On the Economics and Regulation of Smart Transportation Systems

by

Daniel A.M.C. Vignon

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Civil Engineering)
in the University of Michigan
2022

Doctoral Committee:

Professor Yafeng Yin, Chair
Professor Elisabeth Gerber
Professor Henry Liu
Associate Professor Neda Masoud
Professor Romesh Saigal

Daniel A.M.C. Vignon

dvignon@umich.edu

ORCID iD: 0000-0002-8190-3564

To my dad.

ACKNOWLEDGMENTS

I would first like to give thanks, with my whole being, to God Almighty for sustaining me with his grace over these past five years.

I also thank, from the bottom of my heart, my advisor, Professor Yafeng Yin. His availability, kindness, support, and guidance have helped me tremendously during my time at Michigan. They were instrumental in helping me navigate the world of academia and research, with its joys and travails. His trust also gave me the opportunity to grow into my own as a researcher and his intelligence continuously challenges me to learn more and do more. I certainly hope I can become a professor of his stature.

I am also indebted to my committee members, Professors Elisabeth Gerber, Henry Liu, Neda Masoud, and Romesh Saigal. Special shout-out to Liz and Romesh, whose insights and ideas in our regular meetings helped me give shape to the present work.

I am also forever indebted to my professors at Foothill College and MIT. To Dr. Young Hee Park Lee, I owe my heartfelt gratitude for conveying to me the joy and excitement of mathematics and of teaching in a way only she could. To Dr. Scott Lankford, I owe my thanks for kindling in me a love for writing and teaching me to make my paragraphs concise and to the point. I hope this has not gone to waste in the present work. I am also grateful that Jo-Ann Graziano kept that love of writing alive by showing me different ways it could be used. And, above all, heartfelt thanks to Pr. Anthony Patera. I most likely would not have gone to graduate school if not for his trust, advice and encouragement. From my first classes in numerical simulation to the present moment, he has been and outstanding mentor. And I certainly do not forget the warmth of his home where, with his wife Anddie and their beloved Max, I have been welcomed so many times.

To my family, I owe pretty much everything. I am grateful to my parents, Maimouna and Jean-Claude, for raising me, putting up with me, and providing for me. Be it for my love for sports and reading, my attention to detail, or my dedication to learning, they hold the blame. I am also indebted to my sisters, Christelle and Fati, for their gentle and not-so gentle challenges, but also for their care and attention. I admire whom they have become and this certainly motivates me to strive to do better. And to my younger brother, Pascal, who has brought such joy in my life and who keeps making me proud. I am also forever grateful to my uncle Christian, who took a chance on me and not only sent me to the US, but also takes me as his own son.

As far as family is concerned, I must also mention my Filipino family. They have provided for me a home away from home, be it my Uncle Danny, my Auntie Clarissa, my brothers Nate and Pat, or all the cousins, siblings and children. They have been a great blessing and I look forward to many more memories with them.

I would also like to thank the many friends and acquaintances that have supported me over the years. I am grateful to the Ann Arbor Catholic community, especially my parish, St Thomas the Apostle. I have truly found a home in your midst and will dearly miss the Masses, the singing and parties, but also the opportunity to grow closer to God through and with you. I especially want to thank Mark, Chris, Marjorie, Jacob, Dan, Zsiga and Eszter, Jenny, Stephen, Jophin, Chris, Mike, Paul, Jeff, Pablo, Rachel, Ellie, and Frs. Joe Wagner, Bill, Dan, and Gerald. I also am grateful to have found such a dynamic and enticing skating community both at the Ann Arbor and Arctic figure skating clubs. These provided a much needed break from the doctoral life and I am thankful to my coach, Connie, and my ”skating grandma”, Melanie, for helping me progress and enjoy my time on the ice. I also want to thank my trainer, Dean Anderson, for helping me get back into running and understand my body a little better. It has been a tremendous help in keeping me healthy and happy here.

I also cannot omit my dear friends Tiwaloluwa–*l'homme des hommes*–and his wife JGA, whose support, prayers, and friendship I deeply cherish.

Last but not least, I am grateful for the friendships I made within LIMOS, and the department. I have enjoyed growing with and next to so many talented people. Thanks to Mehrdad for wooing me to Michigan and to Mehrdad and Zhibin for taking me under their wings when I first joined LIMOS. Thanks to Zhengtian and Xiaotong for their support through many late nights, joys, and sorrows. Thank you to Zhen Yang: I will miss our lunches, dinners, and banter. Thank you to Sina, for his inspiring hard work and thoroughness. And thank you to Alex, Mojtaba, Tara, Tianming, Zhichen, Minghui, Tian and all other past and present LIMOS members from whom I have learned so much.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

ABSTRACT

Technological advances in sensing, mobile communication, computation, imaging, artificial intelligence and many other fields have brought the city of tomorrow within reach. In that city, infrastructure and vehicles will be connected and automated; mobility services–from ground transportation to airborne services–will be readily accessible from smartphones and highly personalized thanks to an abundance of data; and sensors and cloud infrastructure will allow for constant monitoring and timely maintenance of infrastructure components. While this future is exciting, its concretization will disrupt our current way of life: the technologies and services enabling the city of tomorrow will first emerge in an environment that was not thought out or designed for them. This will create challenges both for the growth of these technologies but also to society as it seeks to integrate them. Naturally, these challenges raise an array of policy-relevant questions: what are the benefits of these new technologies? Do the benefits justify the costs of integration and disruption? Is the current regulatory ecosystem appropriate for the growth of beneficial technologies and, if not, how should it be adapted?

In this dissertation, by bringing together insights from economics, operations research and traffic science, I seek to address these policy concerns along two main thrusts.

In the first thrust, I investigate the optimal regulation of the ride-hailing industry in the age of uberization. Indeed, over the past decade, e-hailing transportation network companies such as Uber and Lyft have entered the ride-hailing industry and quickly grown in popularity. Naturally, this surge in popularity has resulted in the decline of street-hailing services. Faced with dwindling revenues, traditional taxi drivers have been calling for regulatory action, accusing e-hailing companies of unfair competition due to their unregulated status. Moreover, e-hailing services have been

linked to an increase in congestion in a number of metropolitan cities. Regulators have, however, struggled to address both competition and congestion concerns in ways that promote efficiency and consumer welfare. Thus, I propose a model of the ride-hailing market that captures both competition between e-hailing and street-hailing services but also the effect of these services on congestion. My analysis shows that the emergence of e-hailing need not be a death sentence for street-hailing under appropriate regulatory and market conditions. More importantly, I show that a simple mechanism exists to address issues of unfair competition in the industry as well as of congestion, thus providing an avenue to simplify the host of regulations that have historically burdened the industry.

In the second thrust, I focus on the question of automated mobility and infrastructure services. The current approach to driving automation has been primarily vehicle-centric. However, despite tremendous spending on R&D, fully automated vehicles are still far from being a reality. In this context, a vehicle-infrastructure cooperative approach, in which infrastructure and vehicles cooperate to perform the different driving tasks, could emerge and be preferable. To study the implication of such a paradigm, I develop a model of an automated mobility market. In this market, consumers interact with both automakers and with infrastructure support service providers (ISSPs)–entities which provide automation services through road infrastructure–in order to find suitable mobility solutions. This model allows me to study the suitability of vehicle-infrastructure cooperation; the outcome of strategic interactions between ISSPs and automakers; and the implications for safety and liability should that market emerge.

Part I

Regulating the Ride-hailing Market

CHAPTER 1

Introduction

1.1 Motivation

Traditionally, the ride-hailing industry has been dominated by street-hailing services. In metropolitan cities like New York City (NYC), these services have usually been subject to major regulations: limits on the number of cabs operating, licensing requirements, fare controls. However, over the past decade, e-hailing transportation network companies such as Uber and Lyft have entered the ride-hailing industry and quickly grown in popularity while operating largely unregulated. For example, Uber's quarterly earning reports show that between 2016 and 2018 the number of trips served by the platform increased tenfold (Chai, 2019). Naturally, this surge in popularity has resulted in the decline of street-hailing services. For example, in NYC, between April 2016 and April 2020, the market share of street-hail fell from 86% to 14% (Schneider, 2022). Faced with dwindling revenues, traditional taxi drivers have been calling for regulatory action, accusing e-hailing companies of unfair competition due to their unregulated status (Rana, 2022).

While one might argue that the imminent death of the street-hailing industry is the natural result of technological advancement and of e-hailing's superior efficiency, empirical evidence suggests that the picture is more complicated. Indeed, using data from Shenzhen, China, Nie (2017) and Zhang et al. (2019) show that, in very dense settings, using street-hailing will tend to result in lower waiting times than using e-hailing. This seems to indicate that, at least in major urban centers like NYC, society might benefit from sustained operation of street-hailing, even if I ignored the mobility needs from those who don't have access to a smartphone. This is, however, less likely to occur unless policymakers provide for an appropriate regulatory environment that allows for street-hailing's competitiveness.

However, to date, in major US cities, no regulatory action has sought to address the legal discrepancy between e-hailing companies and the street-hailing industry. Moreover, in other parts of the world, regulatory action has often sought to assuage street-hailing drivers by curtailing or outlawing the use of e-hailing (Rana, 2022). However, the fact that e-hailing companies have a user-base more expansive than street-hailing (Contreras and Paz, 2018; Rayle et al., 2016; Clewlow and Mishra, 2017) suggests that there may be a better regulatory approach to the question. Moreover, in other parts of the world and, recently, in NYC, partnerships between e-hailing companies and taxi drivers seem to indicate that synergistic operation between these two services is possible (Rana, 2022). In this context, one might wonder the potential welfare implications of greater synergy and even consolidation between street-hailing and e-hailing.

Moreover, evidence that e-hailing services are a major contributor to increased traffic congestion in cities has been surfacing in recent years. Using the National Household Travel Survey data, Schaller (2018) found that replacing a private vehicle trip with an e-hailing trip at least doubles the number of miles traveled. Moreover, he also reported that only a little over 20% of trips taken with Uber or Lyft in New York City in February 2018 using their pooling service resulted in actual sharing. This puts into question the claim that e-hailing services might positively impact congestion by inducing sharing. After accounting for population, employment growth, and roadway modifications, Castiglione et al. (2018) and Erhardt et al. (2019) showed that e-hailing companies are a major contributor to the drop in travel speed and the increase in vehicle miles travelled in San Francisco between 2010 and 2016. More recently, using exogenous variation provided by Uber and Ola drivers' strike in three major Indian cities, Agarwal et al. (2019) showed that the absence of e-hailing drivers resulted in a reduction in delay nearly equivalent to half of that observed on major holidays[1]. E-hailing vehicles may, in certain instances, alleviate private car use, although available research suggests that a higher proportion of users switch from taxis or higher-occupancy modes, or would forego their trip altogether (Rayle et al., 2016; Clewlow and Mishra, 2017; Hampshire et al., 2017; Schaller, 2018). However, as compared with private car use, in serving their customers, e-hailing vehicles generate massive vacant or empty trips. These vacant trips create additional vehicular traffic demand. Unless pooling plays a larger part and the fleet is efficiently managed, e-hailing will likely worsen traffic conditions, at least in the short term (Beojone and Geroliminis, 2021; Wei et al., 2020).

Cities have already started to take steps to mitigate congestion caused by e-hailing vehicles. For example, in 2019, New York City announced new regulations that impose a cap on new licenses issued to for-hire vehicles, mandate a minimum percent of time e-hailing vehicles must carry a passenger while operating in Manhattan below 96th Street, and collect a congestion surcharge to trips that begin in, end in or pass through the area. While e-hailing companies have generally been favorable to the congestion surcharge, some have criticized that the regulations hurt customers–especially low-income customers–and drivers by locking them out of the app at times of low

[1]These results, however, represent short-term effects and do not necessarily reflect longer term patterns resulting from a stoppage in e-hailing services

demand (Bellon, 2019; Dobbs, 2019; **?**). This debate highlights the importance of understanding the welfare implications of these regulations.

1.2 Contribution and Outline

The aim of the first part of this dissertation is to make policy recommendations that can successfully address the issues arising from the rise and adoption of e-hailing services. I divide this first part into two chapters. In Chapter 2, to address the issue of e-hailing-induced congestion, I develop an analytical model that replicates the inner workings of the e-hailing market and captures the effect of e-hailing vehicles on traffic congestion. Our model accounts for the presence of both a solo service (e.g.: UberX, Lyft...) and a pooling service (e.g.: UberPool, Lyft Line...) and customers' choices between these two services. Our model also captures the effect of e-hailing vehicles on traffic congestion using traditional traffic flow analysis methods. Using that framework, I identify the issues that arise when an e-hailing company is left to operate on its own, especially with regards to congestion. I am also able to investigate the adequacy–or lack thereof–of the current regulatory solutions put forth by cities like New York. Then, I design and analyze the impact of a simple regulatory scheme that preserves the flexibility of the e-hailing platform while appropriately addressing its congestion effects and, thus, improves welfare.

Then in Chapter 3, to address the issue of competition in the ride-hailing industry, I propose a model of competition in the ride-hailing market. After investigating the socially optimal configuration in such a context, I investigate the outcome of e-hailing and street-hailing competition. I also investigate the integration of these two services into a single platform in which fares are jointly determined but e-hailing customers cannot match with street-hailing drivers through an app. Lastly, I am also able to assess the impact of current policies and to propose policies more suited for the age of e-hailing.

CHAPTER 2

E-hailing and Congestion

2.1 Literature review

This chapter contributes to our discussion on how to regulate e-hailing services, especially when accounting for their effect on congestion.

In the literature, several studies have analyzed ridesourcing markets and investigated the welfare implications of potential regulations. Using an aggregate model of the ridesourcing market, Zha et al. (2016) demonstrated that, under the assumption of homogeneous value of time and labor supply, capping the commission that the ridesourcing firm takes on each ride is sufficient to achieve a second-best. Zha et al. (2018a,b) further examined the commission cap in a market with spatial or temporal heterogeneity and demonstrated that the commission cap can, in effect, significantly improve welfare. It appears that the cap may be imposed per trip, by unit distance, or time, and can even vary with respect to location or time of day. The choice of the granularity level will depend on the trade-off between implementation complexity and policy effectiveness. Ke et al. (2020a) compared the pricing equilibrium in two different types of ridesourcing market: a pooling market and a non-pooling market. They show that, at the monopolist, first-best, and second-best equilibria, both markets operate in an efficient regime (as opposed to the wild-goose-chases or WGC regime, first described by Castillo and Weyl (2018)). Additionally, in either equilibrium, the fare in the pooling market is lower than that in the non-pooling market. Their study does not, however, consider the joint operation of both services. Additionally, all the aforementioned studies do not consider traffic congestion. In fact, most of the analyses of regulations accounting for congestion externality in the for-hire vehicle market have focused on the taxi market. Yang et al. (2005) proposed a model of taxi service with a fare structure that implicitly incorporates congestion. Under the proposed framework, they showed that the first-best solution might be sustainable when congestion is high. Yang et al. (2014a) extended this work by incorporating bilateral taxi-customer search frictions and showed that taxi demand may decrease with taxi fleet size in the presence of congestion externality. Moreover, under the assumption of constant returns to scale for the matching function, taxi utilization rates decrease along the Pareto frontier from the system optimal solution to the monopolist solution. Albeit insightful, these results should be re-examined against the ridesourcing market due to its distinctive features, such as its two-sided nature, matching technology and workforce flexibility, etc. A few recent studies seek to address the question of congestion and regulation in the ridesourcing market. Li et al. (2019) investigated different regulations in the ridesourcing market: the minimum wage, the congestion fee, and the driver cap. Their analyses show that a minimum wage can actually result in higher welfare for consumers and drivers but lower profits for the firm. In contrast to Li et al. (2019), Zhang and Nie (2019) modelled both the solo and pooling options offered by the ridesourcing platform. This allows them to capture the trade-offs between maximizing vehicle occupancy and mitigating congestion on the one hand, and ensuring that drivers are incentivized to provide service on the other. However, both analyses of congestion regulation fall short by not incorporating the mechanism through which ridesourcing vehicles affect traffic congestion. In contrast, Ke et al. (2020b) incorporated such a mechanism using the concept of macroscopic fundamental diagram (Geroliminis and Daganzo, 2008). They are able to show that pooling can, under certain circumstances and with an appropriate pairing time window, reduce the total travel cost experienced by ridesourcing customers. Their analysis assumes, however, that pooling does not coexist with solo rides on the same platform. Thus, in the present work, we propose a stylized framework that captures the workings of a ridesourcing market, its two different services, i.e., riding alone or sharing with someone else, and their effects on traffic congestion. We then derive optimal solutions to the monopolist and social-optimum problems, and analyze the impact of a commission cap and a toll on system performance and social welfare.

This chapter is organized as follows. In **Section 2.2**, we present our main assumptions and our model. **Section 2.3** analyzes the monopolist and first-best solutions. **Section 2.4** discusses our proposed mechanism to regulate the market and **Section 2.5** illustrates the results through numerical examples.

2.2 Model

Consider a e-hailing market with one platform offering two types of services: *solo* rides (denoted by s) and *pooling* rides (p), which customers choose to use based on their preferences. However, drivers are required to provide both services, and the assignment of customers will be made by the platform via a matching algorithm. In addition, the platform decides the fares f^s and f^p that customers will pay to use each service as well as the amount that drivers will receive for delivering customers. Thus, by setting the fares and drivers' share, the platform essentially determines the demand rates $\tilde{\lambda}^s$ and $\tilde{\lambda}^p$ for the solo and pooling services.

2.2.1 Pairing and matching

In this paper, we consider a stylized matching process that is consistent in spirit with the matching algorithms implemented by some e-hailing platforms in practice. To avoid confusion, we use *pairing* to refer to the process by which a pooling customer is grouped with other customers with whom they will split their ride. We use *matching* to refer to the process by which both solo and pooling customers are assigned to their drivers.

We consider a pairing process where pooling customers wait for a pairing time window to be matched to other pooling customer. Customers are paired if they depart and arrive within a given pooling radius of each other. Once a target occupancy (the maximum number of customers to be paired) is reached, the pairing process will be terminated and the pooling customers will be subsequently matched with a driver. If the target occupancy is not achieved at the end of the pairing period, currently pooled customers will be assigned to a driver. It implies that if no other customer is found, the waiting customer will ride alone. With this consideration, given the target occupancy, pooling radius and pairing time window exogenously determined by the platform, the average pairing time w^p experienced by a pooling customer is given by:

$$w^p = W^p(\tilde{\lambda}^p)$$

with $W^{p'} < 0$, suggesting that an increase in pooling demand leads to a decrease in pairing time. Correspondingly, the relationship between average occupancy o and the pooling demand can then be described by another function:

$$o = O(\tilde{\lambda}^p)$$

where $O' > 0$.

On the other hand, we assume, as in Castillo and Weyl (2018), that the matching time for customers is zero, owing to instantaneous matching and a sufficiently large matching radius. In other words, customers are matched to drivers as they arrive in the matching queue. Pooling customers enter the matching queue either as a unit (with the other customers they will ride with) or alone (if suitable pooling partners were not found before the expiration of the pairing time window).

2.2.2 Meeting and delivery

Upon being matched to a driver, solo customers experience an expected pickup time w^m and an expected delivery time w^r. In addition to w^m and w^r, pooling customers will also experience an expected detour time Δw that stems from picking up and dropping off their pooling partners. We have:

$$w^r = \frac{d^r}{v}$$
$$w^m = \frac{d^m}{v}$$
$$\Delta w = \frac{\Delta d}{v}$$

where v is an average traffic speed; d^r is the average trip distance for e-hailing trips, which is assumed to be given; d^m is the average pickup distance and is a decreasing function of the density of idle drivers n^I i.e., $d^m = D^m(n^I)$ and $D^{m'} < 0$; and Δd is the expected detour distance. For a given pooling radius, we assume that the detour distance is a decreasing function of pooling demand as follows:

$$\Delta d = \Delta D(\tilde{\lambda}^p)$$

with $\Delta D' < 0$. The exact functional form of ΔD will necessarily depend on the operational decisions of the platform as shown, for example, in **?**.

Following the network macroscopic fundamental diagram approach (Geroliminis and Daganzo, 2008), it is possible to describe the average traffic speed v using the number of e-hailing vehicles n and the number of background vehicles n^b in the network:

$$v = V(\theta \cdot n + n^b)$$

with $V'(\cdot) < 0$. The parameter $\theta \geq 1$ reflects higher marginal effect of e-hailing vehicles on congestion compared to regular background vehicles, because, for example, e-hailing vehicles often drive slower as they await their next assignment.

Assuming that that background traffic trips originate at a rate λ^b with an average travel distance of d^{rb}, we have:

$$w^{rb} = \frac{d^{rb}}{v}$$
$$n^b = \lambda^b \cdot w^{rb}$$

where w^{rb} represents the average travel time of background traffic. The second equation holds as per Little's law. We further note that while it is straightforward to consider λ^b to be congestion-dependent, we treat it to be exogenous for simplicity and clarity.

2.2.3 Demand

Customers face the following costs from taking a trip on the platform:

$$\mu^s = f^s + \beta \cdot (w^m + w^r) - \xi^s$$
$$\mu^p = \frac{f^p}{o} + \beta \cdot (w^p + w^m + w^r + \Delta w)$$

where f^s and f^p denote the total fare collected by the platform for completing a transaction for the solo and pooling services, respectively; β represents customers' value of time; and $\xi^s > 0$ is a constant, reflecting the fact that, all else equal, riding alone provides higher utility to customers. Additionally, customers also have access to an outside option (such as public transit or driving) with cost μ^0. We further assume that customers' value of travel time β is a variable across the customer population with cumulative density function $G(\cdot)$. Thus, each customer chooses which service to use by comparing the costs μ^s, μ^p and μ^0 and choosing the cheapest option. Thus, we have:

1. if $\mu^s \leq \mu^p$ and $\mu^s \leq \mu^0$, then customers with $\beta_2 \leq \beta \leq \beta_1$ choose the solo service. The proportion of these customers is $G(\beta_1) - G(\beta_2)$.
2. if $\mu^p \leq \mu^0$ and $\mu^p < \mu^s$, then customers with $\beta \leq \beta_3 = \min\left\{\beta_2, \frac{\mu^0 - \frac{f^p}{o}}{\left(w^p + w^m + w^r + \Delta w\right)}\right\}$ choose the pooling service, whose proportion is $G(\beta_3)$.

Here:

$$\beta_1 = \frac{\mu^0 - f^s + \xi^s}{w^m + w^r}$$
$$\beta_2 = \frac{f^s - \frac{f^p}{o} - \xi^s}{w^p + \Delta w}$$

If we let λ^0 denote the population size, then demand for each service is given by:

$$\tilde{\lambda}^s = \lambda^0 \cdot \left[G(\beta_1) - G(\beta_2)\right]$$
$$\tilde{\lambda}^p = \lambda^0 \cdot G(\beta_3)$$

2.2.4 Supply

We further assume that drivers decide to provide service if their average hourly earnings during the study period ω exceed their opportunity cost. That is, the supply of drivers is given by:

$$n = S(\omega)$$

where $S'(\cdot) > 0$ and $S(\cdot)$ captures the distribution of drivers' opportunity cost. The above relationship between driver supply and driver earnings can be further simplified if one considers the inverse supply function $S^{-1}(\cdot)$:

$$C(n) = S^{-1}(n) \cdot n = \omega \cdot n$$

where $C' > 0$ by construction. $C(\cdot)$ can be understood as the cost for the platform of using n drivers. Now, let e denote the amount that drivers receive per unit service time [1]. Then:

$$\omega = e \cdot \frac{w^r \cdot (\tilde{\lambda}^s + \frac{\tilde{\lambda}^p}{o})}{n}$$

2.2.5 Equilibrium

At equilibrium, we consider a steady state in the system where the following conservation holds as per Little's law:

$$n = n^I + (\tilde{\lambda}^s + \frac{\tilde{\lambda}^p}{o}) \cdot (w^m + w^r) + \frac{\tilde{\lambda}^p}{o} \cdot \Delta w^d$$

where Δw^d is the additional time that drivers spend on pickup and delivery for a pooling ride relative to a solo ride. As pointed out in Ke et al. (2020a), Δw^d and Δw are correlated, though this correlation, once again, depends on operational decisions from the platform. To simplify our analysis, we also assume that $\Delta w^d = \gamma \cdot \Delta w$.

All of the above considerations yield our model of eqs. **2.1a** to **2.1o**, which is a system of 14 equations and 17 unknowns. By specifying exogenous variables f^s, f^p, and e, we can solve the system to evaluate the performance of the e-hailing market at the steady state.

$$\tilde{\lambda}^s = \lambda^0 \cdot \Big[G(\beta_1) - G(\beta_2)\Big] \tag{2.1a}$$

[1]Here, note that we assume that drivers do not get paid for the extra detour time induced by pooling. Our model could easily be extended to include such a possibility. In this formulation, we however seek to capture the reality that, for many drivers, they work more and earn almost the same (if not less) when serving pooling trips as opposed to solo trips.

$$\tilde{\lambda}^p = \lambda^0 \cdot G(\beta_3) \tag{2.1b}$$

$$\beta_1 = \frac{\mu^0 - f^s + \xi}{w^m + w^r} \tag{2.1c}$$

$$\beta_2 = \frac{f^s - \frac{f^p}{o} - \xi}{w^p + \Delta w} \tag{2.1d}$$

$$\beta_3 = \min\left\{\beta_2, \frac{\mu^0 - \frac{f^p}{o}}{\Big(w^p + w^m + w^r + \Delta w\Big)}\right\} \tag{2.1e}$$

$$o = O(\tilde{\lambda}^p) \tag{2.1f}$$

$$w^p = W^p(\tilde{\lambda}^p) \tag{2.1g}$$

$$w^m = \frac{D^m(n^I)}{v} \tag{2.1h}$$

$$w^r = \frac{d^r}{v} \tag{2.1i}$$

$$\Delta w = \frac{\Delta D(\tilde{\lambda}^p)}{v} \tag{2.1j}$$

$$w^{rb} = \frac{d^{rb}}{v} \tag{2.1k}$$

$$v = V(\theta \cdot n + n^b) \tag{2.1l}$$

$$n = n^I + (\tilde{\lambda}^s + \frac{\tilde{\lambda}^p}{o}) \cdot (w^m + w^r) + \frac{\tilde{\lambda}^p}{o} \cdot \gamma \cdot \Delta w \tag{2.1m}$$

$$e \cdot \Big[w^r \cdot (\tilde{\lambda}^s + \frac{\tilde{\lambda}^p}{o})\Big] = C(n) \tag{2.1n}$$

$$n^b = \lambda^b \cdot w^{rb} \tag{2.1o}$$

2.3 Scenario analysis

Below we analyze scenarios that determine the choice of exogenous variables f^s, f^p, and e given our e-hailing market model presented above. In order to analyze the market and derive policy insights, both in this section and **Section 2.4**, we will make a few assumptions and specifications to simplify the model. Later in Section 5, we will conduct numerical experiments to examine the effectiveness of the policies derived in the general case considered in the previous section.

2.3.1 Model simplification

We first assume that customers are homogeneous in their value of time β. With homogeneous value of time, the demand-side equilibrium can be described by the following set of equations:

$$\tilde{\lambda}^s + \tilde{\lambda}^p = \Lambda(\mu)$$
$$\mu^s - \mu \geq 0 \text{ and } \tilde{\lambda}^s \geq 0$$
$$\mu^p - \mu \geq 0 \text{ and } \tilde{\lambda}^p \geq 0$$
$$\tilde{\lambda}^s \cdot [\mu^s - \mu] = 0$$
$$\tilde{\lambda}^p \cdot [\mu^p - \mu] = 0$$

where $\Lambda(\cdot)$ is a demand function for e-hailing services and is such that $\Lambda' < 0$. The above complementarity condition indicates that, at equilibrium, the cost of services with non-zero demand must be equal, and less than that of non-utilized services.

We also assume, as in Korolko et al. (2018), that pooling customers must walk to and from common meeting and drop-off locations. As such, $\Delta d = 0$. This scenario can be likened to the operation of Uber Express Pool, Uber's low-cost service. In this case, the driver compensation becomes the same for solo and pool services and thus it is convenient to consider the choice of $r = e \cdot w^r$, the compensation per ride, rather than that of e. In light of the above, the e-hailing market model becomes:

$$\mu^s = f^s + \beta \cdot (w^m + w^r) - \xi^s \tag{2.2a}$$
$$\mu^p = \frac{f^p}{o} + \beta \cdot (w^p + w^m + w^r) \tag{2.2b}$$
$$\tilde{\lambda}^s + \tilde{\lambda}^p = \Lambda(\mu) \tag{2.2c}$$
$$\mu^s - \mu \geq 0 \tag{2.2d}$$
$$\mu^p - \mu \geq 0 \tag{2.2e}$$
$$\tilde{\lambda}^s \cdot [\mu^s - \mu] = 0 \tag{2.2f}$$
$$\tilde{\lambda}^p \cdot [\mu^p - \mu] = 0 \tag{2.2g}$$
$$o = O(\tilde{\lambda}^p) \tag{2.2h}$$
$$w^p = W^p(\tilde{\lambda}^p) \tag{2.2i}$$
$$w^m = \frac{D^m(n^I)}{v} \tag{2.2j}$$
$$w^r = \frac{d^r}{v} \tag{2.2k}$$
$$w^{rb} = \frac{d^{rb}}{v} \tag{2.2l}$$

$$v = V(\theta \cdot n + n^b) \tag{2.2m}$$

$$n = n^I + (\tilde{\lambda}^s + \frac{\tilde{\lambda}^p}{o}) \cdot (w^m + w^r) \tag{2.2n}$$

$$r \cdot (\tilde{\lambda}^s + \frac{\tilde{\lambda}^p}{o}) = C(n) \tag{2.2o}$$

$$n^b = \lambda^b \cdot w^{rb} \tag{2.2p}$$

In our analyses, we assume that the above modeling system defines continuously differentiable functions between endogenous variables and exogenous variables f^s, f^p, and r, as per the implicit function theorem. To facilitate the presentation of our analysis results, we define the *vehicle* trip rates for solo and pool services as follows:

$$\lambda^s = \tilde{\lambda}^s$$
$$\lambda^p = \frac{\tilde{\lambda}^p}{o}$$

To conclude our model presentation, below we present a few results that highlight some useful properties of our model and that will be used in analyzing our different scenarios.

From **Equation 2.2n**, we can obtain the following derivative for the total vehicle trip rate $\lambda^s + \lambda^p$ with respect to the number of vacant vehicles:

$$\frac{\partial(\lambda^s + \lambda^p)}{\partial n^I} = -\frac{1 + (\lambda^s + \lambda^p) \cdot \frac{D^{m'}}{v}}{w^m + w^r} \tag{2.3}$$

When $\frac{1+(\lambda^s+\lambda^p)\cdot \frac{D^{m'}}{v}}{w^m+w^r} < 0$, then $\frac{\partial(\lambda^s+\lambda^p)}{\partial n^I} > 0$, which corresponds to the WGC described in Castillo and Weyl (2018).

From **Equations 2.2l**, **2.2m** and **2.2p**, the derivative of the total traffic $n^T = n + n^b$ with respect to the fleet size is:

$$\frac{\partial n^T}{\partial n} = 1 - \frac{\theta \cdot d^{rb} \cdot \lambda^b \cdot V'}{v^2 + \lambda^b \cdot d^{rb} \cdot V'} \tag{2.4}$$

Given our assumption of inelastic background traffic demand, it must be that $\frac{\partial n^T}{\partial n} > 0$ so that $\frac{\theta \cdot d^{rb} \cdot \lambda^b \cdot V'}{v^2+\lambda^b \cdot d^{rb} \cdot V'} < 1$. If $v^2 + \lambda^b \cdot d^{rb} \cdot V' < 0$, then it follows that $v^2 < (\theta - 1) \cdot \lambda^b \cdot d^{rb} \cdot V' < 0$ (given that $\theta \geq 1$), which is absurd. Thus, it follows that when background traffic demand is inelastic, we must have $v^2 + \lambda^b \cdot d^{rb} \cdot V' > 0$.

2.3.2 First-best

Here, a social planner seeks to maximize social welfare by solving the following optimization problem:

$$W = \max_{\substack{f^s \geq 0, \\ f^p \geq 0, \\ r \geq 0}} \overbrace{\int_{\mu}^{\infty} \Lambda(x) \cdot dx}^{\textbf{Consumers' surplus}} + \overbrace{(f^s - r) \cdot \lambda^s + (f^p - r) \cdot \lambda^p}^{\textbf{Platform profit}} + \overbrace{\int_0^{\frac{r \cdot (\lambda^s + \lambda^p)}{n}} S(x) \cdot dx}^{\textbf{Drivers' surplus}} - \overbrace{\gamma^b \cdot \lambda^b \cdot w^{rb}}^{\textbf{Background cost}} \tag{FB}$$

where γ^b is the value of time for the background traffic[2].Note that the dependent variables λ^s, λ^p, n, w^{rb} and μ are functions of the decision variables through the system defined in **Equation 2.2**. Without loss of generality, we do not consider the platform's operation cost.

Assuming that $\lambda^s > 0$ and $\lambda^p > 0$, the first-order optimality conditions (FONC) of the problem yield the following formulae for the price of a solo and a pool ride and drivers' hourly income:

$$f^s = mc \tag{2.5a}$$

$$\frac{f^p}{o} = mc \cdot c^o + \beta \cdot W^{p'} \cdot \lambda^p \cdot o \tag{2.5b}$$

$$\frac{C(n)}{n} = \frac{mc}{w^m + w^r} - \tau^{int} - \tau^b \tag{2.5c}$$

where:

$$mc = -\beta \cdot \frac{D^{m'}}{v} \cdot \frac{(w^m + w^r)}{1 + (\lambda^s + \lambda^p) \cdot \frac{D^{m'}}{v}} \cdot (\lambda^s + \lambda^p \cdot o)$$

$$c^o = \frac{1 - \lambda^p \cdot O'}{o} = \frac{O'}{\frac{\partial o}{\partial \lambda^p}} < \frac{1}{o}$$

$$\tau^b = -\theta \cdot \gamma^b \cdot \frac{d^{rb} \cdot \lambda^b \cdot V'}{v^2 + \lambda^b \cdot d^{rb} \cdot V'} > 0$$

$$\tau^{int} = -\theta \cdot \beta \cdot \frac{V'}{v^2 + \lambda^b \cdot d^{rb} \cdot V'} \cdot \frac{(d^m + d^r)}{1 + (\lambda^s + \lambda^p) \cdot \frac{D^{m'}}{v}} \cdot (\lambda^s + \lambda^p \cdot o) > 0$$

Here, mc represents the marginal cost, to the platform, of providing a ride. Since $f^s > 0$, it follows from **Equation 2.5a** that, at the social optimum, $mc > 0$, which implies that $1 + (\lambda^s + \lambda^p) \cdot \frac{D^{m'}}{v} > 0$. Thus, at the first-best, the equilibrium of the e-hailing market lies in the non-WGC regime.

[2]Truly, this is a *quasi first-best* since we take λ^b as given. Under the first-best, the planner would also be able to determine λ^b by considering the demand function for background traffic.

Additionally, we note that **Equation 2.5b** contains the negative term $\beta \cdot \lambda^p \cdot o \cdot W^{p'}$, which captures the matching externality that each additional pooling customer creates on the platform. It then follows that, since $f^p > 0$, $c^o > 0$ at the first-best equilibrium. Thus, increasing the pooling *vehicle* trip rate increases occupancy for the pooling service ($\frac{\partial o}{\partial \lambda^p} > 0$). By combining **Equations 2.2a**, **2.2b**, **2.5a** and **2.5b**, we obtain:

$$mc \cdot (1 - c^o) = \beta \cdot W^p \cdot (1 + \lambda^p \cdot o \cdot \frac{W^{p'}}{W^p}) + \xi^s \tag{2.6}$$

If we assume that the mode-specific constant $\xi^s \approx 0$, then $1 + \lambda^p \cdot o \cdot \frac{W^{p'}}{W^p} > 0$ from which it follows that the solution to the first-best problem lies in the non-elastic portion of the pairing time function for the pooling service. When $\xi^s > 0$, this condition is relaxed and the market outcome may lie in the elastic region of the pairing time function. Here, the disutility of using the pooling service is so large that the resulting equilibrium number of pooled rides is relatively low. It is also evident from **Equation 2.5** that $f^s - \frac{f^p}{o} > 0$ and $f^s - f^p > 0$. In other words, the fare per customer and the revenue per trip for the single service are higher than their respective counterparts for the pooling service at the first-best.

From **Equation 2.5c**, the average driver hourly income is equalized with the drivers' marginal social benefit. This marginal benefit is composed of three main components:

- the benefit of a marginal driver to the platform $\frac{mc}{w^m + w^r}$, which captures drivers' impact on meeting distance;
- the intra-platform congestion externality τ^{int} that a marginal driver imposes on e-hailing customers;
- the extra-platform congestion externality τ^b that a marginal driver imposes on the background traffic.

From this latest point, we notice that the externality component is independent of occupancy. This suggests that, irrespective of the number of passengers they carry, e-hailing vehicles impose the same externality on background traffic.

We can also rewrite **Equation 2.5** as follows:

$$f^s = \left[\frac{C(n)}{n} + \tau^{int} + \tau^b\right] \cdot (w^m + w^r) \tag{2.7a}$$

$$\frac{f^p}{o} = \left[\frac{C(n)}{n} + \tau^{int} + \tau^b\right] \cdot (w^m + w^r) \cdot c^o + \beta \cdot W^{p'} \cdot \lambda^p \cdot o \tag{2.7b}$$

Thus, at the first-best, the fare only covers the marginal social cost of vehicles that are in pickup or delivery modes. In particular, it does not cover the marginal cost of drivers when they are idle. Now, from Arnott (1996) and Yang et al. (2005), we know that unless congestion is high, the taxi market must be subsidized at the first-best. Additionally, from Zha et al. (2016), without considering congestion, the e-hailing market must be subsidized when the production function of rides is increasing returns to scale. We examine here whether, in the presence of pooling and congestion, the e-hailing market is sustainable at the first-best. Consider the first-best profits π^f:

$$\pi^f = \overbrace{-\frac{mc}{w^m + w^r} \cdot n^I + (\beta \cdot W^{p'} \cdot \lambda^p \cdot o^2 - mc \cdot O') \cdot \lambda^p}^{\textbf{Operating loss} \leq 0} + \left(\tau^b + \tau^{int}\right) \cdot n \tag{2.8}$$

From examining **Equation 2.8**, we notice two main components. On one hand, there is a congestion-independent component that is negative and captures the industry's losses. Those include the cost of idle drivers–which is not covered by the fare at the first-best (Arnott, 1996; Yang et al., 2005)–and a pooling-service-related cost. This latter cost emerges from the positive matching externality that the pooling service enjoys and the fact that pooling reduces the marginal impact of riders on the platform but not the marginal cost of drivers relative to the single service. On the other hand, there is a congestion related component, $\left(\tau^b + \tau^{int}\right) \cdot n$, which is positive. When congestion externality is high, this term may exceed the operating loss, which would result in positive profits for the platform. In other words, similar to the taxi market investigated by Yang et al. (2005), the first-best is sustainable for the e-hailing platform when congestion externality is high.

2.3.3 Monopoly

In this section, we derive the monopolist equilibrium, its properties, and the distortions that arise relative to the social optimum. In this setting, the e-hailing platform determines exogenous variables f^s, f^p, and r to maximize its profits. Thus its revenue-maximizing decision can be obtained by solving the following optimization problem:

$$\pi = \max_{\substack{f^s \geq 0,\\ f^p \geq 0,\\ r \geq 0}} \quad (f^s - r) \cdot \lambda^s + (f^p - r) \cdot \lambda^p \tag{M}$$

The FONC of the optimization problem are given below:

$$f^s = -\frac{\lambda^s + \lambda^p \cdot o}{\Lambda'} + mc \tag{2.9a}$$

$$\frac{f^p}{o} = -\frac{\lambda^s + \lambda^p \cdot o}{\Lambda'} + mc \cdot c^o + \beta \cdot \lambda^p \cdot o \cdot W^{p'} \tag{2.9b}$$

$$C'(n) = \frac{mc}{w^m + w^r} - \tau^{int} \tag{2.9c}$$

Compared to first-best pricing, f^s and f^p include a markup term $-\frac{\lambda^s+\lambda^p \cdot o}{\Lambda'} > 0$ under monopoly pricing. Thus, fares under the monopolist are higher than under the first-best. Moreover, **Equation 2.9c** indicates that, at optimality, the platform equalizes the cost of the marginal driver, $C'(n)$, to its marginal benefit. Thus, while the regulator is concerned with the average cost of the fleet, the platform is only concerned with the cost of the marginal driver. Additionally, by comparing the right-hand sides of **Equations 2.5c** and **2.9c**, it appears that the monopolist only internalizes part of the congestion externality that arises from running the platform. Indeed, under monopoly pricing, customers only bear the congestion cost they impose on each other but do not bear the cost of congestion on the background traffic. Taken together, it follows that, for a convex driver cost function $C(\cdot)$, assuming equal marginal driver and marginal externality costs, the number of drivers under the monopolist is higher than the first-best number of drivers [3]. Since demand is lower under the monopolist, it follows that utilization rates are also lower under the monopolist, a finding similar to that of Ke et al. (2020a) when considering pooling and solo services separately.

Now, the question arises as to whether it may be optimal for the monopolist to operate in the WGC-regime ($mc < 0$) when accounting for congestion. Indeed, unlike in the first-best case, it is not straightforward to rule out a WGC solution at optimality for the monopolist. It is, however, easy to show that for any solution with $mc < 0$, we can construct another solution with identical n, λ^s, and λ^p but higher fares f^s and $\frac{f^p}{o}$. Thus, though they may exist, solutions with $mc < 0$ are dominated, even in the presence of congestion. This is consistent with the findings of Zha et al. (2018b) in a spatial market without congestion.

Lastly, while $f^s > \frac{f^p}{o}$ and $f^s > f^p$ in the first-best, it is not necessary that $f^s > f^p$ under the monopolist. Indeed:

$$f^s - f^p = -\frac{\lambda^s + \lambda^p \cdot o}{\Lambda'} \cdot (1 - o) + mc \cdot (1 - c^o \cdot o) - \beta \cdot \lambda^p \cdot o^2 \cdot W^{p'} \tag{2.10}$$

Since $o \geq 1$, the markup per ride that the platform collects is higher for the pooling service. When occupancy is high and the markup per passenger is significantly higher than the difference

[3] This can be seen by noting that:

- C and C' are both increasing in n and that the right hand side of **Equation 2.5c** is lower than that of **Equation 2.9c**;
- $C'(n) \geq \frac{C(n)}{n}$ for convex functions.

in marginal costs between the solo and the pooling service, the monopolist may earn more revenue from the pooling service than from the solo service. In practice, however, such a situation does not seem to occur: numerous reports suggest that the prices required to make the pooling service profitable tend to induce a demand lower than that necessary to achieve high occupancy (Spotswood, 2017; Anand, 2020). Thus, for the rest of our analysis, we will assume that $f^s > f^p$ at the monopolist equilibrium.

2.4 Policy discussion

Before we dive into the analysis of potential policies, it is important to characterize an efficient policy. First, such a policy must target the two sides of the market, i.e., it must address the monopolist's market power on the demand and supply sides. Second, it must address congestion externality by ensuring that drivers (and by extension the passengers) bear the social cost of the congestion they impose on the background traffic. Lastly, the policy should be easy to implement.

We seek for an optimal policy with which the monopolist problem admits the first-best solution. Suppose that the fares $\hat{f}^s$ and $\hat{f}^p$ and the per-ride driver revenue $\hat{r}$ solve the above first-best problem. An obvious policy would be to regulate the fares and the earnings to be at this first-best level. Such regulations, however, may be unpopular as they would restrict the operational freedom and flexibility of the platform. Since the modeling system of the e-hailing market, eqs. **2.2a** to **2.2p**, enjoys three degrees of freedom (14 equations and 17 unknowns), a policy that can ensure that any three variables stay at their first-best level will be optimal, if the reduced system admits a unique solution. However, this implies that regulating fares, driver per-ride earnings, or the fleet size alone cannot induce the first-best; it is necessary to explore a combination of regulatory instruments.

2.4.1 New York City's regulatory scheme

We first briefly analyze NYC's approach to regulating the for-hire vehicle market to determine whether it meets the aforementioned criteria. We preface our discussion by noting that NYC's e-hailing market operates with multiple companies while our setting only considers a single firm. Thus, we do not comment on the effectiveness of NYC's policy as it pertains to NYC's current market, but rather on the effectiveness of that policy when applied to a market similar to the one in our setting. In 2019, New York City announced new regulations that impose a cap on new licenses issued for for-hire vehicles, mandate a minimum percent of time e-hailing vehicles must carry a passenger while operating in Manhattan below 96th Street, and collect a congestion surcharge on trips that begin in, end in or pass through the area. In December 2018, NYC additionally implemented an effective minimum wage requirement of $15\,\frac{\$}{\text{hr}}$ for e-hailing drivers. We investigate here whether such policies are effective in our setting, where, under these regulations, the problem for the monopolist would be as follows:

$$\begin{aligned} \max_{\substack{\mathbf{F}\geq 0,\\ \omega\geq 0,\\ n\geq 0}} \quad & (f^s - r - \hat{\tau}^s)\cdot\lambda^s + (f^p - r - \hat{\tau}^p)\cdot\lambda^p \\ \text{s.t.} \quad & (\lambda^s + \lambda^p)\cdot w^r \geq \hat{\rho}\cdot n \quad \text{(Occupied time constraint)}, \\ & n \leq \hat{n} \quad \text{(Fleet size constraint)}, \\ & C(n) \leq \hat{\omega}\cdot n \quad \text{(Minimum wage constraint)} \end{aligned} \tag{2.11}$$

where $\hat{\tau}^s$ and $\hat{\tau}^p$ are the congestion surcharges on single and pooled trips respectively; $\hat{\rho}$ is the first-best utilization rate for the e-hailing fleet; $\hat{n}$ is the first-best vehicle fleet cap; and $\hat{\omega}$ is the wage under the first-best.

First, we note that the fleet size constraint and the minimum wage constraint cannot simultaneously be binding at equilibrium. Indeed, the fleet size constraint is only necessary and effective when the congestion externality imposed on the background traffic exceeds the monopolist's market power, so that the monopolist wage is higher than the socially efficient wage. On the other hand, the minimum wage constraint is only effective and necessary when congestion is low and the firm's market power leads to wages lower than socially efficient. We analyze below the two cases.

Suppose that the fleet size constraint is binding so that $n = \hat{n}$ and $C(\hat{n}) = \hat{\omega}\cdot\hat{n}$. Then, $w^r = \frac{d^r}{V(\hat{n}+\hat{n}^b)}$, i.e., the monopolist and first-best travel times are equal and the monopolist solves the following problem:

$$\begin{aligned} \max_{\mathbf{F}\geq 0} \quad & (f^s - r - \hat{\tau}^s)\cdot\lambda^s + (f^p - r - \hat{\tau}^p)\cdot\lambda^p \\ \text{s.t.} \quad & \lambda^s + \lambda^p \geq \hat{\lambda} \end{aligned} \tag{2.12}$$

where $\hat{\lambda} = \hat{\rho}\cdot\frac{\hat{n}}{\hat{w}^r} = \hat{\lambda}^s + \hat{\lambda}^p$. Essentially, the occupied time constraint becomes a minimum on the number of trips served by the platform. Thus, it essentially acts as a way to maximize the demand served by the monopolist. However, the fleet size cap renders the congestion management effect of the congestion surcharge unnecessary. Moreover, the occupied time constraint might be redundant. Indeed, following Xu et al. (2017), when congestion is high, the objectives of the platform and of the planner tend to be aligned, so that the platform inherently seeks to maximize its fleet utilization rate.

Now consider the case in which congestion is low and the monopolist's driver supply is lower than that targeted by the regulator. Then, we consider two situations of interest. In the first, $\hat{n}$ is such that $C(\hat{n}) = \hat{\omega} \cdot \hat{n}$. Then, the monopolist solves the following problem:

$$\begin{aligned} \max_{\mathbf{F} \geq 0, n \geq 0} \quad & (f^s - r - \hat{\tau}^s) \cdot \lambda^s + (f^p - r - \hat{\tau}^p) \cdot \lambda^p \\ \text{s.t.} \quad & (\lambda^s + \lambda^p) \cdot w^r \geq \hat{\rho} \cdot \hat{n} \quad \text{(Occupied time constraint)} \end{aligned} \tag{2.13}$$

This is identical to **Equation 2.12**. Here again, the surcharge is unnecessary (since congestion is not a problem). Moreover, it likely hampers the effectiveness of the minimum trip requirement, since higher prices might discourage consumers from using the service. In the second case, the platform's optimal choice of n is such that $C(n) \leq \hat{\omega} \cdot \hat{n}$: it is more advantageous for the platform to hire fewer drivers. In such situation, not only would the number of drivers be suboptimal, but the rationing mechanism used by the platform could further decrease welfare.

Our brief analysis shows that, while NYC's regulatory scheme might be effective in mitigating the congestive effect of ride-sourcing vehicles, it is unnecessarily burdensome and might not always improve welfare in our setting. In the following section, we propose another solution that not only remedies that issue but is also more parsimonious.

2.4.2 Commission cap regulation and congestion toll

Zha et al. (2016) showed that, when customers are homogeneous in their value of time, regulating the amount of commission that the platform receives can achieve a second-best. Consider such regulation applied to our current framework. First, we note that, given a commission cap on the solo service corresponding to the first-best commission $\hat{p}^s = \hat{f}^s - \hat{r}$, it must be that the firm's choice of pooling commission is such that $p^p \leq \hat{p}^s$[4]. Thus, there exists a natural, non-binding cap on the pooling service commission. It might therefore be possible to regulate the market with a single commission cap. Second, we note that **Equations 2.5c** and **2.9c** differ by τ^b. *A priori*, a policy that increases the cost of drivers by τ^b should be enough to address the congestion externality. Thus, for the regulated monopoly, **Equation 2.2o** becomes:

$$\begin{aligned} r &= \frac{C(n)}{\lambda^s + \lambda^p} + \tau^b \cdot (w^I + w^m + w^r) \\ &= [S^{-1}(n) + \tau^b] \cdot (w^I + w^m + w^r) \end{aligned}$$

with $w^I = \frac{n^I}{\lambda^s + \lambda^p}$. We consider then the following regulated problem for the monopolist:

[4] This is because the driver revenue per ride r is identical for both services. Since $f^s > f^p$, the conclusion follows.

$$\pi = \max_{\substack{f^s \geq 0, \\ f^p \geq 0, \\ r \geq 0}} \quad (f^s - r) \cdot \lambda^s + (f^p - r) \cdot \lambda^p \tag{M-CAPT}$$
$$\text{s.t.} \quad f^s - r \leq \hat{p}^s \quad \text{(Single service cap)}$$

The FONC of M-CAPT satisfy:

$$f^s = -[\lambda^s - \nu_1 + o \cdot \lambda^p] \cdot \frac{1}{\Lambda'} + [C'(n) + \tau^b + \tau^{int}] \cdot (w^m + w^r) + \left[S^{-1}(n) + \tau^b\right] \cdot w^I \cdot \frac{\nu_1}{\lambda^s + \lambda^p} - \tau^{int} \cdot (w^m + w^r) \cdot \frac{\nu_1}{\lambda^s + \lambda^p \cdot o} \tag{2.14a}$$

$$\frac{f^p}{o} = -\frac{\lambda^s - \nu_1 + \lambda^p \cdot o}{\Lambda'} + c^o \cdot \left[[C'(n) + \tau^b + \tau^{int}] \cdot (w^m + w^r) + \left[S^{-1}(n) + \tau^b\right] \cdot w^I \cdot \frac{\nu_1}{\lambda^s + \lambda^p} - \tau^{int} \cdot (w^m + w^r) \cdot \frac{\nu_1}{\lambda^s + \lambda^p \cdot o} \right] + \beta \cdot \lambda^p \cdot o \cdot W^{p'} \tag{2.14b}$$

where $\nu_1 \geq 0$ is the Lagrangian multiplier associated with the commission cap.

Now consider a constructed, constrained social welfare maximization problem below:

$$W = \max_{\substack{f^s \geq 0, \\ f^p \geq 0, \\ r \geq 0}} \quad \int_{\mu}^{\infty} \Lambda(x) \cdot dx + (f^s - r) \cdot \lambda^s + (f^p - r) \cdot \lambda^p - \gamma^b \cdot \lambda^b \cdot w^{rb} \tag{SB}$$
$$\text{s.t.} \quad f^s - r \geq \hat{p}^s$$

It is straightforward to see that the first-best solution solves this constrained social welfare maximization problem. Now, the FONC of SB yield:

$$f^s = -\frac{\eta_1}{\Lambda'} + \left[\frac{C(n)}{n} + \tau^b + \tau^{int}\right] \cdot (w^m + w^r) - S^{-1}(n) \cdot w^I \cdot \frac{\eta_1}{\lambda^s + \lambda^p} + \tau^{int} \cdot (w^m + w^r) \cdot \frac{\eta_1}{\lambda^s + \lambda^p \cdot o} \tag{2.15a}$$

$$\frac{f^p}{o} = -\frac{\eta_1}{\Lambda'} + c^o \cdot \left[\left[\frac{C(n)}{n} + \tau^b + \tau^{int}\right] \cdot (w^m + w^r) - S^{-1}(n) \cdot w^I \cdot \frac{\eta_1}{\lambda^s + \lambda^p} + \tau^{int} \cdot \frac{\eta_1}{\lambda^s + \lambda^p \cdot o} \cdot (w^m + w^r) \right] + \beta \cdot \lambda^p \cdot o \cdot W^{p'} \tag{2.15b}$$

where $\eta_1 \geq 0$ is the Lagrangian multiplier associated with the commission cap.

Now, assuming the first-best is attained by the regulation, does there exist $\nu_1 \geq 0$ so that **Equation 2.14** holds? By analyzing the system of **Equations 2.14** and **2.15**, the answer to this question can be reduced to determining the conditions under which:

$$\nu_1 = (\lambda^s + \lambda^p \cdot o) \cdot \frac{\left[\tau^{int} + \frac{C(n)}{n} - C'(n)\right] \cdot (w^m + w^r) \cdot \bar{\lambda} - S^{-1}(n) \cdot w^I}{\tau^b \cdot w^I} \geq 0 \tag{2.16}$$

where $\bar{\lambda} = \frac{\lambda^s + \lambda^p}{\lambda^s + \lambda^p \cdot o}$. It then follows that the first-best can be replicated with a cap and toll if:

$$\bar{\pi} = \tau^{int} \cdot (w^m + w^r) \cdot \bar{\lambda} - S^{-1}(n) \cdot w^I \geq \left[C'(n) - \frac{C(n)}{n}\right] \cdot (w^m + w^r) \cdot \bar{\lambda} \tag{2.17}$$

Now, $\tau^{int} \cdot (w^m + w^r) \cdot \bar{\lambda}$ represents the platform's revenue per customer (solo and pooled) served; $\left[C'(n) - \frac{C(n)}{n}\right] \cdot (w^m + w^r) \cdot \bar{\lambda}$ represents the opportunity cost of equalizing drivers' benefit to their average cost rather than their marginal cost; and $S^{-1}(n) \cdot w^I$ represents the cost of an idle vehicle per customer served. Thus, **Equation 2.17** simply indicates that the first-best can be replicated when the platform's revenues are sufficiently high to cover the economic cost of its drivers. Can this condition be satisfied, regardless of the composition of the driver pool?

When drivers are homogeneous [5], this is a slightly weaker condition than requiring that the platform's profits be positive since, as **Equation 2.8** indicates, profits include a pooling-service-related cost that does not appear in **Equation 2.17**. Thus, under the assumption of homogeneous drivers, as long as the first-best is sustainable, the proposed regulation will be effective.

Additionally, **Equations 2.16** and **2.17** do not depend on η_1. Thus, any desired second-best equilibrium can be replicated by a regulation with a cap $\hat{p}^s > 0$ and an appropriate toll. This is especially important when the first-best is not sustainable ($\hat{p}^s < 0$) and the cap must be increased beyond its first-best level to ensure platform operation.

When drivers are heterogeneous, assuming the cost function $C(\cdot)$ is convex, then the right hand side of **Equation 2.17** is positive. Then, that the first-best is sustainable may or may not be sufficient to guarantee that the policy can replicate the regulator's objective.

In practice, collecting a toll on e-hailing vehicles can be challenging, especially if there is no preexisting mechanism to toll other vehicles. However, imposing a congestion fee per use time on each rider is more easily implementable (cities and states already impose multiple fees on riders). From **Equation 2.14**, we can easily deduce that the appropriate fee structure per rider is as follows:

$$\tau^s = \tau^b \tag{2.18}$$

[5] That is, $C'(n) = \frac{C(n)}{n} = c$ and $S^{-1}(n) = c$

$$\tau^p = \tau^b \cdot c^o \tag{2.19}$$

where τ^s and τ^p denote the fee imposed on solo and pooled riders, respectively. We note that $\tau^s > \tau^p$ since $c^o < \frac{1}{o} < 1$. Additionally, $\frac{\tau^s}{\tau^p} = \frac{1}{c^o} > o$ so that the ratio between the single fee and the pooling fee is not linear in occupancy. It will rather depend on the extent of congestion and its costs on society.

2.4.3 Commission cap only

Zha et al. (2016) explained that the commission cap incentivizes the monopolist to serve a higher demand than it otherwise would. When congestion is taken into account, maximizing demand served might involve maximizing occupancy for the platform, thus alleviating congestion and achieving the regulator's objective.

To analyze such a regulation, it is convenient to introduce $\hat{\tau}^n \in [0, \tau^b]$, the toll imposed by the regulator on each e-hailing vehicle.

Then, **Equation 2.14** becomes:

$$\begin{aligned} f^s = -[\lambda^s - \nu_1 + o \cdot \lambda^p] \cdot \frac{1}{\Lambda'} + [C'(n) + \tau^{int} + \hat{\tau}^n] \cdot (w^m + w^r) + \\ \left[S^{-1}(n) + \hat{\tau}^n\right] \cdot w^I \cdot \frac{\nu_1}{\lambda^s + \lambda^p} - \tau^{int} \cdot \frac{\nu_1}{\lambda^s + \lambda^p \cdot o} \cdot (w^m + w^r) \end{aligned} \tag{2.20a}$$

$$\begin{aligned} \frac{f^p}{o} = -\frac{\lambda^s - \nu_1 + \lambda^p \cdot o}{\Lambda'} + \Big[[C'(n) + \tau^{int} + \hat{\tau}^n] \cdot (w^m + w^r) + \\ \left[S^{-1}(n) + \hat{\tau}^n\right] \cdot w^I \cdot \frac{\nu_1}{\lambda^s + \lambda^p} - \tau^{int} \cdot \frac{\nu_1}{\lambda^s + \lambda^p \cdot o} \cdot (w^m + w^r)\Big] \cdot c^o + \beta \cdot \lambda^p \cdot o \cdot W^{p'} \end{aligned} \tag{2.20b}$$

Then, a solution to the monopoly problem that satisfies **Equation 2.15** can be obtained if:

$$\nu_1 = (\lambda^s + \lambda^p \cdot o) \cdot \frac{\left[\tau^{int} + \frac{C(n)}{n} - C'(n) + \tau^b - \hat{\tau}^n\right] \cdot (w^m + w^r) \cdot \bar{\lambda} - S^{-1}(n) \cdot w^I}{\hat{\tau}^n \cdot w^I} \geq 0 \tag{2.21}$$

From **Equation 2.21**, we note the following:

- When $\hat{\tau}^n = \tau^b$, we recover the commission cap and toll policy from **Section 2.4.2**;
- For any toll $\hat{\tau}^n \in (0, \tau^b)$, the first-best can be replicated if:

$$\bar{\pi} = (\tau^{int} + \tau^b - \hat{\tau}^n) \cdot (w^m + w^r) \cdot \bar{\lambda} - S^{-1}(n) \cdot w^I \geq \left[C'(n) - \frac{C(n)}{n}\right] \cdot (w^m + w^r) \cdot \bar{\lambda} \tag{2.22}$$

Here, $(\tau^b - \hat{\tau}^n) \cdot (w^m + w^r) \cdot \bar{\lambda} > 0$ represents the additional revenue for the platform due to bearing only a fraction of the congestion externality it imposes on the background traffic. The interpretation of **Equation 2.22** is then similar to that of **Equation 2.17**.

- Using our approach, it is not possible to recover ν_1 when $\hat{\tau}^n = 0$. However, we note that, as $\hat{\tau}^n \to 0$, $\nu_1 \to \infty$ i.e., ν_1 becomes more and more positive. This suggests that it is always possible to find a smaller toll $\hat{\tau}^n$ such that the first-best (or another targeted equilibrium) can be replicated.

From this last point above, it appears that the selection of $\hat{\tau}^n$ does not have a significant impact on the ability of the cap to bring the system to an efficient equilibrium. In fact, the commission cap alone might be able to achieve the desired effect. We will verify this intuition using numerical examples in **Section 2.5**. In practice, the regulating authority might still, however, choose $\hat{\tau}^n > 0$ in order to meet other objectives: improving road infrastructure, satisfying special interests, etc.

2.5 Numerical experiments

Our proposed policies in **Section 2.4** assume that users are identical in their value of time and that pooling users experience no detour time. In this section, we relax this assumption and apply our proposed policies to the full model described in **Equation 2.1**. In this context, the social welfare maximization problem now becomes:

$$\begin{aligned} W = \max_{\substack{f^s \geq 0,\, f^p \geq 0 \\ r \geq 0}} \quad & \overbrace{U^0 \cdot (\lambda^s + \lambda^p \cdot o)}^{\text{Total trip utility}} - \overbrace{\lambda^0 \cdot \int_{\beta_2}^{\beta_1} \beta \cdot (w^m + w^r) \cdot G'(\beta) \cdot d\beta}^{\text{Solo time cost}} \\ & - \overbrace{\lambda^0 \cdot \int_{\underline{\beta}}^{\beta_3} \beta \cdot (w^p + w^m + w^r + \Delta w) \cdot G'(\beta) \cdot d\beta}^{\text{Pooled time cost}} \\ & - \underbrace{r \cdot (\lambda^s + \lambda^p)}_{\text{Driver cost}} - \underbrace{\gamma^b \cdot \lambda^b \cdot w^{rb}}_{\text{Background traffic}} \end{aligned} \tag{FB}$$

where $\underline{\beta}$ is the lower bound of the support of G; and U^0 is the utility of completing the trip for an individual customer. Unlike in the homogeneous value of time case, however, analyzing the above system analytically for policy insights is substantially difficult. Thus, we turn instead to numerical experiments, for which we adopt the following functional forms:

$$D^m(n^I) = A \cdot {n^I}^{-\alpha} \tag{2.23a}$$

$$C(n) = c \cdot n \tag{2.23b}$$

$$W^p(\tilde{\lambda}^p) = \frac{1 - \exp\left(-\delta \cdot \phi \cdot \tilde{\lambda}^p\right)}{\delta \cdot \tilde{\lambda}^p} \tag{2.23c}$$

$$O(\tilde{\lambda}^p) = 2 - \exp\left(\delta \cdot \phi \cdot \tilde{\lambda}^p\right) \tag{2.23d}$$

$$\Delta D(\tilde{\lambda}^p) = \frac{B}{\tilde{\lambda}^p} \tag{2.23e}$$

$$V(\theta \cdot n + n^b) = v^0 - v^c \cdot (\theta \cdot n + n^b) \tag{2.23f}$$

$$G(\beta) = \frac{\beta - \underline{\beta}}{\bar{\beta} - \underline{\beta}} \tag{2.23g}$$

$$c^o = \frac{1 - \exp\left(-\delta \cdot \phi \cdot \tilde{\lambda}^p\right) \cdot \delta \cdot \phi \cdot \frac{\tilde{\lambda}^p}{o}}{o} \tag{2.23h}$$

We assume, as in Korolko et al. (2018), that the maximum occupancy is two and that the probability of being paired is constant. Then, the resulting pairing time and occupancy functions are given in **Equations 2.23c** and **2.23d**. Furthermore, we assume that β follows a uniform distribution on $(\underline{\beta}, \bar{\beta})$. **Equation 2.23a** is a classical representation of the pickup distance between a randomly chosen driver and their closest unmatched customer (Zha et al., 2018b; Korolko et al., 2018; Castillo, 2018). We also borrow from Korolko et al. (2018) for information on δ. The data and parameters used are presented in Appendix A.1.

2.5.1 Comparative analysis

In this section, we compare the first-best and monopoly outcomes. To tease out the effects moving away from the homogeneous value of time assumption, we also vary the variance of the value of time distribution and investigate how it affects both outcomes [6]. Especially, we consider the producer's and social surpluses, the distribution of rides across the two services and the congestion effect.

As **Figure 2.1** shows, the monopolist's behavior results in an inefficient outcome. This inefficient outcome is, however, not driven by congestion, since, as shown in **Figure 2.4**, traffic speed in the unregulated scenario is similar to that in the first-best scenario when congestion is high. Thus, most inefficiencies are the results of the monopolist's exercise of market power, as shown by the fares in **Figure 2.2**. As the variance of the value of time distribution increases, so does the difference between the *laissez-faire* outcome and the first-best outcome. This is due to the fact that, when variance is high, the monopolist can easily maximize profits by catering to high value of time customers. However, as variance decreases, the e-hailing service must become

[6] The effect of the mean of the distribution was also studied. However, the findings were as expected and do not add much regulatory insights to what can be gleaned from studying the variance. Higher value of time implies lower profits for the platform since it needs to hire a larger number of drivers. Lower value of time implies that pooling becomes preferable and that profits are easier to generate.

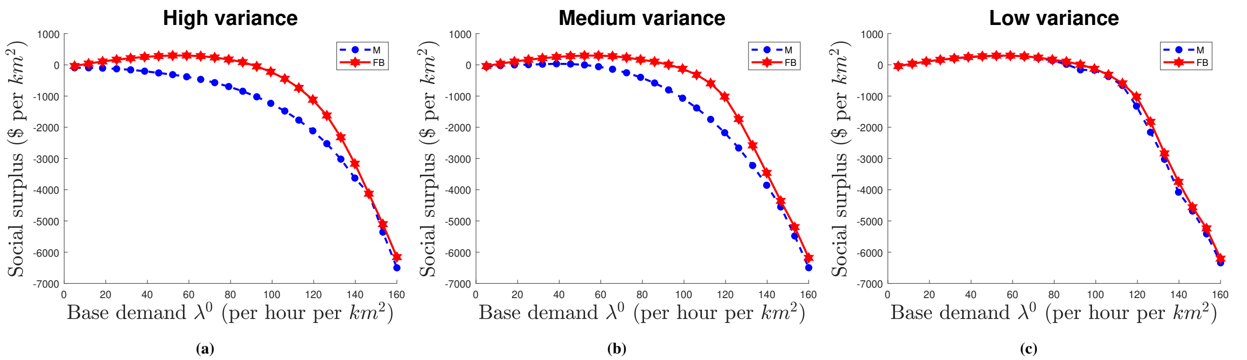


Figure 2.1: *Social surplus comparisons*

less niche to survive and maximize its profits. This is best seen in **Figure 2.3**: as variance in the population decreases, total demand served (across both services) increases. Such a pattern can also easily be understood in terms of price elasticity: as variance in the population increases, the price elasticity decreases, thus making it easier for the firm to charge higher fares (**Figure 2.2**).

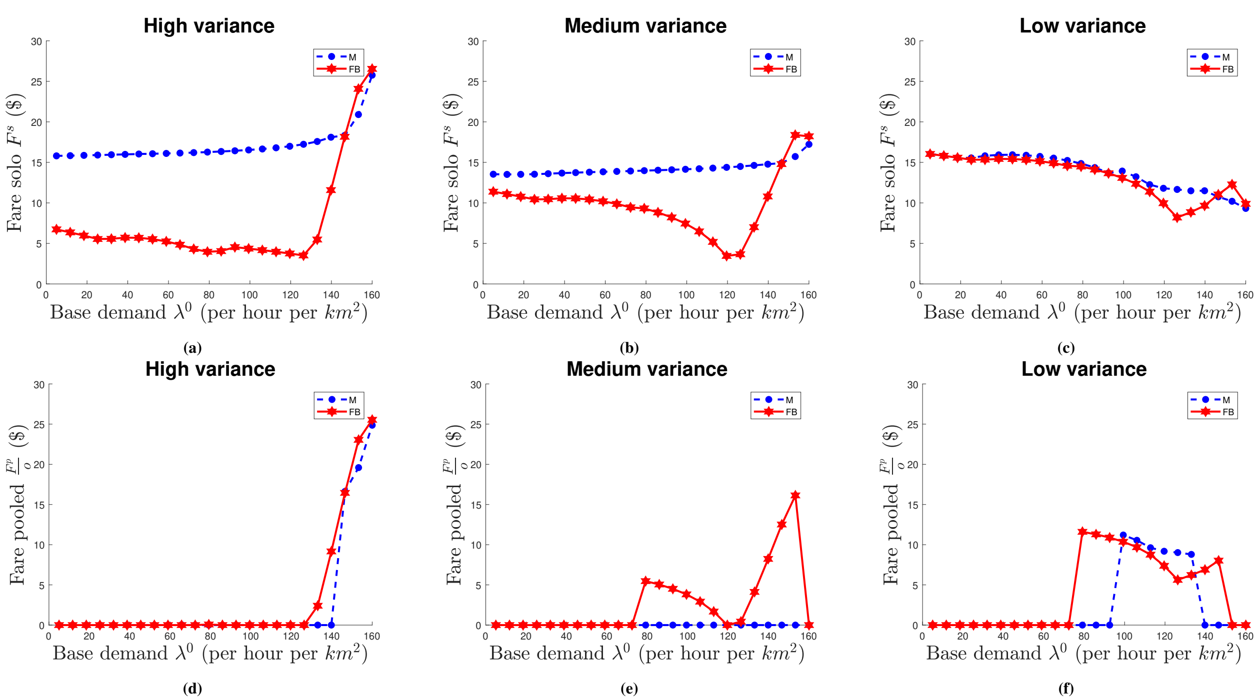


Figure 2.2: *Fare comparisons*

As far as the trip distribution is concerned, **Figure 2.3** shows that, at the first-best, we can discern two regimes. In the first, as base demand increases, the fraction of demand served by the solo service is non-increasing while that served by the pooling service increases. This is because, when demand is low, frictions due to pooling cannot be easily overcome. However, as demand increases, economies of scale lead to a reduction in pooling and detour time costs, thus making pooling the more efficient option to serve the demand. In the second regime, as congestion

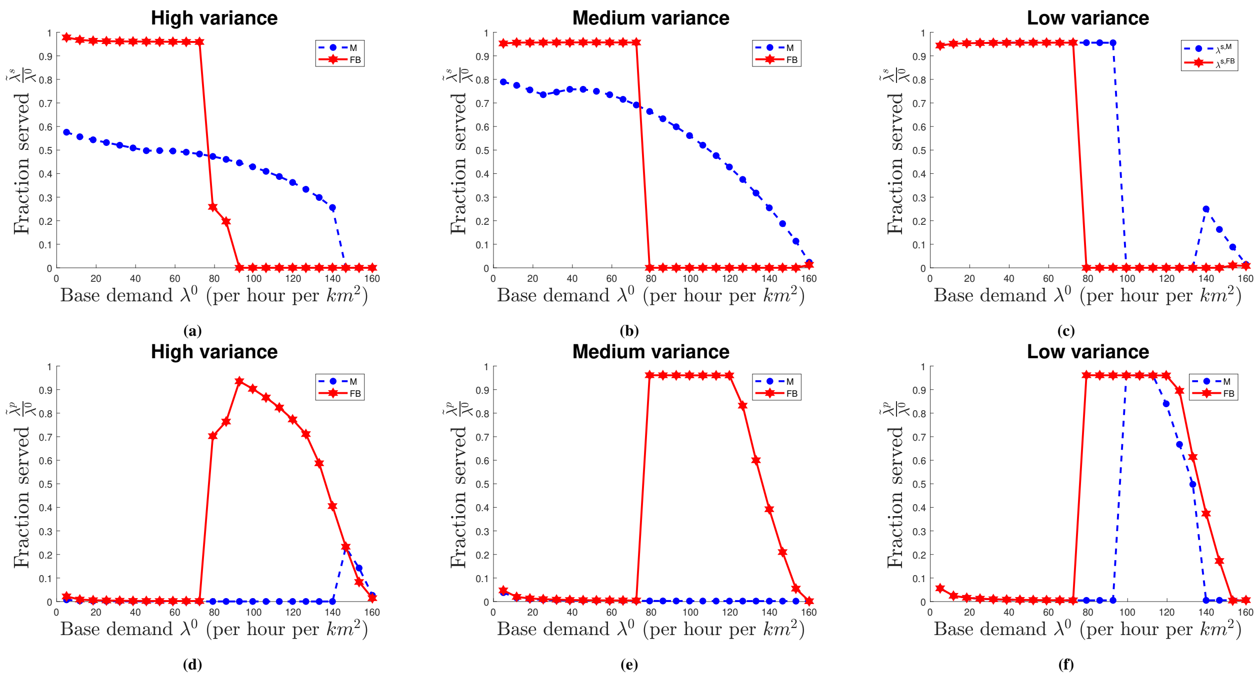


Figure 2.3: *Demand rate comparisons*

increases, pooling (and the e-hailing service overall) become less desirable and the service is eventually eliminated. Under the monopolist, the share of solo trips is non-increasing, just as in the first-best case. However, when variance is high, pooling is only provided when congestion is very high: because it manages less demand than under the first-best, the monopolist is able to maintain a good quality of service for its high value of time customers longer while getting adequately compensated. When variance is low, the distribution pattern under the monopolist becomes similar to that under the first-best, with a caveat: when congestion is very high, the monopolist still provides the solo service, since it does not have to bear its full contribution to congestion.

Interestingly, regardless of the value of time distribution, beyond a certain point, as the base demand increases, the lesser the discrepancy between the first-best and the monopolist. This is because, similar to the findings of Xu et al. (2017), the interests of the platform and the planner become more aligned. As demand rises, both actors are looking to manage their fleet more efficiently in order to serve the growing demand. Thus, both the fleet size and the utilization rates increase (**Figure 2.4**). However, as congestion becomes more problematic, both under the first-best and the monopoly, the number of vehicles and their utilization rate decrease (**Figure 2.4**). This is the result of two forces. Firstly, falling pooling demand served increases detour times. Secondly, the reduction in traffic speed increases pickup and detour times. Both forces thus contribute to reducing the attractiveness of the e-hailing service while increasing the time drivers spend without customers on board.

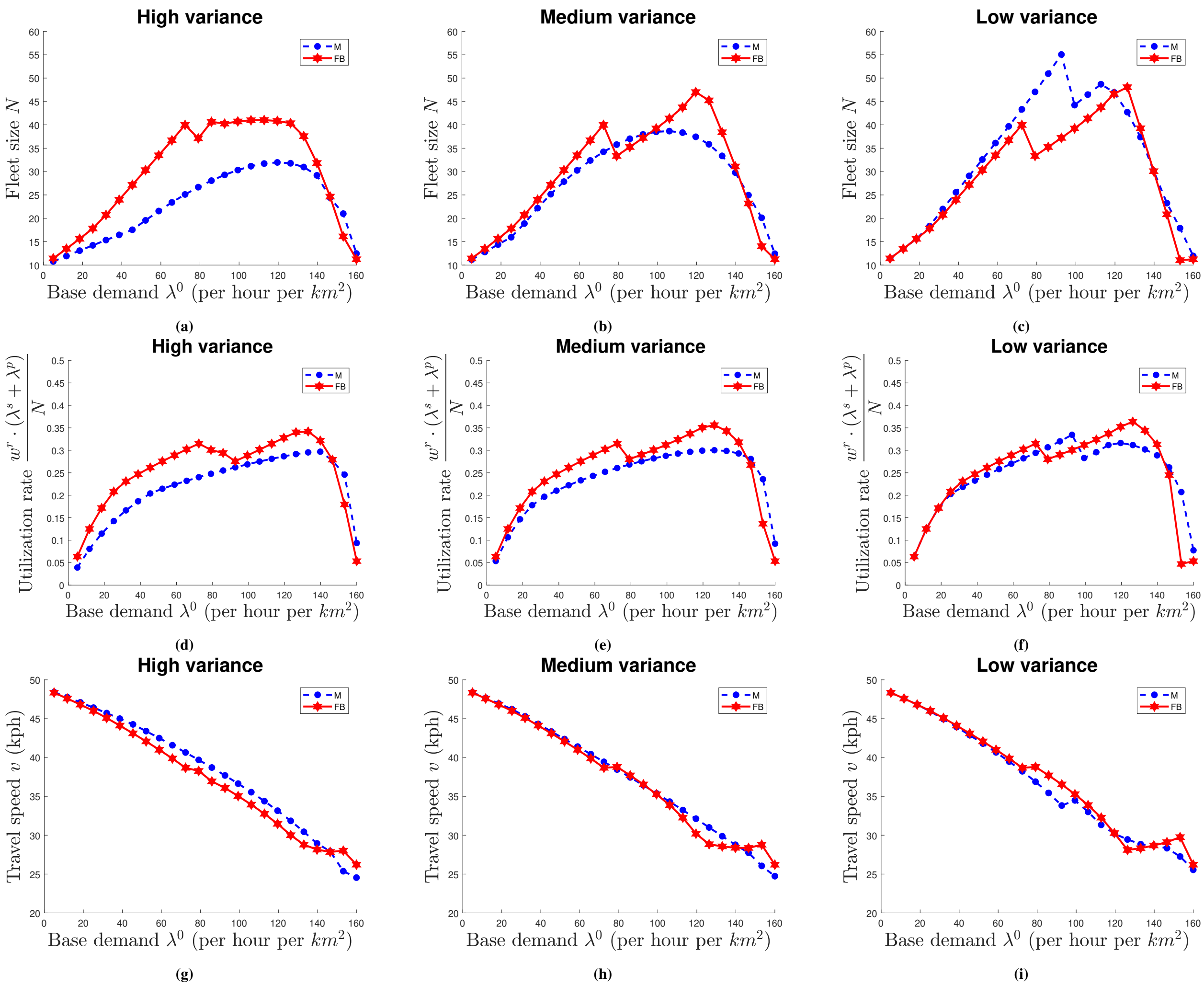


Figure 2.4: *Fleet size, utilization rate and traffic speed comparisons*

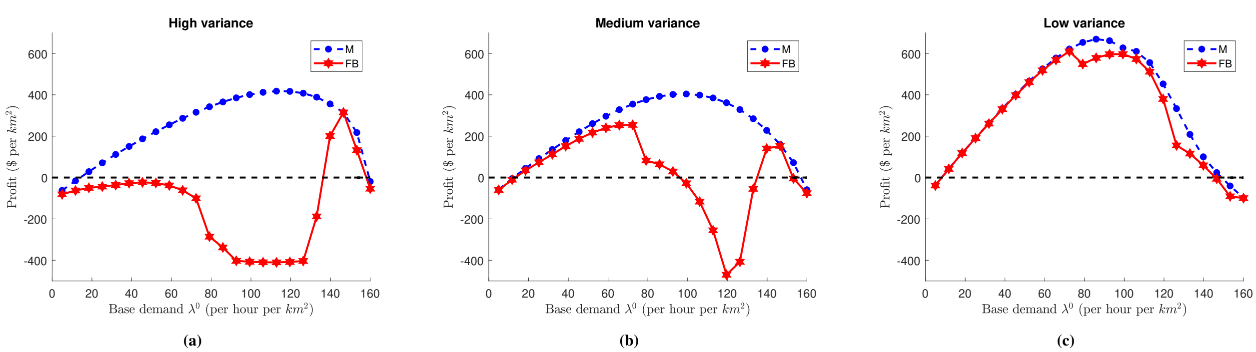


Figure 2.5: *Profit comparisons*

Lastly, we consider the e-hailing company's profits. As predicted in our analysis, under the first-best, the e-hailing platform becomes sustainable in high congestion regimes (**Figure 2.5**). Moreover, as variance in the value of time decreases, the range under which the service is sustainable increases. Intuitively, when variance is high, welfare can be increased mainly through expanding access, so that effects of scale dominate. In that context, average costs and fares decrease with increasing demand, leading to a net loss for the platform. However, when variance is low, demand increases play a lesser role, since if one customers uses the platform, most of them likely also do. The main gains to be made are from efficiency in fleet and demand management, which accrue to the platform under the form of fares.

These numerical examples have a few implications. First, intervention by the planner to reduce congestion might not be necessary, since, as congestion increases, the platform acts similarly to the planner. Second, significant welfare gains can only be made when heterogeneity in the market served by the platform is high. Then, the main contribution from the planner is to increase demand served by expanding access to lower value of time customers. In that context, there may be limited gains to be made from focusing on e-hailing-induced congestion, as pointed out by Tarduno (2021). Rather, applying commission caps with limited tolling should be the preferred regulatory strategy.

2.5.2 Effects of proposed policies

In order to evaluate whether our proposed policies can improve welfare relative to the monopolist solution, we can only consider cases under which the ride-sourcing market is sustainable under the first-best. Therefore, if the first-best is sustainable, it will be our regulatory target. Otherwise, we settle for a second-best in which the monopolist makes some profit. From the previous numerical examples, it is clear that when regulation is needed (i.e. in the high and medium variance cases), one will often have to settle for the second-best. Most importantly, we must determine how the choice of the caps p^s and p^p and of the toll τ^n is made. Our analysis in **Section 2.4** assumed a homogeneous value of time. When the population is heterogeneous in the value of time, selecting the optimal caps and tolls can be modeled as a bilevel program in which the social planner is at the upper-level and seeks to maximize welfare by choosing p^s, p^p and τ^n subject to a profit constraint. We present the problem and our solution method in detail in Appendix A.2.

Upon implementation, our strategy is able to substantially improve welfare (**Figure 2.6**) and results in an outcome close to the second-best. The results suggest that significant welfare gains can be realized from our policy, especially when the welfare gap is significant.

Figure 2.7 shows our chosen caps ($\hat{p}^s$ and $\hat{p}^p$) as well as the commissions implemented by the platform ($\hat{f}^s - \hat{r}$ and $\hat{f}^p - \hat{r}$) in response to the regulation. It is evident that, at most demand levels considered, only one cap is required for the solo service. However, in the highly congested regimes,

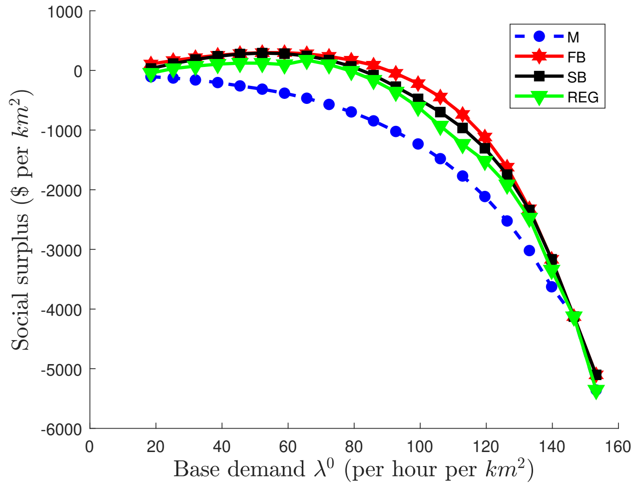
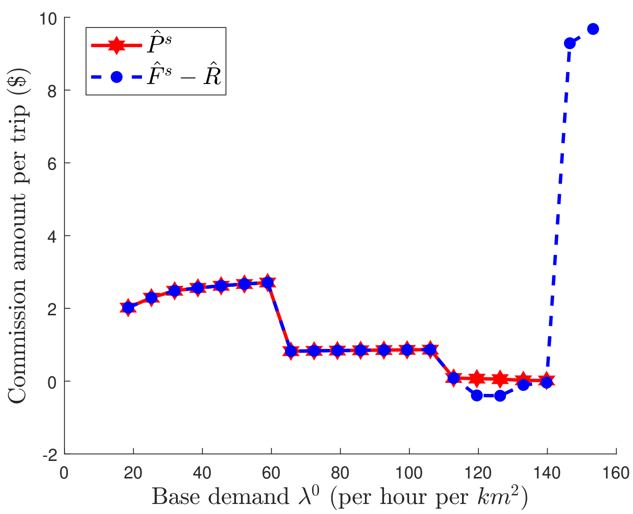
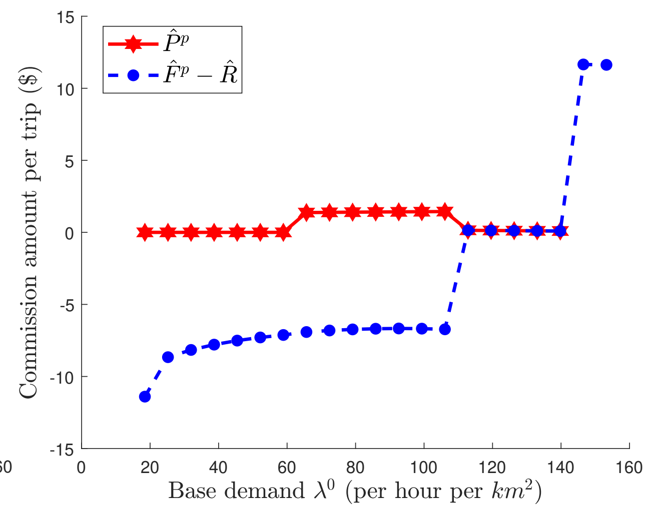


Figure 2.6: *Comparison of regulation results to first-best, second-best and monopoly*

the regulation is not needed, since, as discussed in **Section 2.5.1**, the monopolist's behavior aligns with the regulator's objective. It is interesting to note that, as base demand increases, the cap for the solo service is reduced while that for the pooling service increases.



(a)



(b)

Figure 2.7: *Regulatory and realized commissions*

Finally, **Figure 2.8** shows the regulatory tolls. We note that the tolls first increase with congestion but then decrease. This is in keeping with the fact that the monopolist's objective becomes closer to that of the regulator when congestion becomes very high.

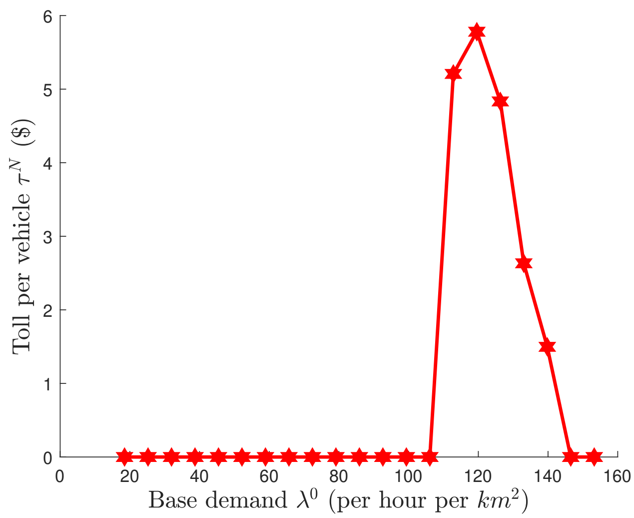


Figure 2.8: *Implemented toll*

2.6 Summary

In this chapter, we present a model of the e-hailing market with the presence of congestion externality, and the integration of the solo and pooling services. In order to derive analytical insights, we then consider a simplified version of our model with a homogeneous value of travel time. Analyzing the market equilibrium under both the first-best and the monopolist, we show that:

- under a socially optimal equilibrium and similar to the taxi market, the e-hailing market may be sustainable when congestion is high;
- a monopoly platform internalizes part of the congestion externality its drivers impose but still employs larger number of vehicles than is socially efficient;
- a regulation coupling a single commission cap and a congestion toll (however small) can replicate any sustainable equilibrium when customers are homogeneous in their value of time;
- in the case that the collection of a toll is impractical, we derive a set of congestion fees to be collected directly from customers of each service. Interestingly, the ratio between the fee for the solo service and the pooling service does not vary linearly with occupancy.

We also briefly apply New York's congestion mitigation policies to our setting and show that, compared to our proposed regulations, they are redundant. In some cases, this redundancy could also potentially lead to inefficiencies. In order to understand how our policy performs in the more realistic setting of heterogeneous value of time, we perform numerical experiments. Our numerical results show that regulatory intervention is only warranted when the population is highly heterogeneous. In those circumstances, however, the main source of inefficiency may not necessarily be congestion and there are limited welfare gains to be made by focusing on that issue. Rather, maximizing demand served would be the best strategy for the regulator. We confirm our intuition by solving a Stackelberg game to choose optimal regulatory caps and tolls. These examples reveal that, when the welfare gap between the unregulated market and the first-best is the highest, it is optimal not to impose a toll but, rather, to impose low commission caps–with lowest caps on the solo service. Moreover, when tolls are applied, their value should decrease with the level of congestion, since the monopolist naturally aligns with the regulator in these cases. Thus, since imposing commission caps is more parsimonious and justifiable with regards to addressing inefficiencies, we favor that approach to regulating the e-hailing market under a monopoly. However, the regulator might still choose to impose a toll on traffic as a whole, rather than singling out e-hailing vehicles.

While our assumption of a monopolistic market might make sense in certain contexts[7], most other markets feature two or more companies competing for customers and drivers. Thus, a few questions may arise in that context. As shown in our analysis of the first-best, marginal cost pricing is not sustainable, except in highly congested instances. Thus, for more than one firm to subsist in a long-term equilibrium, competition must result in a non-efficient pricing pattern and/or significant product differentiation. Whether welfare will be higher than in the monopolist setting and closer to the first-best is, however, unclear. Zha et al. (2016) showed that, in a duopoly setting, welfare might be lower than in the monopoly setting when matching frictions are high or if market size is too small. The inclusion of congestion will likely increase matching frictions, though the extent to which this will degrade profitability is unclear. Additionally, since congestion enhances profitability, one might reasonably ask whether competition for drivers–driven by profit seeking–might worsen traffic conditions at peak times in a manner that is welfare degrading. In-depth analysis of these aspects is left to future work.

Customer characteristics indubitably vary across geographical locations. This leads to more opportunities for product differentiation but also raises the question of the fairness of regulations such as congestion pricing. Indeed, Uber has argued that e-hailing regulations in NYC disproportionately hurt customers from low-income neighborhoods and from areas with poor transit access who disproportionately use their pooling service (Dobbs, 2019). Thus, taking into account these unintended effects might provide an opportunity for spatially differentiated tolling/subsidy strategies that we will investigate. Moreover, such a setting provides an opportunity to analyze the fleet management behavior of the platform as it must contend with the opportunity cost of providing service in one neighborhood as opposed to another. Here again, supplementing our work with an empirical basis might be appropriate and will be done in future work.

[7] Currently Didi Chuxing in some Chinese cities, for example.

E-hailing companies have also been touting their potential to aid and complement public transit, thus making it more accessible. In select cities, for example, users are able to see public transit options alongside UberX and UberPool in the Uber app. Additionally, rides beginning or ending near public transit stations in select cities are now subsidized in an effort to address the first- and last-mile problem. This provides additional opportunities for congestion mitigation and increasing the share of pooled rides on the platform, but also creates additional modeling challenges.

Lastly, we have also taken the background traffic demand to be fixed, thus obviating the possibility that congestion management and social welfare might be better served by encouraging the use and aiding the efficiency of the e-hailing service–rather than discouraging it. Our future work will seek to integrate these substitution effects to provide a more complete picture of the transportation conundrum faced by urban planners.

CHAPTER 3

Competition in the Ride-hailing Market

3.1 Introduction

The previous chapter explored the question of regulating an e-hailing market with a single platform in the presence of congestion externalities. While such discussion can be adequate for certain markets (e.g.: the Chinese market), there exists two or more companies in most US markets. Moreover, both the previous chapter and previous literature (Zha et al., 2016; Buchholz, 2019) points to the fact that, in the presence of competition, welfare may decrease even though the total number of vehicles in the market might increase. Thus, an in-depth exploration of competition in the e-hailing market as well as its effect of congestion and welfare is warranted. Moreover, as mentioned in Chapter 1, competition between and regulation of e-hailing and street-hailing services ought to be considered. In this chapter, we will focus on this latter aspect, since it seems to bear more urgently on policy matters.

3.2 Literature Review

Historically, ride-hailing regulations have been catered to street-hailing services and have spanned three areas: price regulation, entry regulation and quality requirements. Naturally, these regulations have generated a lot of academic and political discussions which are worth exploring when considering how they should evolve with the rise of e-hailing.

Proponents of regulation argue that, in the absence of fare regulations, an equilibrium to the street-hailing market may not exist or may be undesirable due to imperfect information for both customers and drivers (Shreiber, 1975, 1981; Coffman and Shreiber, 1977; Gallick and Sisk, 1987; Cairns and Liston-Heyes, 1996). Indeed, given the spatial nature of the market, the inability of customers to easily collect pricing information and the inability of drivers to easily signal their lower price compared to other drivers, absent regulations, fares might be unnecessarily high or might never stabilize. Evidence also suggests that, at taxi stands with first-in-first-out rules for customers and drivers, bargaining is next to impossible, which results in even higher fares than in the cruising market (Frankena, 1984). Additionally, some advocate that entry restrictions are needed to allow positive profits for the industry and address congestion and pollution externalities (Shreiber, 1975; Arnott, 1996). Absent regulations, the industry's low entry cost would lead to an oversupply of cabs which, absent coordination among drivers, would reduce utilization rates and make taxi operation unsustainable. This oversupply of cabs would also adversely affect traffic and generate more pollution, without either customers or drivers bearing the cost they impose on others. Lastly, there is also an argument that quality requirements in the form of knowledge test, background checks, insurance and vehicle condition and size are needed. Indeed, because of the temporary nature of their interactions with drivers, customers are unable to properly assess the safety and risks associated with their ride before experiencing it.

In general, opponents to regulations argue that price and supply limits lead to higher waiting times and fares than are efficient (Frankena, 1984; Beesley, 1973; Barrett, 2003). This arises in part because these restrictions are set for industry benefits rather than for customers' welfare. These restrictions–coupled with other restrictions such as the inability to provide shared rides–also inhibit the ability of drivers and fleet operators to differentiate their service and thus serve a larger market (Beesley, 1979). Moreover, these price and entry restrictions fail to account for the spatial and temporal nature of the market (Fréchette et al., 2019; Buchholz, 2019). Thus, fares might not reflect the fact that the opportunity cost of a ride depends on its destination, and supply restriction might lead to demand/supply imbalances at different times of day in different geographic locations. Moreover, to address the high prices that arise at taxi stands or in isolated locations where finding a taxi is difficult, the matching process could be tweaked and price caps could be set at these locations instead of widespread industry restrictions (Frankena, 1984). Lastly, externalities could be addressed through tolls and taxes, thus making taxis, their customers and other vehicle users bear the full social cost of their activities.

Evidence from the US and other places are available for either side of the debate, and careful analysis seems to suggest that some regulation might be needed to ensure competitiveness and limit entries in already saturated corners of the industry (Teal and Berglund, 1987; Gaunt, 1995; Dempsey, 1996; Schaller, 2007). With the introduction of e-hailing services, new evidence seems to validate some of the claims for either side. On the one hand, waiting times on the e-hailing services tend to be lower than for street-hailing services. There has also been an increase in the number of rides with e-hailing compared to street-hailing services in some cities, and new services such as pooled rides have been made available. However, there is evidence that both e-hailing (Erhardt et al., 2019; Tarduno, 2021) and street-hailing (Mangrum and Molnar, 2020) contribute to increased congestion levels in cities like NYC and San Francisco. Moreover, the dynamic pricing mechanism used by the platforms to deal with unforeseen surges in demand has occasionally resulted in very high prices, to the dismay of customers.

A number of studies have also investigated the issue of competition in markets with search friction, especially ride-hailing (Zha et al., 2016; Nikzad, 2017; Fréchette et al., 2019; Benjaafar et al., 2020; Zhang and Nie, 2021b). A consensus seems to emerge that extreme competition between these two ride-hailing firms might not necessarily be socially efficient because it may reduce market thickness for each service, thus leading to lower service quality and higher fares than with a single service. More recently, a number of papers have focused on the specific issue of competition between e-hailing and street-hailing (Yu et al., 2019; Daniels and Turcic, 2021; Noh et al., 2021). Yu et al. (2019) considers the effect of regulatory intervention in a market in which e-hailing and street-hailing compete. While their findings seems to validate our premise–better regulations can improve street-hailing's competitiveness against e-hailing–their model does not incorporate one of the crucial differences between the two services: their matching technology and the resulting waiting times. On the other hand, Noh et al. (2021) focuses on the impact of each service's managerial structure (i.e. street-hailing exercises more control on its supply compared to e-hailing) on service quality, prices and welfare. Similarly to Yu et al. (2019), their model does not incorporate the effect of matching technology on the outcome of competition in the industry. Additionally, in the long run, e-hailing drivers' greater flexibility compared to street-hailing drivers should be a minor factor in determining the outcome of competition [1]. Daniels and Turcic (2021) investigates the role of matching technology in determining the outcome of competition between e-hailing and street-hailing. By modeling the matching process, they are able to derive waiting time functions for each service as a function of demand and supply levels. Then, they proceed to show that restricting the service area–as opposed to adopting e-hailing style centralized dispatch–would be a better option for street-hailing to compete against e-hailing.

My work is closest in spirit to that of Daniels and Turcic (2021). I incorporate the differences in matching technology between e-hailing and street-hailing and use the resulting model to derive insights as to the operational settings that favor one service over the other. While Daniels and Turcic (2021) focuses on waiting times, my analyses account for both waiting times and pricing behavior. Additionally, going further than Daniels and Turcic (2021), we evaluate potential regulatory strategies for the ride-hailing industry as a whole and their implications for welfare and congestion.

The rest of this chapter paper is organized as follows. **Section 3.3** presents my main assumptions and our model. Then, in **Section 3.4** I analyze the first-best while in **Section 3.5** I analyze the Nash equilibrium and an integrated monopoly. In **Section 3.6**, I propose and determine sufficient conditions for the market to be regulated parsimoniously. Additionally, in **Section 3.7**, I

[1] Since, in the long run, drivers can decide on which of these two platforms to provide services and, from a managerial perspective, compensation costs can be adjusted accordingly.

extend my model to consider the policy implications of accounting for the effect of ride-hailing on congestion. I conclude in **Section 3.8**.

3.3 Model

We consider a market with a street-hailing and an e-hailing service. It is assumed that customers have access to the prices and waiting times on both services and can choose which one to use by comparing features across both platforms. Drivers also have access to earning information on both services and can choose whether and for which company to provide service. However, no multi-homing is considered in this setting. The e-hailing platform decides the fares and driver earning per ride and earns the difference, i.e., a per-trip commission. The street-hailing company decides its fleet size as well as the leasing fee for its vehicles.

Additionally, we consider the following:

1. The e-hailing platform implements instantaneous matching upon a customer's request with an infinite matching radius. Thus, e-hailing customers experience no (online) matching time (Castillo, 2018; Xu et al., 2019). This tends to be a common practice among e-hailing companies in the US market. Alternatively, some companies (e.g.: DiDi Chuxing in China) implement batch matching that would yield online matching time for customers in addition to the pickup time. We do not consider such a strategy here.

2. Street-hailing customers experience no pickup time. Rather, (physical) matching between customers and drivers occurs as described in Arnott (1996); Chen et al. (2019); Zhang et al. (2019).

3. Each firm only provides a solo service: no pooling is considered.

A description of the main variables used in our model is given in **Table B.1** and a description of relevant parameters is given in **Table B.1**.

3.3.1 Matching, pickup, and delivery

Let w_i^m denote the expected waiting time experienced by customers using service $i \in \{e, s\}$[2]. If service i is an e-hailing company, then w_i^m corresponds to the pickup time experienced by the requesting customer. If, instead, i is a street-hailing company, then w_i^m corresponds to the "matching" time (search frictions) experienced by the requesting customer. Then, we have:

$$w_i^m = \frac{d_i^m}{v}$$

[2] e stands for e-hailing

where d_i^m represents the average distance between a customer and her closest available driver on service i; d^r represents the average trip length, which is given and independent of the service used; and v represents the traffic speed. The average distance d_i^m is a decreasing function of the number of idle drivers n_i^I and is such that:

$$d_i^m = D_i^m(n_i^I) \tag{3.1}$$

where $D_i^m(\cdot)$ is a decreasing and convex function that reflects the matching technology for service i. Then, if we let d^r be the average trip distance, the trip time w^r experienced by customers of either service once they are picked up is given by:

$$w^r = \frac{d^r}{v}$$

3.3.2 Demand

We assume that customers are homogeneous in their value of time β but differ in other aspects that influence choice (e.g., preference for a particular service, inertia and diligence in comparing features across services). Given the distribution of these customer specific attributes, the demand rate λ_i for each service i is given by:

$$\lambda_i = \Lambda_i(\mu_i, \mu_{-i})$$

where $\Lambda_{i,1} < 0$ and $\Lambda_{i,2} > 0$; and μ_i is the average cost of using service i and is such that:

$$\mu_i = f_i + \beta \cdot (w_i^m + w^r)$$

where f_i is the fare on service i.

3.3.3 Supply

When multi-homing is not permitted, drivers must decide which platform to join before beginning service. As such, given the distribution of drivers' reservation costs and their idiosyncratic preferences, the supply of drivers is given by:

$$n_i = S_i(\omega_i, \omega_{-i})$$

where $S_{i,1} > 0$ and $S_{i,2} < 0$; and ω_i is the expected hourly earning on service i.

3.3.4 Equilibrium

At equilibrium, we consider a steady state in the system where the following conservation equation holds as per Little's law:

$$n_i = n_i^I + \lambda_i \cdot (w_i^d + w^r)$$

where w^d represents the pickup time experienced by drivers. For e-hailing drivers, $w_i^d = w_i^m$ and, for street-hailing drivers, $w_i^d = 0$.

We then obtain the system of Equations **3.2a** to **3.2g**:

$$\lambda_i = \Lambda_i(\mu_i, \mu_{-i}) \quad \forall i \in \{s, e\} \tag{3.2a}$$

$$\mu_i = f_i + \beta \cdot (w_i^m + w^r) \quad \forall i \in \{s, e\} \tag{3.2b}$$

$$w_i^m = \frac{d_i^m}{v} \quad \forall i \in \{s, e\} \tag{3.2c}$$

$$w^r = \frac{d^r}{v} \quad \forall i \in \{s, e\} \tag{3.2d}$$

$$d_i^m = D_i^m(n_i^I) \quad \forall i \in \{s, e\} \tag{3.2e}$$

$$n_i = S_i(\omega_i, \omega_{-i}) \quad \forall i \in \{s, e\} \tag{3.2f}$$

$$n_i = n_i^I + \lambda_i \cdot (w_i^d + w^r) \quad \forall i \in \{s, e\} \tag{3.2g}$$

The above system of equations contains 14 equations and 18 unknowns. By specifying a set of 4 exogenous variables and 14 endogenous variables, we can assume that this system defines a continuously differentiable function $G\colon \mathbb{X} \to \mathbb{E}$ where $\mathbb{X} \subseteq \mathbb{R}_+^4$ is a set of exogenous variables and $\mathbb{E} \subseteq \mathbb{R}_+^{14}$ is a set of endogenous variables.

3.4 First-best analysis

In this section, we analyze the first-best to understand when and how regulatory actions can be taken in the ride-hailing market.

3.4.1 Problem equivalence

Before proceeding to analyze the first-best, a simple analysis of firms' independent decisions can help us further simplify the problem by treating e-hailing and street-hailing in a similar fashion.

An e-hailing company maximizes its profits as follows, taking its competitors' decisions as given:

$$\begin{aligned} \pi_e = \max_{\substack{f_e \geq 0, \\ n_e \geq 0, \\ r_e \geq 0}} \quad & (f_e - r_e) \cdot \lambda_e \\ \text{s.t.} \quad & (\omega_e + c_e) \cdot n_e = r_e \cdot \lambda_e \end{aligned} \tag{3.3}$$

where c_e is the (given) per unit time operation cost for drivers who decide to serve on platform e; and r_e is the compensation per ride that drivers earn. The constraint in the above problem indicates that, under free-entry condition and sufficient supply, drivers enter the market until economic profits reach zero. The problem can be rewritten as:

$$\pi_e = \max_{\substack{f_e \geq 0, \\ n_e \geq 0}} \quad f_e \cdot \lambda_e - (\omega_e + c_e) \cdot n_e \tag{3.4}$$

Meanwhile, a street-hailing company solves the following problem:

$$\begin{aligned} \pi_s = \max_{\substack{f_s \geq 0, \\ n_s \geq 0, \\ l_s \geq 0}} \quad & (l_s - c_s^c) \cdot n_s \\ \text{s.t.} \quad & (\omega_s + c_s^d + l_s) \cdot n_s = f_s \cdot \lambda_s \end{aligned} \tag{3.5}$$

where c_s^c is the (given) per unit time cost of vehicle ownership for the company; c_s^d is the (given) cost of operation incurred by drivers; and l_s is the leasing fee that the company charges drivers. By posing $c_s = c_s^c + c_s^d$, it is now possible to rewrite the problem for the street-hailing company as:

$$\pi_s = \max_{\substack{f_s \geq 0, \\ n_s \geq 0}} \quad f_s \cdot \lambda_s - (\omega_s + c_s) \cdot n_s \tag{3.6}$$

Thus, provided that c_s and c_e are properly specified, both street-hailing and e-hailing companies solve mathematically equivalent problems. This greatly simplify the rest of our work.

In our following analyses, both in this and subsequent sections, we assume that the exogenous variables for this system will be the fares f_i and the number of drivers n_i for each company.

3.4.2 Comparison between e-hailing and street-hailing

Using empirical data and a model of the matching technology for both e-hailing and street-hailing, Zhang et al. (2019) show that, in high density settings, waiting times for street-hailing could be lower than those for e-hailing. This is largely attributed to stronger returns to scale for street-hailing. These stronger returns to scale stem from the fact that, by expanding customers' hailing radius, e-hailing also increases inter-customer competition for drivers. Thus, while the radius expansion can lead to lower waiting times for customers in low density areas, it can result in longer waiting times in denser settings. This waiting time advantage of street-hailing in denser settings is corroborated by other researchers, both analytically and empirically (Nie, 2017; Daniels and Turcic, 2021).

In this section, we are interested in examining whether, when considering the generalized cost of service, street-hailing could still possess an advantage over e-hailing. Indeed, lower waiting times could be associated with higher fares and, subsequently, make a service less desirable. Naturally, to capture these trade-offs between higher fares and lower waiting times, we must be able to capture the impact of waiting times on fares. A potential approach to establish that connection would be to consider fares under marginal cost pricing. Indeed, under marginal cost pricing, the fares would capture the waiting time externality that a given customer imposes on others. Moreover, an appropriate comparison between the two services would compare their generalized costs when they operate at their most efficient level. These considerations motivate studying the optimality condition of the first-best problem:

$$W = \max_{\substack{f_i \geq 0, \\ n_i \geq 0}} \quad CS(\mu_s, \mu_e) + \sum_i f_i \cdot \lambda_i - (c_i + \omega_i) \cdot n_i + DS(\omega_1, \omega_2) \tag{FB}$$

where $CS(\cdot, \cdot)$ represents the consumer surplus with $CS_i = -\lambda_i$; and $DS(\cdot, \cdot)$ represents driver surplus with $DS_i = n_i$. We note that, at the first-best, the planner maximizes the joint profit of both companies, not favoring one over the other beyond what efficiency would require. Then, the fare per unit assigned time is given by:

$$f_i = \overline{mc}_i = mc_i \cdot (w_i^d + w^r) \tag{3.7a}$$

$$mc_i = c_i + \omega_i \tag{3.7b}$$

with:

$$mc_s = -\beta \cdot \lambda_s \cdot w_s^{m'} \tag{3.8a}$$

$$mc_e = -\beta \cdot \lambda_e \cdot w_e^{m'} \cdot \frac{1}{1 + \lambda_e \cdot w_e^{m'}} \tag{3.8b}$$

$$w_i^{m'} = \frac{D_i^{m'}(n_i^I)}{v} \tag{3.8c}$$

In the above, $\overline{mc}_i$ represents the marginal service cost for service i i.e. the cost that serving the marginal customer imposes on other users of the service; and mc_i represents the marginal service cost per unit time for service i. Now, let us compare the relative socially optimal costs of using e-hailing and street-hailing. To do this, let us consider a widely adopted form in the literature for the pickup distance (Arnott, 1996; Zhang et al., 2019):

$$D_i^m(n_i^I) = \frac{1}{a_i} \cdot (n_i^I)^{\alpha_i}$$

where a_i is a parameter that captures the efficiency of the meeting process; and α_i is the elasticity parameter with $\alpha_s = -1$ and $\alpha_e = -0.5$.

Thus, using **Equations 3.2b** and **3.8b**, the socially efficient cost of serving a trip can be written as:

$$\mu_s = \beta \cdot w_s^m \cdot \left[\lambda_s \cdot \left(n_s^I\right)^{-1} \cdot w^r + 1 \right]$$
$$\mu_e = \beta \cdot w_e^m \cdot \left[\frac{1}{2} \cdot \lambda_e \cdot \left(n_e^I\right)^{-1} \cdot \left(\frac{w_e^m + w^r}{1 - \frac{1}{a_e \cdot 2 \cdot v} \cdot \lambda_e \cdot \left(n_e^I\right)^{-1.5}} \right) + 1 \right]$$

Then, we obtain **Proposition 3.1**:

Proposition 3.1. *Given a_i, d^r and v, and assuming identical labor costs for both services, there exists supply and demand levels such that street-hailing is more efficient than e-hailing.*

Proof. Street-hailing is more efficient than e-hailing when $\frac{\mu_e}{\mu_s} \leq 1$, assuming identical levels of demand served and identical number of vehicles (i.e $\lambda_e = \lambda_s = \lambda$ and $n_s = n_e = n$). However, it is sufficient to show that $\frac{\mu_s}{\mu_e} \leq 1$ holds when considering identical levels of demand served and identical number of idle drivers ($n_s^I = n_e^I = n^I$). Indeed, when $\lambda_i = \lambda$ and $n_i^I = n^I \; \forall i$, it follows that $n_s < n_e$. If $\frac{\mu_s}{\mu_e} \leq 1$, it implies that it is possible to serve λ with a lower driver cost using street-hailing. Reducing n_2 to n_1 would not alter that result but would increase the customer cost by increasing waiting time for e-hailing customers.

Then, considering $\lambda_i = \lambda$ and $n_i^I = n^I \; \forall i$, it is a matter of algebraic transformations to rewrite $\frac{\mu_s}{\mu_e} \leq 1$ as:

$$-\frac{1}{2 \cdot a_s \cdot v} \cdot (n^I)^{-1} \cdot w^r \cdot \phi^2 - \left[\frac{1}{2 \cdot a_s \cdot v} \cdot (n^I)^{-1} + w^r \cdot \left(\frac{1}{2} - \frac{a_e}{a_s} \cdot (n^I)^{-0.5} \right) \right] \cdot \phi + \left(\frac{a_e}{a_s} \cdot (n^I)^{-0.5} - 1 \right) \leq 0 \tag{3.9}$$

where $\phi = \frac{\lambda}{n^I}$. Now, we have:

- if $n^I \geq \frac{a_e^2}{a_s^2}$, then regardless of the value of ϕ (which is uniquely determined by the level of demand once n^I is set), the above condition is met.

- if $n^I < \frac{a_e^2}{a_s^2}$, the left hand-side of **Equation 3.9** is a concave parabola whose value at $\phi = 0$ is positive and whose smallest root is negative. Thus, if ϕ is greater than the positive root of

Equation 3.9, the condition is met.

□

Proposition 3.1 indicates that if there is a large supply of drivers, or if demand is sufficiently large, then street-hailing will tend to be preferable to e-hailing. This could potentially be explained by stronger increasing returns to scale for street-hailing. A closer examination reveals the underlying mechanism that gives rise to those increasing returns to scale. First, note that even if both services had the same waiting time functions ($\alpha_s = \alpha_e$ and $a_s = a_e$), street-hailing would still be more efficient under certain conditions. This is simply due to the nature of the marginal cost of both services. While the marginal cost of street-hailing is only a function of delivery time, that of e-hailing is a function of both delivery time and pickup time. Indeed, a marginal street-hailing user only monopolizes driving resources for the duration of her trip. In e-hailing, unless reassignments are made, a marginal user also monopolizes a driver for the duration of her pickup time. High density of drivers and customers then has two effects on e-hailing. First, a marginal e-hailing user imposes a pickup time cost on a larger number of users. Second, relative to street-hailing, an increase in both e-hailing supply and demand does not fully translate into an increase in available drivers, since the time dedicated to pickup by the fleet also increases (due to the larger number of people to pickup). Thus, when supply and/or demand are high, street-hailing will have higher market shares than e-hailing if demand is served optimally. It is possible to gain further insights into the relative performance of both services by considering the following:

Corollary 3.1.1. *Given a demand level λ and supply n, there exists $(a_s, a_e, d^r, v) \in \mathbb{R}_+^4$ such that street-hailing is more efficient than e-hailing. In particular:*

- *If $a_e \leq \frac{\sqrt{n^I}}{2} \cdot a_s$, then street-hailing is more efficient regardless of demand level.*
- *If $\frac{\sqrt{n^I}}{2} \cdot a_s < a_e$ and $v \leq \tilde{v}$, then street-hailing is more efficient.*
- *If $\frac{\sqrt{n^I}}{2} \cdot a_s < a_e < \sqrt{n^I} \cdot a_s$, $v > \tilde{v}$ and $d^r \leq \tilde{d}^r$, then street-hailing is more efficient.*
- *If $a_e > \sqrt{n^I} \cdot a_s$, $\tilde{v} < v < \tilde{\tilde{v}}$, and $d^r \leq \tilde{d}^r$, then street-hailing is more efficient.*

In the above:

$$\tilde{v} = \lambda \cdot (n^I)^{-\frac{3}{2}} \cdot \frac{1}{2 \cdot a_e - a_s \cdot \sqrt{n^I}}$$
$$\tilde{\tilde{v}} = \frac{1}{2} \cdot \lambda \cdot (n^I)^{-\frac{3}{2}} \cdot \frac{1}{a_e - a_s \cdot \sqrt{n^I}}$$
$$\tilde{d}^r = v \cdot \frac{n^I}{\lambda} \cdot \frac{1 - \frac{v}{\tilde{v}}}{\frac{v}{\tilde{\tilde{v}}} - 1}$$

Proof. Applying the same principles as in **Proposition 3.1** and rewriting **Equation 3.9** as a quadratic in v, we obtain the following condition for street-hailing dominance:

$$\left(\frac{a_e}{a_s}\cdot(n^I)^{-0.5}-1\right)\cdot v^2-\left[\frac{1}{2\cdot a_s}\cdot(n^I)^{-1}+d^r\cdot\left(\frac{1}{2}-\frac{a_e}{a_s}\cdot(n^I)^{-0.5}\right)\right]\cdot\phi\cdot v-\frac{1}{2\cdot a_s}\cdot(n^I)^{-1}\cdot d^r\cdot\phi^2\leq 0 \tag{3.10}$$

In other words, **Equation 3.10** describes a parabola whose concavity is governed by the relative efficiency $\frac{a_e}{a_s}$ and for which the negative region is determined by the relative efficiency as well as the travel distance d^r. It is then easy to verify that the statements in **Corollary 3.1.1** hold. □

Corollary 3.1.1 indicates that, when traffic speed is low, street-hailing is more efficient. Indeed, low traffic speed implies that e-hailing drivers spend a larger fraction of their time in pickup mode, thus serving less rides [3]. Additionally, settings with low travel distances will also favor street-hailing. Indeed, as trip distance increases, the reduction in available supply depresses the quality of service of street-hailing to a stronger extent than that of e-hailing. For the latter, an increase in travel distance actually reduces the fraction of time drivers spend on pickup, thus reducing the pickup inefficiency. This latter result is consistent with findings from Daniels and Turcic (2021) whose counterfactual show that limiting the service area of street-hailing could help them better compete against e-hailing.

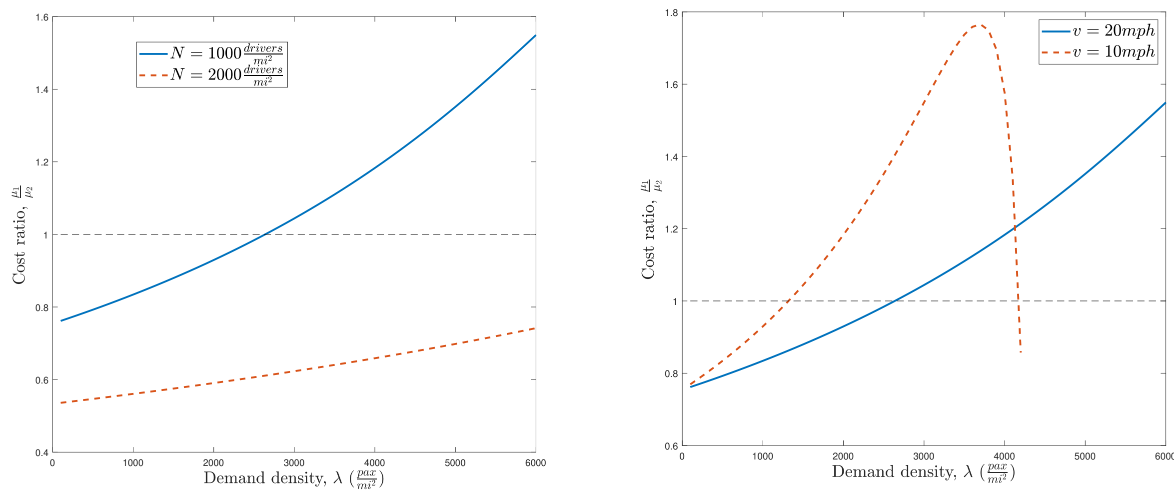
The insights from **Proposition 3.1** and **Corollary 3.1.1** are illustrated in **Figure 3.1** which presents the cost ratio between e-hailing and street-hailing, $\frac{\mu_s}{\mu_e}$ as a function of demand density.

These results illustrate that, depending on market characteristics, one alternative may be preferable to the other. We might wonder under which circumstances it will be optimal to use both services. Indeed, following the literature on ride-hailing, when two companies operate, increased market frictions due to demand/supply splitting could actually reduce welfare (Zha et al., 2016; Zhang and Nie, 2021b). Thus, it would seem that in dense urban settings with large available supply, two services could be supported. In settings with low driver supply or low demand, operating an e-hailing service might be preferable to operating both. When both services operate, **Proposition 3.1** and **Corollary 3.1.1** simply indicate the circumstances under which street-hailing would dominate the market.

Additionally, as noted by Arnott (1996); Yang et al. (2014b); Zha et al. (2016), at the first-best in a single firm environment, ride-hailing services must be subsidized. When two firms operate, the same finding holds. Indeed, using **Equations 3.2g** and **5.4**, it is possible to show that first-best profits are as follows:

$$\pi_i=-(c_i+\omega_i)\cdot n_i^I<0 \tag{3.11}$$

[3]Note that this result has nothing to do with different congestion externality from both services, since we have not yet included the effect of vehicles on congestion



(a) $v = 20$ *mph*

(b) $n = 1000$ $\frac{drivers}{mi^2}$

Figure 3.1: *Equilibrium variables under the first-best and the Nash equilibrium. In (3.1a), we see that when supply increases, the cost of using street-hailing significantly drops. Moreover, economies of scale are stronger for street-hailing, resulting in a better cost advantage for street-hailing as demand and supply increase. However, when supply is constant, increased demand can eventually cause street-hailing to become more expensive than e-hailing unless demand exceeds a certain threshold and traffic speed is low. Past that threshold, e-hailing drivers' utilization becomes more inefficient (3.1b). For all demand levels below that threshold, the cost curves in (3.1b) also capture the effect of increased travel distance* d^r *(as opposed to reduced travel speed) on the relative cost of each service. Due to effects of scale, longer trips have a stronger negative impact on the quality of service of the street-hailing service.*

In essence, the fare only covers the time a driver is assigned to serving a customer but not their idle time, resulting in a loss for both ride-hailing companies. Thus, we now take a look at the second-best.

3.4.3 Second-best

Here, the social planner maximizes welfare subject to profit constraints for both e-hailing and street-hailing. The problem is as follows:

$$\begin{aligned} W = \max_{\substack{f_i \geq 0,\\ n_i \geq 0}} \quad & CS(\mu_s, \mu_e) + \sum_i f_i \cdot \lambda_i - (c_i + \omega_i) \cdot n_i + DS(\omega_s, \omega_e) \\ \text{s.t.} \quad & \pi_i \geq \bar{p}_i \cdot \lambda_i \end{aligned} \tag{SB}$$

where $\bar{p}_i$ is the reservation profit per ride served (or commission) for company i. The FONC for this problem is given by:

$$f_i = mc_i \cdot (w_i^d + w^r) - \frac{\delta_i}{1+\delta_i} \cdot \frac{f_i}{\epsilon_i^{dm}} + \frac{\delta_{-i}}{1+\delta_i} \cdot \frac{f_{-i}}{\epsilon_i^{dm}} \cdot \frac{\lambda_{-i}}{\lambda_i} \cdot \frac{\epsilon_{-ii}}{\epsilon_{-i-i}} + \bar{p}_i \cdot \frac{\delta_i}{1+\delta_i} \tag{3.12a}$$

$$c_i + \omega_i = mc_i - \frac{\delta_i}{1+\delta_i} \cdot \frac{\omega_i}{\eta_i^{dm}} + \frac{\delta_{-i}}{1+\delta_i} \cdot \frac{\omega_{-i}}{\eta_i^{dm}} \cdot \frac{n_{-i}}{n_i} \cdot \frac{\eta_{-ii}}{\eta_{-i-i}} \tag{3.12b}$$

where $\delta_i \geq 0$ is the Lagrangian multiplier associated with the profit constraint for firm i. Essentially, the profit constraint introduces a wedge between the fare and the marginal cost of service (on the customer side) and a wedge between drivers' earnings and the marginal cost. In essence, the second-best realizes a transfer from customers to companies and drivers. We also note that additional earnings beyond marginal costs on service i are tied to the level of service on the other platform. Thus, in essence, the consolidated firm[4] endogenizes competitive externalities that arise under the Nash game studied earlier.

It is also worth noting that **Equation 3.12** defines, given the δ_i, a solution on the Pareto frontier that joins the joint monopoly between e-hailing and street-hailing to the first-best.

Further, it is interesting to consider the case when $\delta_e = 0$, i.e., the e-hailing industry is not subject to a profit constraint but operates close to marginal cost. This scenario is somewhat close to the current status quo in the industry. Indeed, because of the vast amount of their financial resources, e-hailing companies have been operating at a loss, subsidizing customers while offering multiple incentives to drivers. Thus, in reality, fares might be below marginal costs while driver incentives bring their earning in line with their marginal benefit. In this context, it is easy to see that street-hailing vehicles are at a natural disadvantage: at most levels of demand, the fare required to use taxi services would be higher than marginal costs (to guarantee profitability of the industry). Thus, the higher market shares for e-hailing that we observe might not necessarily be the result of their efficiency, but rather the outcome of inefficient pricing at two levels: higher than efficient prices for street-hailing and lower than efficient prices for e-hailing.

In light of the above, we might ask whether removing restrictions on street-hailing could provide a better outcome for the street-hailing industry as well as the overall impacts of such a move on welfare.

3.5 Unregulated market

In this section, we focus on an unregulated market featuring both street-hailing and e-hailing. On the one hand, we consider the case when the two industries compete for customers and drivers and try to understand the implications for the profitability of street-hailing. Then, we consider the case of a more integrated operation between the two services and whether and when it is preferable to competition.

[4] Recall that joint profits are maximized

3.5.1 Nash equilibrium

We define the Nash equilibrium as a set of fares $\{f_s^*, f_e^*\}$ and fleet sizes $\{n_s^*, n_e^*\}$ such that, $\forall i \in \{s, e\}$:

$$(f_i^*, n_i^*) = \underset{\substack{f_i \geq 0, \\ n_i \geq 0}}{\arg\max} \quad f_i \cdot \lambda_i(f_i, f_{-i}^*, n_i, n_{-i}^*) - [\omega_i(f_i, f_{-i}^*, n_i, n_{-i}^*) + c_i] \cdot n_i \tag{NE}$$

The FONC resulting from equation NE is as follows:

$$f_i = mc_i \cdot (w_i^d + w^r) - \frac{f_i}{\epsilon_i^{ne}} \tag{3.13a}$$

$$mc_i = c_i + \omega_i \cdot \left[1 + \frac{\eta_{-i-i}}{\eta_{ii} \cdot \eta_{-i-i} - \eta_{-ii} \cdot \eta_{i-i}}\right] \tag{3.13b}$$

with:

$$\epsilon_s^{ne} = \epsilon_{ss} - \frac{\lambda_e}{f_e} \cdot \frac{\beta \cdot (w_e^m + w^r) \cdot w_e^{m'}}{1 + w_e^{m'} \cdot \left[1 + \beta \cdot (w_e^m + w^r) \cdot \Lambda_{ee}\right]} \cdot \epsilon_{se} \cdot \epsilon_{es} < 0 \tag{3.14a}$$

$$\epsilon_e^{ne} = \epsilon_{ee} - \frac{\lambda_s}{f_s} \cdot \frac{\beta \cdot w^r \cdot w_s^{m'}}{1 + \beta \cdot w_s^{m'} \cdot w^r \cdot \Lambda_{ss}} \cdot \epsilon_{se} \cdot \epsilon_{es} < 0 \tag{3.14b}$$

In the above, ϵ_{ij} represents the elasticity of demand on service i with respect to the cost of service j; and η_{ij} is the elasticity of supply on service i with respect to the hourly wage on service j. Comparing **Equations 3.13** and **5.4**, it is expedient to note that the action of the planner contributes to increasing both the demand and supply for each service compared to their levels in the unregulated case.

By examining **Equations 3.14a** and **3.14b**, we note that the markup on one service is a function of friction imposed by a marginal customer on the other service: the higher the marginal friction on the other service, the higher the markup charged by a given firm, $\frac{1}{|\epsilon_i^{ne}|}$. In particular, because of deadheading trips on the e-hailing platform, marginal frictions are higher for e-hailing when demand is high: the street-hailing platform will be able to command higher prices relative to its marginal cost than e-hailing ($|\epsilon_s^{ne}| < |\epsilon_e^{ne}|$). This insight could also be gotten at by considering that, when demand is high, the marginal cost for e-hailing will be higher than that for street-hailing (again, due to increased friction). Thus, the ability of e-hailing to raise fares above marginal cost is lower compared to that of street-hailing. *Thus, in the absence of supply and price constraints and in high density areas, street-hailing stands to benefit more than e-hailing, from a pricing perspective*.

It is also useful to compare the above results to those under a monopoly (which operates a single e-hailing or street-hailing platform with the same market size), which we recall below from

Zha et al. (2016):

$$f^{sm} = mc \cdot (w^d + w^r) - \frac{f^{sm}}{\epsilon^{sm}} \tag{3.15a}$$

$$mc = c^{sm} + \omega^{sm} \cdot \left[1 + \frac{1}{\eta^{sm}}\right] \tag{3.15b}$$

where the superscript sm indicates that the relevant functions and quantity are that faced by a single service monopoly.

In principle, we expect $|\epsilon_{ii}| \geq |\epsilon^{sm}|$ since, given other options, users become more sensitive to price increases. However, as indicated earlier, the market power term, $\frac{f_i}{\epsilon_i^c}$ under the Nash equilibrium also depends on friction on the competing platform: the larger the friction (low n_{-i}^I or high w^r), the lower, in absolute term, ϵ_i^c. If friction is sufficiently high, higher prices may result under the NE compared to the single firm, single service case: $|\epsilon_i^c| \leq |\epsilon^{sm}|$. This is consistent with numerical experiments from Zha et al. (2016); Zhang and Nie (2021b) and we show in Appendix B.2 that frictions are the main drivers of this phenomenon.

Comparing **Equations 3.13b** and **3.15b**, we note that competition on the driver side can lead to a substantial reduction of supply under the duopoly. Indeed, when the cross-elasticities of substitution are high (high η_{i-i} and η_{-ii}), unless demand served is high, the incentive for hiring supply can decrease, leading to lower supply overall. This corroborates numerical examples from Zhang and Nie (2021b) and indicates that, multi-homing on the driver side might actually result in a lower driver pool and lower quality of service.

3.5.2 Integrated monopoly

In this section, we seek to provide insights into recent developments in cities like NYC where Uber will begin listing street-hail drivers on its platform (Rana, 2022). In the context of this paper, we will assume that a single firm is able to integrate and manage pricing and compensation for both services at the same time. Importantly, we do not incorporate the fact that street-hailing drivers could also be matched to e-hailing customers. Indeed, in our context, we are concerned with efficiency gains/losses that might result from joint pricing and compensation. The case of differentiated supply for e-hailing in the presence of street-hailing will be modeled and discussed in a subsequent work. With this in mind, the profit maximization problem is given by:

$$\pi^{dm} = \max_{\substack{f_i \geq 0, \\ n_i \geq 0}} \quad \sum_i f_i \cdot \lambda_i - (\omega_i + c_i) \cdot n_i \tag{DM}$$

The FONC for equation DM yields:

$$f_i = mc_i \cdot (w_i^d + w^r) - \frac{f_i}{\epsilon_i^{dm}} + \frac{f_{-i}}{\epsilon_i^{dm}} \cdot \frac{\lambda_{-i}}{\lambda_i} \cdot \frac{\epsilon_{-ii}}{\epsilon_{-i-i}} \tag{3.16a}$$

$$c_i + \omega_i = mc_i - \frac{\omega_i}{\eta_i^{dm}} + \frac{\omega_{-i}}{\eta_i^{dm}} \cdot \frac{n_{-i}}{n_i} \cdot \frac{\eta_{-ii}}{\eta_{-i-i}} \tag{3.16b}$$

where

$$\epsilon_i^{dm} = \epsilon_{ii} - \frac{\epsilon_{-ii} \cdot \epsilon_{i-i}}{\epsilon_{-i-i}} \tag{3.17a}$$

$$\eta_i^{dm} = \eta_{ii} - \frac{\eta_{-ii} \cdot \eta_{i-i}}{\eta_{-i-i}} \tag{3.17b}$$

and dm indicates that quantities correspond to those faced by a dual service monopoly. Comparing **Equations 3.13** and **3.16**, we note that integrating the platforms in an unregulated market has two effects. On the one hand, fares for both services increase in tandem since the absence of competition increases the monopolist's market power. On the other hand, the monopolist internalizes the friction-related competitive externality that arises in the Nash equilibrium. For given levels of supply and demands, the direct effect of this internalization can be to either increase or decrease prices for a given service. Indeed, let us compare ϵ_i^{ne} and ϵ_i^{dm}. We have:

$$\epsilon_s^{ne} - \epsilon_s^{dm} = -\frac{\lambda_e}{f_e} \cdot \epsilon_{se} \cdot \epsilon_{es} \cdot \left[\frac{\beta \cdot (w_e^m + w^r) \cdot w_e^{m'}}{1 + w_e^{m'} \cdot \left[1 + \beta \cdot (w_e^m + w^r) \cdot \Lambda_{ee}\right]} - \frac{1}{\Lambda_{ee}} \right] \tag{3.18a}$$

$$\epsilon_e^{ne} - \epsilon_e^{dm} = -\frac{\lambda_s}{f_s} \cdot \epsilon_{se} \cdot \epsilon_{es} \cdot \left[\frac{\beta \cdot w^r \cdot w_s^{m'}}{1 + \beta \cdot w_s^{m'} \cdot w^r \cdot \Lambda_{ss}} - \frac{1}{\Lambda_{ss}} \right] \tag{3.18b}$$

When frictions are high (low n_i^I and high w^r), the quantities in **Equation 3.18** are positive: each individual firm's markup is higher than what it would be under the integrated monopoly. If friction is sufficiently high, then, the price under the Nash equilibrium might even exceed that under the integrated monopoly, and this despite the latter's greater market power. Thus, it cannot be ruled out that integration might actually reduce prices and be profitable for both customers and producers [5].

3.6 Policy analysis

Following our discussions in Sections 3.4 and 3.5 it is indubitable that the regulatory environment must change. This can be done by regulating both fares and entry in both markets. However, such

[5] Again, we show in Appendix B.2 that such an issue does not arise in a friction-less environment.

an approach would be rather heavy-handed and would remove an important feature of ride-hailing's performance: its flexibility. In the following section, we will investigate possible alternatives and their outcomes.

3.6.1 Commission cap on e-hailing, regulation adjustment for street-hailing

Following Zha et al. (2016) and Vignon et al. (2021), it is known that a cap on the commission of an e-hailing monopoly can replicate the second-best when supply is perfectly elastic. In the presence of competition between e-hailing and street-hailing, would such a cap be effective, assuming fares and supply regulations for street-hailing are properly adjusted? Under a commission cap, the problem for the e-hailing platform becomes:

$$\begin{aligned} \pi_e = \max_{\substack{f_e \geq 0, \\ n_e \geq 0, \\ r_e \geq 0}} \quad & (f_e - r_e) \cdot \lambda_e \\ \text{s.t.} \quad & (\omega_e + c_e) \cdot n_e = r_e \cdot \lambda_e, \\ & (f_e - r_e) \cdot \lambda_e \leq \bar{p}_e \cdot \lambda_e \end{aligned}$$

where $\bar{p}_e$ is the commission cap. The problem can be recast as:

$$\begin{aligned} \pi_2 = \max_{\substack{f_e \geq 0, \\ n_e \geq 0}} \quad & \bar{p}_e \cdot \lambda_e \\ \text{s.t.} \quad & (\omega_e + c_e) \cdot n_e \geq (f_e - \bar{p}_e) \cdot \lambda_e \end{aligned} \tag{EH-REG}$$

Thus, the e-hailing platform must maximize demand served on its platform. Can such as regulation induce the platform to choose the socially optimal f_e and n_e? Suppose not. Clearly, the set of decision variables that would be chosen by the platform must be such that $\lambda_e^{dm} > \lambda_e^{sb}$. Such an increase over the socially-optimal market share would have to come at the expense of either the street-hailing service or the outside option. However, those customers switching from either of these services would see their utility increase relative to the second-best[6]. Therefore we must ask what happens to those customers who remain on the street-hailing service. Since both the fares and the fleet size are set, it follows that they will experience lower waiting times. Thus, their utility also increases. This implies that the second-best is actually not Pareto efficient, a contradiction.

Thus, the current regulation is optimal and provides a simple avenue for policymakers to make the ride-hailing market more efficient. Are there other, potentially simpler regulations able to achieve the same objective?

[6]This is why they switch.

3.6.2 Consolidation and commission cap regulation

Since the social planner maximizes the joint producer profit, we might ask whether consolidating the ride-hailing industry and then regulating the resulting monopoly might be a better approach. This approach would lead to a major restructuring of the ride-hailing industry but would also simplify the patchwork of regulation that governs the industry. It would also leave a number of operational features such as pricing and wages into the hands of service operators which could provide for better and more synergistic operation between the two services.

Thus, consider the following problem:

$$\begin{aligned}
\pi_M = \max_{\substack{f_i \geq 0, \\ n_i \geq 0, \\ l_s, r_e \geq 0}} \quad & (l_s - c_s^c) \cdot n_s + (f_e - r_e) \cdot \lambda_e \\
\text{s.t.} \quad & (\omega_s + c_s^d + l_s) \cdot n_s = f_s \cdot \lambda_s, \\
& (\omega_e + c_e) \cdot n_e = r_e \cdot \lambda_e, \\
& (l_s - c_s^c) \cdot n_s \leq \bar{p}_s \cdot \lambda_s, \\
& (f_e - r_e) \cdot \lambda_e \leq \bar{p}_e \cdot \lambda_e
\end{aligned}$$

First, it is easy to note that, unless $\bar{p}_s = \bar{p}_e = \bar{p}$, the regulation might lead to the highest cap service being over-utilized following the monopolist's bid to maximize profits. Thus, for this regulation to be effective, both caps must be identical to ensure that the monopoly favors a service over the other only on efficiency grounds. Then, we can rewrite the above as:

$$\begin{aligned}
\pi_M = \max_{\substack{f_i \geq 0, \\ n_i \geq 0}} \quad & \bar{p} \cdot \sum_i \lambda_i \\
\text{s.t.} \quad & (\omega_i + c_i) \cdot n_i \geq (f_i - \bar{p}) \cdot \lambda_i
\end{aligned} \tag{MONO-REG}$$

Naturally, the monopolist will choose f_i and n_i such that $\lambda_s^{dm} + \lambda_e^{dm} \geq \lambda_s^{sb} + \lambda_e^{sb}$. If $\lambda_s^{dm} + \lambda_e^{dm} > \lambda_s^{sb} + \lambda_e^{sb}$, it could mean that the monopolist is charging lower than second-best prices. Since this behavior must be profit maximizing, it follows that the reduction in fares is also accompanied by a reduction in earnings for either one or both groups of drivers. Thus, essentially, the monopolist would realize a surplus transfer from drivers to consumers in order to increase its profits. When supply is perfectly elastic so that all drivers are identical and drivers have a reservation wage $\underline{\omega}$, such a profitable move from the monopolist does not arise since driver surplus is null and any reduction in that surplus would result in a loss of supply.

On the other hand, a profitable deviation of the monopoly might be to increase driver earnings which would in turn reduce waiting times and result in higher fares for customers but, most importantly, increase the number of customers for the platform[7]. However, if such an approach were preferable to the second-best, it would mean that every agent' surplus increases, which would contradict the Pareto optimality of the second-best. Thus, when supply is homogeneous, consolidating the industry and imposing an identical commission cap is not only welfare improving but can replicate the second-best. *Following Uber's move of opening its app to street-hail drivers, regulating the industry might be simpler and more straightforward than in the case of outright competition.*

When supply is heterogeneous, the distortion created from the second-best may or may not be significant: this becomes an empirical question.

3.6.3 Nash game with commission cap

Would it be possible to keep both industries separate while applying a commission cap? This would avoid the regulatory headache that would arise from trying to determine how stakes in a consolidated ride-hailing company would be allocated among current actors in the ride-hailing industry. Moreover, the consolidated setting is somewhat restrictive since only identical caps can be imposed. Under a commission cap regulation of two competing firms, we would have:

$$\begin{aligned} \pi_i = \max_{\substack{f_i \geq 0,\\ n_i \geq 0}} \quad & f_i \cdot \lambda_i - (c_i + \omega_i) \cdot n_i \\ \text{s.t.} \quad & (\omega_i + c_i) \cdot n_i \geq (f_i - \bar{p}_i) \cdot \lambda_i \end{aligned} \tag{NE-REG}$$

Here, it is not straightforward to demonstrate that neither company has an incentive to deviate from the second-best. Indeed, either company could increase its wage rate and/or decrease its prices to increase its market share at the expense of the other company. However, we are able to show that, so long as the company is sustainable under the second-best, the regulation can achieve the desired outcome. This can be shown in the manner of Vignon et al. (2021). Indeed, let θ_i be the Lagrangian multiplier associated with the commission cap constraint in equation NE-REG. Then, assuming the regulation replicates the second-best, we must have:

$$\theta_i = \frac{f_i^{ne} - f_i^{sb}}{f_i^{ne} - (c_i + \omega_i) \cdot \frac{n_i}{\lambda_i}} \geq 0 \tag{3.19}$$

where f_i^{ne} and f_i^{sb} are the fare formulae derived in **Equations 3.7a** and **3.13a** evaluated at the targeted second-best. Now, at the second-best, it must be that $f_i^{ne} \geq f_i^{sb}$. Otherwise, regulatory intervention is not warranted since it increases prices while there are no externalities. If the second-best is sustainable, then it must also be that $f_i^{sb} \geq (c_i + \omega_i) \cdot \frac{n_i}{\lambda_i}$. Thus, necessarily, $\theta_i \geq 0$. Thus,

[7] Otherwise, profits would not increase.

provided both companies make positive profits at the second-best, the commission cap regulation can regulate the duopoly ride-hailing market.

3.7 Congestion and competition

As mentioned in **Section 3.2**, the rise of e-hailing has been associated with an increase in congestion in major urban centers. Thus, regulators have been looking into ways to address the issue by imposing congestion fees, tolls and minimum fleet utilization rate requirements. As pointed out in Vignon et al. (2021) and in Zhang and Nie (2021a), the latter set of regulations are either redundant or detrimental to welfare. Rather, as shown by Zhang and Nie (2021a), the imposition of tolls on either trips or e-hailing drivers present the best opportunity for regulators to address the issue of congestion. Moreover, as pointed out by Vignon et al. (2021) and Xu et al. (2017), in a monopoly setting, when congestion increases, the monopolist and the social planner tend to behave similarly: they both look to mitigate the negative impact of congestion. Thus, regulatory intervention in that context might not be needed. However, we might wonder whether such a reasoning holds in the context of competition within the ride-hailing industry. Lastly, should competition create significant congestion issues, we must determine whether and to which extent tolls should be differentiated between e-hailing and street-hailing companies.

To answer these questions, we must first extend our model in **Equation 3.2** to incorporate background traffic and congestion.

Assuming that background traffic trips originate at a (given) rate λ^b with an average trip length of d^{rb}, we have:

$$w^{rb} = \frac{d^{rb}}{v} \tag{3.20a}$$

$$n^b = \lambda^b \cdot w^{rb} \tag{3.20b}$$

where w^{rb} represents the average travel time of background traffic; and β^b is the (homogeneous) value of time of background travellers. Then, following the network macroscopic fundamental diagram approach (Geroliminis and Daganzo, 2008), it is possible to describe the average traffic speed v using the accumulation of vehicles in the network, which is the sum of the number of vehicles for each service n_i and the number of background vehicles n^b:

$$v = V(n^b + \sum_i n_i) \tag{3.21}$$

with $V'(\cdot) < 0$.

We show in Appendix B.3 that, under the Nash game, unlike the integrated monopolist which internalizes all but the congestion externality it imposes on background travellers, ride-hailing companies competing against each other benefit from the fact that congestion hurts their opponent (Appendix B.3.1). Thus, they do not fully internalize their congestion externality. This leads to more drivers on the road than in the monopoly case and can have deleterious effects on congestion. Under the first-best, because the planner regulates a consolidated ride-hailing company, the ”competitive” externality is internalized (Appendix B.3.2). More interestingly, from the point of view of the planner, all vehicles impose the same externality on traffic regardless of their status–background vehicle, e-hailing vehicle or street-hailing vehicle. Thus, not only should ride-hailing vehicles be tolled, but so should background traffic vehicles[8].

When it comes to regulatory actions, imposing appropriate tolls and commission caps can readily achieve the first-best (or the second-best). Indeed, when congestion is present, the social planner essentially maximizes demand served after accounting for the higher (social) cost of operating a vehicle. Thus, once an appropriate toll is set, it is straightforward to verify that, given the commission cap, our insights from **Section 3.6** carry over. Most importantly, however, despite congestion, it is possible to regulate the market, regardless of its structure, using only commission caps (Appendix B.3.3). If, however, regulatory authorities decide to impose a toll, e-hailing might end up bearing a higher toll per unit time than street-hailing ($\bar{\tau}_e - \bar{\tau}_s \geq 0$), as shown in **Figure 3.2**. This strategy is consistent with, for example, NYC's surcharge structure which charges e-hailing trips $\$0.25$ higher than street-hailing trips (TLC, 2022).

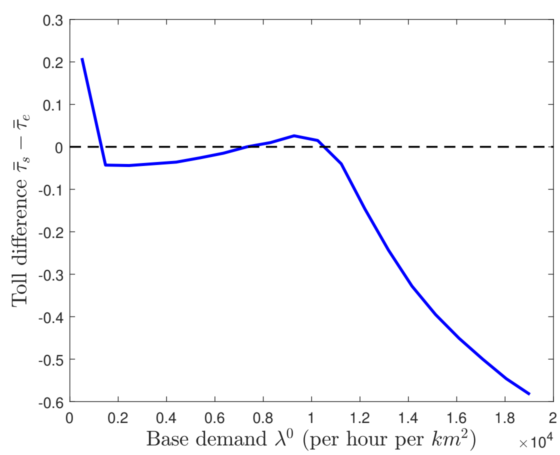


Figure 3.2: *Toll difference between street-hailing and e-hailing*

We illustrate our findings in numerical examples whose results are shown in **Figure 3.3**. Details on the numerical examples are given in Appendix B.4. As shown in Figures 3.3a and 3.3b, the commission cap regulation can significantly improve welfare in both the monopoly and duopoly cases. However, the manner in which the cap achieves that objective differs in both settings. In the

[8]This is shown by making background traffic elastic

monopoly case, the effect of the cap is to increase demand in low density markets (**Figure 3.3c**). Indeed, in these markets, the monopolist's decisions tend to restrict demand with higher than optimal prices. As density increases, the behavior of the monopolist starts to mirror that of the planner, thus resulting in improved welfare (despite a mild increase in congestion relative to the second-best). Under a duopoly, the same demand and traffic patterns compared to the monopoly occur when density is low (Figures 3.3d and 3.3f). However, as density increases, competition between the two services leads to a significant reduction in traffic speed. In that context, by forcing each company to improve its efficiency, the commission cap also contributes to improving traffic speed while increasing consumer welfare.

3.8 Conclusion

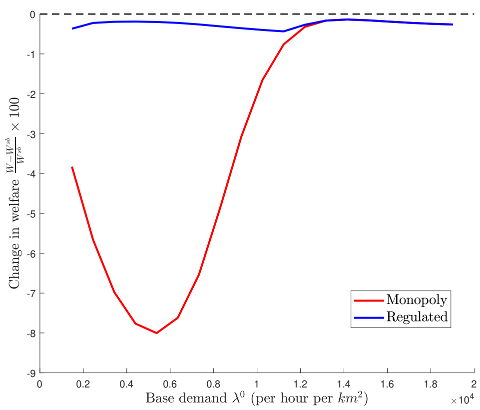
In this chapter, we have sought to inform the development of policies for the ride-hailing industry in the age of uberization. To this effect, we presented a model of competition in the ride-hailing industry and analyzed the impact of current and alternative policies to regulate that market. Some of our key findings are as follows:

- we analytically show that, in denser settings, or when trip distances are low or traffic speed is low, the socially-efficient cost of street-hailing will be lower than that of e-hailing. Thus, in cities like NYC, there is room for a more expanded role of street-hailing in serving the market;

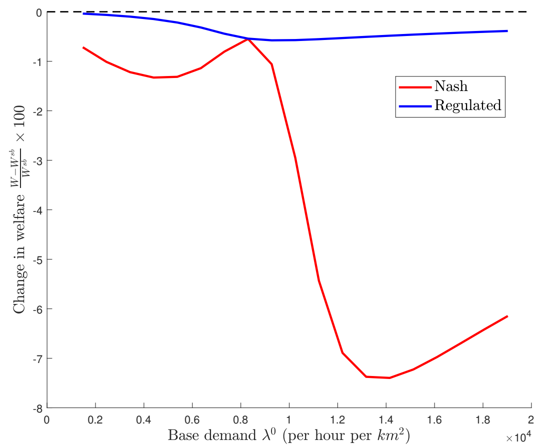
- additionally, we show that in these settings and barring any supply or fare restrictions, the street-hailing industry can have greater market power and thus, better fend for itself. Thus, the industry should seek to relax their current supply restrictions as opposed to trying to curtail e-hailing;

- we also show that, despite the potential benefits of relaxed regulation on street-hailing for that industry, unchecked competition between e-hailing and street-hailing would still result in higher prices and higher congestion than is socially efficient. This latter effect of deregulation on congestion does not, however, arise when the pricing and compensation for both platforms is managed by a single platform;

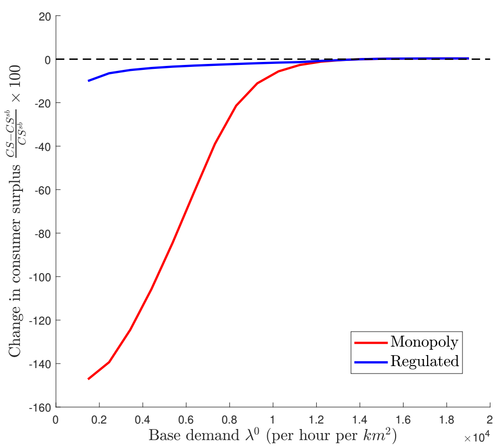
- lastly, we show that, by imposing a commission cap on either an integrated company or the two competing services, both consumer welfare and congestion levels will improve. Especially, under certain assumptions, such regulation can readily achieve the socially optimal configuration. This demonstrates the effectiveness of commission cap regulation, as it is effective even in congested and competitive settings.



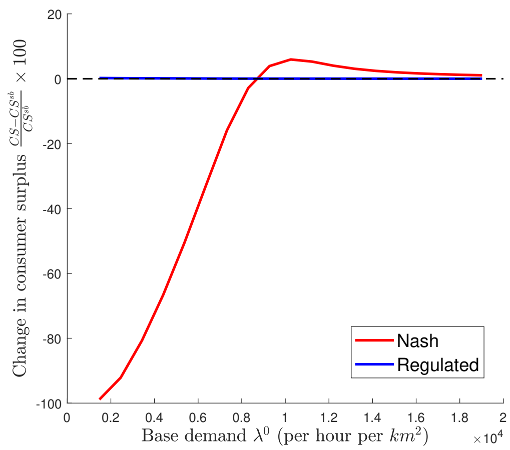
(a) *Welfare under integrated monopoly*



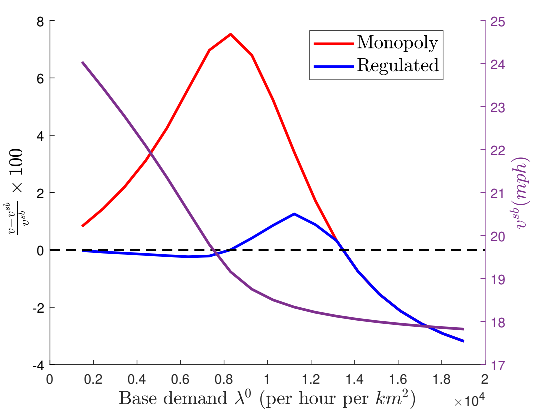
(b) *Welfare under Nash equilibrium*



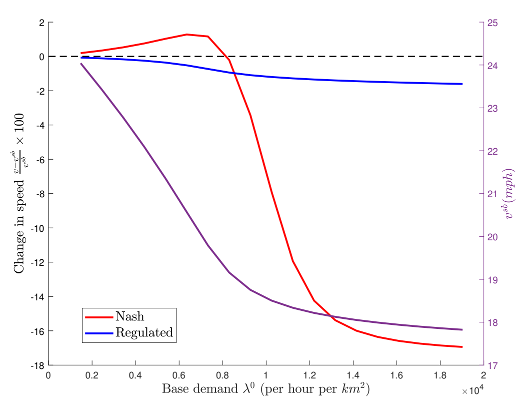
(c) *Consumer surplus under integrated monopoly*



(d) *Consumer surplus under Nash equilibrium*



(e) *Traffic speed under integrated monopoly*



(f) *Traffic speed under Nash equilibrium*

Figure 3.3: *Effect of commission cap regulation under both an integrated monopoly and a Nash game. The commission cap improves welfare and contributes to reducing congestion in both settings. As demand increases and the system becomes more congested under the second-best, the monopolist's behavior leads to slightly more congestion than under the second-best. Thus, the commission cap regulation mostly serves to increase consumer welfare (**Figure** 3.3c). However, in the Nash game, competition can have have dramatic effects on congestion. In that context, the commission cap can reduce congestion (**Figure** 3.3f) while obtaining socially efficient levels of demand (**Figure** 3.3d).*

In our analysis, we have not considered spatially heterogeneous markets. Indeed, we have shown that both services have different efficiency advantages depending on the characteristics of demand (density and distance of trips requested) and of traffic. In a spatially heterogeneous context, we might wonder whether an equilibrium in which the two services do not directly compete with each other but operate in markets in which they have a competitive advantage is socially efficient. We might also wonder whether a commission cap regulation could achieve the socially efficient outcome in a setting of spatial competition. We also have not considered the potential efficiency gains or losses that might come from allowing street-hailing drivers to be matched to street-hailing customers. We will propose a model appropriate for exploring that question. Lastly, our present work will be further enhanced by bringing data to our model to better inform policymakers and help answer practical questions such as the value of an appropriate commission cap in NYC or the distributional effects of such a policy.

Part II

Infrastructure in the Age of Automation

CHAPTER 4

Introduction

4.1 Motivation

By analyzing the investment strategies of e-hailing firms, it becomes apparent that the current state of the ride-hailing industry–in which drivers are major actors–is transitory. Indeed, Uber and Lyft have heavily invested in developing autonomous vehicle technology. For example, Uber's initial public offering documents reveal that, from 2016 to 2018, the company spent close to $\$29$ millions per month on research and development for automated vehicles (Chai, 2019). These investments point to a larger trend in the transportation sector toward automation. Indeed, traditional automakers and other institutional investors have dedicated billions of dollars–$\$80$ billions from 2014 to 2018–to the development of autonomous vehicles technology (Karsten, 2017; Efrati, Amir, 2020). More recently, Amazon acquired self-driving car company Zoox for close to $\$1$ billion (Weise and Griffith, 2020). These investment patterns underscore investors' beliefs that, despite their negative impact on some industries, automated driving technology will increase productivity and profitability by reducing the social and human cost of driving (Clements and Kockelman, 2017).

However, to date, fully automated vehicle technology has failed to materialize (only Level 2 has been commercialized thus far by companies like Tesla and General Motors). Moreover, the enormous amount spent on research and development suggests that significant technological hurdles must be overcome before the automated age of transportation–thus discouraging investors and companies alike (Efrati, Amir, 2020; Metz and Griffith, 2020). Indeed, loosely speaking, drivers perform three tasks when driving: perception, planning and control. Across the industry the effort has been centered on enhancing vehicles to perform all the above tasks, which has been difficult: AVs' perception abilities are still sensitive to weather and lighting conditions (Zhu et al., 2017; Van Brummelen et al., 2018); *a priori* vehicle localization and mapping is not, as of yet, robust to infrastructure changes and simultaneous localization and mapping is subject to perception challenges (Van Brummelen et al., 2018); and algorithms for planning and decision making (e.g., for lane changing) are computationally burdensome for vehicles, thereby limiting their applicability to real-time decision making (Katrakazas et al., 2015; González et al., 2016; Dixit et al., 2018; Schwarting et al., 2018). Additionally, while AV adoption is low, AVs in traffic mixed with conventional vehicles may compromise traffic stream stability and throughput (Seo and Asakura, 2017; Luo et al., 2018). While these are not insurmountable challenges, public tolerance for errors and mistakes may be thin, as demonstrated by the aftermath of recent deadly AV crashes and the reported reservation of a non-negligible segment of the population towards driving automation (Ge et al., 2017). Thus, some investors are starting to question their commitment to AV technology since it is unclear when they will be able to recoup their investment or whether the increased spending is worth the potential gains: there simply is no clear timeline for the deployment and adoption of fully automated vehicles as companies seek to minimize safety risks (Metz and Griffith, 2020).

In contrast to this vehicle-centric approach, researchers have realized that placing some sensors and algorithms on the infrastructure may be a more effective way to enable automated driving. Recent developments primarily focus on the design, modeling and assessment of sensor networks to aid with vehicle perception (Rebsamen et al., 2012; Jun and Markel, 2017; Leone et al., 2017; Bieshaar et al., 2017; Eilbrecht et al., 2017; Reitberger et al., 2018; Jayaweera et al., 2019; Kong, 2020). Others have envisioned a more infrastructure-centric approach in which sensors and algorithms are placed on the infrastructure to perform the tasks of perception and planning, achieving cooperative perception and driving (Gopalswamy and Rathinam, 2018). Note that infrastructure-enabled automation is not a new concept, as the Automated Highway Systems, investigated and demonstrated in the 1990s (Tan et al., 1998; Godbole et al., 1996), can be considered as an early attempt of such an infrastructure-driven approach for driving automation.

Many believe that neither the vehicle- nor the infrastructure-centric approach will prevail in the future (Li et al., 2020). Rather, an infrastructure-vehicle cooperative approach for enabling automated driving would emerge. In this approach, both the level of automation in the future fleet and the level of digitalization in the future infrastructure will be heterogeneous. Similar to the SAE classification on the AVs, Carreras et al. (2018) recently proposed a scheme to classify the readiness of road infrastructure to support and guide AVs . In their classification, at Level A and B the infrastructure will support cooperative driving and perception while at Level C all dynamic and infrastructure information will be provided in digital form to AVs; conventional infrastructure will be at Level D and E. With infrastructure-vehicle cooperation, a Level 3 AV may achieve full automation on a Level A infrastructure. Moreover, the liability associated with automated driving can be shared among OEMs, infrastructure providers, and/or a third-party player, which may substantially accelerate the diffusion of the AV technology. Additionally, the digitalization of infrastructure could provide revenue generating opportunities for road operators. Indeed, smart infrastructure will make the provision of digital services to users as well as the monetization of traffic data easier. Worldwide, governments and supra-governmental agencies are seeing the potential for connectivity and automation to revolutionize transportation systems, especially infrastructure. In the US, the Federal Highway Administration (FHWA)'s CARMA program seeks to encourage and accelerate the research and development of cooperative driving automation (**?**). In Europe, the European Road Transport Advisory Council (ERTRAC), a structure that seeks to encourage cooperation and investment in critical road transportation innovation, supports multiple projects related to infrastructure connectivity and cooperative systems (ERTRAC, 2019). However, despite recent investments in smart road technology in the US (Nordrum, 2018; Calvert, 2020), smart infrastructure investment lags behind the investment in vehicle technology and very few roads are equipped with the necessary digital technology for vehicle-infrastructure communication and cooperation. This could be the consequence of the lack of regulatory coordination in the US–with rules regarding automation and infrastructure development varying across states–and the lack of coordination between automakers and road operators, which forces automakers to rely heavily on automation technology (Ge et al., 2017). This could also be due to a lack of infrastructure funding in the US. However, with the bipartisan push for infrastructure spending in recent years, we might soon enough be faced with the issue of what to spend the money on rather than whether to spend it. In contrast, in China, the question is to determine where to invest. Indeed, the market for smart cities solutions, grown out of the central government's decades long development plan, is estimated at $\$1.1$ trillions. This includes, among other things, investment in automated vehicles and smart transportation infrastructure (Atha et al., 2020).

4.2 Contribution and Outline

The aim of the second part of this dissertation is to lay the groundwork for this vehicle-infrastructure cooperative approach. In Chapter 5, to make a policy and economic case for this vision, we build an analytical model that captures the interactions between different actors of the transportation industry: automakers, road operators and customers. We model both automakers' investment in the research and development of automation technology and road operators' investment decisions in infrastructure sensors and algorithms. We also model customers' choices between different vehicle offerings and their travel patterns across the road network. This stylized model allows us to glean two main insights. First, we show that, from a societal perspective, the vehicle-cooperative approach is superior to the vehicle-centric approach. Second, we show that, when left on their own, automakers and road operators make suboptimal investment decisions. On the one hand, automakers might aim for higher than optimal automation levels because of their inability to rely on infrastructure for assistance. On the other, road operators, unable to capture sufficient returns on their investment from customers, shy away from equipping their infrastructure.

Then, in Chapter 6, we turn to the issue of liability and safety and investigate how they are affected in our proposed automated mobility market. Modeling liability as an insurance contract between automakers and ISSPs, on one hand, and their customers, on the other, we investigate who will bear the burden of AV-involved accidents, from a societal perspective. Then, we show that, when accounting for safety, it is likely that only investment in infrastructure will reduce vehicle ownership costs. We also extent our model to include speed selection and mixed traffic. This extension forms the building blocks for future work on this topic.

CHAPTER 5

Economic Analysis of a Vehicle-Infrastructure Cooperative Approach

5.1 Introduction

The aim of the present chapter is twofold. First, we present a modeling framework to examine whether this vehicle-infrastructure cooperative approach for enabling automated driving makes economical sense. Specifically, we investigate, from a societal perspective, the optimal allocation of investment between the on-board and infrastructure-based sensors. We show that, at the social optimum, the heterogeneous provision of digitalization and automation naturally arises from the heterogeneity in vehicle and road usage. Then, to understand why there is little digitalization investment compared to automation investment, we analyze the outcome of strategic interactions between OEMs and infrastructure operators/third-party providers. Our model especially highlights the negative effects of a lack of coordination between auto manufacturers and road operators on vehicle automation and infrastructure digitalization spending. To the best of our knowledge, this study is the first economic study on the joint provision of vehicle and infrastructure technology for enabling automated driving. Our results offer insights on infrastructure-assisted automated driving and provide both the public and private sectors with additional avenues for cooperation in developing and deploying smart infrastructure. The rest of the paper is organized as follows. **Section 5.2** presents our model setting. Then, in **Section 5.3**, both the social optimum and the Nash equilibrium are derived and compared. We present in **Section 5.4** a numerical example to illustrate our analysis, and then conclude in **Section 5.5**.

5.2 Model

For the convenience of readers, frequently used notations are listed in **Table 5.1**.

Notation	Description
x_i	Vehicle automation level for users of type i
d_i	Number of users of type i
U_i	Utility of users of type i
p	Price per unit of automation level
z_k	Digitalization level for road group k
V_k	Traffic volume on road group k
l_k	Total road length for road group k
v_{ik}	Vehicle miles travelled (VMT) per day by users of type i on road group k
τ_k	fee per mile travelled on road k
π^c	Manufacturer profit
π^r	Service provider profit

(a) *Frequently used variables*

Notation	Description
v_i	Daily VMT by users of type i
κ_c	Amortization parameter for vehicle purchase
κ_r	Amortization parameter for road investment

(b) *Frequently used parameters*

Table 5.1: *Frequently used notations*

We consider a setting with vehicle-infrastructure cooperative deployment of automated driving where sensors, edge-computing devices and intelligence can either reside on the vehicle or infrastructure side to perform various driving tasks. A car manufacturer produces vehicles of various levels of automation and prices them differently. A private automation service provider or ISSP like Cavnue equips roads, whose usage varies, with various types of sensors and devices to assist vehicles in sensing, perception planning and maneuver. To finance the digitalization of these roads, the ISSP will charge a service fee. Additionally, the ISSP will also benefit from the data collected from users' digital footprint on her roads. Such an infrastructure-vehicle cooperation may yield significant benefits for users: reduction in driving opportunity cost, increased safety, etc. The level of benefits will depend on the combination of the automation level of the vehicles that the users are riding and the digitalization level of the roads their vehicle is on. Thus, users are faced with two choices. On one hand, they must decide the automation level of the vehicle they purchase based on the car manufacturer's offerings. On the other hand, they must decide which roads to use to complete their trips based on the digitalization choices of the ISSP.

In our model, users are divided into I user groups based on their vehicle-miles travelled (VMT). While we could instead consider heterogeneity along other dimensions (such as willingness to pay), considering VMT heterogeneity allows us to directly connect customers' decisions regarding automation to road usage since the latter is the main channel through which ISSPs earn a profit from digitalization. Thus, within each group i, users are identical in all aspects, including in their

VMT v_i. However, users' VMT differ across groups. The number of users in group i is d_i. In making their vehicle purchase decisions, users essentially decide the automation level of their car x_i based on the price per unit of automation p_i. Additionally, they take into account the allocation of their VMT across K different groups of road infrastructure equipped with a digitalization level z_k and of a total road length l_k. Using a given road k results in users paying a service fee τ_k per mile. We treat x_i and z_k as continuous variables in $[x_{min}, x_{max}]$ and $[z_{min}, z_{max}]$ respectively.

Users' travel benefit is captured by a function $f_k(x_i, z_k, V_k, v_{ik})$ where v_{ik} denotes the amount driven by a user of group i on road k,and $V_k = \frac{\sum_i v_{ik} \cdot d_i}{l_k}$ denotes the average traffic volume on road k. Thus, a user from group i chooses her automation level x_i and travel pattern $\boldsymbol{v}_i = \{v_{ik}\}_k$ to maximize her utility U_i given by:

$$U_i = \sum_k [f_k(x_i, z_k, V_k, v_{ik}) - \tau_k \cdot v_{ik}] - \kappa_c \cdot p_i \cdot x_i \tag{5.1}$$

where v_{ik} is daily VMT of user i on road group k; and κ_c is a term that amortizes vehicle purchase cost to daily costs. Using $f_{k,i}$ to denote the derivative of f_k with respect to its i^{th} argument while $f_{k,ij}$ denotes the cross partial derivative of f_k with respect to its i^{th} and j^{th} arguments, we make the following assumptions.

Assumption 5.1. *Our assumptions on utility are as follows:*

A1.1 *Utility increases with automation and digitalization:* $f_{k,1},\ f_{k,2}\ > 0$.

A1.2 *Users' utility increases with their amount of travel:* $f_{k,4}\ > 0$.

A1.3 *Congestion decreases a given user's utility:* $f_{k,3}\ < 0$.

A1.4 *Utility is concave in automation, digitalization and individual miles travelled:* $f_{k,11},\ f_{k,22},\ f_{k,44} < 0$.

A1.5 *Utility is concave in travel volume:* $f_{k,33} < 0$.

Assumptions **A1.1** and **A1.3** are readily understood. Assumption **A1.2** captures the fact that users derive a positive benefit from travelling (be it for leisure, work, shopping etc). Note that travel's intrinsic purpose is to complete tasks that improve utility. Assumptions **A1.4** and **A1.5** ensure concavity of the user maximization problem and are intuitive when we consider decreasing marginal utility of consumption.

Vehicles for different user groups are manufactured by a profit-maximizing firm that decides the price per unit of automation p_i for vehicles it produces. More importantly, the firm must decide how much to invest to expand its production capabilities to meet the demand for automation from each user group. We specify the manufacturer's profit function as follows:

$$\pi^c = \sum_i p_i \cdot x_i \cdot d_i - c_c(x_i, d_i) \tag{5.2}$$

where $c_c(\cdot, \cdot)$ is the automation-related manufacturing and R&D costs for a vehicle as a function of its automation level and the number of units produced.

Assumption 5.2. *Our assumptions on the manufacturer's cost function are as follows:*

A2.1 *The cost function is increasing in automation levels and in the quantity of vehicles manufactured:* $c_{c,1},\ c_{c,2} > 0$.

A2.2 *The cost function is convex in automation level:* $c_{c,11} > 0$.

A2.3 *The cost function is concave in the quantity of vehicles manufactured:* $c_{c,22} < 0$.

A2.2 indicates that achieving higher levels of automation becomes increasingly costly for the manufacturer. As detailed in **Section 5.1**, this forms one of the basis for our present inquiry. **A2.3** indicates economies of scale in the manufacture of vehicles.

Lastly, we consider a profit-maximizing ISSP who is interested in digitizing roads to achieve cooperative perception, planning and control of AVs. Note that this private ISSP does not necessarily own these roads. Instead, it partners with the road owner, who is likely a public agency, and is responsible for constructing and maintaining the digital infrastructure. This ISSP thus decides how much to invest to equip each road group with digitalization level z_k. Additionally, she decides the service fee per mile τ_{ik} on each of her roads for each user group. More importantly, this digital infrastructure operator is able to harness some additional benefits for each mile driven on its road via, e.g., the revenue from data monetization and advertising. This non-pricing benefit per mile can be captured by a function $b_r(x_i, z_k, V_k)$ and the profit function for the service provider is given by:

$$\pi^r = \sum_k \sum_i d_i \cdot [\tau_{ik} + b_r(x_i, z_k, V_k)] \cdot v_{ik} - \kappa_r \cdot \sum_k c_r(z_k, l_k) \tag{5.3}$$

where κ_r is a term that amortizes the investment cost to daily costs, and $c_r(\cdot, \cdot)$ captures the digitalization-related investment and maintenance costs per road mile as a function of digitalization level. Our functional form assumptions are given below:

Assumption 5.3. *Our assumptions on the ISSP's cost and benefit functions are as follows:*

A3.1 *The non-pricing benefit function is increasing in automation and digitalization levels:* $b_{r,1},\ b_{r,2} > 0$.

A3.2 *The non-pricing benefit function is increasing in travel volume:* $b_{r,3} > 0$.

A3.3 *The non-pricing benefit function is concave in automation and digitalization levels:* $b_{r,11}, b_{r,22} < 0$.

A3.4 *The cost function is increasing and convex in digitalization levels:* $c_{r,1}, c_{r,11} < 0$.

In essence, higher digitalization and automation allow the ISSP to collect and provide more information to aid in maintenance, data monetization and other services (Assumption **A3.1**). Moreover, the company benefits from higher usage on its roads since this leads to more data collected for monetization purposes (Assumption **A3.2**). Buried in that latter assumption is also that the contribution of road usage to maintenance cost is negligible or always lower than its contribution to the non-pricing benefit. Thus, in essence, $b_r(\cdot,\cdot,\cdot)$ could be thought of as the *net* pricing-benefit. Lastly, the higher the digitalization level, the higher the installation and maintenance costs (Assumption **A3.4**). Indeed, sensors and digital infrastructure will require constant monitoring to ensure their proper operation and reduce the risk of cyber-attacks and other related issues.

5.3 Equilibrium analysis

5.3.1 Social optimum

In this section, we consider the case in which a social planner maximizes social surplus by choosing automation and digitalization levels, in addition to users' travel patterns. The social surplus maximization problem is given by:

$$\begin{aligned}
\max_{x_i, z_k, V_k, v_{ik}} \quad & \sum_i d_i \cdot \sum_k [f_k(x_i, z_k, V_k, v_{ik}) + b_r(x_i, z_k, V_k) \cdot v_{ik}] - \kappa_c \cdot \sum_i c_c(x_i, d_i) - \kappa_r \cdot \sum_k c_r(z_k, l_k) \\
\text{s.t.} \quad & \sum_k v_{ik} = v_i \quad \forall i \in I \quad \text{(VMT constraint for users of group } i\text{)}, \\
& V_k \cdot l_k = \sum_i v_{ik} \cdot d_i \quad \forall k \in K \quad \text{(Flow conservation constraint on} k\text{)}, \\
& v_{ik} \geq 0 \quad \forall k \in K,\ i \in I \quad \text{(Flow positivity constraint)}
\end{aligned} \tag{W}$$

where the social surplus is the sum of consumers' utilities, manufacturer's profits and service provider's profits. At optimality, assuming that $f_k(\cdot,\cdot,\cdot,\cdot)$ and $b_r(\cdot,\cdot,\cdot)$ are strictly concave so that the optimum is an interior point, we obtain:

$$\sum_k \left[f_{k,1}(x_i, z_k, V_k, v_{ik}) + b_{r,1}(x_i, z_k, V_k) \cdot v_{ik} \right] \cdot d_i = \kappa_c \cdot c_{c,1}(x_i, d_i) \quad \forall i \in I \tag{5.4a}$$

$$\sum_i \left[f_{k,2}(x_i, z_k, V_k, v_{ik}) + b_{r,2}(x_i, z_k, V_k) \cdot v_{ik} \right] \cdot d_i = \kappa_r \cdot c_{r,1}(z_k, l_k) \quad \forall k \in K \tag{5.4b}$$

$$\alpha_k + b_r(x_i, z_k, V_k) + f_{k,4}(x_i, z_k, V_k, v_{ik}) - \frac{\gamma_i}{d_i} \leq 0 \quad \forall k \in K, \forall i \in I \tag{5.4c}$$

$$v_{ik} \cdot \left[\alpha_k + b_r(x_i, z_k, V_k) + f_{k,4}(x_i, z_k, V_k, v_{ik}) - \frac{\gamma_i}{d_i} \right] = 0 \quad \forall k \in K, \forall i \in I \tag{5.4d}$$

$$\sum_i \left[f_{k,3}(x_i, z_k, V_k, v_{ik}) + b_{r,3}(x_i, z_k, V_k) \cdot v_{ik} \right] \cdot d_i = \alpha_k \cdot l_k \quad \forall k \in K \tag{5.4e}$$

$$v_{ik} \geq 0 \quad \forall k \in K, \forall i \in I \tag{5.4f}$$

where γ_i is the Lagrangian multiplier associated with the i^{th} VMT constraint; α_k is the Lagrangian multiplier associated with the k^{th} flow conservation constraint. Equations **5.4c** to **5.4e** indicate that, for each user group, the marginal benefit per mile of each used road group is equal, and is more than or equal to the marginal benefit per mile of non-used road groups. From **Equation 5.4a**, the marginal social benefit of automation for users of type i must equal the social marginal cost of providing these users with automation x_i. From **Equation 5.4b**, the marginal social benefit of digitalization for roads of type k must equal the social marginal cost of equipping these roads with digitalization z_k. In other words, allocating some resources to the infrastructure is socially optimal under the assumption of strict concavity of the benefit functions and strict convexity of the cost functions.

This suggests that an infrastructure-vehicle cooperative approach to automated driving deserves more attention. As expected, the levels of automation and digitalization for user and road groups will be determined by equalizing their social marginal cost to their social marginal benefit. Thus, under a set of constraints (budgetary and political etc.), automation and digitalization technologies with a higher marginal return should be given priority. Moreover, because of different VMT and volume distributions and cost functions, the equilibrium will result in a heterogeneous provision of both automation and digitalization. As pointed out by previous research, those with higher VMT will likely benefit from and desire higher levels of automation (Hardman et al., 2019; Hardman, 2021). We now proceed to investigate strategic interactions between automakers and service providers and how such strategic interactions affect the allocation of resources.

5.3.2 Generalized Nash equilibrium

We now explore a case in which the car manufacturer and the service provider act independently from each other, and model it as a noncooperative simultaneous game. We choose to model these interactions as a Generalized Nash Equilibrium problem (GNEP)[1]. An alternative could be to consider a leader-follower game in which the automaker is the leader and the ISSP is the follower. The rationale for this alternative would be that, while vehicles can operate without digitalized infrastructure, the reverse is not true. However, because of the premise of our work–namely, that reaching full vehicle-automation might be infeasible or too costly to society–vehicles are dependent on infrastructure digitalization in our setting, leading to a chicken-and-egg problem. Therefore, imposing the precedence structure inherent in a leader-follower game might not be appropriate. Let $\mathbf{v}_i = \{v_{ik}\}$. $\{\mathbf{p}^*, \mathbf{x}^*, \mathbf{z}^*, \boldsymbol{\tau}^*\}$ constitutes a Generalized Nash Equilibrium (GNE) if there exists $\{\mathbf{v}_i^*\}$ such that:

$$\begin{aligned} \{\mathbf{x}^*, \mathbf{p}^*, \{\mathbf{v}_i^*\}\} = \quad & \underset{x_i, p_i, V_k, v_{ik} \geq 0}{\arg\max} && \sum_i p_i \cdot x_i \cdot d_i - c_c(x_i, d_i) \\ & \text{s.t.} && \{\mathbf{p}, \mathbf{x}, \mathbf{z}^*, \boldsymbol{\tau}^*, \{\mathbf{v}_i\}\} \in X(\mathbf{v}) \end{aligned} \tag{5.5}$$

and

$$\begin{aligned} \{\mathbf{z}^*, \boldsymbol{\tau}^*, \{\mathbf{v}_i^*\}\} = \quad & \underset{z_k, \tau_k, V_k, v_{ik} \geq 0}{\arg\max} && \sum_k [\tau_k \cdot V_k \cdot l_k + \sum_i b_r(x_i, z_k, V_k) \cdot v_{ik} \cdot d_i] - \kappa_r \cdot c_r(z_k, l_k) \\ & \text{s.t.} && \{\mathbf{p}^*, \mathbf{x}^*, \mathbf{z}, \boldsymbol{\tau}, \{\mathbf{v}_i\}\} \in X(\mathbf{v}) \end{aligned} \tag{5.6}$$

where $\mathbf{v} = \{v_i\}$; and $X(\mathbf{v})$ characterizes the set of all $\{\mathbf{p}, \mathbf{x}, \mathbf{z}, \boldsymbol{\tau}, \{\mathbf{v}_i\}\}$ for which $\{\mathbf{x}, \{\mathbf{v}_i\}\}$ is users' response to $\{\mathbf{p},\ \mathbf{z},\ \boldsymbol{\tau}\}$ due to utility maximization. The utility maximization problem for a user of type i is given by:

$$\begin{aligned} \max_{x_i, v_{ik}} \quad & U_i = \sum_k [f_k(x_i, z_k, V_k, v_{ik}) - \tau_k \cdot v_{ik}] - \kappa_c \cdot p_i \cdot x_i \\ \text{s.t.} \quad & \sum_k v_{ik} = v_i, \\ & v_{ik} \geq 0 \quad \forall k \in K \end{aligned} \tag{UM}$$

UM assumes that when a user i routes themselves selfishly in the network, they take the traffic volume V_k as given. This assumption is implicitly made in the literature of traffic network equilibrium analysis (e.g., Sheffi, 1984) and is particularly valid when the number of users is

[1] For distinction between Nash Equilibrium and Generalized Nash Equilibrium problems, please see Facchinei and Kanzow (2010). The basic distinction is that the feasibility set of a given player is affected by the strategies of other players in a Generalized Nash Equilibrium problem but not in a Nash Equilibrium problem.

sufficiently large. The first-order necessary conditions (FONC) of UM yield:

$$\sum_k f_{k,1}(x_i, z_k, V_k, v_{ik}) \cdot v_{ik} = \kappa_c \cdot p_i \quad \forall i \in I \quad \text{(Pricing constraint)} \tag{5.7a}$$

$$v_{ik} \cdot [f_{k,4}(x_i, z_k, V_k, v_{ik}) - \tau_k - \mu_i] = 0 \quad \forall k \in K, \quad \forall i \in I \quad \text{(Complementarity)} \tag{5.7b}$$

$$f_{k,4}(x_i, z_k, V_k, v_{ik}) - \tau_k - \mu_i \leq 0 \quad \forall k \in K, \quad \forall i \in I \quad \text{(Link travel cost condition)} \tag{5.7c}$$

$$\sum_k v_{ik} = v_i \quad \forall i \in I \quad \text{(VMT constraint for user } i\text{)} \tag{5.7d}$$

$$v_{ik} \geq 0 \quad \forall k \in K, \quad \forall i \in I \tag{5.7e}$$

where Equations **5.7a** to **5.7d** indicate that the benefit of all road groups used by user i is equal and greater than the benefit of all other unused road groups. μ_i is the Lagrangian multiplier associated with the i^{th} VMT constraint, capturing the net benefit from miles travelled for a user of type i. Then, X is the set of all $\{\mathbf{p}, \mathbf{x}, \mathbf{z}, \boldsymbol{\tau}, \{\mathbf{v}_i\}\}$ such that Equations **5.7a** to **5.7d** and Equations **5.8a** to **5.8b** below are satisfied:

$$V_k \cdot l_k = \sum_i v_{ik} \cdot d_i \quad \forall k \in K \quad \text{(Flow conservation constraint for road } k\text{)} \tag{5.8a}$$

$$\mu_i \cdot v_i \geq \kappa^c \cdot p_i \cdot x_i \quad \forall i \in I \quad \text{(Individual rationality constraint)} \tag{5.8b}$$

Here, **Equation 5.8b** indicates that, if there exists an equilibrium, then the user benefit from travel must be enough to justify the purchase of a vehicle. We note that **Equation 5.7a** to **Equation 5.8a** make profit maximization for both the automaker and the service provider mathematical programs with equilibrium constraints. To facilitate our analysis, we therefore consider a more restrictive case when all road groups are used by all user groups. Then, for $k \in K$, τ_k is such that:

$$\tau_k = f_{k,4}(x_i, z_k, V_k, v_{ik}) - \mu_i = \frac{1}{V_k \cdot l_k} \cdot \sum_i [f_{k,4}(x_i, z_k, V_k, v_{ik}) - \mu_i] \cdot v_{ik} \cdot d_i \tag{5.9}$$

The problem for the automaker becomes:

$$
\begin{aligned}
\max_{\substack{x_i\\V_k,v_{ik}}} \quad & \sum_i \Big[\Big(\sum_k f_{k,1}(x_i, z_k, V_k, v_{ik})\Big)\cdot x_i \cdot d_i - \kappa_c \cdot c_c(x_i, d_i)\Big] \\
\text{s.t.} \quad & \sum_k f_{k,1}(x_i, z_k, V_k, v_{ik}) = \kappa_c \cdot p_i \quad \forall i \in I, \\
& \tau_k \cdot V_k \cdot l_k = \sum_i (f_{k,4}(x_i, z_k, V_k, v_{ik}) - \mu_i) \cdot v_{ik} \cdot d_i \quad \forall k \in K, \\
& \sum_k v_{ik} = v_i \quad \forall i \in I, \\
& V_k \cdot l_k = \sum_i v_{ik} \cdot d_i \quad \forall k \in K, \\
& \mu_i \cdot v_i \geq \Big(\sum_k f_{k,1}\Big) \cdot x_i \quad \forall i \in I
\end{aligned} \tag{M}
$$

Then, the FONC yield:

$$d_i \cdot \sum_k \left[f_{k,1} \cdot \left(1 + \delta_k^c - \frac{\gamma_i^c}{d_i}\right) + f_{k,11} \cdot \left(1 - \frac{\gamma_i^c}{d_i}\right) \cdot x_i \right] = \kappa_c \cdot c_{c,1}(x_i, d_i) \quad \forall i \in I \tag{5.10a}$$

$$f_{k,14} \cdot \left(1 - \frac{\gamma_i^c}{d_i}\right) \cdot x_i + (f_{k,44} \cdot v_{ik} + f_{k,4}) \cdot \delta_k^c + \alpha_k^c = \frac{\beta_i^c}{d_i} + \delta_k^c \cdot \mu_i \quad \forall i \in I,\ \forall k \in K \tag{5.10b}$$

$$\sum_i \left[f_{k,3} \cdot \delta_k^c + f_{k,13} \cdot \left(1 - \frac{\gamma_i^c}{d_i}\right) \cdot x_i \right] \cdot d_i = (\delta_k^c \cdot \tau_k + \alpha_k^c) \cdot l_k \quad \forall k \in K \tag{5.10c}$$

$$\gamma_i^c \cdot \Big[\sum_k f_{k,1} \cdot x_i - \mu_i \cdot v_i\Big] = 0 \quad \forall i \in I \tag{5.10d}$$

$$\gamma_i^c \geq 0 \quad \forall i \in I \tag{5.10e}$$

where δ_k^c is the Lagrangian multiplier associated with the fee constraint; β_i^c is the Lagrangian multiplier associated with the i^{th} individual VMT constraint; α_k^c is the Lagrangian multiplier associated with the k^{th} traffic volume constraint; γ_i^c is the Lagrangian multiplier associated with the i^{th} individual rationality constraint.

Firstly, we note that, when all road groups are used, $\delta_k^c \geq 0$. Indeed, in the more general case, the manufacturer is faced with $\tau_k \cdot V_k \cdot l_k \geq \sum_i (f_k(x_i, z_k, V_k, v_{ik}) - \mu_i \cdot v_{ik}) \cdot d_i \ \forall k \in K$. This implies that the Lagrangian multiplier δ_k^c would be non-negative for all used road groups. Then, by comparing **Equation 5.4a** and **Equation 5.10a**, we note that there may be under-provision or over-provision of automation under the GNE. On the one hand, the automaker's exercise of market power (captured by $\sum_k f_{k,11} \cdot x_i < 0$ in **Equation 5.10a**), induces a lower automation level than what would happen under the social optimum. Moreover, due to the lack of coordination with the service provider, the automaker does not account for the non-pricing benefit (captured by $\sum_k b_{r,1} \cdot v_{ik} > 0$ in **Equation 5.4a**) when making its production decisions. This, in turns, leads to a lower provision than socially optimal. On the other hand, the ability of the automaker to affect and exploit changing travel patterns for increased gains (captured by $\sum_k f_{k,1} \cdot \delta_k^c > 0$ in **Equation 5.10a**) could lead to more investment than socially optimal. In an environment with relatively high competition and in which infrastructure-related non-pricing benefits are uncertain or inaccessible for the automaker, the net effect of the automaker's decisions might be too much spending on automation.

The problem for the service provider becomes:

$$\begin{aligned} \max_{\substack{z_k,\tau_k,\\ V_k,v_{ik}}} \quad & \sum_k \tau_k \cdot V_k \cdot l_k + \sum_i b_r(x_i, z_k, V_k) \cdot v_{ik} \cdot d_i - \kappa_r \cdot c_r(z_k, l_k) \\ \text{s.t.} \quad & \sum_k f_{k,1}(x_i, z_k, V_k, v_{ik}) = \kappa_c \cdot p_i \quad \forall i \in I, \\ & \tau_k \cdot V_k \cdot l_k = \sum_i (f_{k,4}(x_i, z_k, V_k, v_{ik}) - \mu_i \cdot v_{ik}) \cdot d_i \quad \forall k \in K, \\ & \sum_k v_{ik} = v_i \quad \forall i \in I, \\ & V_k \cdot l_k = \sum_i v_{ik} \cdot d_i \quad \forall k \in K, \\ & \mu_i \cdot v_i \geq \Big(\sum_k f_{k,1} \Big) \cdot x_i \quad \forall i \in I \end{aligned} \tag{O}$$

The FONC then yield:

$$\sum_i \Big(f_{k,2} + b_{r,2} \cdot v_{ik} - \frac{\lambda_i^r + \gamma_i^r \cdot x_i}{d_i} \cdot f_{k,12} \Big) \cdot d_i = \kappa_r \cdot c_{r,1}(z_k, l_k) \quad \forall k \in K \tag{5.11a}$$

$$- \frac{\lambda_i^r + \gamma_i^r \cdot x_i}{d_i} \cdot f_{k,14} + (f_{k,44} \cdot v_{ik} + f_{k,4}) + b_r + \alpha_k^r = \frac{\beta_i^r}{d_i} + \mu_i \quad \forall k \in K, \ \forall i \in I \tag{5.11b}$$

$$\sum_i \Big[f_{k,34} + b_{r,3} \cdot v_{ik} - \frac{\lambda_i^r + \gamma_i^r \cdot x_i}{d_i} \cdot f_{k,13} \Big] \cdot d_i = \alpha_k^r \cdot l_k \quad \forall k \in K \tag{5.11c}$$

$$\gamma_i^r \cdot \Big[\sum_k f_{k,1} \cdot x_i - \mu_i \cdot v_i \Big] = 0 \quad \forall i \in I \tag{5.11d}$$

$$\gamma_i^r \geq 0 \quad \forall i \in I \tag{5.11e}$$

where λ_i^r is the Lagrangian multiplier associated with the i^{th} pricing constraint; β_i^r is the Lagrangian multiplier associated with the i^{th} flow balance constraint; γ_i^r is the Lagrangian multiplier associated with the i^{th} individual rationality constraint; and α_k^r is the Lagrangian multiplier associated with the k^{th} traffic volume constraint. Here too, with arguments similar to those for the positivity of δ_k^c in the automaker's case, it is possible to deduce that $\lambda_i^r \geq 0$ at equilibrium.

Now, considering **Equation 5.11a**, if automation and digitalization are substitutes ($f_{k,12} \leq 0$), then there is over-provision of digitalization relative to the social optimum (with equal provision when $f_{k,12} = 0$). Simply, in order to have a competitive edge and capture users' willingness to pay, the service provider invests heavily in digitalization. If automation and digitalization are complementary ($f_{k,12} > 0$), there is under-provision of digitalization at the Nash equilibrium. Because of a coordination failure, the service provider is reluctant to invest in digitalization: she cannot ensure that the automaker will make the compatible automation investment that will make the digitalization investment worthwhile.

5.3.3 Cooperation

Our analysis above shows that, in the absence of coordination, it is likely that the level of automation and digitalization are suboptimal. This sub-optimality is due, on one hand, to a lack of coordination between service providers and car manufacturers and, on the other, to the car manufacturer's exercise of market power (as attested by the presence of a markup in **Equation 5.10a**). We discuss here how the former issue could be resolved. In order to achieve coordination, a contract that ensures that both the operator and the manufacturer are better off working together can be designed. Such a contract must meet the following criteria:

- The joint profit π^T must be maximized: $\pi^T = \pi^c + \pi^r$
- Each party must be better off than under the GNE: $\pi^{c,GNE} \leq \pi^{c,CE}$ and $\pi^{r,GNE} \leq \pi^{r,CE}$

where $\pi^{a,GNE}$ refers to the profit under the GNE and $\pi^{a,CE}$ refers to the profit under the cooperative equilibrium (CE) with $a \in \{c, r\}$. Following the Nash bargaining, $\pi^{c,CE}$ and $\pi^{r,CE}$ are such that:

$$\begin{aligned} (\pi^{c,CE}, \pi^{r,CE}) = &\underset{\substack{\pi^c \geq \pi^{c,GNE}, \\ \pi^r \geq \pi^{r,GNE}}}{\arg\max} \quad (\pi^c - \pi^{c,GNE}) \cdot (\pi^r - \pi^{r,GNE}) \\ &\text{s.t.} \qquad \pi^c + \pi^r = \pi^{T,CE} \end{aligned} \tag{NB}$$

where:

$$\begin{aligned} \pi^{T,CE} = &\max_{\substack{x_i, p_i, v_{ik}, \\ z_k, \tau_k, V_k \geq 0}} \quad \kappa_c \cdot \sum_i \Big[p_i \cdot x_i \cdot d_i - c_c(x_i, d_i) - m_c(\bar{x})\Big] + \sum_k \tau_k \cdot V_k \cdot l_k - \kappa_r \cdot c_r(z_k, l_k) \\ &+ \sum_k \sum_i b_r(x_i, z_k, V_k, v_{ik}) \cdot v_{ik} \cdot d_i \\ &\text{s.t.} \qquad \{\mathbf{p}, \mathbf{x}, \mathbf{z}, \boldsymbol{\tau}, \{\mathbf{v}_i\}\} \in X(\mathbf{v}) \end{aligned} \tag{TM}$$

It is easy to show that:

$$\pi^{c,CE} = \pi^{c,GNE} + \phi \cdot \left[\pi^{T,CE} - \pi^{c,GNE} - \pi^{r,GNE}\right] \tag{5.12}$$

$$\pi^{r,CE} = \pi^{r,GNE} + (1-\phi) \cdot \left[\pi^{T,CE} - \pi^{c,GNE} - \pi^{r,GNE}\right] \tag{5.13}$$

where $\phi \in (0,1)$ represents the share of excess profits–relative to the GNE–that the automaker will pocket. Now, the question arises as to whether cooperation is welfare-improving relative to the GNE. Assuming, as in M and O, that all road groups are utilized by all user groups, then TM becomes:

$$\begin{aligned}
\pi^{T,CE} = \max_{\substack{x_i,\bar{x},v_{ik},\\ z_k,V_k,v_{ik}\geq 0}} \quad & \sum_i \Big(\sum_k f_{k,1}(x_i, z_k, V_k, v_{ik}) \cdot x_i + b_r(x_i, z_k, V_k) \cdot v_{ik} \Big) \cdot d_i - \kappa_c \cdot \sum_i c_c(x_i, d_i) \\
& + \sum_k \sum_i (f_{k,4}(x_i, z_k, V_k, v_{ik}) - \mu_i) \cdot v_{ik} \cdot d_i - \sum_k \kappa_r \cdot c_r(z_k, l_k) \\
\text{s.t.} \quad & \sum_k v_{ik} = v_i \quad \forall i \in I, \\
& V_k \cdot l_k = \sum_i v_{ik} \cdot d_i \quad \forall k \in K, \\
& \mu_i \cdot v_i \geq \sum_k f_{k,1}(x_i, z_k, V_k, v_{ik}) \cdot x_i \quad \forall i \in I
\end{aligned} \tag{TM}$$

and the FONC yields:

$$d_i \cdot \sum_k \left[f_{k,1} + b_{r,1} \cdot v_{ik}\right] = \kappa_c \cdot c_{c,1}(x_i, d_i) \quad \forall i \in I \tag{5.14a}$$

$$\sum_i \left[f_{k,2} + b_{r,2} \cdot v_{ik}\right] \cdot d_i = \kappa_r \cdot c_{r,1}(z_k, l_k) \quad \forall k \in K \tag{5.14b}$$

$$\sum_i \left[f_{k,3} + b_{r,3} \cdot v_{ik}\right] \cdot d_i = \alpha_k \cdot l_k \quad \forall k \in K \tag{5.14c}$$

$$f_{k,4} + b_r + \alpha_k = \frac{\beta_i}{d_i} + \mu_i \quad \forall k \in K, \forall i \in I \tag{5.14d}$$

Comparing **Equation 5.14** to **Equations 5.10** and **5.11** indicates that, relative to the Nash equilibrium, cooperation:

- reduces the effect of the manufacturer's market power and increases the provision of automation;
- increases (decreases) provision of digitalization when digitalization and automation are complements (substitutes)

Thus, cooperation between the car manufacturer and the service provider increases surplus. Additionally, comparing **Equations 5.4** and **5.14**, cooperation between the manufacturer and the operator will decentralize the social optimum if:

$$\frac{\gamma_i}{d_i} = \frac{\beta_i}{d_i} + \mu_i \tag{5.15}$$

In other words, if the marginal benefit of travel $\frac{\beta_i}{d_i} + \mu_i$ for the combined entity is equal to the marginal benefit of travel for the social planner, cooperation will achieve the first-best. Otherwise, cooperation achieves the second-best: $\frac{\beta_i}{d_i} + \mu_i < \frac{\gamma_i}{d_i}$. This can potentially be the best-case scenario absent the possibility of subsidies (e.g., when the first-best is not sustainable for either or both companies).

5.4 A numerical example

We propose here to illustrate our model's results as well as other properties.

5.4.1 Functions and parameters

We consider $K = 3$ road groups and $I = 3$ different user groups. The road lengths l_k and capacities V_k^{max}, users' daily VMT v_i and population d_i as well as other parameter values and how they were obtained can be found in **Appendix C.1**. It suffices to say, however, that K and I are ordered in an increasing order of capacity and daily VMT respectively.

Automation and digitalization levels vary continuously from 1 to 100. We assume the following cost and benefit functions:

$$f_k(x_i, z_k, V_k, v_{ik}) = \left[f_{0t} \cdot \left[\alpha \cdot (x_i)^\rho + (1-\alpha) \cdot (z_k)^\rho \right]^{\frac{v}{\rho}} - f_{0c} \cdot \left(\frac{V_k}{V_k^{max}} \right)^2 \right] \cdot \sqrt{v_{ik}} \tag{5.16a}$$

$$\sum_i c_c(x_i, d_i) = m_{c,0} \cdot \max_j x_j + \sum_i [c_0 + c_1 \cdot (x_i)^2] \cdot d_i \tag{5.16b}$$

$$b_r(x_i, z_k, V_k) = b_0 \cdot (x_i)^\gamma \cdot (z_k)^\theta \cdot \left(\frac{V_k}{V_k^{max}} \right)^{\eta_r} \tag{5.16c}$$

$$c_r(z_k, l_k) = m_{r,0} \cdot l_k \cdot (z_k)^2 \tag{5.16d}$$

Equation 5.16a indicates that automation and digitalization interact following a constant elasticity of substitution (CES) utility function. Thus, ρ is the substitution parameter: as ρ increases, automation and digitalization become more substitutable in the eyes of customers. Additionally, drivers' benefit from automation and digitalization will be affected by a congestion cost.

Equation 5.16b indicates that the total cost of manufacturing includes both an investment cost $m_{c,0} \cdot \max_j x_j$–the cost the company must pay to develop its highest level of automation–and production costs $[c_0 + c_1 \cdot (x_i)^2] \cdot d_i$ for each automation levels.

The parameters of the model as well as their value are given in **Appendix C.1**.

5.4.2 Effect of substitution parameter

Figure 5.1 and **Figure 5.2** show the level of automation and digitalization, respectively, as a a function of the degree of substitution under the three different scenarios considered in this study: the social optimum (FB), the cooperative equilibrium (CE) and the GNE. First, we note that as automation and digitalization levels become more substitutable, the socially optimal automation level decreases. Simply, because customers are increasingly indifferent between automation and digitalization and because digitalization can serve mutliple classes simultaneously, the need for automation diminishes. Moreover, the road groups with the highest volumes receive the highest levels of digitalization (**Figure 5.2** and **Figure 5.3**). As expected, the CE improves welfare relative to the GNE, though it still falls short from the socially optimal configuration[2]. The improvement of CE over the GNE is more pronounced as substitutabiltiy increases, thus highlighting the crippling effect of competition.

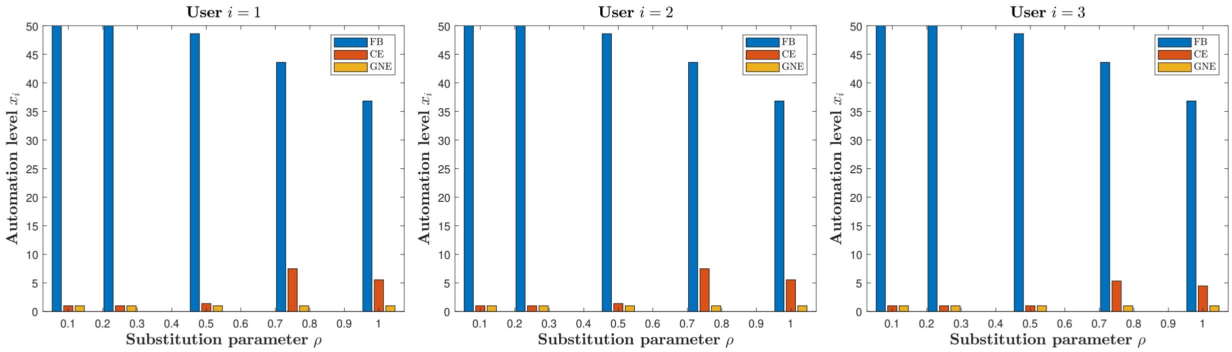
5.4.3 Effect of unit cost of congestion

Here, we evaluate the impact of the cost of congestion, $f_{c,0}$, on equilibrium results. In practice, this can shed light on the difference in automation choices between users with different values of time. As Figures 5.4 and 5.5 show, some of main insights from **Section 5.4.2** still hold. Namely, cooperation usually results in better performance than competition but performs worse than the surplus maximizing configuration. As we would expect, increasing congestion costs leads to an increase in both automation and digitalization investment, though the effect is more pronounced on the infrastructure side. Simply, the higher the cost of congestion, the higher the value of automation and digitalization. Thus, investing in automation and digitalization in highly congested areas would seem like an intuitive first step for both the private and public sector.

5.4.4 Effect of unit monetary value of digitalization

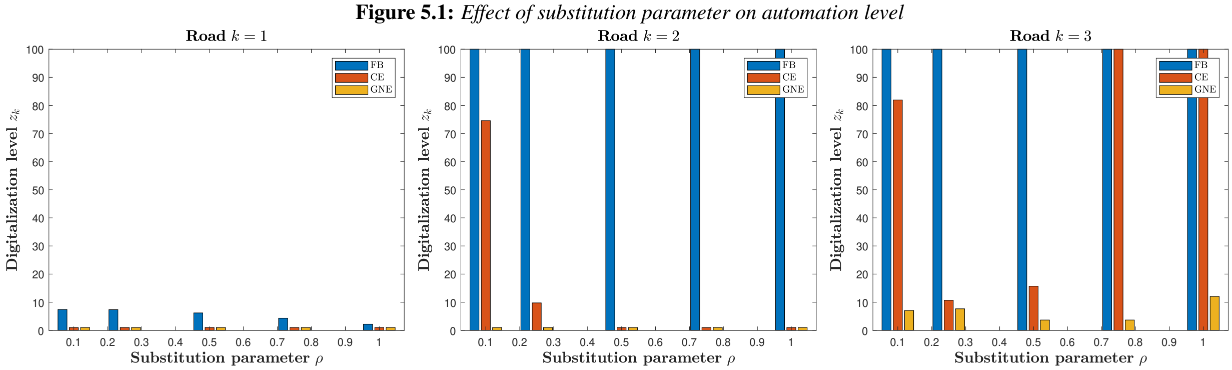
There is uncertainty as to what the monetary benefits of digitalization, b_0, will be for ISSPs. Such benefits will depend, among other things, on the existence of a vibrant market for road

[2]The elasticity of demand will determine the size of the gap between CE and GNE. In our case, because demand is inelastic, the gap will be larger.



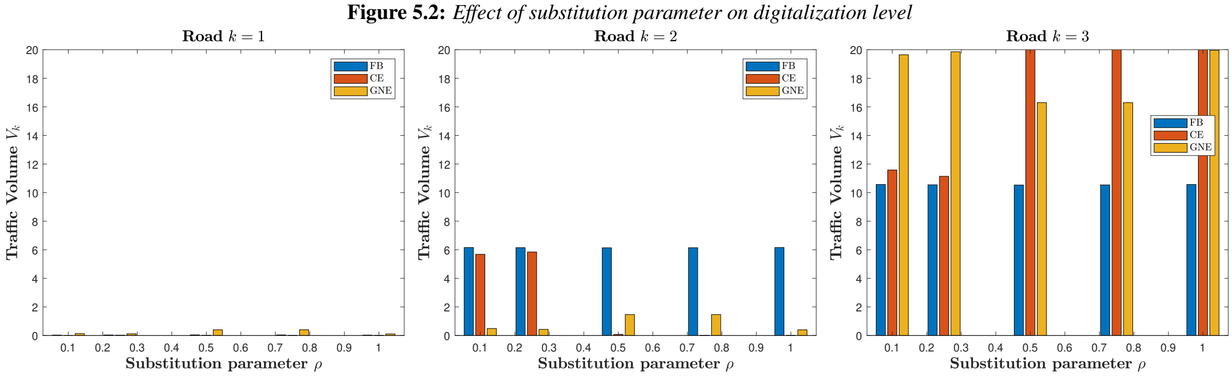
(a) (b) (c)

Figure 5.1: *Effect of substitution parameter on automation level*



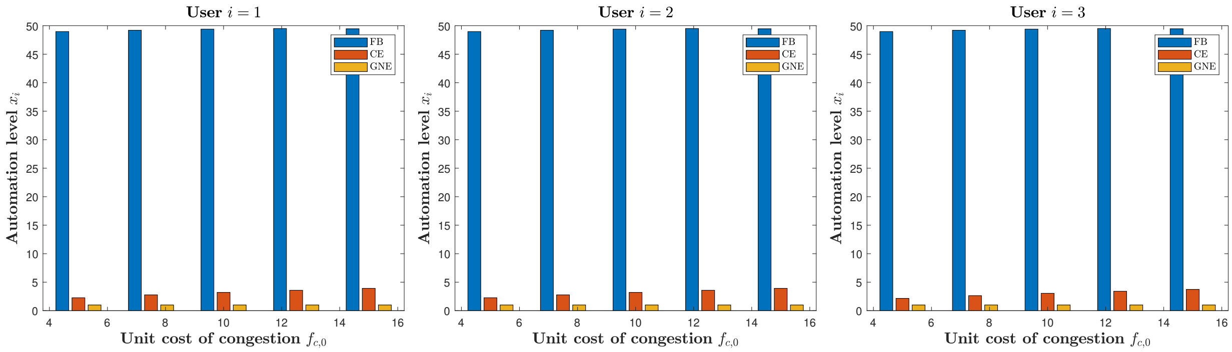
(a) (b) (c)

Figure 5.2: *Effect of substitution parameter on digitalization level*



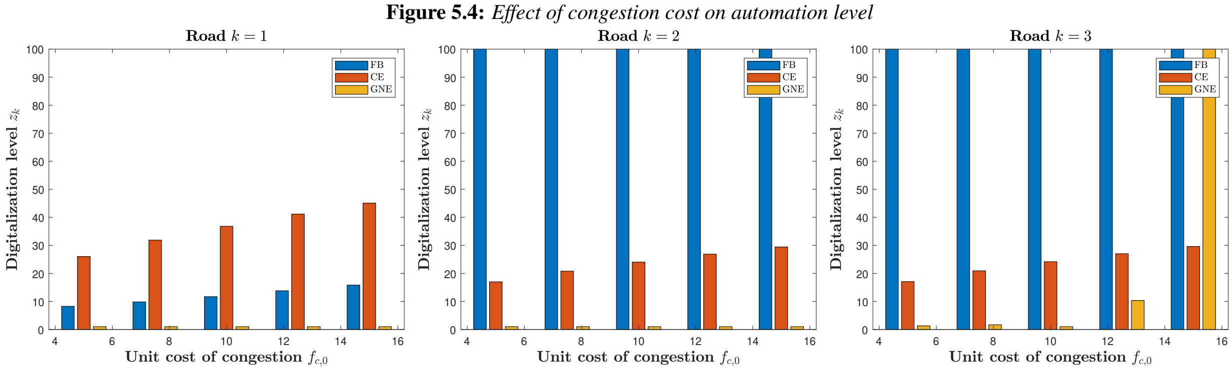
(a) (b) (c)

Figure 5.3: *Effect of substitution parameter on traffic volume distribution*



(a) (b) (c)

Figure 5.4: *Effect of congestion cost on automation level*



(a) (b) (c)

Figure 5.5: *Effect of congestion cost on digitalization level*

data that only infrastructure digitalization could fulfill. To better understand the effect of that monetary value, we vary b_0 across the three scenarios considered. The results are shown in Figures 5.6 and 5.7. It is interesting to note that, in the FB and CE cases, an increase in the value of digitalization leads to an increase in equilibrium automation levels. In essence, because digitalization and automation interact together to generate value, there is an incentive for the social planner and for any joint venture between automaker and ISSP to increase automation levels.

5.4.5 Effect of automation development costs

Lastly, because one of the main motivations for the present work is the high cost of automation, we propose to investigate the effect of development costs, $m_{c,0}$, on the outcome of our scenarios. The results are shown in Figures 5.8 and 5.9. First, we note that, because of the co-dependency between automation and digitalization in generating value, increasing automation costs lead to a reduction in both automation and digitalization for both the social planner and the integrated company. Essentially, the more expensive automation becomes, the lesser the value to both society and the private sector of implementing our vision for vehicle-infrastructure cooperation. Thus, a careful evaluation of the costs and benefits of automation is needed. In the GNE case, the ISSP

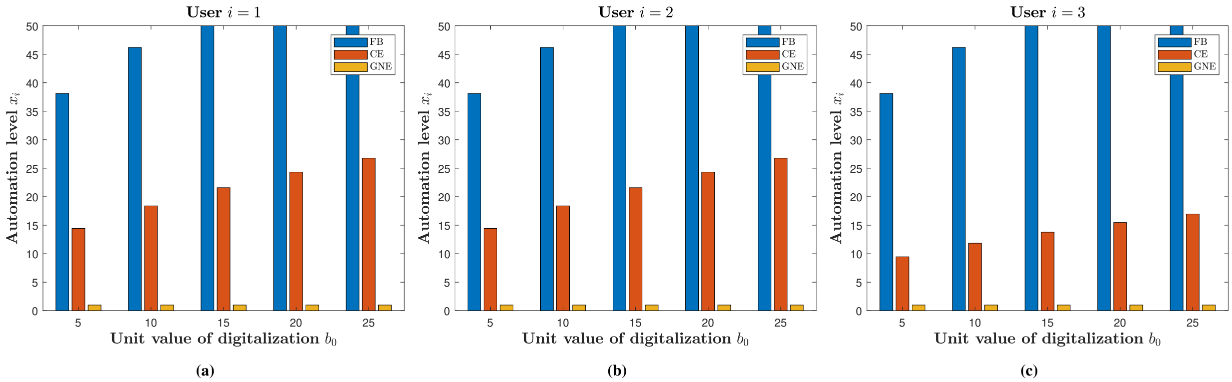


Figure 5.6: *Effect of the value of digitalization on automation level*

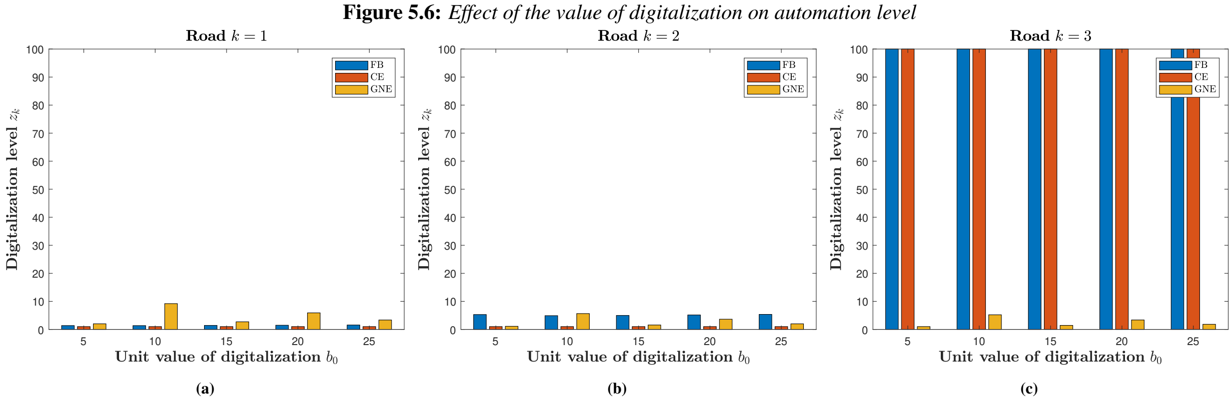


Figure 5.7: *Effect of the value of digitalization on digitalization level*

obviously benefits from the higher automation costs and increases its provision of digitalization beyond the efficient levels to increase profits.

5.5 Conclusion

This paper has investigated vehicle-infrastructure cooperation for enabling automated driving. In this cooperation, the infrastructure can perform driving tasks such as sensing, perception or planning, and essentially becomes an integral part of the driving system of an automated vehicle. By proposing and analyzing a model that captures investment decisions in automation and digitalization and their effect on travellers' purchase and travel decisions, we have shown that such a vehicle-infrastructure cooperative paradigm can be socially optimal. Subsequently, we also show that strategic interactions between a monopolistic automaker and a monopolistic service provider result in suboptimal investment in both automation and digitalization. The suboptimality of automation is due, in part, to the lack of coordination which prevents automakers from enjoying the non-pricing benefits that driving generates for service providers. Whether there is over-investment or under-investment in automation will also depend on the travel behaviors of the different user groups. Users with high VMT will likely see higher than optimal automation while those with

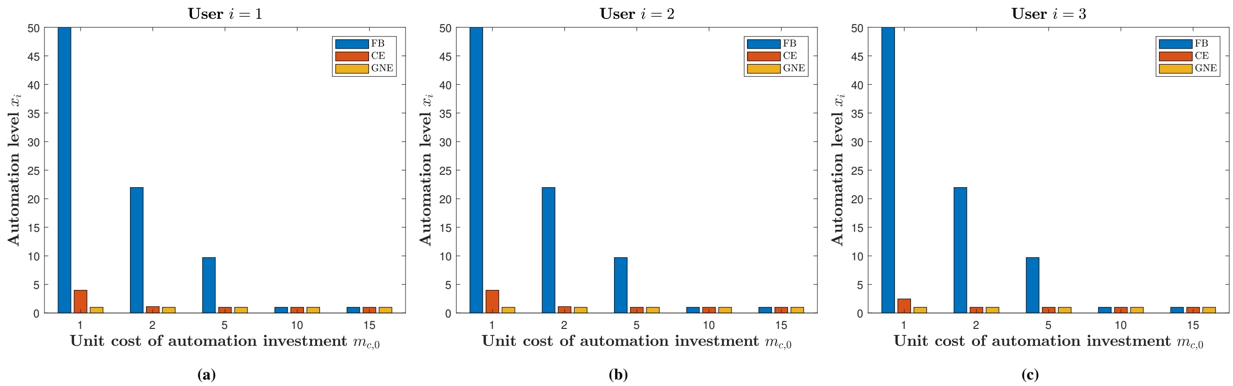


Figure 5.8: *Effect of automation development cost on automation level*

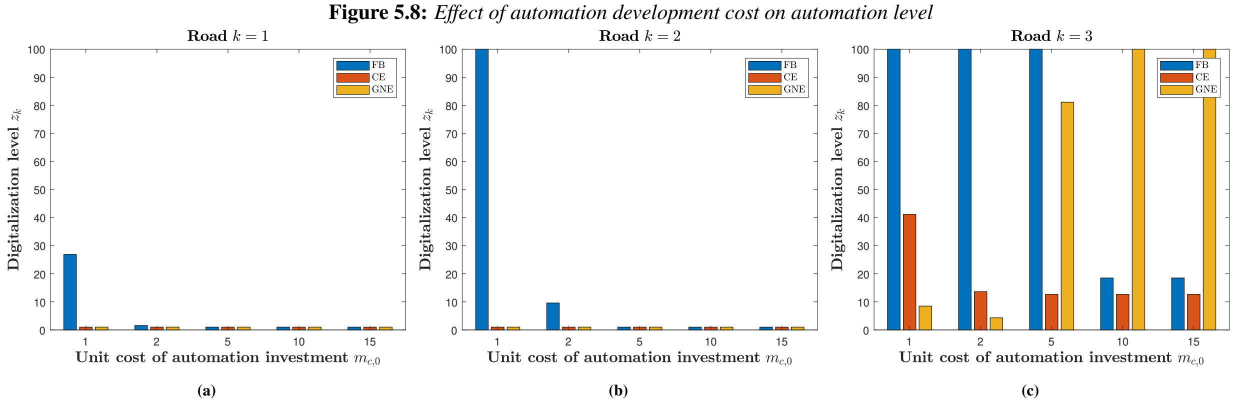


Figure 5.9: *Effect of automation development cost on digitalization level*

low VMT will receive lower than optimal automation. For service providers, when automation and digitalization are substitutes, there is over-investment in digitalization technology as service providers seek to compete with the automation technology. However, when they are complements, service providers are reluctant to invest in digitalization: there is no enforcement mechanism that guarantees that automakers will invest in compatible automation levels. It is then easy to show that, given an appropriate profit-sharing agreement between the two actors, cooperation could yield the socially optimal levels of automation and digitalization. Thus, from a planning perspective, better coordination of infrastructure standards and regulation across states should be a priority. Such coordination across service providers will then provide automakers with the opportunities for economies of scale that would be otherwise lacking when developing the infrastructure-assisted vehicle technology. Finally, it will provide service providers and/or their regulating entities with the value proposition necessary to benefit from vehicle-infrastructure cooperation.

In this work, we assume that users' demand for travel and automation is fixed. However, since automation and digitalization reduce the cost of travel, an increase in VMT is likely after adoption of these technologies and can have two conflicting effects which it will be necessary to investigate. On one hand, by increasing VMT, it could increase the ISSP's ability to generate profits. On the other, that increase in VMT can increase congestion and reduce willingness to pay for road usage.

Moreover, such an increase in VMT can also have a negative social impact. Thus, future iterations will consider the case of elastic travel demands. Moreover, by enabling mobility-as-a-service, automation will also provide an alternative to car ownership for users. As such, automakers face an additional dilemma in providing automation, but also another earning opportunity. The impact of these decisions and their effects on VMT will also be incorporated. Lastly, we have not accounted here for competition among automakers and among service providers. Essentially, there is no product differentiation in either the vehicle or infrastructure market. This makes it difficult to assess the benefits–or lack thereof–that can accrue to different socio-demographic groups. Our model can be extended and made more realistic to include the effects on investment of these new strategic interactions and the relevant incentives to be provided. Lastly, the question of investment is essentially a dynamic problem subject to uncertainty and the different agents involved will make repeated decisions that can significantly alter the trajectory of both automation and digitalization levels. Such rich dynamics is not captured by the current model and will need to be incorporated.

CHAPTER 6

Safety, Liability and Infrastructure

6.1 Introduction

In the previous chapter, we established that vehicle-infrastructure cooperation is socially optimal. We also showed that a cooperative agreement between an automaker and a private road operator can increase welfare and the joint profit between the two entities. However, an important issue arises out of the framework we presented. Our model and approach is unspecific as to the benefits that will come from automation and digitalization. This becomes important when we attempt to evaluate the incentives and gains from a vehicle-infrastructure cooperative approach. Thus, in this chapter, we propose to focus on the question of liability and safety. Indeed, one of the key benefits that will accrue to customers through automation and digitalization is that of safety. For example, using National Highway Traffic Safety Administration (NHTSA) data, Fagnant and Kockelman (2015) identify that 90 $\%$ of traffic accident involve human error. When considering the substantial costs to society of crashes to society–$\$226$ billions in 2005 (2021 dollars) according to Cambridge Systematics (2008)–improved safety resulting from vehicle automation would generate enormous savings to customers and society in general. Moreover, most of the remaining challenges in developing automated vehicles revolve around safety–as discussed in the introduction to Chapter 4–and research shows that customers have high expectations for automated vehicle safety (Shariff et al., 2021; Shariff, 2021). Importantly, after a series of roads incidents involving automated vehicles, regulators are also starting to show concerns about automation technology's safety and readiness for commercialization (Elliott, 2021). Thus, it appears that finding cost effective means to address the remaining safety concerns of automated vehicles is critical if automakers want to encourage adoption.

However, though automated vehicles would increase safety, we are unlikely to see accident-free roads. Indeed, vehicles and infrastructure could still malfunction; pedestrians are still independent agents who can interfere with vehicle operation; and human-driven vehicles (HDVs) will co-exist for a long period of time with AVs, potentially leading to more crashes than in an AV-free environment since HDVs might reduce their level of care (Chatterjee and Davis, 2013; Elvik, 2014; Talley, 2019; Di et al., 2020). In this context, one might ask who should be liable for the accidents in which an AV is involved. While some automakers, like BMW, have decided to shoulder full responsibility for accidents in which their self-driving technology is involved (even when not at Level 5–(Tucker, Sean, 2022)), others such as Tesla, have been content to let current liability rules for drivers prevail and have, mostly, eschewed responsibility for accidents related to their technology (Communications, 2022). Is such a configuration optimal? Should liability be shared between automakers and AV owners? What happens when infrastructure support service providers (ISSPs) enter the fray?

The question of liability and automated vehicles has been explored in a number of works over the past decade. Researchers have explored and discussed ethical problems that arise in the design of AVs (Nyholm and Smids, 2016; Contissa et al., 2017; Thornton et al., 2017; Himmelreich, 2018; Nyholm, 2018; Borenstein et al., 2019; Wu, 2020); and ways in which current tort laws can be adapted in the era of AVs (Lohmann, 2016; Talley, 2019; Di et al., 2020). While insightful, these works all deal with the issue of equipping individual vehicles and responsibility resulting from the failure of these individual vehicles. However, these works have not explored the ways in which liability design might affect the provision of AV-related infrastructure, something the present chapter will attend to.

An important concept that we will exploit in addressing that problem is that of insurance. Indeed, insurance contracts are means through which an agent can offload a part or a totality of its liability burden on another agent. Thus, optimal liability design between different parties could be thought of, in and of itself, as the design of an insurance contract between these parties. These contracts have been studied in detail over the years. Some articles have dealt with the issue of insurance in competitive markets (Ehrlich and Becker, 1972; Rothschild and Stiglitz, 1976; Cook and Graham, 1977; Schlesinger, 1983) and monopolistic markets (Stiglitz, 1977; Ligon and Thistle, 1996) both under full information and information asymmetry; others have looked at insurance under the threat of adverse selection and moral hazard and sought to evaluate their effects on insurance markets and insurance provision (Pauly, 1974; Dionne, 1982). Some have even tried to estimate the effect of connectivity on insurance cost, accounting for issues like privacy (Jin and Vasserman, 2021). However, we are not aware of any work in that literature that seeks to model and address issues of insurance and liability in a vehicle-infrastructure cooperation context. Nevertheless, a survey of the relevant literature provides an insight that will be important for our work: under full information, regardless of market structure and without the threat of moral hazard from the insured, if any insurance is provided, then the contract provides full insurance priced at actuarial odds.

This chapter is divided as follows. **Section 6.2** introduces our model and its basic components.

Notation	Description
z^M	Vehicle automation quality
z^I	Infrastructure automation quality
$\phi^{j,a}$	Liability amount for accident on road $j \in \{r, s\}$ for agent $a \in \{M, I\}$
c^a	Investment cost for $a \in \{M, I\}$
τ	Vehicle cost
V	Value of car ownership
λ	Total demand for vehicle ownership

(a) *Frequently used variables*

Notation	Description
η^j	Share of road $j \in \{r, s\}$
l	Total road length
p^j	Crash probability per mile driven on road $j \in \{r, s\}$
$\tilde{p}^k$	Probability of state $k \in \{0, r, s\}$
W^k	Wealth in state $k \in \{0, r, s\}$
$\tilde{W}^0$	Initial wealth level
s^j	Accident severity on road $j \in \{r, s\}$

(b) *Frequently used parameters*

Table 6.1: *Frequently used notations*

Then, **Section 6.3** derives the socially optimal provision of automation and digitalization and the associated optimal liability design. This discussion is then followed by a discussion of the resulting allocation in an unregulated market. In **Section 6.5**, we revisit our analysis of **Section 6.3** to understand how the introduction of speed choice affects automation and digitalization provision.

6.2 Model

Consider a roadway used exclusively by Level 5 AV owners. These owners purchase their vehicles from a vehicle manufacturer who decides the quality of the technology (e.g.: sensors, algorithms, other hardware and software components...) with which to equip these vehicles. The manufacturer's quality choices affect the crash probability of his vehicles and, therefore, his bottom line. Indeed, customers will shy away from an accident-prone product. We consider that all losses and accidents are the results of vehicle technology. Importantly, we do not consider owners' efforts in maintaining vehicles in fully functional order: thus, we do not account for moral hazard and/or adverse selection.

A portion of the roadway is managed by an ISSP. This ISSP installs and operates AV-related technology on the roadway. While this contributes to reducing the crash probability on the roadway, it could also expose the ISSP to similar accident losses and reputational harms as the

AV manufacturer. The principal goal of this work is to understand how liability for accidents should be shared between manufacturer, ISSP, and customer.

AV owners face three states of the world while using an AV. In the first state, they travel without accident on the full length of the road. In the second state, they are involved in an accident on the regular (r) portion of the road and face potential losses as a result. In the third state, they are instead involved in an accident on the smart (s) portion of the road. Formally, then, user preferences can be described by the following equations:

$$V = \tilde{p}^0 \cdot U(W^0) + \tilde{p}^r \cdot U(W^r) + \tilde{p}^s \cdot U(W^s)$$
$$W^k = \begin{cases} \tilde{W}^0 - \tau & \text{if } k = 0 \\ \tilde{W}^0 - \tau - s^r + \phi^r & \text{if } k = r \\ \tilde{W}^0 - \tau - s^s + \phi^s & \text{if } k = s \end{cases}$$
$$\tilde{p}^k = \begin{cases} 1 - \tilde{p}^r - \tilde{p}^s & \text{if } k = 0 \\ 1 - (1 - p^r)^{(1-\eta^s)\cdot l} & \text{if } k = r \\ 1 - (1 - p^s)^{\eta^s \cdot l} & \text{if } k = s \end{cases}$$

In the above, $\tilde{p}^k$ represents the probability of occurrence of state $k \in \{0, r, s\}$; p^j represents the probability of a crash per mile driven on portion $j \in \{r, s\}$ of the roadway; $\tilde{W}^0$ is consumers' auto budget; s^j represents the severity–or monetary losses–incurred as a result of the accident on portion j of the roadway; ϕ^j represents total payments from the ISSP and/or the manufacturer received by the victim as a result of the accident on portion j of the roadway; l is the length of the roadway; τ represents the price of the AV; η^s represent the fraction of the road operated by the ISSP; and $U(\cdot)$ describes utility as a function of wealth. Essentially, there exists a form of insurance from the producers (manufacturer and ISSP) to the AV owners. Heterogeneity in valuation for car ownership (but not in risk profile) gives rise to a demand function $\lambda = \Lambda(V)$ for AVs. We make the following standard assumptions on $U(\cdot)$ and $\Lambda(\cdot)$:

Assumption 6.1. *We assume the following:*

A1.1 *Utility is strictly increasing in wealth:* $U' > 0$

A1.2 *Customers are risk averse:* $U'' < 0$

A1.3 *The demand function is strictly increasing in the value of AV ownership:* $\Lambda' > 0$

The crash probabilities are influenced by the decisions of the manufacturer and the ISSP as follows:

$$p^r = P(z^M, 0)$$
$$p^s = P(z^M, z^I)$$

where z^M is the quality level of AV technology; z^I is the quality level on the smart portion of the roadway; and $P(\cdot,\cdot)$ is a probability function such that:

Assumption 6.2. *We assume the following:*

A2.1 *The crash probability is strictly decreasing in technology quality:* $P_1,\ P_2 < 0$

The manufacturer and the ISSP face cost c^M per vehicle and c^I per mile of roadway equipped, respectively. These costs depend on technology level as follows:

$$c^M = C^M(z^M)$$
$$c^I = C^I(z^I)$$

We make the following assumptions on $C^M(\cdot)$ and $C^I(\cdot)$:

Assumption 6.3. *We assume the following:*

A3.1 *The cost functions are strictly increasing in technology quality:* $C^{M'},\ C^{I'} > 0$

A3.2 *The cost functions are strictly convex in technology quality:* $C^{M''},\ C^{I''} > 0$.

We can then describe our problem using the system of equations below:

$$\lambda = \Lambda(V) \tag{6.1a}$$
$$V = \tilde{p}^0 \cdot U(W^0) + \tilde{p}^r \cdot U(W^r) + \tilde{p}^s \cdot U(W^s) \tag{6.1b}$$
$$W^k = \begin{cases} \tilde{W}^0 - \tau & \text{if } k = 0 \\ \tilde{W}^0 - \tau - s^r + \phi^r & \text{if } k = r \\ \tilde{W}^0 - \tau - s^s + \phi^s & \text{if } k = s \end{cases} \tag{6.1c}$$
$$\tilde{p}^k = \begin{cases} 1 - \tilde{p}^r - \tilde{p}^s & \text{if } k = 0 \\ 1 - (1 - p^r)^{(1-\eta^s)\cdot l} & \text{if } k = r \\ 1 - (1 - p^s)^{\eta^s \cdot l} & \text{if } k = s \end{cases} \tag{6.1d}$$
$$p^r = P(z^M, 0) \tag{6.1e}$$
$$p^s = P(z^M, z^I) \tag{6.1f}$$

$$c^M = C^M(z^M) \tag{6.1g}$$

$$c^I = C^I(z^I) \tag{6.1h}$$

6.3 Scenario analysis

In this section, we consider three different scenarios to understand how liability will be shared in the age of automation and digitalization. Especially, we would like to understand how liability is shared without policy intervention and whether the outcome of lack of regulation and coordination warrants new rules be put in place.

6.3.1 First-best

We now consider the social welfare maximization problem:

$$\begin{aligned} \max_{\substack{\tau, z^M, z^I \\ \phi^r, \phi^s}} \quad & \int_0^V \Lambda(x) \cdot dx + (\tau - c^M - \tilde{p}^r \cdot \phi^r - \tilde{p}^s \cdot \phi^s) \cdot \lambda - c^I \cdot \eta^s \cdot l \\ \text{s.t.} \quad & \text{Equations } \mathbf{6.1a} \text{ to } \mathbf{6.1h} \end{aligned} \tag{SO}$$

In this problem, given the fraction of the roadway allocated for smart infrastructure, a social planner essentially minimizes the social cost of driving automation and smart infrastructure quality as well as the liability borne by the traveller. One might argue that the problem should contain two prices: one for the AV sale and another for the road usage. However, it is straightforward to show that both fees can be coalesced into a single τ and that multiple solutions would exist should two prices enter our problem. This also highlights another important point: because joint producer profit is maximized, the planner is agnostic as to whom bears liability on the smart roadway.

The first-order necessary conditions (FONCs) for optimality yield:

$$\phi^j = s^j \quad \forall j \in \{r, s\} \tag{6.2a}$$

$$\tau = c^M + \phi^r \cdot \tilde{p}^r + \phi^s \cdot \tilde{p}^s + \frac{\lambda}{\Lambda' \cdot U'(W^0 - \tau)} \cdot [1 - U'(W^0 - \tau)] \tag{6.2b}$$

$$\frac{C^{M\prime}(z^M)}{l} = -P_1(z^M, 0) \cdot \frac{1 - \tilde{p}^r}{1 - p^r} \cdot \phi^r - \eta^s \cdot \left[P_1(z^M, z^I) \cdot \frac{1 - \tilde{p}^s}{1 - p^s} - P_1(z^M, 0) \cdot \frac{1 - \tilde{p}^r}{1 - p^r}\right] \cdot \phi^s \tag{6.2c}$$

$$\frac{C^{I\prime}(z^I)}{\lambda} = -P_2(z^M, z^I) \cdot \frac{1 - \tilde{p}^s}{1 - p^s} \cdot \phi^s \tag{6.2d}$$

As expected from the classical literature on insurance markets, the optimal allocation of liability from a social perspective is for producers to provide full coverage (**Equation 6.2a**). **Equation 6.2b** indicates that the cost of vehicle ownership is adjusted to reflect producers' risk exposure and now includes the expected producer losses from vehicle operation. Thus, this indicates that, even when automakers decide to shoulder losses resulting from AV accident, vehicle ownership cost might not decrease, except if both crash probability and/or crash severity significantly decrease. Moreover, it is apparent that when users are risk neutral ($U'(\cdot) = 1$), the last term in **Equation 6.2b** vanishes. Then, producers' profits at the optimum, π^*, become:

$$\pi^* = -c^{I*} \cdot \eta^s \cdot l < 0 \tag{6.3}$$

In other words, at the first-best with risk-neutral users, the provision of smart infrastructure occurs at a loss for the providers. Especially, contrary to other infrastructure problems with capacity management,–such as the optimal provision of capacity on congested roads (Verhoef and Rouwendal, 2004)–self-financing does not hold. This also implies that, in a decentralized framework, either the provision of infrastructure services will be subsidized by a public agency or, just as for other infrastructure services, a regulated monopoly will form.

Additionally, we note that longer travel and higher demand will lead to increased investment in automation and digitalization technology, respectively (**Equations 6.2c** and **6.2d**). This is in keeping with the results from (**?**). More travel increases the risk of accident and, thus, in response, the planner invests more in vehicle technology. Moreover, more demand increases the expected losses from an accident, thus inducing the planner to invest more on infrastructure technology.

Looking at the relationship between vehicle and infrastructure technology, we note that when an increase in infrastructure technology reinforces the safety potential of AVs ($P_1(z^M, z^I) - P_1(z^M, 0) < 0$ or $P_{12} < 0$), then the introduction of infrastructure ($\eta^s > 0$) increases spending c^M on vehicle technology. Simply and naturally, the introduction of infrastructure technology only reduces vehicle costs when infrastructure acts as a substitute for vehicle technology.

6.3.2 Unregulated environment

6.3.2.1 Integrated monopoly

We now consider what happens when a single entity manages both the automaker and the ISSP. As alluded to in (**?**) and **??**, such a configuration is more optimal than the Nash game. Thus, to isolate effects that might arise from competition from those that arise from lack of established liability rules, considering a single entity is the best course. This monopolist maximizes profits according to the following:

$$\begin{aligned}\max_{\substack{\tau, z^M, z^I \\ \phi^r, \phi^s}} \quad & (\tau - c^M - \tilde{p}^r \cdot \phi^r - \tilde{p}^s \cdot \phi^s) \cdot \lambda - c^I \cdot \eta^s \cdot l \\ \text{s.t.} \quad & \text{Equations } \textbf{6.1a} \text{ to } \textbf{6.1h}\end{aligned} \tag{MO}$$

Assuming an interior solution exists, the FONC for (MO) yields:

$$\phi^j = s^j \quad \forall j \in \{r, s\} \tag{6.4a}$$

$$\tau = c^M + \phi^r \cdot \tilde{p}^r + \phi^s \cdot \tilde{p}^s + \frac{\lambda}{\Lambda' \cdot U'(W^0 - \tau)} \tag{6.4b}$$

$$\frac{C^{M'}(z^M)}{l} = -P_1(z^M, 0) \cdot \frac{1 - \tilde{p}^r}{1 - p^r} \cdot \phi^r - \eta^s \cdot \left[P_1(z^M, z^I) \cdot \frac{1 - \tilde{p}^s}{1 - p^s} - P_1(z^M, 0) \cdot \frac{1 - \tilde{p}^r}{1 - p^r}\right] \cdot \phi^s \tag{6.4c}$$

$$\frac{C^{I'}(z^I)}{\lambda} = -P_2(z^M, z^I) \cdot \frac{1 - \tilde{p}^s}{1 - p^s} \cdot \phi^s \tag{6.4d}$$

Here too, in keeping with the classical literature on insurance markets, if the monopolist provides coverage at all, it must cover all potential damages (**Equation 6.4a**). However, this naturally occurs at a higher vehicle cost than that chosen by the planner (**Equation 6.4b**).

6.3.2.2 Generalized Nash Equilibrium

What happens, however, when we allow for the possibility of shared liability between the two producers operating independently? To study this question, we introduce τ^M and τ^I, the prices charged by the manufacturer and the ISSP, respectively. We also introduce $\phi^{s,M}$ and $\phi^{s,I}$, the coverage provided by the manufacturer and the ISSP on the smart road, respectively. This allows us to study the Nash game between the two entities. If we let $\mathbf{a}^k = \{\boldsymbol{\tau}^k, \boldsymbol{\phi}^k, \mathbf{z}^k\}$ denote the action vector of player $k \in \{M, I\}$, $\{\mathbf{a}^{M,*}, \mathbf{a}^{I,*}\}$ constitutes a Generalized Nash Equilibrium (GNE) if $\mathbf{a}^{M,*}$ solves the following problem, taking $\mathbf{a}^{I,*}$ as given:

$$\begin{aligned}\max_{\substack{\tau^M, z^M, \\ \phi^r, \phi^{s,M}}} \quad & (\tau^M - c^M - \tilde{p}^r \cdot \phi^r - \tilde{p}^s \cdot \phi^{s,M}) \cdot \lambda \\ \text{s.t.} \quad & \text{Equations } \textbf{6.1a} \text{ to } \textbf{6.1h}\end{aligned} \tag{N-M}$$

and $\mathbf{a}^{I,*}$ solves the following problem, taking $\mathbf{a}^{M,*}$ as given:

$$\begin{aligned}\max_{\substack{\tau^I, z^I, \\ \phi^{s,I}}} \quad & (\tau^I \cdot \eta^s \cdot l - \tilde{p}^s \cdot \phi^{s,I}) \cdot \lambda - c^I \cdot \eta^s \cdot l \\ \text{s.t.} \quad & \text{Equations } \textbf{6.1a} \text{ to } \textbf{6.1h}\end{aligned} \tag{N-I}$$

Using the FONCs of (N-M) and (N-I), we derive the following equations that characterize the equilibrium:

$$\phi^r = s^r \tag{6.5a}$$

$$\phi^{s,M} + \phi^{s,I} = s^s \tag{6.5b}$$

$$\tau^M = c^M + \phi^r \cdot \tilde{p}^r + \phi^{s,M} \cdot \tilde{p}^s + \frac{\lambda}{\Lambda' \cdot U'(W^0 - \tau^M - \tau^I)} \tag{6.5c}$$

$$\tau^I \cdot \eta^s \cdot l = \phi^{s,I} \cdot \tilde{p}^s + \frac{\lambda}{\Lambda' \cdot U'(W^0 - \tau^M - \tau^I)} \tag{6.5d}$$

$$\frac{C^{M\prime}(z^M)}{l} = -P_1(z^M, 0) \cdot \frac{1-\tilde{p}^r}{1-p^r} \cdot \phi^r - \eta^s \cdot \left[P_1(z^M, z^I) \cdot \frac{1-\tilde{p}^s}{1-p^s} - P_1(z^M, 0) \cdot \frac{1-\tilde{p}^r}{1-p^r}\right] \cdot \phi^{s,M} \tag{6.5e}$$

$$\frac{C^{I\prime}(z^I)}{\lambda} = -P_2(z^M, z^I) \cdot \frac{1-\tilde{p}^s}{1-p^s} \cdot \phi^{s,I} \tag{6.5f}$$

Here too, full coverage is provided to AV owners (**Equations 6.5a** and **6.5b**). However, this occurs at a higher cost than that charged by a single entity because of double marginalization (**Equations 6.5c** and **6.5d**). Moreover, who bears most of the cost of providing that coverage on the smart road is unclear. As such, multiple equilibria might exist, ranging from one of the producers providing no coverage to both of them sharing the liability equally. In this context, it appears that both the automaker and the ISSP could offer different coverage options to customers. Then, the cost of vehicles and their quality, on the one hand, and of smart road use, on the other, could differ among customers. If infrastructure technology acts as a substitute for vehicle technology, then vehicles with lower coverage $\phi^{s,M}$ from the automaker on the smart road would naturally be of lower quality. In the case of complementarity between the two technologies, higher coverage from the manufacturer on the smart road would result in higher vehicle quality. In the context of the Nash game, the planner might need to establish a liability rule, in the manner of Di et al. (2020), to ensure that no undesirable equilibrium is reached (e.g.: one of the agents free-riding from the other's quality investment). We will discuss the design of these rules in depth in the next subsection.

6.3.3 Nash game with liability rule

Here, we wish to investigate whether a liability rule can prevent the Nash game to settle at an undesirable equilibrium. To answer this question, we introduce $\alpha^M(\cdot,\cdot)$, the share function for the manufacturer. $\alpha^M(\cdot,\cdot)$ determines the fraction of total damages paid that the manufacturer must shoulder when an accident occurs on the smart portion of the road. If we let $z^{M,**}$ and $z^{I,**}$ denote any target equilibrium, then $\{\mathbf{a}^{M,*}, \mathbf{a}^{I,*}\}$ constitutes a Generalized Nash Equilibrium (GNE) if $\mathbf{a}^{M,*}$ solves the following problem, taking $\mathbf{a}^{I,*}$ as given:

$$\begin{aligned} \max_{\substack{\tau^M, z^M, \\ \phi^r, \phi^{s,M}}} \quad & (\tau^M - c^M - \tilde{p}^r \cdot \phi^r - \tilde{p}^s \cdot \phi^{s,M}) \cdot \lambda \\ \text{s.t.} \quad & \text{Equations } \mathbf{6.1a} \text{ to } \mathbf{6.1h}, \\ & \phi^{s,M} \cdot \lambda \geq \alpha^M \left(z^{M,**} - z^M, z^{I,**} - z^I \right) \cdot s^s \cdot \lambda \end{aligned} \tag{N-MR}$$

and $\mathbf{a}^{I,*}$ solves the following problem, taking $\mathbf{a}^{M,*}$ as given:

$$\begin{aligned} \max_{\substack{\tau^I, z^I, \\ \phi^{s,I}}} \quad & (\tau^I \cdot \eta^s \cdot l - \tilde{p}^s \cdot \phi^{s,I}) \cdot \lambda - c^I \cdot \eta^s \cdot l \\ \text{s.t.} \quad & \text{Equations } \mathbf{6.1a} \text{ to } \mathbf{6.1h}, \\ & \phi^{s,I} \cdot \lambda \geq \left[1 - \alpha^M \left(z^{M,**} - z^M, z^{I,**} - z^I \right) \right] \cdot s^s \cdot \lambda \end{aligned} \tag{N-IR}$$

:

(In the above, pre-multiplying by λ simplifies a lot of the subsequent calculus). We make the following assumptions on $\alpha^M(\cdot, \cdot)$:

Assumption 6.4. *We assume the following:*

A4.1 *The share function is strictly decreasing in the automaker's choice of automation:* $\alpha_1^M > 0$.

A4.2 *The share function is strictly increasing in the ISSP's choice of digitalization:* $\alpha_2^M < 0$.

A4.3 *The share function is strictly positive and below 1:* $0 < \alpha^M < 1$

It is important to note that the shape and functional form of $\alpha(\cdot, \cdot)$ will differ depending on the equilibrium targeted by the planner. In some instances, for example, the desired target might yield $\alpha(0, 0) = \frac{1}{2}$, so that the equal sharing in liability is the effective desired targets. At others, this might change. However, by our assumptions, both agents will always hold some liability, however infinitesimal.

By deriving the FONC for N-MR and N-IR and assuming **Equation 6.2c** and **Equation 6.2d** hold, we can derive sufficient conditions for a share function to replicate the desired quality levels. Indeed, let $\delta^M \geq 0$ and $\delta^I \geq 0$ denote the Lagrangian multiplier associated with the share constraints for the automaker and the ISSP, respectively. Then, sufficient conditions for the GNE to yield $z^{M,**}$, $z^{I,**}$ are that there exists $\delta^M > 0$ and $\delta^I > 0$ satisfying the following two equations:

$$\alpha^M(0, 0) \cdot P_2(z^M, z^I) \cdot \frac{1 - \tilde{p}^s}{1 - p^s} - \delta^I \cdot \alpha_2^{M\prime}(0, 0) = 0 \tag{6.6a}$$

$$\eta^s \cdot [1 - \alpha^M(0,0)] \cdot P_1(z^M, z^I) \cdot \frac{1 - \tilde{p}^s}{1 - p^s} + \delta^M \cdot \alpha_1^{M\prime}(0,0) = 0 \tag{6.6b}$$

By Assumptions **A4.1**, **A4.2** and **A4.3**, the sufficiency condition is always met. Thus, an appropriate liability rule can reproduce the first-best automation and digitalization levels. In practice, however, implementing such a rule would be costly. Indeed, it would involve expanding resources to determine the quality levels of sensors used by both entities. In the presence of certain standards, this cost could be reduced and quality could be ensured *ex ante*, *a priori*.

6.4 Numerical examples

6.4.1 Setup

In order to garner more insights from our model, we now proceed to carry out numerical experiments. Parameters and their values are provided in an appendix (**Table D.1**). We assume the following functional forms:

$$U(W) = \log W \tag{6.7a}$$

$$\Lambda(V) = \lambda^0 \cdot \frac{\exp V}{\exp V + \exp V^0} \tag{6.7b}$$

$$P(z^M, z^I) = \frac{\tilde{p}^0}{1 + b^M \cdot z^M + b^I \cdot z^I} \tag{6.7c}$$

$$C^k(z^k) = a^{0,k} + a^{1,k} \cdot (z^k)^2 \quad \forall k \in \{M, I\} \tag{6.7d}$$

$$\tilde{p}^r = 1 - (1 - p^r)^{(1-\eta^s)\cdot l \cdot \zeta} \tag{6.7e}$$

$$\tilde{p}^s = 1 - (1 - p^s)^{\eta^s \cdot l \cdot \zeta} \tag{6.7f}$$

6.4.2 Results

We can now discuss the results obtained from these numerical experiments, focusing on the first-best [1].

6.4.2.1 Effect of base crash probability

Here, we consider the effect of increasing $\tilde{p}^0$, the base crash probability (i.e. the probability of accident for a HDV) in **Figure 6.1**. This is useful when considering the effect of automation and digitalization on different types of accident. Naturally, because of increased

[1] We found results from the monopoly to be similar to the first-best. Meanwhile, results from the Nash game are not stable and thus provide no discernible trends.

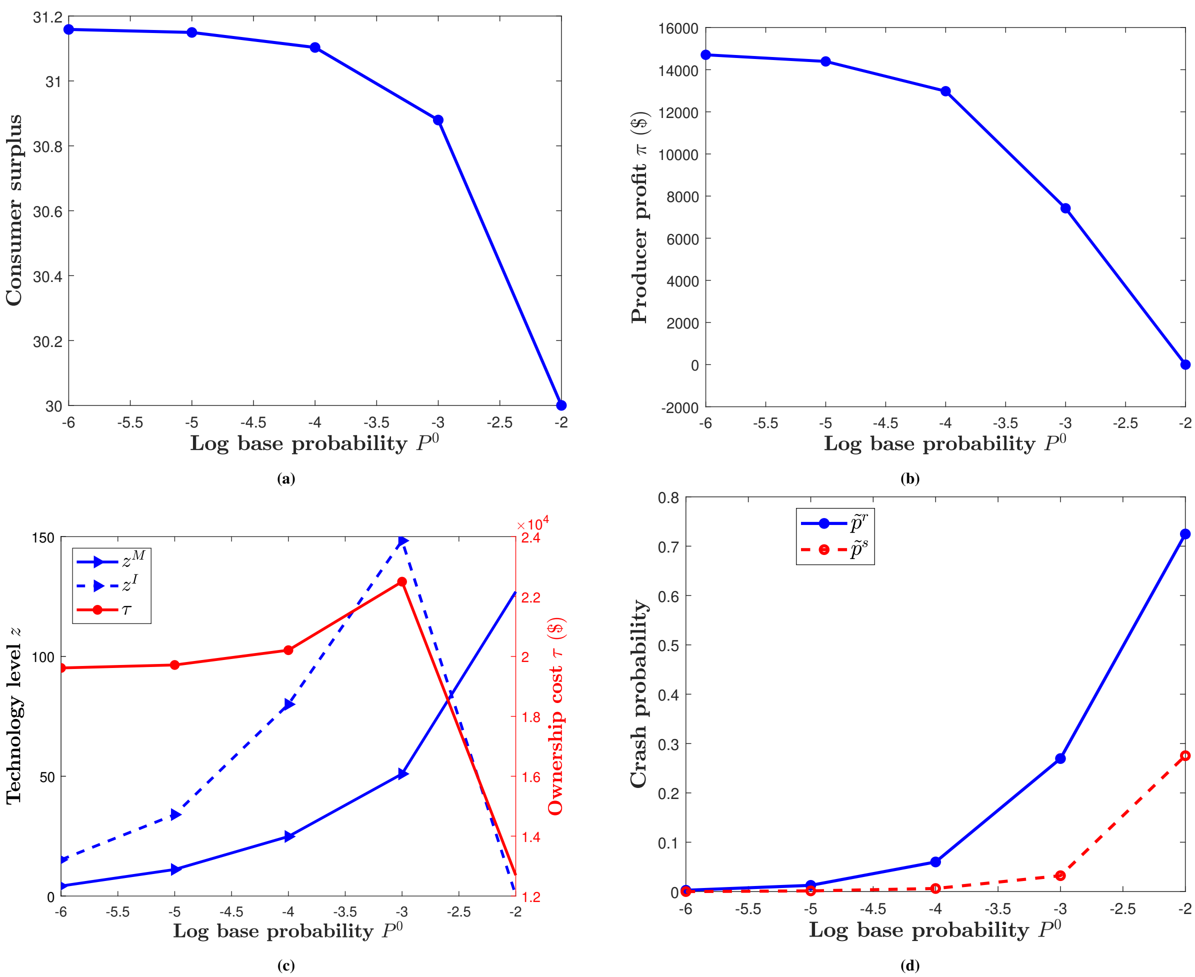


Figure 6.1: *Effect of base crash probability* $\tilde{p}^0$

crash risk and exposure to losses, both consumer surplus (**Figure 6.1a**) and producer surplus (**Figure 6.1b**) decrease. Meanwhile, both vehicle technology and infrastructure technology increase (**Figure 6.1c**). This leads to an increase in the price of the vehicle (**Figure 6.1c**). We also note that, as one should anticipate, providing both automation and digitalization (i.e. on the smart portion of the road) leads to lower crash probability than with automation alone (**Figure 6.1d**).

6.4.2.2 Effect of mileage l

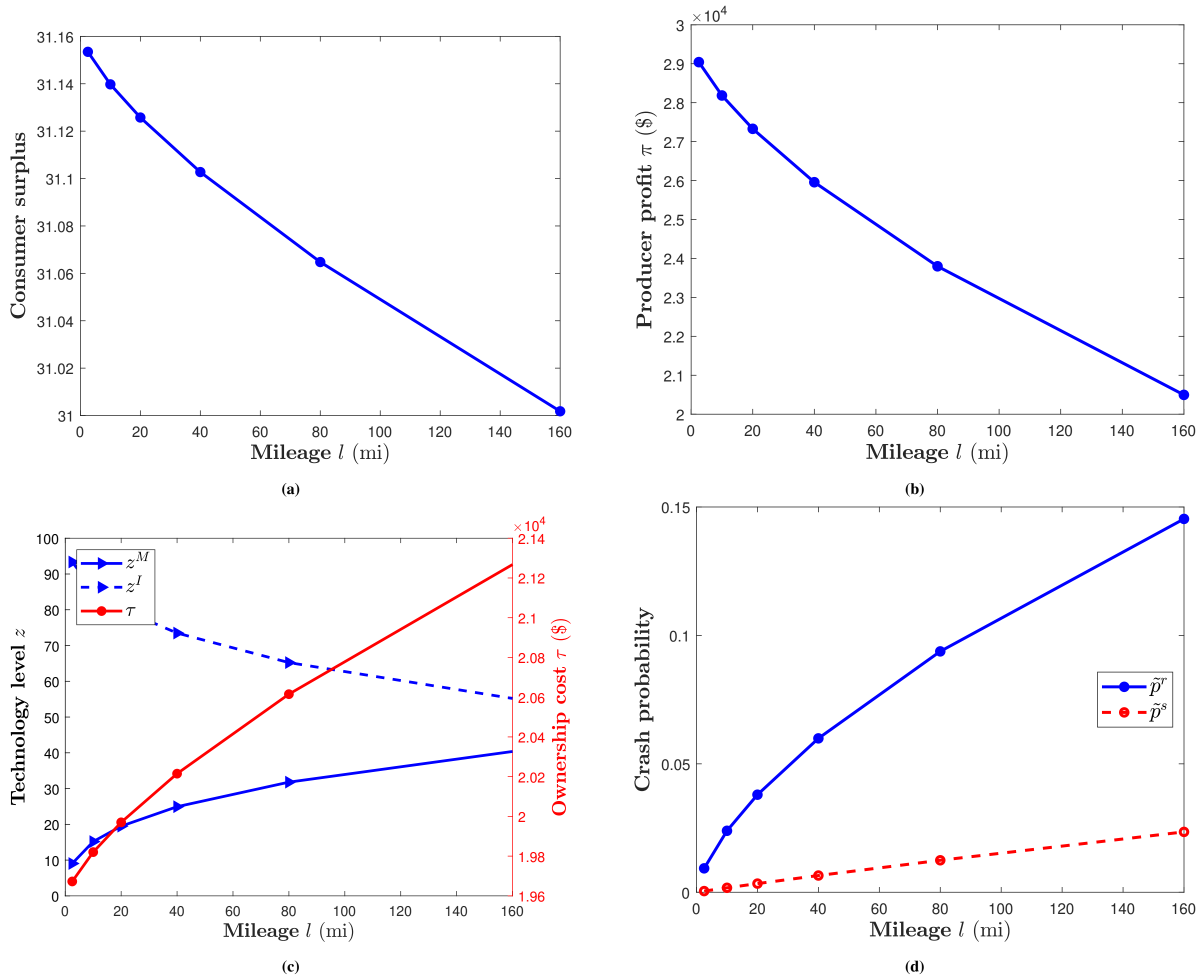


Figure 6.2: *Effect of mileage l*

What happens, in our model, when the length of the roadway–or, equivalently, the number of miles travelled–increases? Because driving only increases the risk of being involved in a crash (and we have not included benefits to driving in our model), increased mileage reduces total welfare (both consumer and producer surplus Figures 6.2a and 6.2b). Similarly to the effect of base crash probability, increased mileage also raises vehicle costs as well as investment in vehicle technology (**Figure 6.2c**). However, as predicted in our analysis, it reduces investment in infrastructure technology.

6.4.2.3 Effect of smart road share

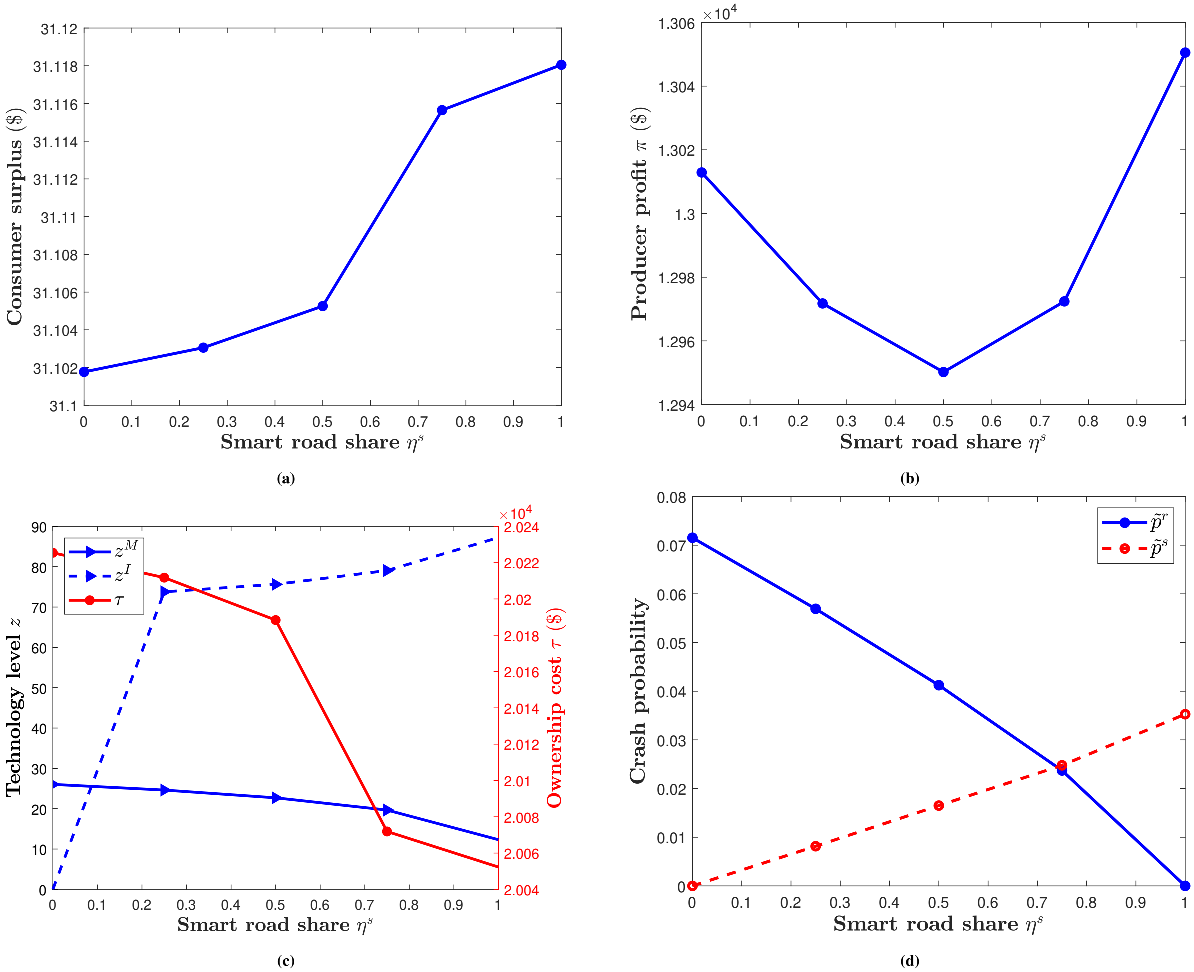


Figure 6.3: *Effect of smart road share* η^s

We now consider the effect of the fraction of roads equipped with smart infrastructure. Here, it is interesting to note that both consumer and producer surplus–and therefore total welfare–eventually increase (Figures 6.3a and 6.3b). This occurs because, when the share of smart roads increase, equipping vehicles becomes less and less important (**Figure 6.3c**). This reduces the marginal cost of production for vehicles, since equipping the roadway is independent of demand. Naturally, then, the utility of vehicle ownership increases for consumers. Moreover, increasing the share of smart roads limits the exposure to accidents on the regular portion of the road (**Figure 6.3d**). Since those tended to occur more frequently, the substitution away from regular roads also improves producer profits.

6.4.2.4 Effect of market size λ^0

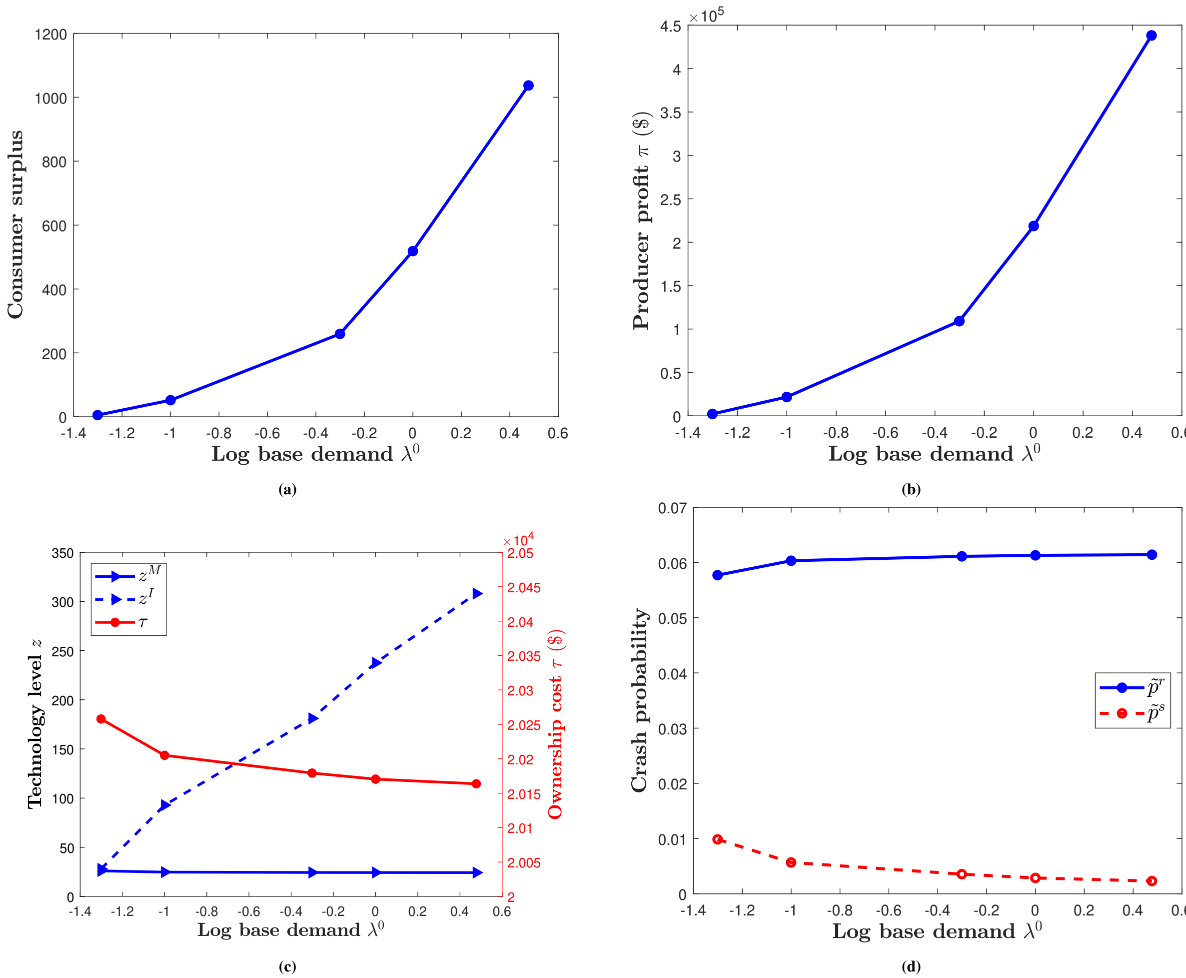


Figure 6.4: *Effect of market size* λ^0

Lastly, we consider the effect of market size on the optimal allocation. Here, as predicted by our analysis, an increase in the number of users increases the investment on infrastructure, though marginally. Thus, accident probability barely changes. However, the cost of vehicle ownership decreases–likely as a result of the marginal increase in infrastructure spending–and partially contributes to the increase in consumer surplus (Figures 6.4a and 6.4c)[2]. Additionally, because demand can grow independent from infrastructure costs in our example, profits also increase (**Figure 6.4b**).

[2]Higher market size being the other contributor

6.4.2.5 Discussion

From our analysis, a key takeaway seems to be that a reduction of car ownership cost through reduced or total elimination of insurance will likely only occur if infrastructure is involved. This is because of low accident probability in general and because infrastructure investment is independent of demand. Thus, marginal cost pricing ensures that vehicle owners do not pay for infrastructure investment, but only their liability coverage. Otherwise, without any infrastructure investment, the costs of any improvements in safety are directly passed on to the vehicle owner. Because accident probabilities are relatively low, there ensues that the total cost of ownership would actually increase. This can be clearly seen in **Figure 6.5**. Only when the base accident probability is unrealistically high (the annual crash rate in the US is well below 500 accidents per 100 million miles travelled (**?**)) does the ownership cost significantly decrease. The decrease is even more significant when the smart road share increases. Of course, one must bear in mind that there

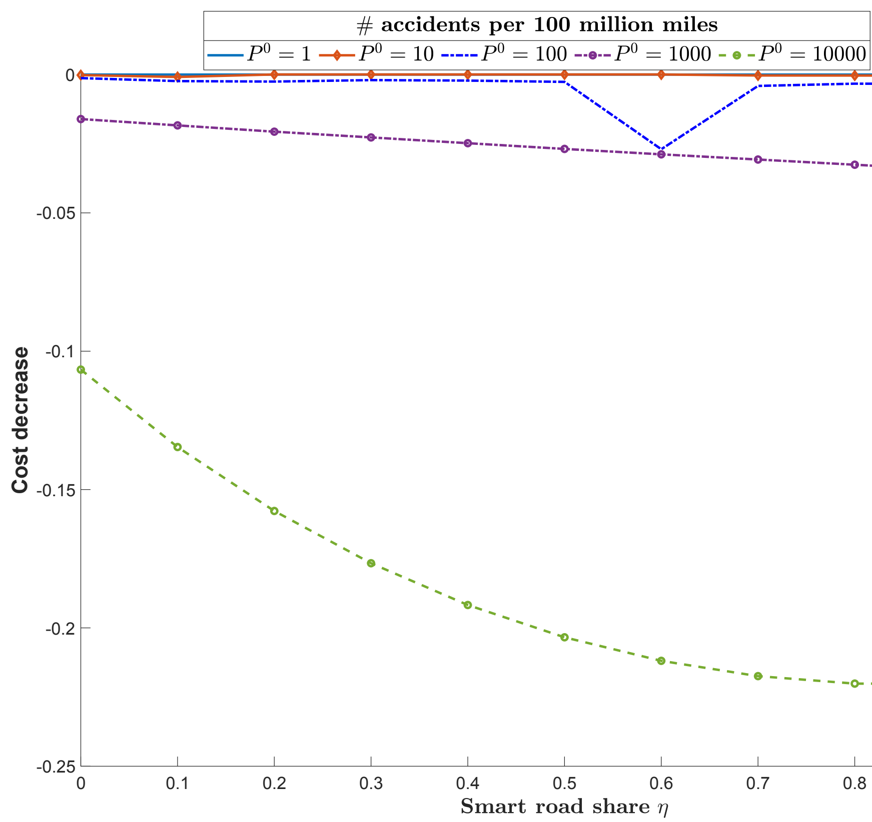


Figure 6.5: *Cost reduction relative to HDV as a function of road share and crash risk*

is no such thing as a free lunch and we are not calling for a return to the automated highway system idea[3]: infrastructure investment will likely be taxpayer-funded either directly or through

[3] An introduction to the AHS

concessions. Thus, it might not necessarily be, when fully considered, a suitable alternative.

Lastly, We must bear in mind that our analysis has, thus far, only considered a homogeneous traffic environment and costly means of increasing safety. In a mixed environment and assuming high crash probabilities, vehicle-side investment might have more value. Additionally, speed limits could be a *seemingly* costless alternative to smart infrastructure. Or, by reducing the traveling speeds and further reducing crash probabilities, speed limits could favor infrastructure even further. We have also not included the effect of density on accidents, and this could tilt the balance slightly in favor of infrastructure investment. We will consider the potential implications of including these in the following sections.

6.5 Future extensions

6.5.1 Speed selection

In the above, we considered a setting in which improving the safety of their products necessarily demanded costly investment from producers. However, by tailoring the settings of their offerings, both automaker and ISSP have an *a priori* costless mean of reducing their exposure. In particular, the automaker can limit the speed of its vehicles while the ISSP can limit the speed at which vehicles travel on the smart road. Consumers must then face the natural trade-off of speed vs safety and said trade-off will influence their choices. Denoting AVs' speed on portion j of the roadway by v^j, we introduce the following equations:

$$V = \tilde{p}^0 \cdot U(W^0) + \tilde{p}^r \cdot U(W^r) + \tilde{p}^s \cdot U(W^s) \tag{6.8a}$$

$$W^0 = \tilde{W}^0 - \tau - \beta \cdot \sum_j t^j \tag{6.8b}$$

$$W^j = \tilde{W}^0 - \tau - s^j - \beta \cdot \sum_k t^k + \phi^j \tag{6.8c}$$

$$t^j = \frac{\eta^j \cdot l}{v^j} \tag{6.8d}$$

$$s^j = S(v^j) \tag{6.8e}$$

$$p^r = P(z^M, 0, v^r) \tag{6.8f}$$

$$p^s = P(z^M, z^I, v^s) \tag{6.8g}$$

where t^j is the travel time on portion j; β is the value of time for AV users; $S(\cdot)$ is a function that relates crash severity to travel speed; and $P(\cdot,\cdot,\cdot)$ is now also a function of speed. We make the following assumptions on $S(\cdot)$ and $P(\cdot,\cdot,\cdot)$:

Assumption 6.5. *We assume the following:*

A5.1 *Crash severity is strictly increasing and strictly convex in travel speed:* $S' > 0$, $S'' > 0$.

A5.2 *Crash probability is strictly increasing and strictly convex in speed:* $P_3 > 0$ *and* $P_{33} > 0$

With speed selection, optimization problems (SO), (MO), (N-M), and (N-I) maintain the same objective function but now face the constraints outlined by **Equation 6.1a**, Equations **6.1d** to **6.1f**, Equations **6.1g** to **6.1h**, and Equations **6.8a** to **6.8g**.

The first-order necessary conditions (FONC) for optimality for these three problems remain unchanged except for the addition of two equations describing the equalization of the marginal time benefit of speeding to its marginal crash cost and a change in the coverage equation to include the effect of travel time:

$$\phi^j = s^j + \beta \cdot \sum_k t^k \quad \forall j \in \{r, s\} \tag{6.9a}$$

$$\beta \cdot \frac{\eta^j \cdot l}{(v^j)^2} = S'(v^r) + \frac{1}{\tilde{p}^j} \cdot \frac{\partial \tilde{p}^j}{\partial v^j} \cdot \phi^j \tag{6.9b}$$

Here, it is worth noting that where liability is assigned matters (beyond technology quality considerations). Indeed, higher liability for accidents on the regular road would lead to a decrease in speed on regular roads. Higher liability for accidents on smart roads would lead to lower speeds on smart roads. Moreover, the shorter the length of travel considered, the lower the equilibrium speed.

6.5.2 Mixed traffic environment

We now introduce human-driven vehicles (HDVs) and explore the potential implications for our findings. To this effect, we introduce the set of subscripts $\{a, h\}$ to differentiate quantities related to AVs and HDVs, respectively. Then, **Equation 6.8** becomes, $\forall i \in \{a, h\}$:

$$V_i = \tilde{p}_i^0 \cdot U(W_i^0) + \tilde{p}_i^r \cdot U(W_i^r) + \tilde{p}_i^s \cdot U(W_i^s) \tag{6.10a}$$

$$W_i^0 = \tilde{W}^0 - \tau_i - \phi_{p,i}^r - \phi_{p,i}^s - \beta_i \cdot \sum_j t_i^j \tag{6.10b}$$

$$W_i^j = \tilde{W}^0 - \tau_i - \phi_{p,i}^j - \phi_{p,i}^j - s_i^j - \beta_i \cdot \sum_k t_i^k + \phi_i^j \tag{6.10c}$$

$$t_i^j = \frac{\eta^j \cdot l}{v_i^j} \tag{6.10d}$$

$$s_i^j = S(v_i^j) \tag{6.10e}$$

$$p_i^r = \begin{cases} P(0, 0, v_i^r, v_{-i}^r, \rho^r) & \text{for } i = h \\ P(z^M, 0, v_i^r, v_{-i}^r, \rho^r) & \text{otherwise} \end{cases} \tag{6.10f}$$

$$p_i^s = \begin{cases} P(0, 0, v_i^s, v_{-i}^s, \rho^s) & \text{for } i = a \\ P(z^M, z^s, v_i^s, v_{-i}^s, \rho^s) & \text{otherwise} \end{cases} \tag{6.10g}$$

where ρ^j represents road density; $\phi_{p,i}^j$ are premiums paid by users; and $P(\cdot,\cdot,\cdot,\cdot,\cdot)$ is now a function of the speeds of all agents involved as well as of road density. For simplicity, we will assume that τ_h is fixed and exogenous and that there exists an insurance market providing insurance for HDVs. Moreover, $\phi_{p,a}^j = 0$ as before, since insurance costs are absorbed in the vehicle costs for AVs. The social welfare maximization problem now becomes:

$$\begin{aligned} \max_{\substack{\tau_i, z^M, z^I \\ \phi_i^j, v_i^j, \phi_{p,i}^j}} \quad & CS(V_a, V_h) + (\tau_a - c^M - p_{2,a} \cdot \phi_a^r - p_{3,a} \cdot \phi_a^s) \cdot \lambda_a + \\ & (\phi_{p,h}^r + \phi_{p,h}^s - p_{2,h} \cdot \phi_h^r - p_{3,h} \cdot \phi_h^s) \cdot \lambda_h - c^I \cdot \eta^s \cdot l \\ \text{s.t.} \quad & \text{Equations } \mathbf{6.1a} \text{ to } \mathbf{6.1h} \end{aligned} \tag{SO-M}$$

where $CS(\cdot,\cdot)$ is such that its i^{th} partial derivative $CS_i = \lambda_i$. At the optimal, full coverage naturally prevails for both HDVs and AVs. For HDVs, coverage on road j is obtained at a premium $\phi_{p,h}^j = p_h^j \cdot \phi_h^j$. Then, the FONC for the (SO-M) is the FONC from **Section 6.5.1** supplemented by the following equations:

$$\beta_i \cdot \frac{\eta^j \cdot l}{(v_i^j)^2} = \tilde{p}_i^j \cdot S'(v_i^j) + \frac{1}{\lambda_i} \cdot \sum_{k \in \{a,h\}} \frac{\partial \tilde{p}_k^j}{\partial v_i^j} \cdot \phi_k^j \cdot \lambda_k \quad \forall i \in \{a, h\}, \forall j \in \{r, s\} \tag{6.11a}$$

As can be seen from **Equation 6.11a**, speeds are further decreased, on both portions of the road, in mixed traffic. Should the value of time for AV users be extremely low, then speeds can be made even lower, thus reducing the risk for accidents and obviating a need for costly investments.

6.6 Conclusion

In this work, we have considered the potential implications of vehicle-infrastructure cooperation on safety and liability. We have first shown that, absent any expected efforts from drivers and assuming homogeneous risk profiles, a consortium of automakers and ISSPs will provide full coverage and endorse full liability for accidents involving AVs. When both automakers and ISSPs operate independently, however and even though full coverage is provided, it is unclear which of the two actors will bear the burden: the Nash game involves mutliple equilibria. In that context, an appropriate liability rule must be designed to ensure that undesirable equilibria do not occur. We show that, with a rule based on a share function, this can be readily achieved. However, implementing such a rule is non-trivial, thus favoring cooperation rather than competition in this automated mobility market. Numerical examples also offer a number of insights on the effects of automation and digitalization on vehicle costs. Namely, only when infrastructure technology is significantly involved in the provision of automated services and when accident probability without technology are significantly high does vehicle ownership costs decrease. Moreover, infrastructure investment seems to be more advantageous, from the customer perspective, than vehicle ownership: customers do not bear any marginal cost from infrastructure investments and accident probabilities are low enough that vehicle side investments would provide a low return.

While our model provides useful basic insights, we have also proposed extensions that need to be investigated further. Indeed, how does liability evolve in mixed traffic environments when both automakers and ISSPs can control travel speeds and when value of time is low for AV users? Answering this question will be the goal of our future work on this topic.

CHAPTER 7

Conclusions and Future Research

In this dissertation work, we have attempted to address policy issues surrounding two important technologies: e-hailing and automated vehicles.

We first showed that, when appropriately designed, a commission cap policy can effectively regulate the ride-hailing industry, even when accounting for competition between e-hailing and street-hailing. Coupled with our analysis of competition between e-hailing and street-hailing and of congestion externality, this work provides useful insights and direction for policy makers seeking to simplify the host of regulations that have historically been a feature of ride-hailing.

Then, we proceeded to investigate vehicle-infrastructure cooperation and assess whether it could be a viable alternative to a vehicle-centric approach to automated driving. This demanded the development of a model of an automated mobility market, the first of its kind. This model allowed us to analyze interactions between automakers and ISSPs and provided us the means to study the impact of market structure on automation, digitalization, and liability between these two entities.

The work in this dissertation lays the groundwork for a number of other exciting research avenues: optimal timing of infrastructure investment; evaluation of the safety benefits of ADAS systems; and evaluation of the welfare effects of automaker's entry in the insurance market.

First, our work on vehicle-infrastructure cooperation has focused on a static setting. While such a framework is useful, it fails to provide any insight to policy makers as to how to address the chicken-and-egg problem of digitization and automation. Indeed, on the one hand, without users of smart infrastructure technology, or without quick adoption and penetration of compatible AVs, investing in infrastructure would provide poor returns on investment. On the other, because infrastructure is meant to help make AVs a reality, lack of infrastructure might mean significantly delaying or foregoing altogether the benefits of AVs. In this context, we must understand how provision of specific infrastructure technology will affect adoption of AVs and the appropriate timing for these investments. This will demand the development and analysis of appropriate models, but also a more empirical approach than the present work.

The need for this empirical approach leads into the next possible research avenue. Since AVs have not yet entered the market, it is difficult to assess whether the benefits they promise justify the economic and social costs of their adoption. However, it is possible to use certain technologies as proxies to evaluate the safety benefits of AVs: advanced driver assistance systems (ADAS). These systems have slowly but readily penetrated the new vehicle markets and evaluating a) their effects of accident risk and b) customers' willingness to pay for these services could help us determine a somewhat reliable estimate of the effect AV adoption on consumer welfare through safety.

Lastly, as touched upon in this work, auto companies have started to use their vehicle technology to enter the insurance market. Over the past couple of years, some auto manufacturers (e.g.: GM, Ford...), have begun partnering with financial institutions to directly offer insurance to their customers while others—such as Tesla—have decided to become insurers themselves. Such a move has been enabled, in part, by automakers' access to usage data from their vehicles and customers rising interest in usage-based insurance (UBI). Such developments could lead to better competition and reduced insurance rates for consumers. Additionally, automakers' understanding of their proprietary technologies could also help drive down repair costs, thus lowering premiums further. However, evidence seems to suggest that safer drivers tend to self-select into monitoring programs. Thus, we could observe an environment in which insurer-automakers attract the safest drivers in their customer pool, while regular insurers are left with two types of monitoring-adverse drivers: safer drivers who experience disutility from being monitored and riskier drivers who prefer to keep their driving behavior hidden. This would limit regular insurers' ability to hedge against adverse selection, offer lower prices and compete. Lastly, to reduce their exposure to accident risks as insurers, automakers will likely aim to make a larger fraction of their products safer by equipping them with various safety and automation features. While such increased provision of safety features would contribute to making driving safer, it may also lead to higher prices for vehicles and could offset the reduction in ownership costs effected by lower insurance premiums.

APPENDIX A

Appendices for Chapter 1

A.1 Numerical examples

A.1.1 Demand data

Base demand λ^0 is obtained from Korolko et al. (2018). Background traffic demand is obtained by scaling λ^0 by 10^1 so that as λ^0 increases, so does congestion and the externality of ridesourcing (as would happen in practice). We show the base demand pattern in **Figure A.1**.

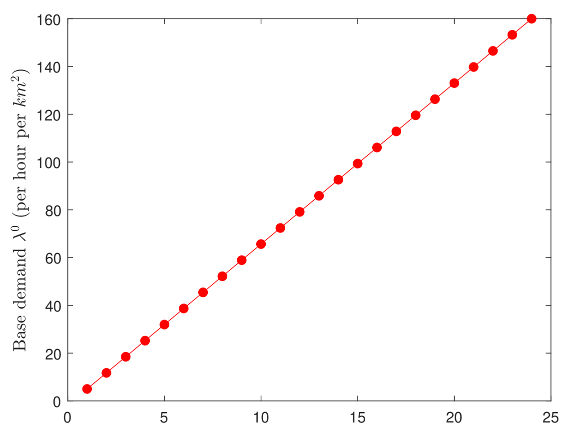


Figure A.1: *Base demand and background demand data*

A.1.2 Parameter values

All the parameters (except the bounds on the support of the value of time distribution) and their values are described in **Table A.1**.

The parameter values used for the different value of time distributions are given in **Table A.2**.

[1]Our results show that, accross demand levels, the number of ride-sourcing vehicles N is *at least* 10% of the number of background vehicles N^b, which seems reasonable.

Notation	Interpretation	Value
α	Meeting distance elasticity	0.5
ξ^s	Mode specific constant	0
γ	Correlation coefficient between customer and driver detour time	1
γ^b	Value of travel time for background traffic	$30\frac{\$}{\text{hr}}$
δ	Pairing probability	0.1
ϕ	Matching time window	5 min
θ	Marginal effect of ridesourcing vehicles on traffic	1
A	Scaling parameter for meeting time function	25
B	Scaling parameter for detour distance function	5
d^r	Average distance of ridesourcing trips	7 km
$d^{r,b}$	Average distance of background traffic trips	3.5 km
V^c	Slope of speed function	$-0.11\ \frac{\text{kph}\cdot\text{km}^2}{\text{veh}}$
V^0	Free-flow speed	50 kph
U^0	Trip utility	\$50
c	Cost per driver	\$10
μ^0	Cost of outside option	\$30

Table A.1: *Parameter values for numerical examples*

	$\bar{\beta}$	$\underline{\beta}$
High variance	75	5
Medium variance	60	20
Low variance	45	35

Table A.2: *Uniform distribution parameter values*

A.2 Finding regulatory policy

As outlined in **Section 2.5**, we consider a Stackelberg game between the planner (leader) and the monopolist (follower). The planner solves (REG) below[2]:

$$
\begin{aligned}
W = \max_{\substack{P^s \geq 0,\, F^s \geq 0,\\ P^p \geq 0,\, F^p \geq 0,\\ \tau^N \geq 0,\, R \geq 0}} \quad & U^0 \cdot (\lambda^s + \lambda^p \cdot o) - \lambda^0 \cdot \int_{\beta_2}^{\beta_1} \beta \cdot (w^m + w^r) \cdot G'(\beta) \cdot d\beta \\
& - \lambda^0 \cdot \int_{\underline{\beta}}^{\beta_3} \beta \cdot (w^p + w^m + w^r + \Delta w) \cdot G'(\beta) \cdot d\beta \\
& - R \cdot (\lambda^s + \lambda^p) + \tau^N \cdot N - \gamma^b \cdot \lambda^b \cdot w^{rb} \\
\text{s.t.} \quad & (F^s, F^p, R) = S(P^s, P^p, \tau^N)
\end{aligned}
\tag{REG}
$$

In the above, $S(P^s, P^p, \tau^N)$ is the best response function of the firm when subject to the regulation (P^s, P^p, τ^N). This best response is obtained by solving (M-CAPT) below:

$$
\begin{aligned}
\pi = \max_{\substack{F^s \geq 0,\, F^p \geq 0,\\ R \geq 0}} \quad & (F^s - R) \cdot \lambda^s + (F^p - R) \cdot \lambda^p \\
\text{s.t.} \quad & F^s - R \leq P^s, \\
& F^p - R \leq P^p
\end{aligned}
\tag{M-CAPT}
$$

Moreover, the positivity constraint on P^s and P^p ensure that the regulation is implementable (the platform can generate a profit and thus operates). The positivity constraint on τ^N enforces a no-subsidy constraint on the outcome of the regulation.

As formulated, the problem is a bilevel program, a class of optimization problems difficult to solve. For our numerical experiments, we used a heuristic solution procedure. In the process, we solve (M-CAPT) with the current regulation policy (P^s_k, P^p_k, τ^N_k). Then, holding the firm's response constant, we solve (REG) and use the method of successive averages (MSA) to update the regulation policy for the next step. Since our algorithm may not necessarily converge and given that we do not consider whether the updated policy effectively increases welfare, we keep track of the best solution obtained up to the current step. Thus, after solving for the firm's best response to (P^s_k, P^p_k, τ^N_k), we compare the welfare resulting from the monopolist's response, W_k, to the maximum realized welfare up to k, W_{best}.

The above solution procedure yielded reasonably good solutions, as demonstrated in **Figure 2.6**. The solutions may be further improved by applying a derivative-free method such

[2] Note that the revenue from tolls needs to be added to the social welfare function. Otherwise there would be a missing transfer in our system.

as Nelder–Mead or pattern search, as the decision variables of REG are of a low dimension.

APPENDIX B

Appendices for Chapter 2

B.1 Nomenclature

Variable	Description	Unit
μ_i	Generalized cost of service i	$\$$
λ_i	Demand density rate for service i	$/mi^2/hr$
f_i	Trip fare for service i	$\$$
w_i^m	Average customer waiting time on service i	hr
w^r	Average trip time	hr
d_i^m	Distance between customer and closest available driver on service i	mi
d^r	Average trip distance	mi
v	Traffic speed	mi/hr
n_i^I	Density of idle/cruising drivers on service i	$/mi^2$
n_i	Driver density on service i	$/mi^2$
ω_i	Hourly driver earnings on service i	$\$$
w_i^d	Average pickup time experienced by drivers on service i	hr
$\overline{mc}_i$	Marginal cost for service i	$\$$
mc_i	Marginal cost per unit time for service i	$\$/hr$
$\bar{p}_i$	commission cap imposed on service i	$\$$
$\bar{\tau}_i$	Toll per unit time imposed on service i vehicles	$\$/hr$
λ^b	Trip density rate for background vehicles	$/mi^2/hr$
w^{rb}	Average travel time for background vehicles	hr
n^b	Density of background vehicles	$/mi^2$

Table B.1: *Variable description*

B.2 Competition in friction-less environment

In order to show that, in the absence of friction, competition necessarily leads to lower fares, we will consider a perfectly elastic labor supply with reservation wage c. Then, both services are identical and the system of equations that describes the operation of this friction-less hailing service is as follows:

$$\lambda_i = \Lambda_i(\mu_i, \mu_{-i}) \tag{B.1a}$$

$$\mu_i = f_i + \beta \cdot w^r \tag{B.1b}$$

$$w^r = \frac{d^r}{v} \tag{B.1c}$$

$$n_i = \lambda_i \cdot w^r \tag{B.1d}$$

The above system is actually determined by two exogenous variables which we select to be the fares. Then, the Nash equilibrium can be described by the following set of equations:

$$f_i = -\frac{f_i}{\epsilon_{ii}} + c \cdot w^r \quad \forall i \tag{B.2}$$

It is trivial to show that the pricing equation for a monopoly operating a single service is identical to **Equation B.2**:

$$f^{sm} = -\frac{f^{sm}}{\epsilon^{sm}} + c \cdot w^r$$

However, at any level of supply and demand combination, the elasticity of demand faced by the monopoly is lower, in absolute terms, than that faced by a duopolist: $|\epsilon^{sm}| < |\epsilon_{ii}|$. It follows that, at equilibrium, the fare is always lower in the duopoly case for this friction-less service. If the monopoly operates both services, the same conclusion holds, since $\epsilon_{ii} - \epsilon_i^{dm} = \frac{\epsilon_{-ii} \cdot \epsilon_{i-i}}{\epsilon_{-i-i}} < 0$.

B.3 Derivation under congestion

We apply the same analysis techniques as in **Section 3.4** and **Section 3.5**.

B.3.1 Nash game and integrated monopoly

Under congestion, the Nash equilibrium can be described with the following set of equations:

$$f_i = mc_i \cdot (w_i^d + w^r) - \frac{f_i}{\epsilon_i^{ne}} \tag{B.1a}$$

$$mc_i = c_i + \omega_i \cdot \left[1 + \frac{\eta_{-i-i}}{\eta_{ii} \cdot \eta_{-i-i} - \eta_{-ii} \cdot \eta_{i-i}}\right] + \tau_i^o + \tau_i^{ne} \tag{B.1b}$$

where

$$\tau_s^o = -\beta \cdot \lambda_s \cdot [d_s^m - d^r \cdot \lambda_s \cdot w_s^{m'}] \cdot \delta^c > 0 \tag{B.2a}$$

$$\tau_e^o = -\beta \cdot \lambda_e \cdot \frac{d_e^m - d^r \cdot \lambda_e \cdot w_e^{m'}}{1 + \lambda_e \cdot w_e^{m'}} \cdot \delta^c > 0 \tag{B.2b}$$

$$\tau_s^{ne} = -\frac{f_s}{\epsilon_s^{ne}} \cdot [d_e^m - d^r \cdot \lambda_e \cdot w_e^{m'}] \cdot \frac{\Lambda_{se}}{1 + w_e^{m'} \cdot \Big[1 + \beta \cdot (w_e^m + w^r) \cdot \Lambda_{ee}\Big]} \cdot \delta^c < 0 \tag{B.2c}$$

$$\tau_e^{ne} = -\frac{f_e}{\epsilon_e^{ne}} \cdot [d_s^m - d^r \cdot \lambda_s \cdot w_s^{m'}] \cdot \frac{\Lambda_{es}}{1 + \beta \cdot w^r \cdot d_s^{m'} \cdot \Lambda_{ss}} \cdot \delta^c < 0 \tag{B.2d}$$

$$\delta^c = \frac{V'}{v^2 + \lambda^b \cdot d^{rb} \cdot V'} < 0 \tag{B.2e}$$

From **Equation B.1a**, it is easy to note that a competing ride-hailing platform internalizes the congestion externality that its marginal customer imposes on its other customers (τ_i^o). However, the term τ_i^{ne} and the absence of τ_{-i}^o in firm's i pricing equation also indicate that the platform considers the externality that this marginal customer imposes on customers on the competing platform. Such externality is advantageous from a competitive standpoint and incentivizes the company to price its rides less than it would in the absence of a competitor. Indeed, if we consider an integrated company, the pricing equations become:

$$f_i = mc_i \cdot (w_i^d + w^r) - \frac{f_i}{\epsilon_i^{dm}} + \frac{f_{-i}}{\epsilon_i^{dm}} \cdot \frac{\lambda_{-i}}{\lambda_i} \cdot \frac{\epsilon_{-ii}}{\epsilon_{-i-i}} \tag{B.3a}$$

$$mc_i = c_i + \omega_i \cdot \left[1 + \frac{1}{\eta_i^{dm}}\right] - \frac{\omega_{-i}}{\eta_i^{dm}} \cdot \frac{n_{-i}}{n_i} \cdot \frac{\eta_{-ii}}{\eta_{-i-i}} + \sum_j \tau_j^o \tag{B.3b}$$

Comparing **Equations 3.16b** and **B.1b**, it is easy to note that τ_i^{ne} disappears from the company's pricing and that τ_{-i}^o is also accounted for. Thus, it appears that competition has the effect of worsening traffic congestion.

B.3.2 First-best

The solution to the first-best problem:

$$W = \max_{\substack{f_i \geq 0,\\ r_i \geq 0}} \quad CS(\mu_1, \mu_2) + \sum_i \Big[f_i \cdot \lambda_i - \omega_i \cdot n_i\Big] + DS(\omega_1, \omega_2) - \beta^b \cdot \lambda^b \cdot w^{rb} \tag{FBC}$$

can be readily derived as:

$$f_i = mc_i \cdot (w_i^d + w^r) \tag{B.4a}$$

$$mc_i = c_i + \omega_i + \sum_j \tau_j^o + \tau^b \tag{B.4b}$$

where

$$\tau^b = -\beta^b \cdot \lambda^b \cdot d^{rb} \cdot \delta^c > 0 \tag{B.5}$$

As the inclusion of τ^b indicates, ride-hailing customers must now pay for the externality they impose on the background traffic. Moreover, just as in the congested monopoly case of Vignon et al. (2021), the first-best is sustainable when congestion is high enough:

$$\pi_i = -(c_i + \omega_i) \cdot n_i^I + (\tau_i^o + \tau^b) \cdot (w_i^d + w^r) \cdot \lambda_i \tag{B.6}$$

B.3.3 Regulation

We focus on the case of regulation using a cap $\bar{p}_i$ and a toll $\bar{\tau}_i \in [0, \tau^b]$ on ride-hailing vehicles. The case for the integrated monopoly is similar to the case of a monopoly with product differentiation considered in Vignon et al. (2021). Under our regulation, the Nash game becomes:

$$\begin{aligned} \pi_i = \max_{\substack{f_i \geq 0, \\ n_i \geq 0}} \quad & f_i \cdot \lambda_i - (c_i + \omega_i + \bar{\tau}_i) \cdot n_i \\ \text{s.t.} \quad & (\omega_i + c_i + \bar{\tau}_i) \cdot n_i \geq (f_i - \bar{p}_i) \cdot \lambda_i \end{aligned} \tag{NEC-REG}$$

Let θ_i be the Lagrangian multiplier associated with the commission cap constraint. Then, a sufficient condition for the cap to replicate the first-best is given by:

$$\theta_i = \frac{1}{f_i^{ne} - (c_i + \bar{\tau}_i) \cdot \frac{n_i}{\lambda_i}} \cdot \left[\frac{f_i^{ne}}{\epsilon_i^{ne}} - (\tau^b - \bar{\tau}_i) \cdot \frac{n_i - n_i^I}{\lambda_i} \right] \geq 0 \tag{B.7}$$

where f_i^{ne} is the fare according to **Equation 3.13a** and all quantities are evaluated at the first-best. Now, because the first-best is sustainable, it must be that $f_i^{fb} > \tau^b \cdot \frac{n_i - n_i^I}{\lambda_i}$ where f_i^{fb} is the fare of service i under the first-best. Then, it follows that $f_i^{ne} > \tau^b \cdot \frac{n_i - n_i^I}{\lambda_i}$ and $f_i^{ne} - (c_i + \bar{\tau}_i) \cdot \frac{n_i}{\lambda_i} > 0$. It follows that the above condition is met, regardless of the value of $\bar{\tau}_i$ so long as the first-best is sustainable.

B.4 Numerical experiments

Functionally, we also assume that demand follows a logit model with dispersion parameter κ:

$$\lambda_i = \lambda^0 \cdot \frac{\exp(-\kappa \cdot \mu_i)}{\sum_j \exp(-\kappa \cdot \mu_j) + \exp(-\kappa \cdot \mu^0)} \tag{B.1}$$

We also assume that speed is a linear function of the number of vehicles in the system:

$$v = v^0 - v^c \cdot \left(\sum_i n_i + n^b \right) \tag{B.2}$$

All the parameters and their values are described in **Table B.1**. We obtain a_1 and a_2 following Zhang et al. (2019). We also obtain β, and κ from Zhang and Nie (2021b). Our market sizes range from 500 to 19025. We assume that supply is perfectly elastic with cost c^0 and we set $\lambda^{b0} = 10000$.

Notation	Interpretation	Value
β	Value of travel time	$27.69\frac{\$}{\text{hr}}$
β^b	Value of travel time for background traffic	$20\frac{\$}{\text{hr}}$
κ	Demand logit dispersion	0.5
κ^b	Background elasticity	0.01
a_1	Efficiency parameter street-hail	38.69
a_2	Efficiency parameter e-hail	1.625
d^r	Average trip distance	4 mi
$d^{r,b}$	Average background trip distance	5 mi
v^c	Slope of speed function	$0.0025\ \frac{\text{mph}\cdot\text{mi}^2}{\text{veh}}$
v^0	Free-flow speed	30 mph
μ^0	Cost of outside option	$\$14.68$
c^0	Driver opportunity cost	$\$23$

Table B.1: *Parameter values for numerical examples*

APPENDIX C

Appendices for Chapter 3

C.1 Parameter values

C.1.1 User groups

We identify user groups and their population distribution using data from the 2017 National Household Travel Survey (NHTS) (McGuckin and Fucci, 2018). We also use the total VMT implied by the data to compute a VMT distribution suitable for our numerical examples. The resulting parameters can be seen in **Table C.1**.

i	User type	Population size (millions)	Daily VMT v_i (miles)
1	Age under 20	60.75	75
2	Over 65	61.80	45
3	Age $21-65$	167.15	15

Table C.1: *Parameter values for numerical examples*

C.1.2 Road groups

We identify road types in our example using classification from the Federal Highway Administration (FHWA) (NYDOT). We then identify the relevant parameters in **Table C.2** using FHWA 2018 statistics for VMT, total miles built, number of lane-miles for each category. Capacity for each road type is determined using guidelines from Margiotta and Washburn (2017).

C.1.3 Other parameters

The other parameters used in our numerical example are listed in **Table C.3**. f_0 and b_0 are approximated by using estimates of the total savings at full automation in Clements and Kockelman

k	Road type	Length (ten thousands)	Capacity V_k^{max} (thousands)
1	Local	290.69	4.79
2	Collector	79.2	5.01
3	Arterial	46.15	7.07

Table C.2: *Parameter values for numerical examples*

(2017) and the total yearly VMT in the US. The base manufacturing cost is obtained by using the average selling price of a car in 2016 (Statista, 2019) and assuming $\$4,000$ in profits for the car manufacturer. κ_c and κ_r are calculated by assuming a discount factor of 5% and assuming that users will own their car for 7 years while operators will operate the roads for 25 years. $m_{c,0}$ and $m_{r,0}$ will likely be higher in practice, but these are values that make our numerical experiments easier.

Notation	Interpretation	Value
f_0	User monetary value of a mile driven	¢5
b_0	Operator monetary value of a mile driven	¢15
$f_{c,0}$	Cost of congestion	¢10
c_0	Base vehicle manufacturing cost	$\$30,000$
c_1	Manufacturing cost per unit of automation	$\$5,000$
$m_{c,0}$	Vehicle investment cost per unit of automation	$\$155,000$
$m_{r,0}$	Road unit investment per unit of digitalization	$\$5,000$
α	Utility automation share	0.6
ρ	Substitution parameter	1
v	Degree of homogeneity	2
η_c	Customer congestion elasticity	0.2
η^r	Operator congestion elasticity	0.4
γ	Operator automation elasticity	0.2
θ	Operator digitalization elasticity	0.4
κ_c	Amortization parameter for vehicle purchase	0.15
κ_r	Amortization parameter for road investment	0.07

Table C.3: *Parameter values for numerical examples*

APPENDIX D

Appendices for Chapter 4

D.1 Parameter values

Notation	Interpretation	Value
$a^{0,k}$	Base investment cost for $k \in \{M, I\}$ (\$)	$\{0, 25000\}$
$a^{1,k}$	Unit investment cost for $k \in \{M, I\}$ (\$)	$\{5, 20\}$
b^k	Probability function parameter for $k \in \{M, I\}$	$\{0.5, 0.25\}$
l	Roadway length / Mileage per trip (mi)	40
η^s	Smart portion of the roadway	0.2
ζ	Number of times driven on roadway	260
λ^0	Base travel demand	3000
P^0	Base crash probability $(/mi)$	10^{-4}
V^0	Value of the outside alternative to car ownership	10
W^0	Initial budget (\$)	30000
s^j	Accident severity for $j \in \{r, s\}$ (\$)	$\{18000, 35000\}$
κ_c	Amortization parameter for vehicle purchase	0.17
κ_r	Amortization parameter for road investment	0.07

Table D.1: *Parameter values for numerical examples*

BIBLIOGRAPHY

Agarwal, S., Mani, D., and Telang, R. (2019). The Impact of Ride-hailing Services on Congestion: Evidence from Indian Cities. SSRN Scholarly Paper ID 3410623, Social Science Research Network, Rochester, NY.

Anand, P. (2020). The Information's 411 — Uber Drowning in the Pool.

Arnott, R. (1996). Taxi Travel Should Be Subsidized. *Journal of Urban Economics*, 40(3):316–333.

Atha, K., Callahan, J., Chen, J., Drun, J., Green, K., Lafferty, B., McReynolds, J., Mulvenon, J., Rosen, B., and Walz, E. (2020). China's Smart City Development. Technical report, US-China Economic and Security Review Commission.

Barrett, S. D. (2003). Regulatory Capture, Property Rights and Taxi Deregulation: A Case Study. *Economic Affairs*, 23(4):34–40. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0270.2003.00441.x.

Beesley, M. E. (1973). Regulation of Taxis. *The Economic Journal*, 83(329):150.

Beesley, M. E. (1979). Competition and Supply in London Taxis. *Journal of Transport Economics and Policy*, 13(1):102–131. Publisher: [London School of Economics, University of Bath, London School of Economics and University of Bath, London School of Economics and Political Science].

Bellon, T. (2019). Uber to limit drivers' app access to comply with NYC regulation. *Reuters*.

Benjaafar, S., Xiao, S., and Yang, X. (2020). Do Workers and Customers Benefit from Competition Between On-Demand Service Platforms? *SSRN Electronic Journal*.

Beojone, C. V. and Geroliminis, N. (2021). On the inefficiency of ride-sourcing services towards urban congestion. *Transportation Research Part C: Emerging Technologies*, 124:102890.

Bieshaar, M., Zernetsch, S., Depping, M., Sick, B., and Doll, K. (2017). Cooperative starting intention detection of cyclists based on smart devices and infrastructure. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*, pages 1–8. ISSN: 2153-0017.

Borenstein, J., Herkert, J. R., and Miller, K. W. (2019). Self-Driving Cars and Engineering Ethics: The Need for a System Level Analysis. *Science and Engineering Ethics*, 25(2):383–398.

Buchholz, N. (2019). Spatial Equilibrium, Search Frictions and Dynamic Efficiency in the Taxi Industry. page 55.

Cairns, R. D. and Liston-Heyes, C. (1996). Competition and regulation in the taxi industry. *Journal of Public Economics*, 59(1):1–15.

Calvert, S. (2020). Utah and Georgia Install Tech That Lets Roads Talk. *Wall Street Journal*.

Cambridge Systematics (2008). Crashes vs Congestion: What's the Cost to Society. Technical report, AAA.

Carreras, A., Daura, X., Erhart, J., and Ruehrup, S. (2018). Road infrastructure support levels for automated driving. page 10, Copenhagen.

Castiglione, J., Cooper, D., Sana, B., Tischler, D., Chang, T., Erhardt, G. D., Roy, S., Chen, M., and Mucci, A. (2018). TNCs and Congestion. Final Report, San Francisco County Transportation Authority.

Castillo, J. C. (2018). Who Benefits from Surge Pricing? *SSRN Electronic Journal*.

Castillo, J. C. and Weyl, E. G. (2018). Surge Pricing Solves the Wild Goose Chase. page 29.

Chai, N. (2019). Uber Technologies Inc. Form S-1.

Chatterjee, I. and Davis, G. A. (2013). Evolutionary Game Theoretic Approach to Rear-End Events on Congested Freeway. *Transportation Research Record*, 2386(1):121–127. Publisher: SAGE Publications Inc.

Chen, H., Zhang, K., Nie, M., and Liu, X. (2019). A Physical Model of Street Ride-Hail. *SSRN Electronic Journal*.

Clements, L. M. and Kockelman, K. M. (2017). Economic Effects of Automated Vehicles. *Transportation Research Record*, 2606(1):106–114. Publisher: SAGE Publications Inc.

Clewlow, R. R. and Mishra, G. S. (2017). Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States.

Coffman, R. B. and Shreiber, C. (1977). The Economic Reasons for Price and Entry Regulation of Taxicabs (Comment and Rejoinder). *Journal of Transport Economics and Policy*, 11(3):288–304. Publisher: [London School of Economics, University of Bath, London School of Economics and University of Bath, London School of Economics and Political Science].

Communications, N. W. (2022). When a Tesla on Autopilot Kills Someone, Who Is Responsible?

Contissa, G., Lagioia, F., and Sartor, G. (2017). The Ethical Knob: ethically-customisable automated vehicles and the law. *Artificial Intelligence and Law*, 25(3):365–378.

Contreras, S. D. and Paz, A. (2018). The effects of ride-hailing companies on the taxicab industry in Las Vegas, Nevada. *Transportation Research Part A: Policy and Practice*, 115:63–70.

Cook, P. J. and Graham, D. A. (1977). The Demand for Insurance and Protection: The Case of Irreplaceable Commodities. *The Quarterly Journal of Economics*, 91(1):143–156. Publisher: Oxford University Press.

Daniels, K. M. and Turcic, D. (2021). Matching Technology and Competition in Ride-hailing Marketplaces. page 33.

Dempsey, P. S. (1996). Taxi Industry Regulation, Deregulation & Reregulation: the Paradox of Market Failure. 24:48.

Di, X., Chen, X., and Talley, E. (2020). Liability design for autonomous vehicles and human-driven vehicles: A hierarchical game-theoretic approach. *Transportation Research Part C: Emerging Technologies*, 118:102710.

Dionne, G. (1982). Moral Hazard and State-Dependent Utility Function. *The Journal of Risk and Insurance*, 49(3):405–422. Publisher: [American Risk and Insurance Association, Wiley].

Dixit, S., Fallah, S., Montanaro, U., Dianati, M., Stevens, A., Mccullough, F., and Mouzakitis, A. (2018). Trajectory planning and tracking for autonomous overtaking: State-of-the-art and future prospects. *Annual Reviews in Control*, 45:76–86.

Dobbs, C. (2019). TLC regulations risk leaving lower-income New Yorkers behind.

Efrati, Amir (2020). Money Pit: Self-Driving Cars' $16 Billion Cash Burn. *The Information*.

Ehrlich, I. and Becker, G. S. (1972). Market Insurance, Self-Insurance, and Self-Protection. *Journal of Political Economy*, 80(4):623–648. Publisher: University of Chicago Press.

Eilbrecht, J., Bieshaar, M., Zernetsch, S., Doll, K., Sick, B., and Stursberg, O. (2017). Model-predictive planning for autonomous vehicles anticipating intentions of vulnerable road users by artificial neural networks. In *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–8. ISSN: null.

Elliott, R. (2021). WSJ News Exclusive | Congressmen, Consumer Reports Raise Concerns Over Tesla's Autopilot. *Wall Street Journal*.

Elvik, R. (2014). A review of game-theoretic models of road user behaviour. *Accident Analysis & Prevention*, 62:388–396.

Erhardt, G. D., Roy, S., Cooper, D., Sana, B., Chen, M., and Castiglione, J. (2019). Do transportation network companies decrease or increase congestion? *Science Advances*, 5(5):eaau2670.

ERTRAC (2019). Connected Automated Driving Roadmap-2019 Update.

Facchinei, F. and Kanzow, C. (2010). Generalized Nash Equilibrium Problems. *Annals of Operations Research*, 175(1):177–211.

Fagnant, D. J. and Kockelman, K. (2015). Preparing a nation for autonomous vehicles: opportunities, barriers and policy recommendations. *Transportation Research Part A: Policy and Practice*, 77:167–181.

Frankena, M. W. (1984). An economic analysis of taxicab regulation / Mark W. Frankena, Paul A. Pautler. page 193.

Fréchette, G. R., Lizzeri, A., and Salz, T. (2019). Frictions in a Competitive, Regulated Market: Evidence from Taxis. *American Economic Review*, 109(8):2954–2992.

Gallick, E. C. and Sisk, D. E. (1987). A Reconsideration of Taxi Regulation. *Journal of Law, Economics & Organization*, 3(1):117–128.

Gaunt, C. (1995). The impact of taxi deregulation on small urban areas: some New Zealand evidence. *Transport Policy*, 2(4):257–262.

Ge, Y., Liu, X., Tang, L., and West, D. M. (2017). Smart transportation in China and the United States. page 20.

Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9):759–770.

Godbole, D. N., Eskafi, F. H., and Varaiya, P. P. (1996). Automated Highway Systems*. *IFAC Proceedings Volumes*, 29(1):5506–5511.

González, D., Pérez, J., Milanés, V., and Nashashibi, F. (2016). A Review of Motion Planning Techniques for Automated Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 17(4):1135–1145.

Gopalswamy, S. and Rathinam, S. (2018). Infrastructure Enabled Autonomy: A Distributed Intelligence Architecture for Autonomous Vehicles. *arXiv:1802.04112 [cs]*. arXiv: 1802.04112.

Hampshire, R., Simek, C., Fabusuyi, T., Di, X., and Chen, X. (2017). Measuring the Impact of an Unanticipated Disruption of Uber/Lyft in Austin, TX. SSRN Scholarly Paper ID 2977969, Social Science Research Network, Rochester, NY.

Hardman, S. (2021). Investigating the decision to travel more in a partially automated electric vehicle. *Transportation Research Part D: Transport and Environment*, 96:102884.

Hardman, S., Lee, J. H., and Tal, G. (2019). How do drivers use automation? Insights from a survey of partially automated vehicle owners in the United States. *Transportation Research Part A: Policy and Practice*, 129:246–256.

Himmelreich, J. (2018). Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations. *Ethical Theory and Moral Practice*, 21(3):669–684.

Jayaweera, N., Rajatheva, N., and Latva-aho, M. (2019). Autonomous Driving without a Burden: View from Outside with Elevated LiDAR. In *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*, pages 1–7. ISSN: 1090-3038.

Jin, Y. and Vasserman, S. (2021). Buying Data from Consumers: The Impact of Monitoring Programs in U.S. Auto Insurance. Working Paper 29096, National Bureau of Economic Research. Series: Working Paper Series.

Jun, M. and Markel, A. J. (2017). Infrastructure-Based Sensors Augmenting Efficient Autonomous Vehicle Operations: Preprint. Technical Report NREL/CP-5400-62390, National Renewable Energy Lab. (NREL), Golden, CO (United States).

Karsten, C. F. K. a. J. (2017). Gauging investment in self-driving cars. Library Catalog: www.brookings.edu.

Katrakazas, C., Quddus, M., Chen, W.-H., and Deka, L. (2015). Real-time motion planning methods for autonomous on-road driving: State-of-the-art and future research directions. *Transportation Research Part C: Emerging Technologies*, 60:416–442.

Ke, J., Yang, H., Li, X., Wang, H., and Ye, J. (2020a). Pricing and equilibrium in on-demand ride-pooling markets. *Transportation Research Part B: Methodological*, 139:411–431.

Ke, J., Yang, H., and Zheng, Z. (2020b). On ride-pooling and traffic congestion. *Transportation Research Part B: Methodological*, 142:213–231.

Kong, P.-Y. (2020). Computation and Sensor Offloading for Cloud-Based Infrastructure-Assisted Autonomous Vehicles. *IEEE Systems Journal*, pages 1–11. Conference Name: IEEE Systems Journal.

Korolko, N., Yan, C., Woodard, D., and Zhu, H. (2018). Dynamic Pricing and Matching in Ride-Hailing Platforms. *SSRN Electronic Journal*.

Leone, G. R., Moroni, D., Pieri, G., Petracca, M., Salvetti, O., Azzarà, A., and Marino, F. (2017). An Intelligent Cooperative Visual Sensor Network for Urban Mobility. *Sensors*, 17(11):2588. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

Li, S., Tavafoghi, H., Poolla, K., and Varaiya, P. (2019). Regulating TNCs: Should Uber and Lyft set their own rules? *Transportation Research Part B: Methodological*, 129:193–225.

Li, Y., Chen, Z., Yin, Y., and Peeta, S. (2020). Deployment of roadside units to overcome connectivity gap in transportation networks with mixed traffic. *Transportation Research Part C: Emerging Technologies*, 111:496–512.

Ligon, J. A. and Thistle, P. D. (1996). Information Asymmetries and Informational Incentives in Monopolistic Insurance Markets. *The Journal of Risk and Insurance*, 63(3):434–459. Publisher: [American Risk and Insurance Association, Wiley].

Lohmann, M. F. (2016). Liability Issues concerning Self-Driving Vehicles Special Issue on the Man and the Machine. *European Journal of Risk Regulation (EJRR)*, 7(2):335–340.

Luo, Q., Saigal, R., Chen, Z., and Yin, Y. (2018). Responsive Regulation on Disruptive Technology: A Continuous-Time Principal-Agent Approach. SSRN Scholarly Paper ID 3275580, Social Science Research Network, Rochester, NY.

Mangrum, D. and Molnar, A. (2020). The marginal congestion of a taxi in New York City. page 78.

Margiotta, R. and Washburn, S. (2017). Simplified Highway Capacity Calculation Method for the Highway Performance Monitoring System. Final Report PL-18-003, Federal Highway Administration.

McGuckin, F. and Fucci, A. (2018). Summary of Travel Trends 2017 National Household Travel Survey. Technical Report ORNL/TM-2004/297, 885762, Federal Highway Administration.

Metz, C. and Griffith, E. (2020). This Was Supposed to Be the Year Driverless Cars Went Mainstream. *The New York Times*.

Nie, Y. M. (2017). How can the taxi industry survive the tide of ridesourcing? Evidence from Shenzhen, China. *Transportation Research Part C: Emerging Technologies*, 79:242–256.

Nikzad, A. (2017). Thickness and Competition in Ride-Sharing Markets. *SSRN Electronic Journal*.

Noh, D., Tunca, T. I., and Xu, Y. (2021). Evolution of Ride Services: From Taxicabs to Ride Hailing and Self-Driving Cars. page 51.

Nordrum, A. (2018). Colorado Prepares to Install "Smart Road" Product by Integrated Roadways - IEEE Spectrum. Library Catalog: spectrum.ieee.org.

NYDOT. Highway functional classifications.

Nyholm, S. (2018). The ethics of crashes with self-driving cars: A roadmap, II. *Philosophy Compass*, 13(7):e12506. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/phc3.12506.

Nyholm, S. and Smids, J. (2016). The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem? *Ethical Theory and Moral Practice*, 19(5):1275–1289.

Pauly, M. V. (1974). Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection. *The Quarterly Journal of Economics*, 88(1):44.

Rana, P. (2022). WSJ News Exclusive | Uber Reaches Deal to List All New York City Taxis on Its App. *Wall Street Journal*.

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

Rebsamen, B., Bandyopadhyay, T., Wongpiromsarn, T., Kim, S., Chong, Z. J., Qin, B., Ang, M. H., Frazzoli, E., and Rus, D. (2012). Utilizing the infrastructure to assist autonomous vehicles in a mobility on demand context. *MIT web domain*.

Reitberger, G., Bieshaar, M., Zernetsch, S., Doll, K., Sick, B., and Fuchs, E. (2018). Cooperative Tracking of Cyclists Based on Smart Devices and Infrastructure. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, pages 436–443. ISSN: 2153-0009.

Rothschild, M. and Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics*, 90(4):629.

Schaller, B. (2007). Entry controls in taxi regulation: Implications of US and Canadian experience for taxi regulation and deregulation. *Transport Policy*, 14(6):490–506.

Schaller, B. (2018). The New Automobility: Lyft, Uber and the Future of American Cities. Technical report, Schaller Consulting.

Schlesinger, H. (1983). Nonlinear Pricing Strategies for Competitive and Monopolistic Insurance Markets. *The Journal of Risk and Insurance*, 50(1):61–83. Publisher: [American Risk and Insurance Association, Wiley].

Schneider, T. W. (2022). Taxi and Ridehailing App Usage in New York City.

Schwarting, W., Alonso-Mora, J., and Rus, D. (2018). Planning and Decision-Making for Autonomous Vehicles. *Annual Review of Control, Robotics, and Autonomous Systems*, 1(1):187–210.

Seo, T. and Asakura, Y. (2017). Endogenous market penetration dynamics of automated and connected vehicles: Transport-oriented model and its paradox. *Transportation Research Procedia*, 27:238–245.

Shariff, A., Bonnefon, J.-F., and Rahwan, I. (2021). How safe is safe enough? Psychological mechanisms underlying extreme safety demands for self-driving cars. *Transportation Research Part C: Emerging Technologies*, 126:103069.

Shariff, I. R. a. A. (2021). Self-Driving Cars Could Save Many Lives. But Mental Roadblocks Stand in the Way. *Wall Street Journal*.

Sheffi, Y. (1984). *Urban transportation networks: equilibrium analysis with mathematical programming methods*. Prentice-Hall, Englewood Cliffs, N.J.

Shreiber, C. (1975). The Economic Reasons for Price and Entry Regugation of Taxicabs. page 13.

Shreiber, C. (1981). The Economic Reasons for Price and Entry Regulation of Taxicabs: A Rejoinder. *Journal of Transport Economics and Policy*, 15(1):81–83. Publisher: [London School of Economics, University of Bath, London School of Economics and University of Bath, London School of Economics and Political Science].

Spotswood, B. (2017). Not Surprisingly, UberPool Has Been A Major Loss Leader: SFist.

Statista (2019). Average selling price of new vehicles - U.S. 2019. Library Catalog: www.statista.com.

Stiglitz, J. E. (1977). Monopoly, Non-Linear Pricing and Imperfect Information: The Insurance Market. *The Review of Economic Studies*, 44(3):407–430. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

Talley, E. (2019). Automatorts: How Should Accident Law Adapt to Autonomous Vehicles? Lessons from Law and Economics. page 37.

Tan, H.-S., Rajamani, R., and Zhang, W.-B. (1998). Demonstration of an automated highway platoon system. In *Proceedings of the 1998 American Control Conference. ACC (IEEE Cat. No.98CH36207)*, volume 3, pages 1823–1827 vol.3. ISSN: 0743-1619.

Tarduno, M. (2021). The Congestion Costs of Uber and Lyft. *Journal of Urban Economics*, page 103318.

Teal, R. F. and Berglund, M. (1987). The Impacts of Taxicab Deregulation in the USA. *Journal of Transport Economics and Policy*, 21(1):37–56. Publisher: [London School of Economics, University of Bath, London School of Economics and University of Bath, London School of Economics and Political Science].

Thornton, S. M., Pan, S., Erlien, S. M., and Gerdes, J. C. (2017). Incorporating Ethical Considerations Into Automated Vehicle Control. *IEEE Transactions on Intelligent Transportation Systems*, 18(6):1429–1439. Conference Name: IEEE Transactions on Intelligent Transportation Systems.

TLC (2022). Congestion Surcharge - TLC.

Tucker, Sean (2022). Mercedes: We'll Be Liable for Self-Driving Cars.

Van Brummelen, J., O'Brien, M., Gruyer, D., and Najjaran, H. (2018). Autonomous vehicle perception: The technology of today and tomorrow. *Transportation Research Part C: Emerging Technologies*, 89:384–406.

Verhoef, E. T. and Rouwendal, J. (2004). Pricing, Capacity Choice, and Financing in Transportation Networks. *Journal of Regional Science*, 44(3):405–435. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.0022-4146.2004.00343.x.

Vignon, D. A., Yin, Y., and Ke, J. (2021). Regulating ridesourcing services with product differentiation and congestion externality. *Transportation Research Part C: Emerging Technologies*, 127:103088.

Wei, B., Saberi, M., Zhang, F., Liu, W., and Waller, S. T. (2020). Modeling and managing ridesharing in a multi-modal network with an aggregate traffic representation: A doubly dynamical approach. *Transportation Research Part C: Emerging Technologies*, 117:102670.

Weise, K. and Griffith, E. (2020). Amazon to Buy Zoox, in a Move Toward Self-Driving Cars. *The New York Times*.

Wu, S. S. (2020). Autonomous vehicles, trolley problems, and the law. *Ethics and Information Technology*, 22(1):1–13.

Xu, Z., Yin, Y., and Ye, J. (2019). On the supply curve of ride-hailing systems. *Transportation Research Part B: Methodological*.

Xu, Z., Yin, Y., and Zha, L. (2017). Optimal parking provision for ride-sourcing services. *Transportation Research Part B: Methodological*, 105:559–578.

Yang, H., Ye, M., Tang, W. H., and Wong, S. (2005). Regulating taxi services in the presence of congestion externality. *Transportation Research Part A: Policy and Practice*, 39(1):17–40.

Yang, T., Yang, H., and Wong, S. C. (2014a). Taxi services with search frictions and congestion externalities. *Journal of Advanced Transportation*, 48(6):575–587.

Yang, T., Yang, H., Wong, S. C., and Sze, N. N. (2014b). Returns to scale in the production of taxi services: an empirical analysis. *Transportmetrica A: Transport Science*, 10(9):775–790. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23249935.2013.794174.

Yu, J. J., Tang, C. S., Max Shen, Z.-J., and Chen, X. M. (2019). A Balancing Act of Regulating On-Demand Ride Services. *Management Science*. Publisher: INFORMS.

Zha, L., Yin, Y., and Du, Y. (2018a). Surge pricing and labor supply in the ride-sourcing market. *Transportation Research Part B: Methodological*, 117:708–722.

Zha, L., Yin, Y., and Xu, Z. (2018b). Geometric matching and spatial pricing in ride-sourcing markets. *Transportation Research Part C: Emerging Technologies*, 92:58–75.

Zha, L., Yin, Y., and Yang, H. (2016). Economic analysis of ride-sourcing markets. *Transportation Research Part C: Emerging Technologies*, 71:249–266.

Zhang, K., Chen, H., Yao, S., Xu, L., Ge, J., Liu, X., and Nie, M. (2019). An Efficiency Paradox of Uberization. *SSRN Electronic Journal*.

Zhang, K. and Nie, M. (2019). To Pool or Not to Pool: Equilibrium, Pricing and Regulation. SSRN Scholarly Paper ID 3497808, Social Science Research Network, Rochester, NY.

Zhang, K. and Nie, M. (2021a). Mitigating traffic congestion induced by Transportation Network Companies: a policy analysis. *SSRN Electronic Journal*.

Zhang, K. and Nie, Y. M. (2021b). Inter-platform competition in a regulated ride-hail market with pooling. *Transportation Research Part E: Logistics and Transportation Review*, 151:102327.

Zhu, H., Yuen, K.-V., Mihaylova, L., and Leung, H. (2017). Overview of Environment Perception for Intelligent Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 18(10):2584–2601.