4), of which 4.4 (1.7) minutes are spent in finding initial feasible solutions (step 1 of the algorithm).

To further understand the system impact of SAVs, we compute an additional scenario with only HVs, which represents the present-day condition. Under this scenario, $Q_{ij} = Q_{ij}^{HV}, \forall (i,j) \in \Omega$. HVs are assigned to the network using UE conditions (i.e., Eqs. (3.39) and (3.57) in P1). We solve an equivalent NLP to P1 (the NLP will be formed in a similar way as P5 in Appendix 3.3) using CONOPT. The solution time is only 4.1 seconds.

3.6.1. Baseline

The baseline scenario is set up with 4-seat sedan HV/AV/SAV with main modeling parameters listed in Table 3.2. We consider the prices of an HV and an (S)AV to be $35k and $42k, the latter 20% higher to reflect the self-driving technology cost (Bösch et al., 2018). The cost estimates in Bösch et al. (2018) are in terms of Swiss Franc (CHF), which is almost equivalent to the US$ used in this study. The $35k estimate by the authors is close to the average market price of $33.5 for an HV in the US (USA Today, 2015). For HVs and AVs, the operating costs ($\theta^m$) are respectively $0.48 and $0.56 per mile. The numbers are calculated as the sum of the distance-based fixed cost (due to depreciation of the vehicle price) and the distance-based variable costs (due to insurance, maintenance, and fuel). The former is derived by assuming
the average annual VMT at 15000 (AAA 2015) and a 10% annual depreciation rate (Bösch et al. 2018) over a 5-year vehicle lifetime (AAA 2015). The latter is computed based on an annual variable cost of $3750 for HVs (AAA 2015). For (S)AVs, the variable cost is modified assuming 30% insurance discount, 10% fuel efficiency, and an additional $1000/year in maintenance cost due to the automated driving technology. The calculated distance-based variable cost for each vehicle type is listed in Table 3.2. For SAVs, the time-based fixed ownership cost ($\mu_F$) equals $4.6 per hour assuming a 40% annual depreciation rate (Bizfluent, 2017) and a 10-hour effective daily service time. We follow Nowakowski et al. (2010) and assume vehicle headways ($\tau_m$) of 1.6, 1.1, and 0.6 seconds for HVs, AVs, and SAVs. The baseline assumes 20% saving in in-vehicle VOT ($\gamma_{1m}^F$) for (S)AV users compared to that for HV users, which is $14.2/hour (US DOT, 2016). $\gamma_1$ is also set at $14.2/hour. The logit scale factor ($\varphi$) is set to 0.5. Following Yang and Yang (2011), the matching function parameters in Eq. (3.12) are $B_i = 0.2$ and $b_1 = b_2 = 0.75 \cdot (b_1 + b_2 > 1$ suggests increasing returns-to-scale of matching. The free-flow parking search time in each zone ($PT_{0,i}$) and parking search speed ($s_i$) are set to be 0.1 hr (Lam et al., 2006) and 7 mi/hr (Belloche, 2015). To prevent illegal parking, the zonal parking capacity ($PK_i$) equates the hourly demand destined to a zone. The parking fee ($PF_i$) is assumed $4 for the most demand-attractive zones (i.e., 10, 11, 15, and 20), and $0 for other zones. We consider the parameters of the BPR functions in Eqs. (3.35) and (3.43) to be $\alpha_i = 0.3, \beta_i = 4$ (Lam et al., 2006) and $\alpha_a = 0.15, \beta_a = 4$. Finally, the income tax rate in the TNC’s profit function (3.65) is set to 20%.

### Table 3.2. Main modeling parameters in the baseline scenario

<table>
<thead>
<tr>
<th>Vehicle type</th>
<th>Vehicle attributes</th>
<th>Traveler attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price ($k)</td>
<td>Variable cost ($/mi)</td>
</tr>
<tr>
<td>HV</td>
<td>35</td>
<td>0.25</td>
</tr>
<tr>
<td>AV</td>
<td>42 (20% ↑)</td>
<td>0.28</td>
</tr>
<tr>
<td>SAV</td>
<td>42 (20% ↑)</td>
<td>0.28</td>
</tr>
</tbody>
</table>

* Values are assumed following Nowakowski et al. (2010). Shorter headway of AVs may be analogous to the adaptive cruise control (ACC) system. Shorter headways and connectedness of centrally operated SAVs may be analogous to the cooperative ACC (CACC) system.
Table 3.3 presents the baseline results. The TNC charges each SAV traveler a fixed amount of $3.45 plus $0.31/mile and earns about $130,500 in profit in a one-hour period. The optimal SAV fleet size, which is 9,413, equals 4.8% of the total hourly travel demand (i.e., $\sum_{(i,j) \in \Omega} Q_{ij}$). The fleet utilization — which measures the percent time an SAV is in in-service state ($s = I$), calculated as $\sum_{r \in R} y_r^I t_r / F$ based on Eq. (3.72) — is 45%. We find that the SAV hours spent in pickup state ($s = II$) significantly outnumbers that of relocation state ($s = III$). The systemwide market shares of HV, AV, and SAV, calculated as $\sum_{(i,j) \in \Omega} Q_{ij}^{m} / \sum_{(i,j) \in \Omega} Q_{ij}$, $\forall m \in M$, are 17%, 53%, and 30%. The average occupancy of an SAV — which is a weighted average of $(q_l + \tilde{q}_l)/\pi_l$ over all sections $l$ with $\pi_l \neq 0$, weighted by SAV traveler flow $(q_l + \tilde{q}_l)$ — is 3.82 travelers. The average traveler group size — which is a weighted average of $\sigma^q_l$ over all zones, weighted by the SAV traveler flow originating from each zone $n^q_i$ — is 2.86. The average matching times of an SAV and a traveler are 13.7 and 3.2 minutes respectively.

<table>
<thead>
<tr>
<th>TNC</th>
<th>Travelers</th>
<th>Other system performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant fare ($/ride)</td>
<td>3.45</td>
<td>Market share (%)</td>
</tr>
<tr>
<td>Distance-based fare ($/mi)</td>
<td>0.31</td>
<td>HV</td>
</tr>
<tr>
<td>Fleet size (% of total demand)</td>
<td>4.8</td>
<td>AV</td>
</tr>
<tr>
<td>SAV hours in each state (%)</td>
<td></td>
<td>SAV</td>
</tr>
<tr>
<td>In-service</td>
<td>45.0</td>
<td>Avg. occupancy of an SAV</td>
</tr>
<tr>
<td>Pickup</td>
<td>46.0</td>
<td>Avg. traveler group size</td>
</tr>
<tr>
<td>Relocation</td>
<td>9.0</td>
<td>Avg. matching time (min)</td>
</tr>
<tr>
<td>Profit ($k)</td>
<td>130.5</td>
<td></td>
</tr>
<tr>
<td>Fleet ownership cost ($k)</td>
<td>43.3</td>
<td></td>
</tr>
<tr>
<td>Operating cost ($k)</td>
<td>31.5</td>
<td></td>
</tr>
<tr>
<td>Parking cost ($k)</td>
<td>9.0</td>
<td></td>
</tr>
<tr>
<td>Avg. matching time (min)</td>
<td>13.7</td>
<td></td>
</tr>
</tbody>
</table>

* Compared to the scenario with only HVs.
In the last column of Table 3.3, we also compare the baseline to the scenario with only HVs. We find that each SAV replaces 6.2 HVs, which is the ratio of SAV traveler demand to the SAV fleet size, assuming that each unit of travel demand would require one HV if SAVs are absent. As a result of the replacement, parking demand and occupied VMT are significantly reduced (by 28.7% and 14.6%). However, total VMT is increased by 4.5% due to empty (S)AV trips. In particular, AV self-parking trips account for 90.8% of the vacant VMT. The increase in total VMT aggravates congestion, which leads to 33.3% increase in total traveler en-route time despite 39% increase in network capacity. Due to in-vehicle VOT saving of (S)AV users, AV self-parking, and ridesharing, (S)AVs reduce the total generalized travel cost by 9.6%.

3.6.2. Scenario analysis

We further investigate how system performance will change if we do not allow ridesharing (section 3.6.2.1) or AV self-parking (section 3.6.2.2), consider alternative SAV sizes (section 3.6.2.3), and consider alternative vehicle headways (section 3.6.2.4). The results are compared mostly with the baseline scenario. For some aspects of the system performance, namely travel time and cost, VMT, and (S)AV parking demand, it is more sensible to make comparison with the present-day scenario with only HVs. Thus the scenario with only HVs is used for benchmarking these aspects of system performance.

3.6.2.1. No ridesharing

When ridesharing is not allowed, an SAV has at most one traveler at any time point. Figure 3.4 and Figure 3.5 present the results. The system performance is worse off than the baseline in terms of TNC profit, traveler en-route time, generalized travel cost, total and occupied VMT, and parking demand. Particularly, the TNC will suffer a deficit of $10K. The TNC would not be able to sustain without ridesharing unless, for example, tax exemptions or subsidies are offered.

Because SAV seats are greatly underutilized without ridesharing, operating SAVs becomes less desirable. As a result, the SAV fleet size is almost halved. Each SAV replaces only 2.4 HVs compared with
6.2 under the baseline scenario. For SAV fare, while the constant fare per ride reduces by 50%, the distance-based fare increases by 50%. The matching time of SAV travelers almost quintuples. Partly because of this, the SAV market share drops from 30% to 6%. Most of the SAV travelers will shift to using AVs, leading to significant increase in parking demand (from 29% to only 5% reduction compared to the scenario with only HVs). By contrast, the change in SAV matching time is much smaller, with a reduction by about 10%.

3.6.2.2. No AV self-parking

In the baseline, an AV can self-park after its user exits at the destination zone. Self-parking may be in a different zone, incurring additional VMT and congestion. To understand the impact of self-parking, we examine a scenario in which AV self-parking is not allowed. Under the scenario, it is assumed that an AV must park in its user’s destination zone. In addition, the AV user exits after the AV is parked. The results of the scenario are illustrated also in Figure 3.4 and Figure 3.5.

We find that TNC profit, traveler en-route time, generalized travel cost, total and occupied VMT, and parking demand are all better than the baseline. This is not surprising as disallowing self-parking prevents selfish AV users from reducing their own parking and travel costs while imposing externalities to the rest of the system. Because AV users would need to spend more time for parking in the vehicle, the AV market share is dropped by 22%. The loss of AV market share is almost equally absorbed by HVs and SAVs. Vacant VMT from AV self-parking is eliminated, leading to reduction in total VMT. Note that there is still empty VMT remaining which comes all from SAVs. Occupied VMT decreases even further due to mode shift to SAV which makes more travelers share rides. By rising the constant fare by 20% and halving the distance-based fare, the TNC maintains a similar SAV fleet size and comparable traveler matching time to the baseline. Because the SAV fleet size does not change much while the market share of SAV significantly increases, a larger SAV-HV replacement ratio is expected (8.4 vs. 6.2 as in the baseline).
3.6.2.3. **Alternative SAV sizes**

The baseline considers 4-seat sedan SAVs at the price of $42k. To understand the effect of SAV size, two alternative scenarios with different SAV sizes are investigated: 1-seat solo SAVs and 8-seat van SAVs. Their prices are assumed $13k and $66k respectively, following the calculation procedure in Bösch et al.
(2018) plus a 20% markup to reflect the self-driving technology cost. Figure 3.6 and Figure 3.7 present the results.

**Figure 3.6.** Changes in system performance compared to the baseline scenario, with different SAV sizes

**Figure 3.7.** Changes in system performance compared to the scenario with only HVs, with different SAV sizes

The results show clear economies of SAV size. More specifically, using larger SAVs will increase TNC profit and reduce traveler en-route time, generalized travel cost, and total and occupied VMT. This is because more travelers can be accommodated in larger SAVs, which contributes to reducing the distance-based fare, the number of required SAVs to meet demand, and the occupied and total VMT. However, the

---

1 The time-based fixed ownership cost ($\mu_F$) of solo and van SAVs are calculated in a similar way to that of sedan SAV in the baseline. The variable costs are calculated by scaling that of sedan SAV, according to Table B.1 in Bösch et al. (2018).
economies of SAV size diminishes as we increase SAV size from solo to sedan to van, as reflected in the extent of change in TNC profit and SAV/traveler matching times.

3.6.2.4. Alternative vehicle headways

The baseline assumes vehicle headways of 1.6, 1.1, and 0.6 sec for HVs, AVs, and SAVs. In this section, two alternative headway scenarios are considered. The first scenario considers that SAVs have the same headway as AVs (1.1 sec). The second scenario assumes that all vehicles have the same headway as HVs (1.6 sec). Figure 3.8 and Figure 3.9 present the results (the two scenarios are termed “New headway 1” and “New headway 2”).

We find that larger headways, which is equivalent to degraded road capacity, worsen system performance. TNC profit is reduced. Traveler en-route time, generalized travel cost, total and occupied VMT, and parking demand all increase. As the headway benefit diminishes, SAVs capture a smaller market share (from 30% to 25% and 13%) and operate with a larger fleet size (from 4.8% to 6.4% and 7.5% of total hourly travel demand) while maintaining a comparable traveler matching time.

Figure 3.8. Changes in system performance compared to the baseline scenario, with different vehicle headways
Figure 3.9. Changes in system performance compared to the scenario with only HVs, with different vehicle headways

3.7. Conclusions

Motivated by understanding the impacts of shared and privately owned autonomous vehicles (SAVs and AVs) on urban mobility, this chapter develops an integrated model encompassing SAV ridesharing and AV self-parking to determine the market share of (S)AVs in mixed traffic with human-driven vehicles (HVs). A novel multimodal autonomous ridesharing and parking user equilibrium (MARPUE) model is developed to characterize the traveler adoption of SAV/AV/HV with two major contributions. First, we propose a novel one (SAV)-to-many (travelers) matching to characterize the waiting times of an SAV and multiple travelers who share rides in the SAV while being matched. Theoretical insights are derived by comparing the waiting times of an SAV/traveler due to matching with and without ridesharing. Second, we introduce a section-based formulation for SAV ridesharing equilibrium. The notion of sections is introduced to represent traveler flows, which prevents undesired traveler en-route transfer and frees up the seat after a traveler leaves an SAV, while respecting the SAV seat capacity constraint. The MARPUE model also incorporates inter- and intra-zone self-parking of vacant AVs after dropping off users. MARPUE further considers (S)AV traveler in-vehicle VOT saving and road capacity increase due to reduced headway of (S)AVs. We formulate MARPUE as nonlinear complementarity and variational inequality problems with a proof of its existence, and develop its equivalent nonlinear program using the primal-dual gap function. Anticipating traveler reactions as characterized by MARPUE, the supply decisions of SAVs, assumed to be made by a centralized transportation network company (TNC) and including SAV fleet size, fare rates, and vehicle allocation/relocation, are optimally determined and thoroughly investigated. The overall
problem is formulated as a mathematical program with complementarity constraints (MPCC) and solved using the active-set algorithm.

Several original insights are obtained from numerical implementation of the model. First, we find that SAV ridesharing brings much reduced traveler waiting time while being matched, increased TNC profit, and larger SAV market share compared to a system in which ridesharing is not allowed in using SAVs. Second, the overall transportation system will benefit from increasing SAV size, but the marginal effect is diminishing. Third, vacant SAV miles will incur, in which miles due to picking up travelers substantially outnumber miles due to SAV relocation. Fourth, AV self-parking significantly degrades the overall system performance by increasing traveler en-route time and generalized cost, total and occupied VMT, parking demand, and reduced TNC profit. In fact, the vacant VMT from AV self-parking turns out to be much larger than the vacant VMT from SAVs. This suggests the negative externalities imposed by selfish AVs to reduce their own parking cost at the expense of greater system cost. These insights, which are first-of-its-kind in a mixed environment with SAV/AV/HV, advance the understanding of how transportation system performance can be reshaped by SAV ridesharing.

As shared autonomous mobility is expected to play a major role in future urban travel, this chapter makes one of the first attempts to develop equilibrium models toward understanding the systemwide impacts of SAV use. Further research can be extended in many directions. Here we list a few. First, competition among multiple TNCs and competition/alliance of TNCs and car manufacturers could be considered, which would lead to lower fare and greater service quality of SAVs thus encouraging more travelers to embrace this new service. Second, in the current model AV self-parking trips are assumed to take the shortest route between zones, which could be relaxed. We may also consider congestion pricing for AVs self-parking to reduce its negative effects. Third, ridesharing inconvenience cost could be explicitly incorporated into the model. Finally, the assumption of fixed total demand can be relaxed to capture potentially induced demand after (S)AVs are introduced, as may occur to travel-restrictive population with HVs (seniors, youth, etc.).
Appendix 3.1. Notations

Sets

\(M\) vehicle types \(m \in M = \{HV, AV, SAV\}\)

\(S\) SAV movement states \(s \in S = \{I, II, III\}\), \(I\): occupied, \(II\): pickup, and \(III\): relocation

\(N\) nodes of road network \(n \in N\)

\(A\) links of road network \(a \in A\)

\(Z\) zones \(i \in Z \subseteq N\)

\(\Omega\) OD pairs \((i, j) \in \Omega\)

\(R\) routes (including one intra-zone route for each zone) \(r \in R\)

\(L\) sections \(l \in L\)

Decision variables of TNC

\(F\) SAV fleet size

\(f_d\) distanced-based part of SAV fare ($/mi)

\(f_0\) constant part of SAV fare ($/ride)

\(y_r^s\) SAV flow on route \(r\) in state \(s\) (veh/hr)

Decision variables of travelers

\(Q_{ij}^m\) travel demand of vehicle type \(m\) for OD pair \((i, j)\) (travelers/hr)

\(q_l\) SAV traveler flow of section \(l\) (travelers/hr)

\(\bar{q}_l\) a portion of SAV traveler flow of competing sections of section \(l\) and using routes that traverse \(l\) (passenger/hr)

\(q_{lr}\) SAV traveler flow of section \(l\) and using SAV route \(r\) (travelers/hr)

\(V_{ij}\) flow of vacant self-parking AVs from zone \(i\) where they drop off users to parking zone \(j\) (veh/hr)

\(x_{ai}^m\) occupied vehicle flow of type \(m\) on link \(a\) and destined to zone \(i\) (veh/hr)

\(x_{a}^m\) occupied vehicle flow of type \(m\) on link \(a\) (veh/hr)

Cost variables

\(U_{ij}^m\) perceived generalized travel cost using vehicle type \(m\) for OD \((i, j)\)
\( h_l \) generalized cost of an SAV traveler on section \( l \) ($)

\( H_{ij} \) minimum generalized travel cost of an SAV traveler for OD \((i,j)\) ($)

\( e_{ij} \) generalized cost of self-parking of an AV from zone \( i \) where it drops off its user to parking zone \( j \) ($)

\( E_i \) minimum generalized cost of self-parking of an AV whose user destination is zone \( i \) ($)

\( c_{am}^m \) generalized travel cost using vehicle type \( m \) on link \( a \)

\( c_{ni}^m \) minimum generalized travel cost using vehicle type \( m \) from node \( n \) to zone \( i \)

\( \omega_{iy}^i \) waiting time of an SAV in zone \( i \) while being matched (hr)

\( \omega_i^q \) waiting time of an SAV traveler group in zone \( i \) while being matched (hr)

\( \bar{\omega}_i \) average waiting time of an SAV traveler originating from zone \( i \) due to matching and meeting (hr)

\( g_{ri} \) variable indicating the possibility that travelers waiting in zone \( i \) can be picked up by in-service SAVs on route \( r \) which goes through \( i \)

\( \eta_l \) extra waiting time for SAV travelers of section \( l \) due to limited SAV seat capacity on the section (hr)

\( \bar{\eta}_l \) extra waiting time for SAV travelers of section \( l \) due to limited SAV seat capacity on the competing sections (hr)

\( f_{lr} \) SAV fare on a portion of route \( r \) that is section \( l \) ($)

\( f_i \) average SAV fare on section \( l \) ($)

\( t_{lr} \) in-vehicle travel time on a portion of route \( r \in R \) that is section \( l \) (hr)

\( t_l \) average in-vehicle travel time on section \( l \) (hr)

\( t_{ri} \) SAV travel time from the origin of route \( r \) to zone \( i \) on route \( r \) (hr)

\( t_r \) SAV travel time on route \( r \) (hr)

\( t_a \) vehicle travel time on link \( a \) (hr)

\( PT_i \) parking search time in zone \( i \) (hr)

\( PD_{im} \) parking demand of vehicle type \( m \) in zone \( i \) (veh)

\( PC_{im} \) operating cost of vehicle type \( m \) during parking search in zone \( i \) ($)

Network correspondence indicators

\( \lambda_{lr} \) section-route correspondence indicator (=1 if route \( r \) traverses section \( l \); =0 otherwise)

\( \chi_{li} \) competing section indicator (=1 if section \( l' \) is a competing section of section \( l \setminus l' \); =0 otherwise)

\( \epsilon_{ii} \) section-zone correspondence indicator (=1 if zone \( i \) is an intermediate zone on section \( l \); =0 otherwise)

\( \zeta_{ri} \) route-zone correspondence indicator (=1 if zone \( i \) is an intermediate zone on route \( r \); =0 otherwise)

\( \delta_{ar} \) link-route correspondence indicator (=1 if link \( a \) lies on route \( r \); =0 otherwise)
\( \delta_{ar} \) link-route-zone correspondence indicator (=1 if: 1) link \( a \) lies on route \( r \); 2) zone \( i \) lies on route \( r \); 3) the end of link \( a \) is no later than \( i \) on route \( r \); =0 otherwise)

**Parameters**

- \( Q_{ij} \) total travel demand between OD pair \((i, j)\) (travelers/hour)
- \( d_{ij} \) shortest distance between OD pair \((i, j)\) (miles)
- \( d_a \) length of link \( a \) (miles)
- \( \gamma_1 \) SAV traveler value of time (VOT) for waiting ($/hour)
- \( \gamma_2^m \) traveler in-vehicle VOT using vehicle type \( m \) ($/hour)
- \( \theta^m \) unit distance-based operating cost of vehicle type \( m \) ($/mile)
- \( \mu_d \) SAV operating cost per unit distance ($/mile)
- \( \mu_F \) capital cost of an SAV per unit time ($/hour)
- \( \kappa \) number of seats in an SAV
- \( PF_i \) parking fee in zone \( i \) ($)
- \( PK_i \) parking capacity in zone \( i \) (veh/hour)
- \( PT_{0,i} \) free-flow parking search time in zone \( i \) (hours)
- \( t_{0,a} \) free-flow travel time on link \( a \) (hours)
- \( \tau_m \) average headway of vehicle type \( m \) (hours)
- \( K_a^{HV} \) base capacity with only HVs on link \( a \) (veh/hour)
- \( \phi \) logit scale factor

**Appendix 3.2. Proof of Proposition 3**

Since the feasible region \( \Phi \) of the VI problem (P2) as defined in Eqs. (3.3), (3.32), (3.40), and (3.60)-(3.64) is polyhedron, the VI has an equivalent KKT system, written as follows (see proposition 1.2.1 in Facchinei and Pang (2003)).

\[
\begin{align*}
    h_l - H_{ij} - \xi_{q,l} &= 0 \quad \forall l \in L_{ij}; (i, j) \in \Omega \\
    (\kappa \pi_l - q_l - \bar{q}_l) - \xi_{\pi,l} &= 0 \quad \forall l \in L \quad (A.1) \\
    e_{ij} - E_l - \xi_{V,ij} &= 0 \quad \forall i, j \in Z \quad (A.3)
\end{align*}
\]
\[ c^m_a + C^m_{D(a),i} - C^m_{0(a),i} - \xi_{x,aim} = 0 \quad \forall a \in A; \, i \in Z; \, m = HV,AV \quad (A.4) \]

\[ (\gamma_1 PT_j + PC_j^m + PF_j + \frac{1}{q} \ln Q^m_{ij}) + C^m_{ij} + \theta_{ij} = 0 \quad \forall (i,j) \in \Omega; \, m = HV \quad (A.5) \]

\[ \frac{1}{q} \ln Q^m_{ij} + E_j + C^m_{ij} + \theta_{ij} = 0 \quad \forall (i,j) \in \Omega; \, m = AV \quad (A.6) \]

\[ \frac{1}{q} \ln Q^m_{ij} + H_{ij} + \theta_{ij} = 0 \quad \forall (i,j) \in \Omega; \, m = SAV \quad (A.7) \]

\[ 0 \leq \xi_q,i \perp q_i \geq 0 \quad \forall l \in L \quad (A.8) \]

\[ 0 \leq \xi_{\eta,l} \perp \eta_l \geq 0 \quad \forall l \in L \quad (A.9) \]

\[ 0 \leq \xi_{V,ij} \perp V_{ij} \geq 0 \quad \forall i,j \in Z \quad (A.10) \]

\[ 0 \leq \xi_{x,aim} \perp x^m_{ai} \geq 0 \quad \forall a \in A; \, i \in Z; \, m = HV,AV \quad (A.11) \]

flow and demand conservation constraints (3.3), (3.32), (3.40), (3.64)

where \( H_{ij}, E_j, C^m_{ij} \), and \( \theta_{ij} \) are the Lagrangian multipliers associated with constraints Eqs. (3.3), (3.32), (3.40), and (3.64). \( \xi_{q,i}, \xi_{\eta,l}, \xi_{V,ij}, \) and \( \xi_{x,aim} \) are the Lagrangian multipliers associated with non-negativity constraints (3.60)-(3.63).

If a section \( l \) is used by SAV travelers (i.e., \( q_l > 0 \)), then \( \xi_q,i = 0 \) by complementarity constraint (A.8), and consequently \( h_l - H_{ij} = 0 \) according to Eq. (A.1). If \( q_l = 0 \), (A.8) enforces \( \xi_q,i \geq 0 \) and (A.1) implies \( h_l - H_{ij} \geq 0 \).

If extra waiting time is incurred by SAV travelers on a section \( l \) (i.e., \( \eta_l > 0 \)), then \( \xi_{\eta,l} = 0 \) by complementarity constraint (A.9), and consequently \( \kappa \pi_l - q_l - \bar{q}_l = 0 \) according to Eq. (A.2). If \( \eta_l = 0 \), (A.9) enforces \( \xi_{\eta,l} \geq 0 \) and (A.2) implies \( \kappa \pi_l - q_l - \bar{q}_l \geq 0 \Rightarrow q_l + \bar{q}_l \leq \kappa \pi_l \).

If AVs destined to a zone \( i \) self-park at zone \( j \) (i.e., \( V_{ij} > 0 \)), then \( \xi_{V,ij} = 0 \) by complementarity constraint (A.10), and consequently \( e_{ij} - E_i = 0 \) according to Eq. (A.3). If \( V_{ij} = 0 \), \( \xi_{V,ij} \geq 0 \) and \( e_{ij} - E_i \geq 0 \).

For vehicle type \( m = HV,AV \) , we sum Eq. (A.4) over all links on a route \( r \) to obtain

\[ \sum_{a \in A} \delta_{ar} (c^m_a + C^m_{D(a),i} - C^m_{0(a),i} - \xi_{x,aim}) \]

If route \( r \) carries flow of vehicle type \( m \) and \( \delta_{ar} = 1 \), then \( x^m_{ai} > 0 \) and consequently \( \xi_{x,aim} = 0 \ (i = D(r)) \) by complementarity constraint (A.11). Following (A.4),
\[ \sum_{a \in A} \delta_{ar} \left( c_a^m + C_{O(a),i}^m - C_{D(a),i}^m \right) = 0, \] which implies \( c_a^m + C_{D(a),i}^m = C_{O(a),i}^m \) that if link \( a \) carries flow from any zone to zone \( i \) \( (x^m_{ai} > 0) \), then travel cost on the link plus the minimum generalized travel cost from the end of the link \( D(a) \) to zone \( i \) should equal the minimum generalized travel cost from the beginning of the link \( O(a) \) to zone \( i \).

Recalling utility functions (3.49)-(3.51), Eqs. (A.5)-(A.7) can be unified as \( \frac{1}{\varphi} \ln Q_{ij}^m + U_{ij}^m + \theta_{ij} = 0, \) or \( Q_{ij}^m = \exp \left( -\varphi (U_{ij}^m + \theta_{ij}) \right), \forall (i,j) \in \Omega; m \in M. \) Summing \( Q_{ij}^m \)'s over vehicle types yields \( \sum_{m \in M} Q_{ij}^m = \exp \left( -\varphi \theta_{ij} \right) \Sigma_{m \in M} \exp \left( -\varphi U_{ij}^m \right). \) Together with demand conservation constraint (3.64), we obtain \( \exp \left( -\varphi \theta_{ij} \right) = Q_{ij} / \Sigma_{m \in M} \exp \left( -\varphi U_{ij}^m \right). \) Plugging this back to \( Q_{ij}^m = \exp \left( -\varphi (U_{ij}^m + \theta_{ij}) \right) \) gives the logit-based market shares. □

**Appendix 3.3. Solving VI-MARPUE (P2)**

To solve VI-MARPUE (P2), we reformulate it as a regular NLP (P5) below using a primal-dual gap function (Aghassi et al., 2006). P5 can be solved efficiently using commercial NLP solvers.

(P5): NLP-MARPUE

\[
\min_{(q, \eta, V, x, Q, H, E, C, \vartheta)} \sum_{l \in L} \left[ h_l q_l + (\kappa \pi_l - q_l - \tilde{q}_l) \eta_l \right] + \sum_{j \in Z} \sum_{e \in E} e_{ij} V_{ij} + \sum_{m=HV, AV} \sum_{i \in I} \sum_{j \in J} c_{ai} x_{ai}^m \ni^m \ni^m \\
+ \sum_{(i,j) \in \Omega} \left[ \sum_{m=HV} \left( y_1 PT_j + PC_j^m + PF_j + \frac{1}{\varphi} \ln Q_{ij}^m \right) Q_{ij}^m + \frac{1}{\varphi} \sum_{m \in M \setminus HV} \ln(Q_{ij}^m) Q_{ij}^m \right] \\
- \sum_{(i,j) \in \Omega} Q_{ij} \theta_{ij}
\]

s.t.

Flow and demand conservation constraints (3.3), (3.32), (3.40), (3.64)

Non-negativity constraints (3.60)-(3.63)

Definitional constraints of problem (P1)

\[ H_{ij} \leq h_l \quad \forall l \in L_{ij}; (i,j) \in \Omega \]  

(A.12)
The objective function of P5 excluding the last term represent the objective of a primal linear program (LP) resulting from P2. The last term in the objective function is the objective of its dual LP. Thus, the objective function of P5 gives the gap between the primal and the dual solutions, which should become zero at optimality by LP strong duality.

**Appendix 3.4. Sensitivity analysis of the results to key model parameters**

This appendix presents the sensitivity of the results to key model parameters, including: matching elasticity, (S)AV price, and in-vehicle VOT saving. For consistency, the performance measures presented below are similar to those in section 3.6.2.

**Matching elasticity**

Recall from Eq. (3.12) that $b_1$ and $b_2$ are the matching elasticities with respect to SAV traveler demand and SAV supply. Values for the two parameters reflect the efficiency of the underlying matching algorithm in the TNC’s platform and apps used by SAV travelers. Larger $b_1$ and $b_2$ imply more efficient matching. We rerun the model with $b_1$ and $b_2$ taking values from 0.5 to 1 in 0.1 intervals. Figures A.1-A.2 exhibit the results.

With increased $b_1$ and $b_2$, SAV and traveler matching times will decrease as expected. The matching efficiency improvement also leads to greater SAV market share and TNC profit. However, SAV market share appears to be “capped” when the matching elasticity is above 0.75. The changes in total and occupied
VMT, en-route travel time, and generalized travel costs compared to the scenario with only HVs are not monotonic, which are associated with significant fluctuations in vacant VMT generated by AVs and SAVs as we increase matching efficiency.

![Graphs](image)

**Figure A.1.** Impacts of matching elasticities on system performance compared to the baseline scenario

![Graphs](image)

**Figure A.2.** Impacts of matching elasticities on system performance compared to the scenario with only HVs

**(S)AV price**

The price markup of (S)AVs is expected to decrease with technological development and industrial scale production. This section varies the price markup of (S)AVs from 0% to 60% of the HV price in 10% increments, to understand the impact on the TNC, travelers, and system performance. Figures A.3-A.4 demonstrate the results.
As (S)AV becomes cheaper, travelers will be less inclined to use HVs. The shifted travelers will be attracted to use AVs more than SAVs. We find that the SAV market share remains largely constant for the range of SAV price considered. This may be attributed to the overall increase in distance-based fare, which adversely affects the attractiveness of SAV. Nonmonotonic changing trends are observed for other aspects of system performance, including TNC profit, SAV fleet size, SAV-HV replacement ratio, and changes in parking demand, total and occupied VMT, and total en-route travel time.

**Figure A.3. Impact of (S)AV price on system performance compared to the baseline scenario**

**Figure A.4. Impact of (S)AV price on system performance compared to the scenario with only HVs**
**In-vehicle VOT saving**

In-vehicle VOT saving is a major benefit of using (S)AVs. However, the exact extent of saving is unclear. In this section, we examine the impact of in-vehicle VOT saving on the TNC, travelers, and system performance by changing it from 0 to 80% in 10% intervals. Figures A.5-A.6 present the results.

As expected, fewer travelers will use HVs as in-vehicle VOT saving increases. The market share shift from HVs mainly goes to AVs rather than SAVs. One possible explanation is that the TNC’s responses in fare and fleet size, which are nonmonotonic, counteract the benefit brought by a larger VOT saving. It is worth noting a peaked SAV fleet size when VOT saving is 50%, suggesting a high SAV ownership cost which negatively affects the profitability of the TNC. The nonmonotonic changing trends of SAV fare and fleet size lead to similar variations in parking demand and VMT (both total and occupied). As in-vehicle VOT saving increases, travelers care less about the time spent in vehicles. As a result, total en-route travel time will increase, but total generalized travel cost will become smaller.

Figure A.5. Impact of in-vehicle VOT saving on system performance compared to the baseline scenario
Figure A.6. Impact of in-vehicle VOT saving on system performance compared to the scenario with only HVs
References


Appendix A. Permission for reuse of the previously published materials

Title: Roads in transition: Integrated modeling of a manufacturer-traveler-infrastructure system in a mixed autonomous/human driving environment

Author: Mohamadhossein Noruzoliaee, Bo Zou, Yang Liu

Publication: Transportation Research Part C: Emerging Technologies

Publisher: Elsevier

Date: May 2018

© 2018 Elsevier Ltd. All rights reserved.

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights
VITA

Mohamadhossein Noruzoliaee

2095 Engineering Research Facility
University of Illinois at Chicago
842 W. Taylor St., Chicago, IL 60607

Email: mnoruz2@uic.edu
Phone: (312) 647-3901

EDUCATION

- **Doctor of Philosophy** (Civil Engineering - Transportation) 9/2013 - 8/2018
  University of Illinois at Chicago
  Dissertation: *Supply-demand equilibrium of private and shared mobility in a mixed autonomous/human driving environment* (Advisor: Dr. Bo Zou)

- **Master of Science** (Civil Engineering - Transportation) 9/2009 - 2/2012
  Tarbiat Modares University, Iran

- **Bachelor of Science** (Civil Engineering) 9/2004 - 5/2009
  Iran University of Science and Technology, Iran

PROFESSIONAL POSITIONS

- **Graduate Research Assistant**, University of Illinois at Chicago 9/2013 - 8/2018
- **Research Associate**, Parseh Transportation Research Institute, Iran 2/2012 - 7/2013
- **Research Assistant**, Tarbiat Modares University, Iran 5/2010 - 2/2012
- **Teaching Assistant**, Tarbiat Modares University, Iran 9/2010 - 1/2012

PUBLICATIONS


**PRESENTATIONS**


**TECHNICAL REPORTS**


HONORS AND AWARDS

- Chancellor’s Student Service and Leadership Award, University of Illinois at Chicago (UIC), 2018.
- George Krambles Transportation Scholarship Award, Urban Transportation Center, UIC, 2016-2018.
- Christopher B. and Susan S. Burke Civil Engineering Graduate Student Award, Department of Civil and Materials Engineering, UIC, 2017.
- David Boyce Graduate Award, Department of Civil and Materials Engineering, UIC, 2015 and 2017.
- Graduate College Student Presenter Award, UIC, 2017-2018.
- Graduate Student Council Presenter Award, UIC, 2017-2018.
- American Society of Civil Engineers (ASCE) Transportation & Development Institute Illinois Chapter Scholarship, 2016.
- Selected as exceptional talented student and exempt from Ph.D. entrance exam, Tarbiat Modares University, Iran, 2013.

LEADERSHIP

- **President**, Institute of Transportation Engineers (ITE) Student Chapter 5/2017 - 8/2018
- **Vice President**, ITE Student Chapter 6/2016 - 4/2017
- **Student Leadership Council**, National University Rail (NURail) Center 7/2016 - 1/2017

AFFILIATIONS

- **Member**, Institute for Operations Research and Management Sciences (INFORMS) 2016 - present
- **Member**, Institute of Transportation Engineers (ITE) 2016 - present
- **Member**, American Society of Civil Engineers (ASCE) 2016 - present