NORTHWESTERN UNIVERSITY

A General Matching Theory of Ride-hail

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

for the degree

DOCTOR OF PHILOSOPHY

Field of Civil and Environmental Engineering

By

Hongyu Chen

EVANSTON, ILLINOIS

December 2020

Abstract

A General Matching Theory of Ride-hail

Hongyu Chen

Mobile internet has brought a disruptive innovation to the ride-hail industry. The technology introduced by Uber, referred to as *e-hail*, matches passengers with drivers through their smart phones, while integrating transaction and feedback in a single app. In comparison to taxis hailed off street, or *street-hail* (s-hail hereafter), e-hail has been widely praised for not only improving user experience, but also boosting productivity. This dissertation focuses on deepening our understanding on the nature and limit of both ride-hail modes by proposing a general matching theory of ride-hail.

This theory, along with calibration methods and empirical evidences, not only shows that the passenger-driver matching process in ride-hail is indeed dictated by two primary physical limitations: the passengers' ability to access distant vacant vehicles and the drivers' (or the platform's) preference for certain locations, but further discovers that the revolution of e-hail is a tale of two markets:

On the one hand, by expanding passengers access to vacant vehicles and vice versa, e-hail dramatically improves matching efficiency. In *low-density markets* particularly, where both demand and supply for ride-hail are low, this advantage substantially lowers the likelihood of unpleasantly long waits. On the other hand, connecting a large number of waiting passengers to the same pool of unmatched vacant vehicles induces competition among passengers, which severely limits e-hail's ability to exploit economies of scale in matching. The impact of this loss in scalability with e-hail will become more prominent in *high-density* markets, where arguably efficiency matters the most.

Acknowledgements

When I was too young to remember things, curiosity had already found me. At that time, it simply meant the desire to discover and understand what our world is and how it works. "You always asked questions one after one until we could not give any answer, and you would then try to find it by yourself", that is my mother's comments on my childhood. My parents noticed that nature of mine, then helped me further develop it, and encouraged me to discover not only what the world is, but also who I want to be—a discoverist. They have supported me since then, both materially and emotionally. I am grateful for having my father Chunxiang Chen and my mother Li Kang standing by my side over the years.

However, that simple curiosity met challenges when it turned out to be difficult to actually verify my new discoveries. It has happened three times during my Ph.D. career, when I found my first research project purely theoretical without any means to verify, when my dissertation project discovered something that lots of people disagreed, and when few people in my organization saw the potential and supported our empirically validated theory. It was Marco's help that made me understand, that curiosity also means the belief in objective evidences other than subjective opinions, and the courage to take on challenges to make theoretical discoveries come true. I appreciate my Ph.D. advisor, Yu Marco Nie, for not only mentoring me on scholarly research, but also on wisdoms of life.

Table of Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Thanks to the advance of mobile technology, activities that once depended on physical interactions can now be accomplished at one's fingertips. This innovation enables consumers and providers of a service to match on a digital platform in real-time with a reduced transaction cost (Coase, 1937; Davis, 2015). The launch of Uber set in motion the rapid adoption of this new business model—known as uberization (Daidj, 2018)—in many aspects of our daily life (Zervas et al., 2017; Caldieraro et al., 2018). Users are attracted to an uberized service platform by the prospect of greater convenience and efficiency, and their accumulation on the platform triggers a *cross-side network effect*; that is, a larger number of users on one side of the market makes the service more appealing to users on the other side (Yao and Mela, 2008; Rysman, 2009; Tucker and Zhang, 2010; Halaburda et al., 2017). In turn, this effect creates a positive feedback that keeps growing the platform and projects a prospect of "winner-take-all." Blessed with this virtuous cycle, uberization is expected to transform and even dominate the way many services are provided and consumed, which would have a far-reaching societal impact (Davis, 2015).

Although uberization continues to gain momentum around the world, whether and when it will deliver on the promises aforementioned remain unclear. Uber has

yet to turn a profit,[1] and has so far failed to secure dominance over its competitors around the world. Uber's much smaller domestic rival, Lyft, actually went public a few months earlier and had steadily gained market share at Uber's expense.[2] Even the seemingly antiquated taxi service are still roaming on streets in big cities, and despite substantial losses, are able to hold on to certain market segments (Nie, 2017; Kim et al., 2018). In this dissertation, we attempt to find out if these troubles for uberization are growing pains of a new economic paradigm, or come from some fundamental limitation of the ride-hail matching mechanism itself.

Our study in this dissertation shows the revolution of e-hail is a tale of two markets. On the one hand, by expanding passengers access to vacant vehicles and vice versa, e-hail dramatically improves matching efficiency. Particularly in *low-density markets*, where both demand and supply for ride-hail are low, this advantage substantially lowers the likelihood of unpleasantly long waits. On the other hand, connecting a large number of waiting passengers to the same pool of unmatched vacant vehicles induces competition among passengers, which severely limits e-hail's ability to exploit economies of scale. Indeed, based on our empirical analysis we find that overall e-hail displays a third less returns to scale than s-hail. The impact of this loss is most prominent in *high-density* markets, where arguably efficiency matters the most.

The findings summarized above was made by developing and empirically validating a general theory of the ride-hail matching mechanism. The theory seeks to link the spatiotemporal features of a local market to the density of vacant vehicles ready

---

[1]In 2019 alone, it lost \$8.5 billion, see e.g., `https://www.theverge.com/2020/2/6/21126965/uber-q4-earnings-report-net-loss-revenue-profit-2019` (accessed on July 2, 2020).

[2]See e.g., `https://secondmeasure.com/datapoints/rideshare-industry-overview/` (accessed on July 2, 2020).

to be matched with a ride request (called *matchable vacant vehicles*). The model yields the wait time distribution, including the expected wait time of a request, as well as the relationship between the number of the trips produced and the densities of vacant vehicles and passengers. For s-hail, the model relies on two key parameters: *effective hail distance* (EHD)—the maximum distance from which a passenger can see and hail a vacant taxi, and *local area attractiveness* (LAA)—a scalar that measures the attractiveness of a given area to nearby vacant taxis. For e-hail, the competition among passengers dictates the proportion of matchable vacant vehicles. This competition is regulated by the matching algorithm used by e-hail, whose efficiency is incorporated into the model as the key parameter. The models are calibrated using taxi and e-hail operational data collected in Shenzhen, a megacity in China. A comparative study of the calibrated models suggests that our theory satisfactorily explains why the success of e-hail is a tale of two markets, as supported by the data.

When specifying the street-hail (s-hail) model, a peculiar difficulty has to do with reliably estimating the average wait time of s-hail passengers. Even though the pickup time and location may be observed (e.g., from taxi GPS trajectory data), there seems no way to find out when the passenger began to wait. As a result, we cannot even answer such basic questions as whether and (if so) by how much e-hail outperforms s-hail in a given local market. To address this open question, we propose to extract *maximum possible wait time*, defined as the maximum time the passenger could have waited given EHD of the local market, by tracking the movements of vacant vehicles cruising around the pickup location. We then prove that, for a given EHD, the extracted

maximum possible wait time follows the same distribution as the real passenger wait time, which enables the construction of reliable estimators for EHD and LAA.

For the remainder, the next chapter reviews related studies. Part 1 presents the general matching theory and derives the main analytical results, with Chapter 3 looking into microscopic matching mechanisms for both ride-hail modes, and Chapter 4 focusing on macroscopic system performances including wait time distribution and scalability formulas. Part 2 proceeds to discuss calibration method developed for the matching theory using real-world operational data for both ride-hail modes, with Chapter 5 describing the data sets and Chapter 6 showing the detailed calibration procedure. Part 3 provides empirical evidences supporting the proposed theory from various aspects. Chapter 7 provides microscopic evidences on the spatiotemporal parameters in s-hail, observing their changes during natural experiments such as metro station opening and road closure. Chapter 8 validates the statistical properties of those parameters with empirical evidences. Chapter 9 compares and explains the relative performance of e-hail and s-hail revealed in empirical data, and completes the tale of two markets through a regression analysis. The last chapter concludes with a summary of findings, limitations and possible directions for future research.

CHAPTER 2

Literature review

The matching problem in ride-hail is often studied along with the demand-supply equilibrium of the overall market (e.g., Douglas, 1972; De Vany, 1975; Beesley and Glaister, 1983; Cairns and Liston-Heyes, 1996; Arnott, 1996; Castillo et al., 2018). The two well-known studies are Douglas (1972) and Arnott (1996), who analyze street-hail (s-hail) and radio-dispatch taxi service markets, respectively. To establish market equilibrium, they propose simple models to describe the relationship between passenger wait time and vehicle supply. For s-hail, Douglas (1972) argues the expected passenger wait time should be inversely proportional to the *line density* of vacant vehicles. In the case of radio-dispatch, Arnott (1996) shows that the expected wait time for radio-dispatch is inversely proportional to the square root of spatial density of vacant vehicles. The matching theory proposed in this dissertation will lead to expected wait time functions consistent with these known results, while offering a sound physical foundation for interpretation, calibration and refinement.

Matching in ride-hail services has been modeled using different approaches. Some studies (e.g., Lagos, 2000, 2003; Bimpikis et al., 2019) simply assume it is frictionless, i.e., the number of pickups always equals to the demand or the supply, whichever is smaller. Some others prefer to describe the relationship between inputs (vacant vehicles and waiting passengers) and outputs (e.g., pickups) using an aggregate matching function, such as the *Cobb-Douglas function* (e.g., Yang et al., 2010; Yang and Yang, 2011;

He and Shen, 2015; Wang et al., 2016; Zha et al., 2016), the *urn-ball matching function* (e.g., Shapiro, 2018; Buchholz, 2019), or an empirical function estimated from simulations (e.g., Frechette et al., 2019). The urn-ball function treats matching as Bernoulli trials, whereas the Cobb-Douglas function draws an analogy between matching and production.

Another popular approach is to view the matching process as a queuing system (e.g., Banerjee et al., 2015; Afeche et al., 2018; Xu et al., 2019). Most models based on the queuing theory implicitly assume: (i) passengers are picked up immediately after being matched with a vacant vehicle, and (ii) passengers are first-come-first-served. Of course, neither is true in the case of general spatial matching. To address (i), Besbes et al. (2018) and Feng et al. (2017) modify $M/M/n$ queue by incorporating the pickup time into the service time and having it determined by the time-varying supply-demand relationship. However, both studies rely on the assumption that the total number vehicles in the system remains constant. In this way, the number of vacant vehicles, which directly determines pickup time, can be simply replaced with the difference between the number of waiting passengers and the fleet size. In addition, the model proposed by Feng et al. (2017) only applies to a stylish *circular* road.

A few recent studies start to pay closer attentions to the nuances of matching process in ride-hail. Zha et al. (2018) propose a geometric matching model to estimate the average matching and pickup time, though it is still relied on equilibrium conditions and a presumed matching function. Yang et al. (2020) propose a physical matching model based on the notion of the "dominant-zone", within which there is only one waiting passenger, and thus he is always matched to the closest vacant vehicle. Since

their model includes detailed matching strategy (e.g., matching radius and matching interval), applying it to traditional ride-hail services may not be easy.

Empirical analysis of ride-hail markets have received much attention recently. Using the taxi data collected in New York City, Frechette et al. (2019) and Buchholz (2019) study the impact of regulations on search frictions in the taxi market. While Frechette et al. (2019) adopt an aggregate model of the taxi market, Buchholz (2019) develops a network model to analyze the spatial allocation of the labor supply. The entry of e-hail is discussed as a hypothetical scenario in both studies. Frechette et al. (2019) argues, although e-hail is more efficient in terms of matching, its competition with taxis would reduce the effective market density, and thus negatively affect the overall service performance. Buchholz (2019) compares the welfare improvement under dynamic pricing and frictionless matching (the best-case scenario). The study concludes that location-based pricing can fix the inefficiency created by the spatial imbalance of supply and demand. With scraped Uber data, Shapiro (2018) further investigates the competition between taxi and e-hail, concluding that the advantage of e-hail in high-density areas mostly attributes to its lower price and less regulatory burden, rather the matching technology.

A salient feature of the traditional taxi service is that a portion of the supply has to be "wasted," in the form of vacant vehicle time, in order to maintain a desired level of service (often measured by the average wait time) (e.g., Douglas, 1972; Arnott, 1996). Ride-hail is widely considered as an industry with decreasing-average-cost because it reduces such wasted supply when demand increases. It subsequently implies that ride-hail should display increasing returns to scale and thereby is subject

to natural monopoly (Hotelling, 1938; Arnott, 1996). Few, however, had empirically demonstrated this property. Schroeter (1983) calibrated a Cobb-Douglas matching function using the data of a radio-dispatch taxi service in Minneapolis. Although the results support increasing returns to scale, they fail to reject the hypothesis of constant returns to scale statistically. Calibrating a Cobb-Douglas function using yearly aggregate taxi data in Hong Kong, Yang et al. (2014b) concludes the taxi service (s-hail) displays a mild increasing return to scale (with an elasticity of $1.14 > 1$). Based on simulated data, Frechette et al. (2019) find the returns to scale of s-hail depends on the market density: it is strongly increasing in low-density markets but constant in high-density ones.

Since the launch of e-hail service, a growing number of empirical studies have been conducted to examine its impact on ride-hail markets and traffic congestion (e.g., Nie, 2017; Kim et al., 2018; Erhardt et al., 2019), to compare its efficiency with that of traditional taxis (e.g., Cramer and Krueger, 2016; Castillo et al., 2018), and to analyze its operational strategies (e.g., Hall et al., 2015; Chen et al., 2015; Chen and Sheldon, 2016; Hall et al., 2019). Using Uber data, Yan et al. (2019) estimate a number of parameters that measure the matching efficiency and demand and supply sensitivities. Specifically, they find the pickup time is proportional to the vacant vehicle time to the power of -0.515. This last empirical finding agrees well with the prediction based on our spatial matching theory described in this dissertation.

Özkan and Ward (2020) formulates a matching queue for ride-sharing and proposes an optimal matching policy that maximizes the total number of pickups. However, their model relies on some critical assumptions, including the exogenous vehicle

arrival rate over the space and pickup time, without a comprehensive consideration of the interactions between supply and demand. Hu and Zhou (2018) proposes a more general dynamic matching model that optimizes the bipartite matching in each period anticipating the future demand and supply. Again, they assume the matching cost is exogenous and the current matching does not affect the distribution of supply and demand in future periods, thus fail to capture the impact of passenger wait time on the system performance. It is worth noting that we are not dealing with detailed matching policy in this dissertation. Yet, the proposed theory should capture the dominating characteristics of matching among ride-hail services that ultimately determine their economic properties.

Part 1

The theory

CHAPTER 3

Matching mechanism

In this chapter, we look into the matching mechanism of ride-hail and describe it with a simple model that captures all the key factors. We will first setup the model with several mild assumptions, then incorporate spatiotemporal factors into the matching process for both ride-hail modes. These analyses will lay the foundation for better understanding of the passenger-driver dynamics in matching, and for further investigation of the system performance in Chapter 4.

3.1. Matchable vacant vehicle

Consider a passenger who enters a ride-hail market where vacant vehicles are cruising around, see Figure 3.1. Upon his entry, let Λ denote the density of vacant vehicles and Π denote the density of waiting passengers in the market.[1] The following assumptions apply to both s-hail and e-hail markets.

Assumption 1. *Vacant vehicles and waiting passengers are all uniformly distributed in space. In addition,*

(1) *all vacant vehicles are cruising at the same speed v, and*

(2) *passengers keep waiting at the same location before pickup.*

---

[1]To avoid notation cluttering, we omit the subscription index for market throughout the dissertation.
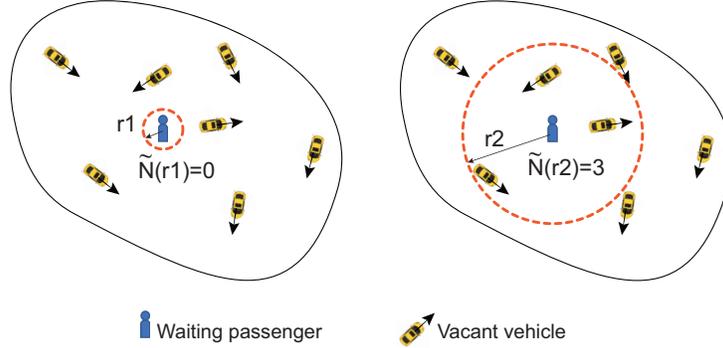
Figure 3.1. Vacant vehicles around a waiting passenger in a ride-hail market

Let $\tilde{N}(r)$ be the number of vacant vehicles within a distance r from the passenger (see Figure 3.1). With Assumption 1, the following result is obtained.

Proposition 1. *The counting process $\tilde{N}(r)$ is an Inhomogeneous Poisson Process with an intensity function of $\eta(r) = 2\pi\Lambda r$.*

Proof: First, we have $\tilde{N}(0) = 0$ by definition. Since vacant vehicles are uniformly distributed over the area, $\tilde{N}(r)$ has independent increments. Consider a ring as the intersection of two concentric circles with radii r and $r + \Delta r$. Equally cut the ring into n small pieces with area Δs, then the number of vacant vehicles in the ring follows binomial distribution with the probability of one vacant vehicle in each piece being $p = \Lambda\Delta s$. As n approaches infinity, such a binomial distribution can be approximated by a Poisson distribution with rate

$$(3.1) \qquad np = \frac{\pi(r + \Delta r)^2 - \pi r^2}{\Delta s}\Lambda\Delta s = \pi\Lambda(2r + \Delta r)\Delta r.$$

Hence,

(3.2)

$$\lim_{\Delta r \to 0} \frac{Pr\{\tilde{N}(r + \Delta r) - \tilde{N}(r) = 1\}}{\Delta r} = \lim_{\Delta r \to 0} \pi \Lambda(2r + \Delta r) \exp[-\pi \Lambda(2r + \Delta r)\Delta r]$$

$$= 2\pi \Lambda r,$$

(3.3)

$$\lim_{\Delta r \to 0} \frac{Pr\{\tilde{N}(r + \Delta r) - \tilde{N}(r) \geq 2\}}{\Delta r} = \lim_{\Delta r \to 0} \frac{1 - \exp[-\pi \Lambda(2r + \Delta r)\Delta r]}{\Delta r} - 2\pi \Lambda r$$

$$\approx \lim_{\Delta r \to 0} \frac{1 - [1 - \pi \Lambda(2r + \Delta r)\Delta r + o(\Delta r)]}{\Delta r} - 2\pi \Lambda r$$

$$= 0,$$

which are equivalent to

(3.4) $$Pr\{\tilde{N}(r + \Delta r) - \tilde{N}(r) = 1\} = 2\pi \Lambda r \Delta r + o(\Delta r),$$

(3.5) $$Pr\{\tilde{N}(r + \Delta r) - \tilde{N}(r) \geq 2\} = o(\Delta r).$$

Therefore, the counting process $\tilde{N}(r)$ is an Inhomogeneous Poisson Process with intensity function $\eta(r) = 2\pi \Lambda r$. □

At the core of the matching model is to determine which vacant vehicle will be matched with the passenger. To this end, we introduce *matchable vacant vehicles* as a subset of vacant vehicles. Whether or not a vehicle is matchable is independent from other vehicles. Thus, the counting process $\tilde{N}(r)$ can be split into two subprocesses, and the one corresponding to the matchable vacant vehicles, denoted as $\tilde{N}_{mv}(r)$, has

intensity function $\eta_{mv}(r) = 2\pi\Lambda r p(r)$, where $p(r)$ is the fraction of matcheable vehicles. Below we analyze the matching mechanism and derive $p(r)$ for s-hail in Section 3.2 and e-hail in Section 3.3.

3.2. Spatial matching in an s-hail market

In an s-hail market, "matching" between a waiting passenger and a vacant vehicle occurs when they are in close proximity so that they can see each other. In light of this observation, we define *effective hail distance* (EHD, denoted as *d*) as the threshold distance beyond which the driver and the passenger cannot see each other and thus cannot be matched. Accordingly, an *EHD area* is the circle defined by the passenger's waiting location (center) and the EHD (radius). To simplify the analysis, we further assume

Assumption 2. *After the passenger enters the market, all vacant vehicles will continue on their respective straight-line path defined by their current heading until one of them is matched with the passenger.*

This assumption, however, requires mitigation because, instead of driving in straight lines, vacant vehicles tend to turn towards areas where they are more likely to meet waiting passengers. To this end, we introduce another parameter to measure the attractiveness of the passenger's waiting location to the surrounding vacant vehicles. Hereafter, this parameter will be referred to as the *local area attractiveness* (LAA) and denoted as σ. We standardize σ such that $\sigma = 1$ if the passenger's waiting location does not affect vacant vehicles' directions and they continue on their current paths;
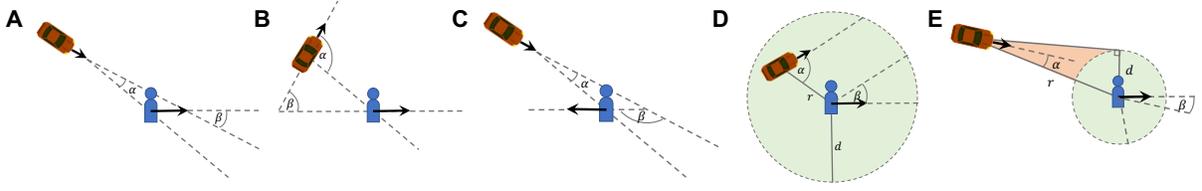
Figure 3.2. Illustration of matchable vacant vehicle in an s-hail market

and $\sigma > 1$ (< 1) if vacant vehicles are more likely to move towards (away from) the passenger's waiting location. Both d and σ are key parameters of the spatial matching model for s-hail and will be calibrated from data.

With the above settings, we can see that matching could only take place if a vacant vehicle would eventually enter the passenger's EHD area. The criteria of matchable vacant vehicle in s-hail is formally given as follows.

Definition 1 (Matchable vacant vehicle in s-hail). *A vacant vehicle in an s-hail market is matchable for the waiting passenger if (1) it is cruising towards the passenger; i.e., its heading has an acute angle with its direction to the passenger ($\alpha < \pi/2$ in Figure 3.2A); (2) it is traveling on the same side of the road as the passenger; i.e., its heading has an acute angle with the passenger's travel direction ($\beta < \pi/2$ in Figure 3.2A); and (3) either it is in the passenger's EHD area (Figure 3.2D), or it will eventually enter the passenger's EHD area if it continues on its current path (Figure 3.2E).*

Figure 3.2 illustrates: (A) A vehicle that satisfies conditions (1) and (2) in Def. 1; (B) A vehicle that violates condition (1) in Def. 1; (C) A vehicle that violates condition (2) in Def. 1; (D) A matchable vehicle inside EHD area; (E) A matchable taxi outside EHD area.

Lemma 1. *Given Assumption 2 and Definition 1, the fraction of matchable vacant vehicles at a distance r from the passenger can be approximated as $p(r) = \frac{\sigma d}{2\pi r}$.*

Proof: We consider two cases.

(1) Distance r smaller than or equal to EHD

Let θ be the angle between the passenger's travel direction and the line connecting the taxi and the passenger (Figure 3.3A). The probability that the taxi's heading satisfies both directional conditions of matchable taxis is thus $\frac{\theta}{2\pi}$. Due to the symmetry of such probability for $\theta \leq \pi$ and $\theta > \pi$, the fraction of matchable taxis within the passenger's EHD area is given by

(3.6)
$$p(r) = \frac{\int_0^\pi \frac{\theta}{2\pi} d\theta}{\pi} = \frac{1}{4}.$$

(2) Distance r larger than EHD

Let ϕ be the angle between the taxi heading direction and the line connecting the taxi and the passenger. Given the passenger's travel direction, there are two critical values θ_1 and θ_2 (Figure 3.3B), such that: (1) when $\theta < \theta_1$, all taxis with $\phi \in [-\arcsin(d/r), \arcsin(d/r)]$ are matchable taxis; (2) when $\theta > \theta_2$, there is no matchable taxi; and (3) when $\theta_1 \leq \theta \leq \theta_2$, the fraction of matchable taxis is $2\frac{\theta-\theta_1}{\theta_2-\theta_1} \arcsin(d/r)$.

Let $\phi_0 = \arcsin(d/r)$, then

(3.7)
$$\theta_1 = \frac{\pi}{2} - \phi_0, \quad \theta_2 = \pi + \phi_0,$$
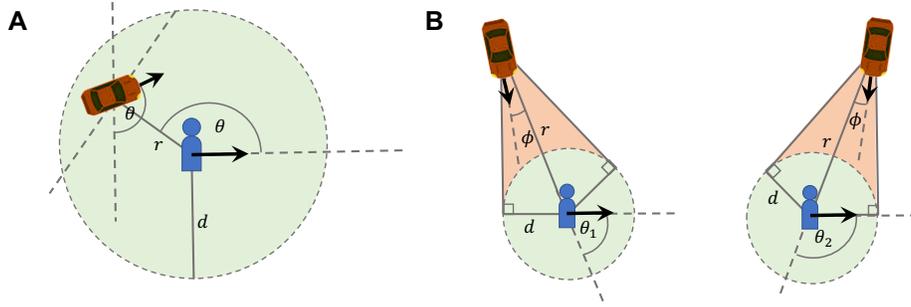
Figure 3.3. Fraction of matchable taxis at distance r

which yields

(3.8)
$$2\frac{\theta - \theta_1}{\theta_2 - \theta_1}\phi_0 = \theta - (\frac{\pi}{2} - \phi_0).$$

Since the symmetry holds as in the case of taxis within EHD area, the fraction of matchable taxis outside EHD area is given by

$$
\begin{aligned}
p(r) &= \frac{1}{\pi}\left[\frac{2\phi_0(\frac{\pi}{2} - \phi_0)}{2\pi} + \frac{1}{2\pi}\int_{\frac{\pi}{2}-\phi_0}^{\frac{\pi}{2}+\phi_0} \theta - (\frac{\pi}{2} - \phi_0)\mathrm{d}\theta\right] \\
&= \frac{1}{4\pi^2}[(\frac{\pi}{2} + \phi_0)^2 - (\frac{\pi}{2} - \phi_0)^2] \\
&= \frac{\phi_0}{2\pi} = \frac{1}{2\pi}\arcsin(d/r).
\end{aligned}
$$

(3.9)

To get a well-behaved analytical model, we further assume that all vacant vehicles are outside the EHD area when a passenger arrives. Our analysis of empirical data show that EHD ranges between 15 to 40 meters. Supposing that the EHD is 30 m and the vacant-taxi density is 50 /km^2 (close to the highest value observed in the data), the probability that a matchable vacant vehicle is inside the passenger's EHD area at a given time is around 0.035 (computed by $p(r)\Lambda\pi d^2$ for $r \leq d$). Hence, it is reasonable

to ignore these low probability events for the sake of analytical tractability. Hence, we focus on the second case above ($r > d$). For this case, $\arcsin(d/r) \approx d/r$ based on the first-order Taylor expansion, which leads to $p(r) = d/(2\pi r)$.

Finally, the above result is derived based on Assumption 2. As vacant taxis may not continue on their current path, we propose to correct $p(r)$ by scaling it with LAA, which leads to $p(r) = \sigma d/(2\pi r)$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

We are now ready to present a main result concerning the distribution of matchable vacant vehicles around the focal passenger.

Proposition 2. *In an s-hail market, the counting process of the number of matchable vehicles $\tilde{N}^s_{mv}(r)$ (superscript 's' denotes 's-hail') can be approximated by a Homogeneous Poisson Process with intensity $\eta^s_{mv} = \sigma \Lambda d$.*

Proof: Per Lemma 1, a vacant vehicle at a distance r from the passenger has a probability $p(r) = \frac{\sigma d}{2\pi r}$ of being matchable. Accordingly, the intensity function of $\tilde{N}^s_{mv}(r)$ is given by $\eta^s_{mv}(r) = p(r)\eta(r) = (\sigma d/2\pi r) \cdot (2\pi \Lambda r) = \sigma \Lambda d$, which no longer depends on r. Hence, $\tilde{N}^s_{mv}(r)$ is reduced to a Homogeneous Poisson Process with a constant intensity $\eta^s_{mv} = \sigma \Lambda d$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

It is reasonable to expect that, of all the matchable vacant vehicles, the one that is closest to the passenger when he enters the market would finally be matched with her. This assertion, however, rules out the possibility that the closest matchable vehicle may first enter the EHD area of another waiting passenger. If such a competition between passengers is taken into consideration, we may have to either further reduce $p(r)$ with another parameter or assign a "pickup probability" to each of the matchable

vehicles. Neither approach, however, is analytically tractable. For simplicity, inter-passenger competitions are ignored in s-hail markets. While this simplification tends to overestimate the performance of s-hail, the magnitude of the overestimation is expected to be small. Specifically, we note that the highest pickup rate observed in our data ranges between 600 and 800 per hour per km^2. At any given moment, the number of passengers waiting within an area of 1 km^2 is likely no more than 50 (assuming an average wait time of 3 - 5 minutes). If $d = 0.03$ km (based on the calibration results discussed below), the total EHD area of 50 passengers amounts to about 0.14 km^2 (14% of all space). Thus, the probability that the closest matchable vehicle runs into another passenger's EHD area is an event with relatively low likelihood, even in such an extremely high density area.

3.3. Spatial matching in an e-hail market

The e-hail market brings two notable differences. First, matching between the passenger and the driver no longer relies on physical proximity. Rather, it is arranged online through the e-hail platform, which assigns a vacant vehicle to the passenger according to certain matching algorithm. This means that theoretically the passenger's effective hailing distance $d \to \infty$. Second, in e-hail vacant vehicles consist of *unmatched* ones, with a density $\Lambda_0 = b_\Lambda \Lambda$, and *matched* ones, with a density $\Lambda_1 = (1 - b_\Lambda)\Lambda$. Similarly, waiting passengers can be divided into *unmatched* passengers, with a density $\Pi_0 = b_\Pi \Pi$, and *matched* passengers, with a density $\Pi_1 = (1 - b_\Pi)\Pi$. An unmatched passenger (vehicle) is waiting to be matched, whereas a matched passenger

(vehicle) is waiting to be picked up (en-route to pick up the passenger). In contrast, for s-hail, $b_\Lambda = b_\Pi = 1$.

Assumption 3. *In an e-hail market, waiting passengers and vacant vehicles, matched or unmatched, are all uniformly distributed in space.*

Accordingly, the number of unmatched vacant vehicles within distance r from the passenger is a Poisson Process with intensity function $2\pi\Lambda_0 r$. At first glance, one is inclined to consider all these vehicles matchable, since theoretically, any of them can indeed be assigned to the passenger by the platform. However, a moment of reflection suggests that the platform can not reserve all unmatched vehicles for the passenger, precisely because it also has to take care of other unmatched passengers in the market. While e-hail increases d dramatically, it also intensifies the inter-passenger competitions so much that they can no longer be ignored. To capture the impact of such competitions, we first introduce the following assumption.

Assumption 4. *The unmatched vacant vehicles are evenly distributed among unmatched waiting passengers.*

Assumption 4 implies that any unmatched passenger has an equal access to vacant vehicles around her. Specifically, because the waiting passengers and vacant vehicles are uniformly mixed, the expected number of vacant vehicles allocated to each passenger would be roughly the same. Thus, each passenger should receive an equal share $(1/(\Pi_0 \times 1))$ of unmatched vacant vehicles. Note here we convert the passenger density Π_0 to the number of the passengers in a unit area by multiplying the density by 1.

We can now characterize matchable vacant vehicles in an e-hail market as follows.

Proposition 3. *For the waiting passenger in an e-hail market, the counting process of the number of matchable vacant vehicles $\tilde{N}^e_{mv}(r)$ (superscript 'e' denotes 'e-hail') can be approximated by an Inhomogeneous Poisson Process with the intensity function*

$$(3.10) \qquad \eta^e_{mv}(r) = \frac{2\pi\Lambda_0 r}{\Pi_0 \times 1} = \frac{2\pi r b_\Lambda \Lambda}{b_\Pi \Pi \times 1} \triangleq \frac{2\pi r k \Lambda}{\Pi}.$$

.

Proof: The result directly follows from Assumptions 3 and 4. $\qquad\qquad\qquad \square$

The parameters b_Λ and b_Π are closely related to how the platform manages the matching process in real time. Accordingly, $k \triangleq b_\Lambda / (b_\Pi \times 1)$ in Eq. (3.10) is essentially a measure of the matching efficiency of the platform's algorithm. It is worth emphasizing here that the unit of k is the reciprocal of area. Like d and σ for s-hail, k is treated as a model parameter to be calibrated from data.

CHAPTER 4

System performance

Based on the matching mechanism, we derive important system performances including passenger wait times and matching rates for different ride-hail modes. Our theory predicts the key differences between taxi and e-hail in terms of passenger wait time distributions, expected wait time and stationary state pickup rate, and discovers a tale of two markets in economies of scale.

4.1. Distribution of passenger wait time

The following assumption is necessary to simplify the analysis of passenger wait time (see e.g. Arnott, 1996).

Assumption 5. *In both s-hail and e-hail markets, the passenger would always be picked up by the closest matchable vacant vehicle.*

Hereafter the closest matcheable vacant vehicle will be simply referred to as the *pickup vehicle*. Let the line distance between the pickup vehicle and the waiting location be denoted as $\tilde{D}$. The actual distance travelled by the pickup vehicle before meeting the passenger is likely longer than $\tilde{D}$. The ratio between the traveled distance and $\tilde{D}$ is known as the *detour factor* (denoted as δ) in quantitative geography (see e.g., Fairthorne, 1964; Boscoe et al., 2012; Yang et al., 2018). For simplicity, we assume that δ is a constant. Hence, the passenger's wait time $\tilde{w}$ is given by $\delta\tilde{D}/v$.

The probability that $\tilde{w}$ is longer than t equals the probability that there is no matchable vacant vehicles within $r = vt/\delta$ from the passenger (i.e., $\tilde{N}_1(vt/\delta) = 0$). Thus, we have

$$(4.1) \qquad Pr(\tilde{w} \leq t) = 1 - Pr(\tilde{N}_{mv}(r) = 0) = 1 - \exp\left[-\int_0^r \eta_{mv}(x)\mathrm{d}x\right].$$

For s-hail, recall that $\eta_{mv}^s(x) = \sigma d\Lambda$ (Proposition 2) and let $x = vw/\delta$. Accordingly, we have the cumulative distribution function (CDF) of $\tilde{w}$ as

$$(4.2) \qquad F^s(t) = Pr(\tilde{w} \leq t) = 1 - \exp\left[-\int_0^t \sigma d\Lambda \frac{v}{\delta}\mathrm{d}w\right] = 1 - \exp\left(-\frac{\sigma d\Lambda v}{\delta}t\right).$$

The above result shows that under the given assumptions the passenger wait time in an s-hail market follows a standard exponential distribution with a rate

$$(4.3) \qquad \lambda = \frac{\sigma d\Lambda v}{\delta}.$$

For e-hail, $\eta_{mv}^e(x) = 2\pi xk\Lambda/\Pi$ (Proposition 3). Thus, the CDF of passenger wait time is

$$(4.4) \qquad F^e(t) = 1 - \exp\left[-\int_0^t \frac{2\pi k\Lambda v^2}{\Pi\delta^2}w\mathrm{d}w\right] = 1 - \exp\left(-\frac{\pi k\Lambda v^2 t^2}{\delta^2\Pi}\right);$$

that is, the wait time in e-hail follows a Rayleigh distribution with a mode

$$(4.5) \qquad \theta = \frac{\delta}{v}\sqrt{\frac{\Pi}{2\pi k\Lambda}}.$$

Consequently, the expected passenger wait time for s-hail is

(4.6)
$$E[\tilde{w}^s] = \bar{w}^s = \frac{1}{\lambda} = \frac{\delta}{\sigma d \Lambda v}.$$

For e-hail, the expected passenger wait time is

(4.7)
$$E[\tilde{w}^e] = \bar{w}^e = \theta \sqrt{\frac{\pi}{2}} = \frac{\delta}{2v} \sqrt{\frac{\Pi}{k\Lambda}}.$$

Eq. (4.6) indicates that the mean passenger wait time is inversely proportional to the product of the vacant vehicle density and the cruising speed for taxis. While Douglas (1972) proposes a similar relationship, his matching model does not allow for location-specific parameters such as LAA and EHD. In comparison, Eq. (4.7) bears similarities with the formula give by Arnott (1996) for radio-dispatch taxi service. This is not surprising given e-hail can be viewed as a more sophisticated form of radio-dispatch service. Notably, they both state the mean wait time is inversely proportional to the square root of the vacant vehicle density. The features added in Eq. (4.7) are the parameters accounting for the passenger competition (Π) and the matching efficiency (k). Thus, the models of Douglas (1972) and Arnott (1996) may be viewed as special cases of the generalized spatial matching model proposed herein.

4.2. Stationary state pickup rate

A ride-hail market reaches a *stationary state* when the vacant vehicle density and waiting passenger density are held constant over time. At this state, the passenger arrival rate (demand) equals the pickup rate (the number of successful matching per

unit time), denoted as *m* in this dissertation. We note this dissertation does not investigate how the market arrives at the stationary state. Instead, our focus is to compare the performance of e-hail and s-hail services after each has reached their respective stationary state. To this end, the spatial matching model will be calibrated against data collected when the stationary conditions are approximately satisfied in various local markets.

Focusing on the stationary state allows us to establish useful relationships between the key variables using *queuing theory*. In particular, as per Little's formula (Little, 1961), we have

$$\Pi = m\bar{w},$$
(4.8)

where Π is the number of waiting passenger per unit area (analogous to the queue length), *m* is the passenger arrival rate (i.e., the pickup rate) and $\bar{w}$ is the mean wait time.

For s-hail, Eq. (4.8) yields

$$\Pi = m\bar{w}^s = m\frac{\delta}{\sigma d \Lambda v},$$
(4.9)

which implies the vacant vehicle density Λ should increase proportional to the arrival rate *m* to hold Π constant. Moreover, when an increase in *m* draws more vacant vehicles to the market, the wait time w^s always improves as per Eq. (4.6), even if the increase in Λ is unable to keep up with that of *m*.

For e-hail, we have

$$(4.10) \qquad \Pi = m\bar{w}^e = m\frac{\delta}{2v}\sqrt{\frac{\Pi}{k\Lambda}} \quad \Rightarrow \quad \Pi = m^2\frac{\delta^2}{4v^2}\frac{1}{k\Lambda}.$$

That is, when m doubles, Λ has to be quadrupled in order to prevent waiting passengers accumulate in the system indefinitely. Eq. (4.10) further yields

$$(4.11) \qquad \bar{w}^e = \frac{\delta^2}{4v^2}\frac{m}{k\Lambda}.$$

Hence, when m doubles but Λ fails to increase as much, the wait time worsens. In other words, a surge in m could in theory lead to a longer wait time, if it does not attract enough vacant vehicles. This feature is markedly different from s-hail.

Finally, by replacing $\bar{w}$ in Eq. (4.8) using Eq. (4.6) and Eq. (4.7), we arrive at a relationship between m, Π and Λ as follows:

$$(4.12) \qquad \text{s-hail:} \quad m = \frac{\sigma dv}{\delta}\Pi\Lambda,$$

$$(4.13) \qquad \text{e-hail:} \quad m = \frac{2v\sqrt{k}}{\delta}\sqrt{\Pi\Lambda}.$$

Following the literature (see e.g., Yang et al., 2010; Yang and Yang, 2011; Yang et al., 2014a; He and Shen, 2015; Wang et al., 2016; He et al., 2018; Nourinejad and Ramezani, 2019; Wang and Yang, 2019), we interpret the above relationship as the classical Cobb-Douglas production function (Cobb and Douglas, 1928). Specifically, the pickup rate m is analogous to the total production, Λ and Π are inputs, and their exponents are output elasticities. The coefficient independent of Λ and Π ($\sigma dv/\delta$ for s-hail

and $2v\sqrt{k}/\delta$ for e-hail) is interpreted as the total factor productivity (TFP), which is determined by the respective production technology of s-hail and e-hail.

4.3. Economies of scale: a tale of two markets

The classic economic theory states that a production displays constant/increasing/ decreasing returns to scale, when the sum of the two output elasticities is equal to/greater than/less than 1. Based on Eqs. (4.12)–(4.13), s-hail displays a strong increasing returns to scale (with the sum of the output elasticities equal to 2) and e-hail displays a constant returns to scale (with the sum equal to exactly 1). Thus, our spatial matching model suggests, quite unexpectedly, it is s-hail that enjoys greater economies of scale.

How do we make sense of e-hail's disappointing economies of scale? The answer hinges on the very technology that has enabled passengers to reach faraway vacant vehicles that are "invisible" to them in the s-hail context.

Figure 4.1 tells a tale of two markets about the performances of s-hail and e-hail. In a low density s-hail market (Figure 4.1(A)), a passenger surrounded by five vacant vehicles may end up meeting none, because his relatively small EHD (the red circle) may not intercept the trajectory of any vehicle (dotted lines in the figure). In a low density e-hail market (Figure 4.1(B)), e-hail's matching technology expands the passenger's hail area dramatically (the small red cycle now can be seen as filling the entire space). Hence, of the five vacant vehicles, the four unmatched ones can all be matched with the passenger. As the density in the market increases, however, the tide begins to turn. Figure 4.1(C) shows high density improves s-hail passengers' matching
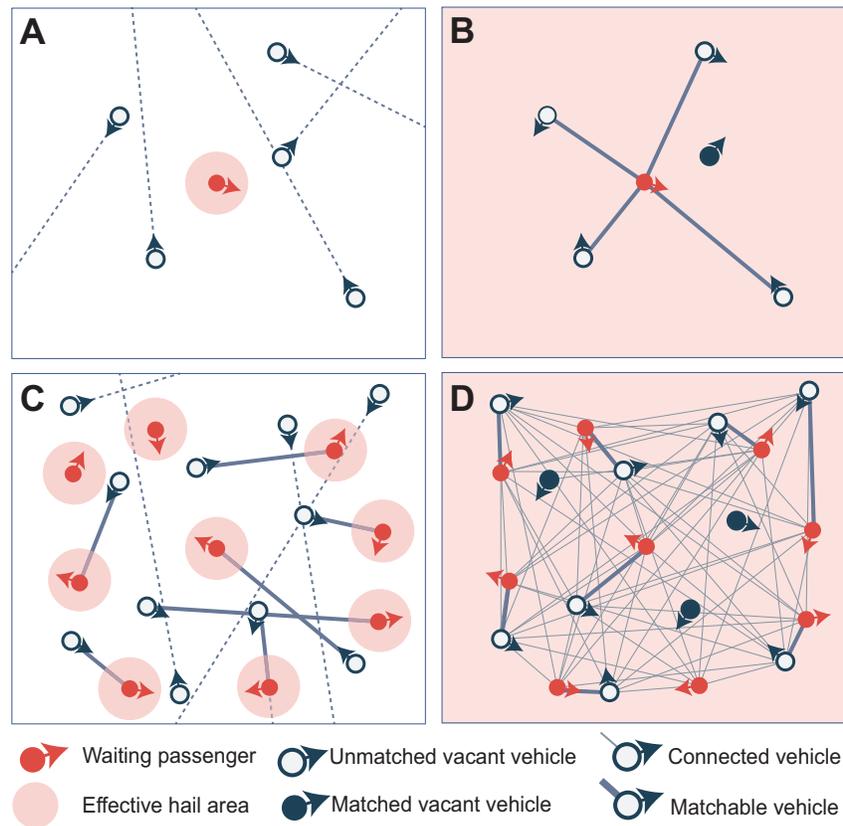
Figure 4.1. A tale of two markets

probability through the cross-side network effect. That is, although each passenger's EHD remains small, the proximity ensures a high level of successful matching. In the figure, only two out of nine passengers fail to find a matched vehicle. Yet, e-hail passengers in a high density environment (Figure 4.1(D)) find themselves locked in intense competitions with each other for the same set of vacant vehicles. As a result, each can only be matched with a small subset of all vacant vehicles. As illustrated in the figure, the nine e-hail passengers compete for eight unmatched vacant vehicles, and one of them eventually failed to find a match.

Thus, as the market scales up, s-hail benefits from rising density in general while e-hail does not, an insight consistent with Eqs. (4.12) and (4.13). This is not to say that s-hail markets are completely immune to negative impact of inter-passenger competition. Beyond certain threshold (e.g., at a so-called hot spot), when passengers' EHD areas begin to move closer to or even overlap with each other, the efficiency of s-hail would suffer too. However, those extremely dense scenarios are rare, and when they do occur, they tend to concentrate narrowly in space and time.

Part 2

Calibration methodology

CHAPTER 5

Data

We proceed to empirically calibrate the spatial matching theory put forth in Part 1. In this chapter, we first describe the data sets (Section 5.1), then define local markets (Section 5.2) and explain how key input parameters are approximated for each market (Section 5.3).

5.1. Description

The s-hail taxi data used in this dissertation were collected in Shenzhen, China for a full week in 2011, 2012 and 2016, respectively. The e-hail data were for the same week and city in 2016 as in the s-hail case. We will focus on the five weekdays in our analysis. Shenzhen is a megacity in China with more than 13 million residents and a dense urban core surrounded by an expansive metropolitan area. In addition to the ride-hail data, we also obtain the map data of traffic analysis zones (TAZ) that define the basic spatial unit in this dissertation.

The s-hail data consist of GPS trajectories of *all* registered taxis in the city with an average inter-record interval of 20 seconds. Each GPS record contains, among other information, taxi license ID, time stamp, coordinates, instantaneous speed, heading, and passenger occupancy status (0/1 variable). Following the procedure described in Nie (2017), trajectories are segmented into occupied and vacant trips. Each occupied trip defines a pickup event with a timestamp and location.

The e-hail data contains a random sample of trips collected by a large e-hail provider operating in the city, in the same week as the taxi data. At a 6% sampling rate, the data set includes about 45K trips per day. Each trip record includes coordinates of the trip origin and destination, order placement time, driver arrival time, trip starting and ending time, pickup distance, etc. Besides the trip records, the e-hail data also include the cruising time of unmatched vacant vehicles aggregated in each time period and local markets (see Section 5.2).

5.2. Local ride-hail market

In this dissertation, we focus on the five weekdays. A local ride-hail market, or a *local market* in short, is defined by a combination of a core area and a time period. In sum, 245 core areas, each corresponding to one or several TAZs, and two time periods, a morning off-peak period (9 AM–12 PM) and an evening peak period (5 PM–8 PM), are selected. Hence, there are a total of 490 local markets each for s-hail and e-hail. Figure 5.1 illustrates the selected core areas. These core areas belong to six municipal districts of the city. Based on the population density and urban function, the six districts are further classified as downtown (D-1 and D-2), urban (U-1 and U-2), and suburban (S-1 and S-2). The two periods introduced to define local markets are selected so that the travel pattern within each would be relatively stable. Accordingly, when averaged over the five weekdays, the observed market conditions can be viewed as approximately satisfying the definition of the stationary state discussed in Section 4.2. Table 5.1 shows the counts and sizes of these local markets in each region.

Region	# core areas	Min. area (km^2)	Max. area (km^2)	Med. area (km^2)
D-1	81	0.0409	1.6540	0.2664
D-2	47	0.0607	0.7113	0.1301
U-1	51	0.0712	1.5247	0.1673
U-2	21	0.1147	0.6709	0.2567
S-1	26	0.0955	1.5614	0.4111
S-2	19	0.1546	0.8746	0.3337

Table 5.1. Counts and sizes of local markets
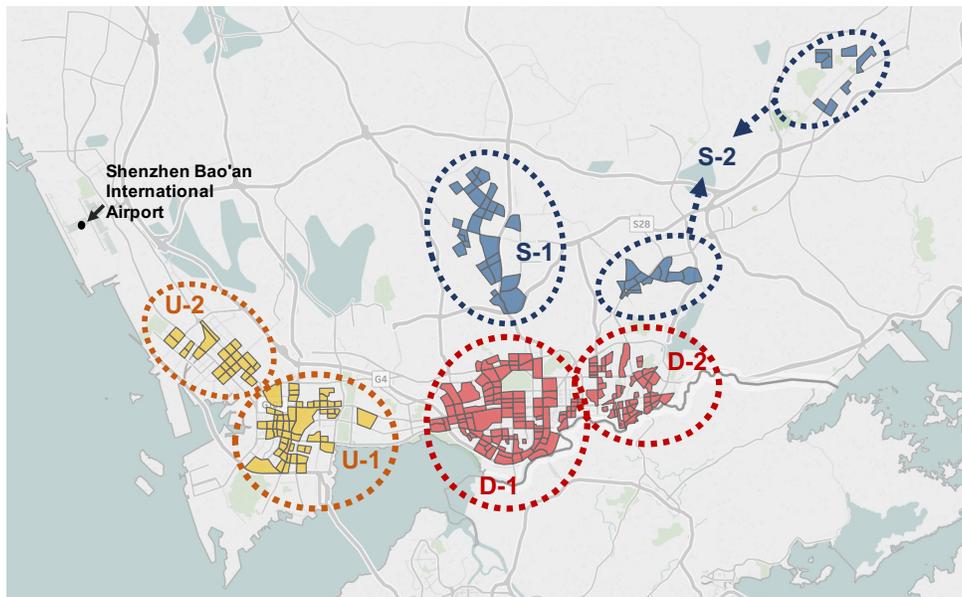
Figure 5.1. Selected core areas in Shenzhen

Important measures such as pickup rate and passenger wait time are computed based on observed pickup events in each local market. However, it should be noted that the demand and supply associated with each local market are not necessarily confined by the physical boundary of the core area. Instead, they are specified differently for s-hail and e-hail, according to the operational characteristics of these services.

For s-hail, we consider the demand as all pickups in the core area. and for each pickup, we define a "supply area" centered at the pickup location. The size of the

Λ_i

Π

Core area   Trajectory contributed

Supply area   Trajectory not contributed
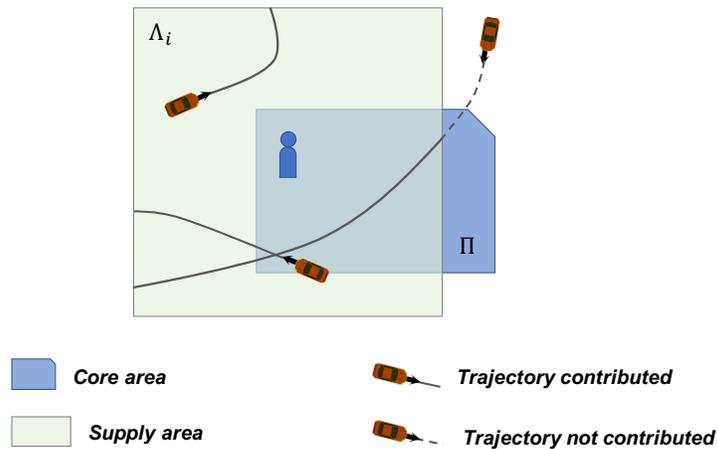
Figure 5.2. Supply area and trajectories associated with s-hail pickup i

supply area is selected to cover matchable vehicles within a reasonable distance, thus often extending beyond the core area. For simplicity, in this study we set the supply area to be a square centered at the pickup location with an area of about 1 km^2. Accordingly, the vacant vehicle density is computed using the trajectories of vacant taxis within the supply area. Figure 5.2 illustrates the supply area and trajectories associated with an s-hail pickup observed in a core area.

To define the supply area for e-hail, we first find the centroid of all pickup locations recorded inside the core area, and then, centering at that point, draw a circle with a radius equal to 90th percentile of all pickup distances (the green circle in Figure 5.3). Defining the demand is more complicated, because a passenger waiting in the core area may compete with those both inside and outside the supply area. Figure 5.3(B) portrays a passenger outside the supply area competing with passengers in the study area as he is close to a vacant vehicle inside the supply area. Consequently, we need to consider waiting passengers in an even larger area, called the "competing area" (the
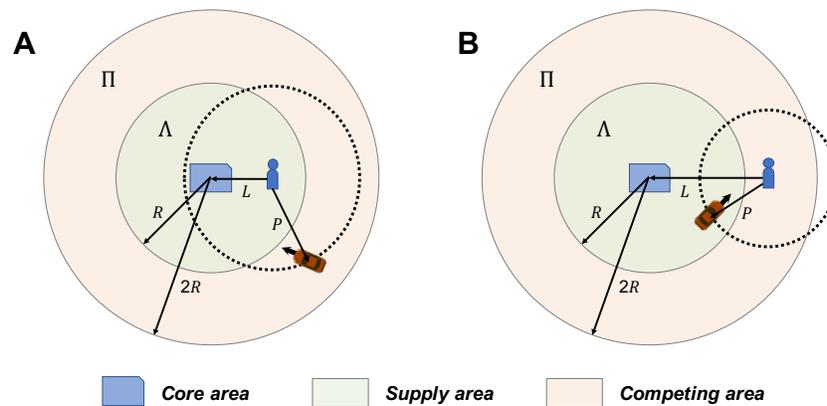
Figure 5.3. Supply area and competing area associated with an e-hail local market

orange circle in Figure 5.3). The competing area is defined as a circle that is concentric with the supply area but doubles its radius. Accordingly, waiting passengers in the competing area who compete for vacant vehicles in the supply area contribute to the demand in this local market.

5.3. Constructing input parameters from data

In each local market, we extract four input parameters directly from data for the matching models of s-hail and e-hail, respectively. The four input parameters are pickup rate, vacant vehicle density, cruising speed, and waiting passenger density. We note these parameters are important inputs to the model calibration considered in the next chapter. They themselves are also key performance metrics of ride-hail services, and will be discussed further in Part 3.

5.3.1. Pickup rate

Let A be the area of the core area, T be the length of the analysis horizon, N be the number of pickups observed in the core area within T, and μ be the sample rate. Then, for both taxi and e-hail, the pickup rate associated with the core area is given by

$$(5.1) \qquad\qquad m_{\text{core}} = \frac{N}{\mu A T}.$$

As mentioned in Section 5.2, the demand associated with an e-hail local market goes beyond the core area. We propose to approximate the pickup rate by

$$(5.2) \qquad\qquad m_{\text{supply}} = \frac{1}{\mu \tilde{A} T} \sum_{i=1}^{\tilde{N}} p^i,$$

where $\tilde{A}$ and $\tilde{N}$ are the supply area and the number of pickups observed within the supply area, respectively, and p^i the probability that pickup i contributes to the demand in the local market. Specifically, $\tilde{A}$ is evaluated as πR^2 according to Figure 5.3, and p^i will be specified in Section 5.3.4.

For taxi local market, $m_{\text{core}} = m_{\text{supply}}$ due to the limited matching radius.

5.3.2. Vacant vehicle density and cruising speed

For s-hail, the vacant vehicle density is computed from GPS trajectories. For pickup i, we query the trajectories within the supply area and a time window $[t^i - \bar{t}, t^i]$, where t^i is the timestamp of pickup and t_0 is an upper bound of passenger wait time (set to be 1000s in this study). Let m be the number of vacant vehicles passing through the

supply area during $[t^i - \bar{t}, t^i]$, τ_j and l_j be the time and distance traveled by vehicle j. The vacant vehicle density and cruising speed with respect to pickup i are respectively approximated by

$$(5.3) \qquad \Lambda^i = \frac{1}{\mu a \bar{t}} \sum_{j=1}^{m} \tau_j, \quad v^i = \frac{\sum_{j=1}^{m} l_j}{\sum_{j=1}^{m} \tau_j},$$

where a in the definition of Λ^i is the supply area (the larger square in Figure 5.2). Accordingly, the vacant vehicle density and cruising speed for the local market are

$$(5.4) \qquad \Lambda = \Lambda_0 = \frac{1}{N} \sum_{i=1}^{N} \Lambda^i, v = \frac{1}{N} \sum_{i=1}^{N} v^i.$$

For e-hail, the unmatched vacant vehicle density Λ_0 is directly obtained from the data. Approximating the matched vacant vehicle density Λ_1, however, has to be addressed in the competing area. For each pickup observed in the competing area, we need to determined how much the pickup distance contribute to Λ_1. Let w^i be the pickup time, P^i be the pickup distance and $\tilde{P}^i$ be the pickup distance associated with the local market (see Section 5.3.4). Then the matched vacant vehicle density is approximated by

$$(5.5) \qquad \Lambda_1 = \frac{1}{\mu \tilde{A} T} \sum_{i=1}^{\tilde{N}} \frac{\tilde{P}^i}{P^i} w^i.$$

Finally, the cruising speed for e-hail is computed as the average speed of pickup trip, i.e.,

$$(5.6) \qquad v = \frac{1}{N} \sum_{i=1}^{N} \frac{P^i}{w^i}.$$

5.3.3. Waiting passenger density

Since the exact passenger wait time is not available from the s-hail data, we simulate the waiting process of all pickups in the local market based on the calibrated model (see Section 6.1). For each simulation, we compute the waiting passenger density as

$$(5.7) \qquad \Pi = \frac{1}{\mu A T} \sum_{i=1}^{N} \hat{w}^i,$$

where $\hat{w}^i$ is the simulated wait time of pickup i. The final approximation is taken as the average among multiple simulations.

As for e-hail, again, the waiting passenger density need to be computed in the competing area, which requires the probability of each pickup being associated with the local market, i.e.,

$$(5.8) \qquad \Pi = \frac{1}{\mu \tilde{A} T} \sum_{i=1}^{\tilde{N}} p^i w^i.$$

5.3.4. Competing passengers in e-hail local markets

Consider a pickup i in the competing area. Let P^i be the pickup distance, L^i be the distance from the pickup location to the center of the local market, and R be the radius of the supply area. Our goal is to derive the probability that the pickup belongs to the local market, p^i, and the pickup distance contributing to the matched vacant vehicle density $\tilde{P}^i$. We will drop the subscript i from L^i, P^i and p^i to simplify notation. Two scenarios are considered, each having three subcases, as illustrated in Figs. 5.4 and 5.5.

(1) The pickup is located inside the supply area, i.e., $L \leq R$.
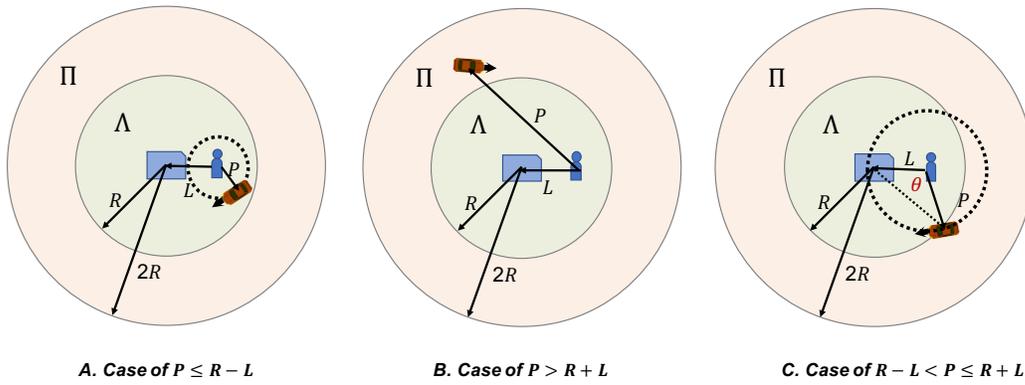
Figure 5.4. Cases for pickup inside the supply area
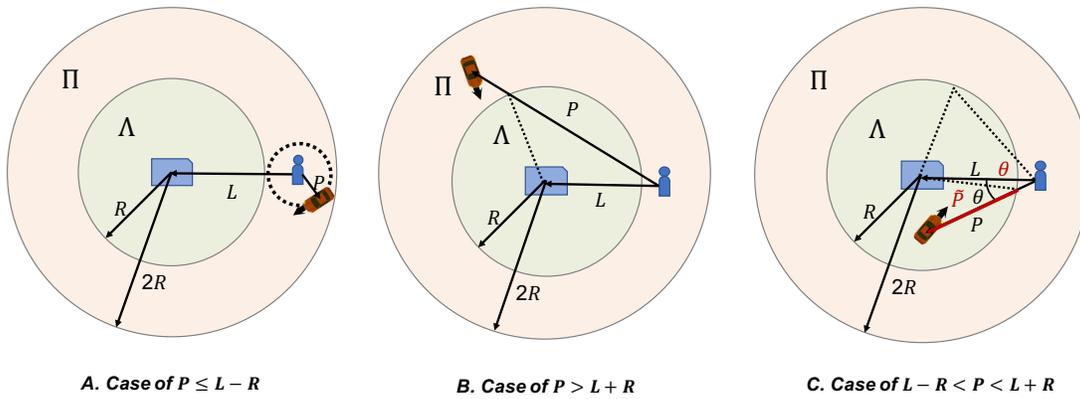
Figure 5.5. Cases for pickup outside the supply area

(a) $P \leq R - L$ (Figure 5.4(A)): In this subcase, the pickup vehicle must start inside the supply area. Hence, the passenger associated with the pickup is definitely part of the demand in the market and the entire pickup distance should be counted in the matched vacant vehicle density, i.e., $p = 1$ and $\tilde{P} = P$.

(b) $P > R + L$ (Figure 5.4(B)): In this subcase, the pickup vehicle starts out-
side the supply area. Hence, the passenger did not compete for the sup-
ply inside the supply area of the market. Then pickup event has no
contribution to either demand or supply. $p = 0$ and $\tilde{P} = 0$.

(c) $R - L < P \leq R + L$ (Figure 5.4(C)): In this subcase, the pickup may
or may not start inside the supply area. The probability that it starts
inside can be computed as $p = \theta/\pi$, where θ can be solved by the law of
cosine (i.e., $R^2 = L^2 + P^2 - 2LP \cos\theta$). Accordingly, the contribution of
the pickup distance to the matched vacant vehicle density is $\tilde{P} = P\theta/\pi$.

(2) The pickup is located outside the supply area, i.e., $L \leq R$.

(a) $P \leq L - R$ (Figure 5.5(A)): In this subcase, the pickup vehicle must start
outside the supply area. As per the argument in 1(b), $p = 0$ and $\tilde{P} = 0$.

(b) $P > L + R$ (Figure 5.5(B)): $p = 0$ and $\tilde{P} = 0$ for the same reason as in 1(b).

(c) $L - R < P \leq L + R$ (Figure 5.5(C)): Similar to 1(c), the probability that
the pickup vehicle starts inside the supply area is θ/π. However, only
a fraction of the pickup distance is covered by the supply area. To be
consistent with the matched vacant vehicle density, we count a cropped
distance $P' = P - L\cos\theta + \sqrt{R^2 - (L\sin\theta)^2}$, and approximate the aver-
age of P' by

(5.9)
$$\bar{P}' = \frac{1}{\pi} \int_0^\theta \left(P - L\cos u + \sqrt{R^2 - (L\sin u)^2} \right) du$$
$$\approx \frac{\theta}{\pi} P - \frac{\sin\theta}{\pi} L + \frac{1}{\pi} \sqrt{\theta^2 \left(R^2 - \frac{L^2}{2} \right) + \frac{\theta \sin 2\theta}{4} L^2}.$$

Hence, the contribution of the pickup distance to the matched vacant vehicle density is given by $\tilde{P} = \bar{P}'\theta/\pi$.

CHAPTER 6

Calibration methods

In this chapter, we describe the procedure developed to empirically calibrate a set of key input variables from data for the matching models of both s-hail (Section 6.1) and e-hail (Section 6.2). These input variables are crucial for empirically testing the matching models in Part 3.

6.1. Calibration for s-hail

For each taxi local market, four variables need to be calibrated, namely, the vacant vehicle density Λ, the cruising speed v, EHD d, and LAA σ. Obtaining Λ and v is straightforward as they are directly observed in the data as described in Chapter 5. The other two parameters, however, are more difficult to calibrate because neither is directly observable from the data.

Because the matching model ties EHD d and LAA σ to the wait time distribution (see Section 4.1), we may estimate them using observed passenger wait time. However, passenger wait time is not directly observable either. To tackle this difficulty, we define the maximum possible passenger wait time of each pickup event, denoted as $\tilde{w}_M$, as the maximum time the passenger could have waited given d of the local market. It can be proved that, conditional on d, $\tilde{w}_M$ follows exactly the same distribution as the passenger wait time.

Proposition 4. *Assume that (1) the EHD is fixed in the local taxi market, and (2) passengers enter the market randomly. Then, the distribution of passengers' wait time is the same as that of their maximum possible wait time.*

Proof: The key observation is that, given a fixed d, the maximum possible wait time extracted from Figure 6.3 is essentially the second last appearance of vacant vehicle inside the hail area (the last one is the pickup vehicle). Recall that $\tilde{N}_1(r)$ is the number of matchable vehicles within a distance of r from a passenger upon his arrival. Since $\tilde{N}_1(r)$ follows a Homogeneous Poisson Process for $r > d$ and the cruising speed is a constant shared by all vacant vehicles, the number of matchable vehicles entering the passenger's effective hail area (EHA) is also a Homogeneous Poisson Process. Accordingly, the inter-arrival time of matchable vehicles into that area follows exponential distribution.

Figure 6.1 illustrates four key time points on the reverse time axis in a pickup event: the pickup time ($t = 0$); the time when the pickup vehicle enters the passenger's EHA ($t = d/v$); the time when the passenger begins to wait ($t = w$); and the time when the last matchable vehicle enters the EHA ($t = w_M$).

Let s be the inter-arrival time of matchable vehicles entering EHA and s' be the time elapsed from the second last appearance of matchable vehicle in EHA to the passenger's arrival. Then, s follows exponential distribution. The passenger wait time and maximum possible wait time can be represented as

(6.1)
$$w = d/v + s - s', \quad w_M = d/v + s.$$
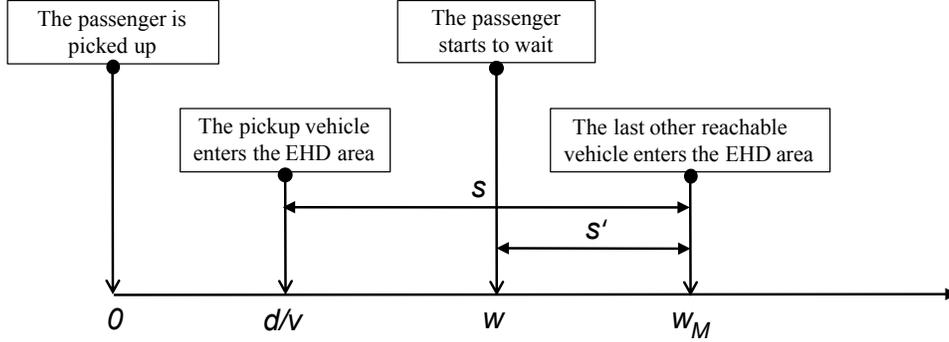
Figure 6.1. Relationship between maximum possible wait time and passenger wait time

Therefore, w_M follows exponential distribution as per the Memoryless Property. Since the passenger's arrival is independent of the movements of matchable vehicles (i.e., $s \perp s'$), w and w_M follow the same distribution as per the Strong Memoryless Property. The proof is completed. $\qquad\square$

6.1.1. Empirical mapping between maximum possible wait time and EHD

The maximum possible wait time $\tilde{w}_M$ is defined as the maximum time the passenger could have waited given the EHD d of the local market. In what follows, we empirically construct a function $G^i : d \to w_M$, which, for each pickup event i, builds a one-to-one mapping between d and w_M (an observation of $\tilde{w}_M$).

The queried vacant vehicle trajectories (see Figure 5.2) contain both trajectories of the pickup vehicle and other vacant vehicles. For the pickup vehicle, we compute the line distance from each GPS point to the pickup location and plot it against the time before the pickup. As for other vacant vehicles, we first extract the segments in their trajectories when they are matchable (see criteria in Section 3.2), and then compute

the minimum distance among them to the pickup location, again, backtracking the time from the pickup time. Figure 6.2 illustrates the derived line charts of two pickup events. From these charts we could identify two important timestamps: (1) t_1, beyond which the pickup vehicle is no longer monotonically approaching the pickup location, and (2) t_2, beyond which the pickup vehicle is no longer the closest vacant vehicle to the pickup location.
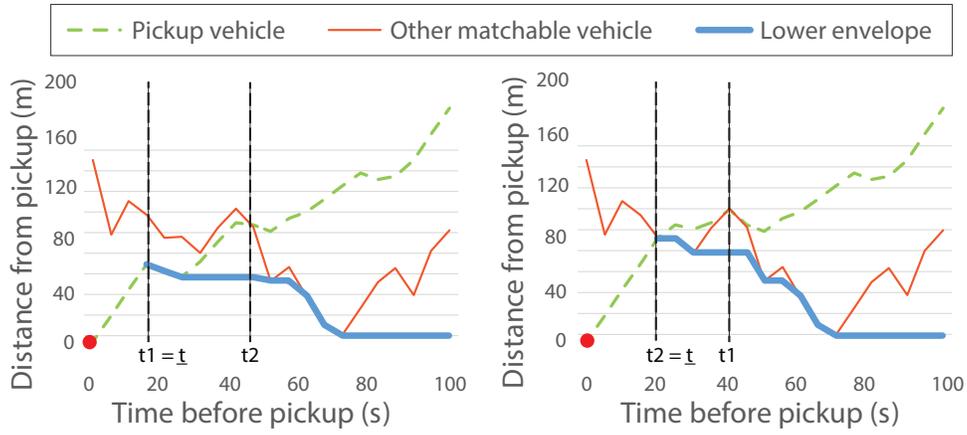
Figure 6.2. Distance from pickup location against time before pickup

Let $\underline{t} = \min(t_1, t_2)$ and for each pickup event i define

(6.2)
$$g^i(t) = \min(l_1(t), l_2(t), \forall t \in [\underline{t}, \bar{t}],$$

where $\bar{t}$ is the maximum backtracking time, $l_1(\cdot)$ and $l_2(\cdot)$ represent the two curves corresponding to the "pickup vehicle" (green dashed line) and "other matchable vehicle" (red solid line) in Figure 6.2, respectively. Accordingly, $g^i(t)$ is the bold blue line in Figure 6.2.

We claim that, if a passenger starts to wait at t before the pickup time, EHD must be smaller than $g^i(t)$. To see this, note that if $d \geq g^i(t)$, either the pickup vehicle or

another vacant vehicle would have picked him up earlier than t as per the definition of $g^i(t)$. Therefore, $g^i(t)$ gives the maximum d, denoted as d_M, corresponding to a wait time t. Conversely, for each feasible d, $g^i(t)$ also bounds the maximum possible wait time, which enables the construction of $G^i(d)$.

Figure 6.3 illustrates the empirical function $G^i(d)$ for two pickup events. To avoid unrealistic results, we bound the feasible range of d between $\underline{d}$ and $\bar{d}$, resulting two boundary points for $G(d)$. Take Event 1 as example. First, point $(d^1, \underline{w}_M^1)$ corresponds to the left end of $g(t)$ (see Figure 6.2). Since we cannot further bound the passenger wait time beyond d^1, $\underline{w}_M^1$ will be set as the maximum possible wait time for all $d \in [d^1, \bar{d}]$. The second boundary condition corresponds to $\underline{d}$. If we find, within the feasible range of $g^i(t)$, a point such that $d \leq \underline{d}$, then the corresponding wait time is set as the maximum possible wait time for all $d \in [0, \underline{d}]$. This is what happens in Event 1 (see Figure 6.3). Otherwise, we set the maximum possible wait time as the time associated with $\bar{d}$, denoted as $\bar{t}$ (see pickup event 2 in Figure 6.3).

6.1.2. Calibration algorithm

Another challenge encountered in the calibration is that d and σ interact in the distribution of passenger wait time (see Eq. (4.2)). To decouple them, we propose to iterate between estimating one of the two while hold the other fixed until convergence is achieved. The procedure is formally described as follows.

Given v and Λ and a set of empirically constructed functions $G^i(\cdot), \forall i \in \mathcal{P}$, where $\mathcal{P}$ is the set of pickup events used in calibration, Proposition 4 implies that, for each
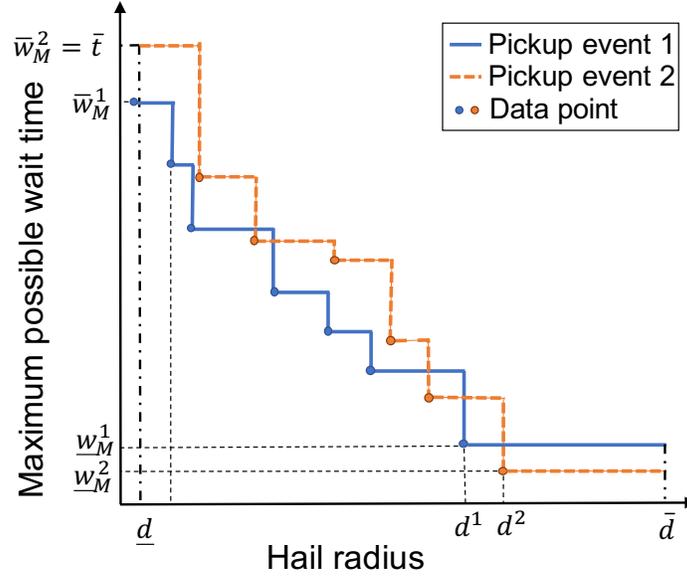
Figure 6.3. Illustration of empirical mapping between maximum possible wait time and EHD

d, we have

(6.3) $$f_{\tilde{w}_M}(G^i(d)) = f_{\tilde{w}}(G^i(d)) = \frac{\sigma d \Lambda v}{\delta} \exp\left[-\frac{\sigma d \Lambda v}{\delta} G^i(d)\right].$$

Suppose we have an estimate of EHD, denoted as $\hat{d}$, σ can be estimated by maximizing the following objective function

(6.4) $$\mathcal{L}(\sigma) = \log\left[\prod_{i=1}^{N} \left(f_{\tilde{w}}(G^i(\hat{d}))\right)^{b_i} (1 - F_{\tilde{w}}(\bar{t}))^{1-b_i}\right],$$

where $\bar{t}$ is a limit on how far back in time we trace vehicle movements to determine the maximum wait time, and b_i is 1 if $G^i(\hat{d})$ is less than $\bar{t}$, or equals 0 otherwise. Because of $\bar{t}$ (which is exogenously selected), the samples with maximum possible wait time longer than $\bar{t}$ would not be observed. Therefore, the likelihood of these samples is

represented by $P(\tilde{t}_w > \bar{t}) = 1 - F_{\tilde{w}}(\bar{t})$. This leads to the form of Eq. (6.4), which has been widely used in the statistical estimation with censored survival data (Aitkin and Clayton, 1980; Kay, 1977; Laird and Olivier, 1981).

The estimator of LAA is thus $\hat{\sigma} = \arg\max_{\sigma} \mathcal{L}(\sigma)$, which is obtained by setting

$$(6.5) \qquad \frac{\partial \mathcal{L}_{\sigma}}{\partial \sigma} = 0 \quad \Rightarrow \quad \hat{\sigma} = \frac{\delta \sum_{i=1}^{N} b_i}{\sum_{i=1}^{N} \left(\hat{d} \Lambda v \left[b_i G^i(\hat{d}) + (1 - b_i)\bar{t} \right] \right)}.$$

To estimate d based on $\hat{\sigma}$, we construct the following objective function:

$$(6.6) \qquad \mathcal{P}(d) = \log \left[\prod_{i=1}^{N} P(\tilde{w} \le G^i(d)) \right]$$

$$= \log \left[\prod_{i=1}^{N} F_{\tilde{w}}(G^i(d)) \right]$$

$$= \sum_{i=1}^{N} \log \left[1 - \exp \left(-\frac{\hat{\sigma} d \Lambda v}{\delta} G^i(d) \right) \right].$$

Technically, $\mathcal{P}(d)$ is not a log-likelihood function, but may be considered as a surrogate for the fitness of the model with the observations. The best estimate of d should maximize the fitness, or formally, $\hat{d} = \arg\max_d \mathcal{P}(d)$. Unlike Eq. (6.4), maximizing Eq. (6.6) does not yield a closed form solution. Rather, a one-dimensional search is needed to locate the maximum. Algorithm 1 below describes in detail how the above iterative procedure is executed.

Algorithm 1 Calibration of d and σ

1: **Input**: a given set of pickup events $\mathcal{P}$, v, Λ, ϵ_σ and ϵ_d.

2: **Output**: $\hat{d}$, $\hat{\sigma}$.

3: Initialization. Set $k = 0, \hat{\sigma}^k = 1, \tilde{d}^k = 0, g_\sigma = \infty, g_d = \infty$.

4: **while** $k < M$ and ($g_\sigma > \epsilon_\sigma$ or $g_d > \epsilon_d$) **do**

5:     **for** each feasible $d \in [\underline{d}, \overline{d}]$ **do**

6:         Compute $\mathcal{P}(d)$ by Eq. (6.6), using v, Λ and $\hat{\sigma}^k$ as inputs.

7:     **end for**

8:     Set $\tilde{d}^{k+1} = \text{argmax}_d\{\mathcal{P}(d)\}$.

9:     Update $\hat{\sigma}^{k+1}$ by solving Eq. (6.5), using v, Λ and $\tilde{d}^{k+1}$ as inputs.

10:     Update $g_\sigma = |\hat{\sigma}^k - \hat{\sigma}^{k+1}|$ and $g_d = |\tilde{d}^k - \tilde{d}^{k+1}|$. Set $k = k + 1$.

11: **end while**

12: **Return**: $\hat{d} = \tilde{d}^k$, and $\hat{\sigma} = \hat{\sigma}^k$

6.1.3. Moving waiting passengers

Besides missing data, another practical issue arises when we consider moving waiting passengers: it is unclear if the calibration method will still provide unbiased estimation for wait times if the passenger was moving before he was picked up and recorded in our dataset. The following proposition deals with this issue with a positive answer.

    Proposition 5. *Eq. 4.6 offers an unbiased estimator for the expected wait time of a passenger who moves before being picked up, provided that there exists a time t_0 before the pickup time, after which the probability of meeting a vacant taxi is always equal to or higher than that at time t_0.*

Proof: Since the passenger wait time follows negative exponential distribution in a street-hail taxi market, for any location (x, y), Eq. 4.6 implies that a location-specific intensity rate

$$(6.7) \qquad h(x, y) = \sigma(x, y)v(x, y)\Lambda(x, y)d(x, y)/\delta.$$

Without loss of generality, we assume the detour ratio $\delta \equiv 1$ to simplify notation. Suppose that the passenger starts waiting at time $t = 0$ and then moves on the trajectory $(x(t), y(t))$ before pickup. In the process, the distribution of passenger's wait time changes over time with a time-varying rate parameter

$$(6.8) \qquad \begin{aligned} h(t) &\equiv h(x(t), y(t)) \\ &= \sigma(x(t), y(t))v(x(t), y(t))\Lambda(x(t), y(t))d(x(t), y(t)) \\ &= \sigma(t)v(t)\Lambda(t)d(t). \end{aligned}$$

Assume that within a small time interval Δt, the rate parameter remains the same. Then the probability that the passenger fails to meet a vacant taxi in the time interval $[t, t + \Delta t]$ is given by

$$(6.9) \qquad \bar{p}(t) = \exp(h(t)\Delta t).$$

Hence, the probability of the passenger's wait time $\tilde{w}_m = t$ is the product of the probability of failing to meet a vacant taxi within $[0, t]$ and the probability of meeting

a vacant taxi within $[t, t + \Delta t]$, i.e.,

$$(6.10) \quad P(\tilde{w}_m = t) = h(t) \exp(h(t)\Delta t) \prod_{i=1}^{n-1} \exp(h(i\Delta t)\Delta t) = h(t) \exp\left(\sum_{i=1}^{n} h(i\Delta t)\Delta t \right),$$

where $\Delta t = t/n$. As $n \to \infty$, the PDF of moving passenger wait time is given by

$$(6.11) \qquad\qquad f_{\tilde{w}_m}(t) = h(t) \exp\left(- \int_0^t h(x)\mathrm{d}x \right).$$

Let $H(t) = \int_0^t h(x)\mathrm{d}x$. Given the second condition in the proposition statement, $H(t)$ is bounded from below for all $t \geq t_0$, which reads

$$(6.12) \qquad\qquad \begin{aligned} H(t) &= \int_0^{t_0} h(x)\mathrm{d}x + \int_{t_0}^t h(x)\mathrm{d}x \\ &\geq \int_0^{t_0} h(x)\mathrm{d}x + \int_{t_0}^t h(t_0)\mathrm{d}x \\ &= C_0 + h(t_0)(t - t_0) = C + h(t_0)t; \end{aligned}$$

where C_0 and C are both constants. Eq. 6.12 means $H(t)$ is at least linearly increasing with positive rate $h(t_0)$ after t_0, which further yields

$$(6.13) \qquad\qquad \lim_{t \to +\infty} \{t \exp(-H(t))\} = 0.$$

Finally, the expected moving passenger wait time is given by

$$(6.14)$$
$$\begin{aligned} E[\tilde{w}_m] &= \int_0^{+\infty} t f_{\tilde{w}_m}(t)\mathrm{d}t = \int_0^{+\infty} t[h(t)\exp(-H(t))]\mathrm{d}t = - \int_0^{+\infty} t\,\mathrm{d}[\exp(-H(t))] \\ &= -t\exp(-H(t))|_0^{+\infty} + \int_0^{+\infty} \exp(-H(t))\mathrm{d}t = \int_0^{+\infty} \exp(-H(t))\mathrm{d}t. \end{aligned}$$

We are now ready to show that Eq. 4.6 offers an unbiased estimator for $E[\tilde{w}_m]$. As per the proposition, $\tilde{w}_m$ can be estimated by the intensity rate at the pickup location, i.e.,

$$
(6.15) \qquad \hat{w}_m = \frac{1}{h(x(\tilde{w}_m), y(\tilde{w}_m))} = \frac{1}{h(\tilde{w}_m)}.
$$

The expectation of the estimator is obtained by integrating over its support $\in [0, \infty)$, i.e.,

$$
\begin{aligned}
(6.16) \qquad E[\hat{w}_m] &= \int_0^{+\infty} \frac{1}{h(x(t), y(t))} f_{\tilde{w}_m}(t) \mathrm{d}t = \int_0^{+\infty} \frac{1}{h(t)} h(t) \exp(-H(t)) \mathrm{d}t \\
&= \int_0^{+\infty} \exp(-H(t)) \mathrm{d}t = E[\tilde{w}_m].
\end{aligned}
$$

This completes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$.

6.2. Calibration for e-hail

Since the passenger wait time is directly observable for each pickup event in e-hail, it is much easier to calibrate the model for e-hail.

The matching efficiency k is the only parameter that requires estimation in Eq. (4.7). Denote the observed passenger wait time for each pickup as w^i, then we can construct the following log-likelihood function with respect to k using the PDF of passenger wait time for e-hail derived from Eq. (4.4):

$$
(6.17) \qquad \mathcal{L}(k) = \sum_{i=1}^{N} \log f(w^i) = \sum_{i=1}^{N} \log \left\{ \frac{2\pi k \Lambda v^2 w^i}{\delta^2 \Pi} \exp\left[-\frac{\pi k \Lambda v^2}{\delta^2 \Pi} (w^i)^2 \right] \right\}.
$$

Maximizing Eq. (6.17) yields a closed form estimator for k as

$$\hat{k} = \frac{\delta^2 \Pi}{\pi \Lambda v^2} \frac{N}{\sum_{i=1}^{N} (w^i)^2}.$$

(6.18)

Part 3

Empirical evidences

CHAPTER 7

Microscopic evidences

In this chapter, a full sample of continuous taxi trajectory data collected in 2011 and 2012 Nie (2017) is used. We consider microscopic road environmental changes as natural experiments in 11 local markets in this evidence finding (Figure 7.1). Table 7.1 shows a more detailed description for these local markets. Four models are calibrated for each local market, corresponding to two time windows in each of the two data collection periods.

Table 7.1. Description of the eleven markets selected in microscopic validation for s-hail

Market	District	Description
A	downtown	Close to Shenzhen Convention Center and a metro station
B1	downtown	Well-developed residential area with a metro station
B2	downtown	Well-developed residential area
C1	downtown	Employment center with a cluster of technology companies, adjacent a metro station opened between 2011 and 2012
C2	downtown	Same as C1
C3	downtown	Commercial area, adjacent a metro station opened between 2011 and 2012
C4	downtown	Same as C3
D	downtown	Residential area with two metro stations
E	urban	Major residential area with a regional long-distance bus station and close to interstate expressway.
F1	suburban	Mix land use area in the northern suburb of Shenzhen, with a new terminal metro station opened in June of 2011
F2	suburban	Mix land use area in the north suburb of Shenzhen
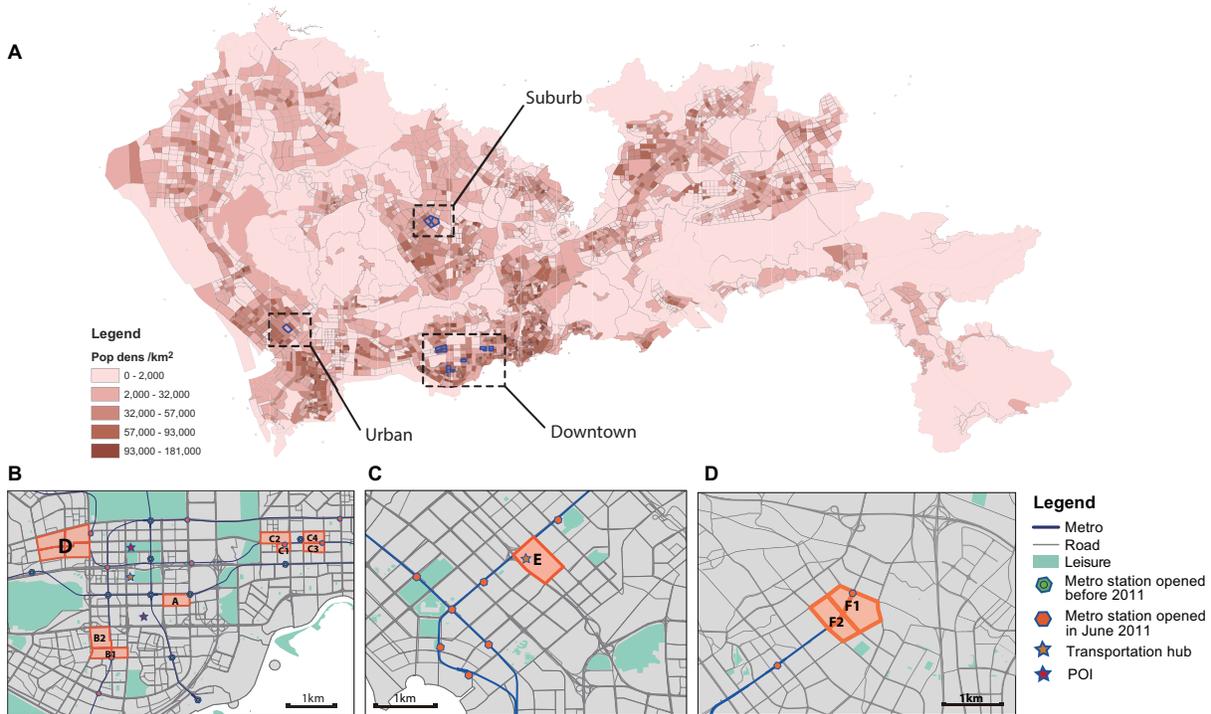
Figure 7.1. Selected local markets for microscopic validation for s-hail

We first validate the assertion that the maximum possible wait time ($\tilde{w}_M$) follows the same distribution as passenger wait time ($\tilde{w}$). Figure 7.2 compares the distributions of observed $\tilde{w}_M$ and calibrated $\tilde{w}$ in six local markets. The six cases presented here cover the best and the worst results in all 44 tests. In each case, the top figure presents the cumulative distribution function (CDF) of the calibrated model and empirical observations; the bottom figure shows PDFs of the calibrated model, empirical observations and an exponential distribution directly fitted from the observations (i.e., "Fitted PDF"). A statistical test is performed to evaluate the model fitness in Table 7.2. Overall, the calibrated model matches the empirical observations reasonably well. The most notable deviation exists on the tails: the observed distribution tends to display

a longer tail than the calibrated distribution. Note that, for some pickup events, the maximum possible wait time cannot be correctly determined, due largely to errors in the trajectory data (e.g., missing GPS points, mislabeled passenger status). Whenever this failure occurs, the maximum possible wait time is always set to the predefined upper bound T, which contributes to a longer tail in the empirical distribution.

Table 7.2. Statistics of wait time results for the six markets

Market	Year	Time	μ_{obs}	σ_{obs}	μ_{cal}	μ_{fit}	K-S
A	2012	9:00–12:00	296.9828	304.9441	327.1559	297.4828	0.0312
B2	2012	9:00–12:00	363.7376	343.7217	426.8684	364.2376	0.0612
C1	2011	9:00–12:00	157.5464	210.3421	162.7242	158.0664	0.0567
D	2012	17:00–20:00	284.7222	304.1247	313.1092	285.2222	0.0728
E	2011	17:00–20:00	143.0136	214.8127	147.8147	143.5136	0.1020
F1	2012	17:00–20:00	159.5759	239.5378	169.1387	160.0760	0.1378

Note: 'obs', 'cal', 'fit' denote empirical, calibrated and fitted values, respectively; K-S refers to Kolmogorov-Smirnov statistic.
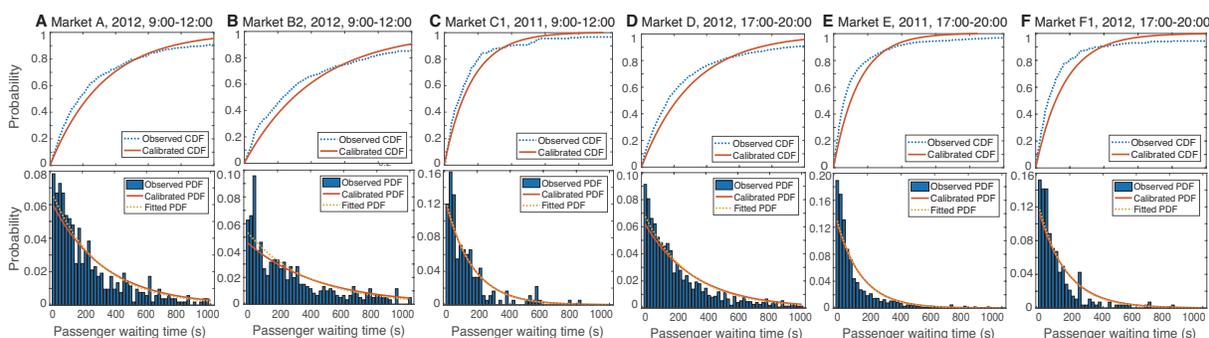
Figure 7.2. Results of wait time validation for s-hail

Figure 7.3 reports parameter estimates for all 44 cases (11 markets in four time periods): (A) average cruising speed; (B) vacant-taxi density; (C) LAA; (D) EHD; (E) pickup rate; and (F) expected passenger wait time. Temporal variations in vacant-taxi density and cruising speed are noticed across all 11 markets. As expected, the

estimated EHD values are generally small, ranging from 15 to 40 m. EHD estimates in downtown (e.g., local markets A, B1, B2) are consistently below 20 m, lower than those in other districts. This finding supports our hypothesis that well-developed urban areas with higher road densities and shorter block lengths may shorten the effective sighting distance for taxi hail. Furthermore, the estimates remain relatively stable across all time periods in each market.

Unlike those of EHD, estimates of LAA appear to be sensitive to both spatial and temporal influences. In particular, they highlight the appeal of metro stations to nearby vacant taxis. An interesting case involves local markets F1 and F2: they are next to each other, and a metro station was opened in F1 in June of 2011. The results reveal the dramatic impact of the new metro station in the peak period. In particular, LAA in F1 almost doubled from 2011 to 2012 in the evening peak. In contrast, F2 became much less attractive in the same period, having its LAA cut in half. Hence, in this case the gains in F1 appear to come at the expense of the loss in F2 (highlighted in Figure 7.3C). On the other hand, the metro opening seems to have little influence on estimated LAAs in the off-peak period.

More significant discrepancies are found in the pickup rate and the expected passenger wait time across the markets (Figs. 7.3E and F). The markets located in the urban and suburban districts (e.g., E, F1, and F2) enjoy a shorter expected wait time, whereas passengers in downtown (e.g., B2, C3, and C4) suffer the worst experience, sometimes having to wait more than 15 min. This finding may be surprising at first glance, especially for those who live in countries with a much lower taxi density in the suburbs. However, the suburban areas in Shenzhen, though far away from the
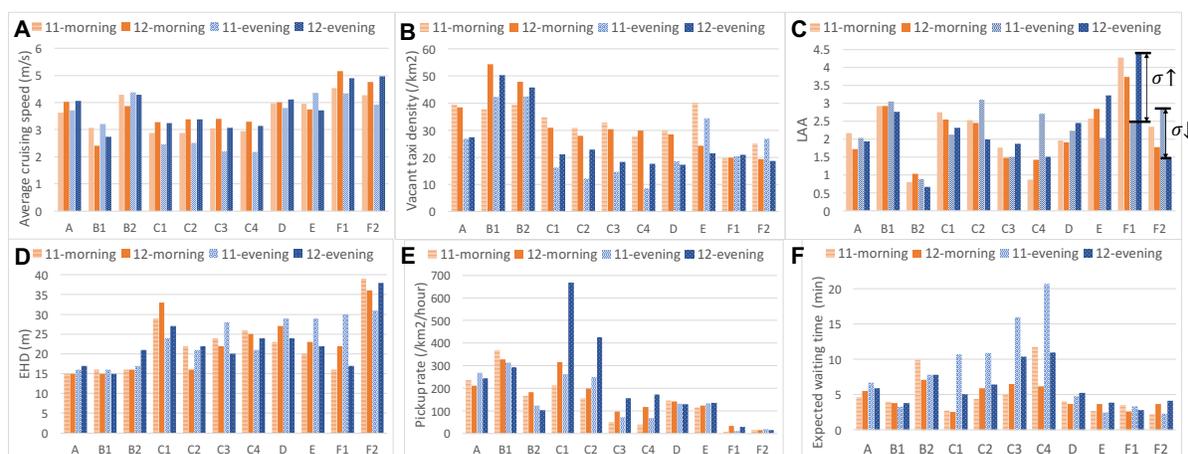
Figure 7.3. S-hail parameter estimates for validation

city center, are densely developed with a high level of commercial activities. There-

fore, these areas still enjoy an adequate supply of vacant taxis (see Figure 7.3B). More

importantly, they tend to have a larger EHD and better traffic conditions compared to

downtown areas (see Figs. 7.3A and D).

Another interesting observation arises from the downtown markets C1–C4. From

2011 to 2012, there was a dramatic increase in the pickup rate accompanied by a sig-

nificant drop in the expected wait time, particularly in the peak period. Moreover,

substantial variations are found in the LAA estimates. Together, these observations

indicate that some major change must have occurred between January 2011 and Jan-

uary 2012. Our search for news revealed that a main road (Zhenhua Road) in this

region was closed for the construction of Shenzhen Metro Line 2 in 2008[1] and re-

mained closed until June 2011[2]. Clearly, the surge in pickup rates is likely related

---

[1]See e.g., http://jt.sz.bendibao.com/news/2008519/69944.htm (in Chinese and accessed on July 2, 2020).

[2]See e.g., http://ditie.mapbar.com/shenzhen/news/124939.html (in Chinese and accessed on July 2, 2020).

to passengers brought by the new metro line. The reduction in the expected passenger wait time can be attributed to two factors: (*i*) improved traffic conditions; and (*ii*) a growing number of vacant taxis attracted by more potential passengers as well as less traffic congestion. These two effects contribute to a 30% and 50% increase in the cruising speed and vacant-taxi density during the evening peak, respectively (see Figs. 7.3A and B).
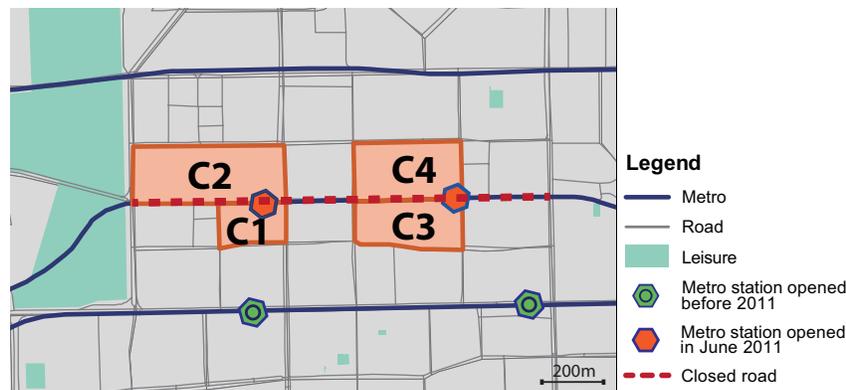
Figure 7.4. Zoom-in map of local markets C1-C4

The results also show that, after the new metro line opened, C1 and C3—located on the south side of Zhenhua Road—became more attractive to vacant taxis, whereas those on the north side of the road became much less attractive (see Figure 7.3C). We conjecture that this has to do with the direction of the one-way road. Because passengers generated by the metro line must travel east, they could only board a taxi on the south side of Zhenhua Road. Figure 7.4 shows a zoom-in map of these local markets. The bold dash line illustrates the road closed from 2008 to 2012. The two orange hexagons indexed metro stations opened in 2012 when the road was also reopened. As a result, it is considerably easier for taxis to get passengers exiting from

the metro station if they cruise on the south side (i.e., in C1 and C3). After the metro line opens, therefore, surrounding taxis are more likely to cruise towards C1 and C3 than towards C2 and C4. This seemingly rather microscopic behavior is captured remarkably well by the LAA, in a way that is consistent with its physical meaning.

CHAPTER 8

Statistical evidences

In this chapter, the calibration methods discussed are applied to all local markets in the six districts for both ride-hail modes in the 2016 data. Figs. 8.1 and 8.2 report the distribution, median and variance of EHD and LAA corresponding to each area type and period in 2016. The probability density functions (PDFs) are constructed using kernel density estimation. EHD values mostly fall in the range between 15 m and 60 m, with a peak around 25–45 m and a median around 32 - 41 m. While the shape of the PDFs of EHD does vary with spatial (area type) and temporal (peak vs. off-peak) features, the variations are modest, indicating relatively strong regularity, which is in line with their physical meaning in the theory. In addition, the distributions of EHD during the peak period tend to have greater variances than those in the off-peak period, an indication that different local markets are subject to more uneven conditions during the peak period than during the off-peak period.

LAA in downtown and urban areas concentrates around 1.2, with a quite strong peak. This means the vacant vehicles do not have preference on the selected local markets for passenger search in these areas. In contrast, the distribution in suburban areas has a quite distinctive shape, with much larger median (almost twice as much) and variances. Thus, the selected suburban local markets seem much more "attractive" to nearby vacant vehicles than downtown and urban markets. This also meets the physical meaning of LAA since most suburban local markets are constructed around
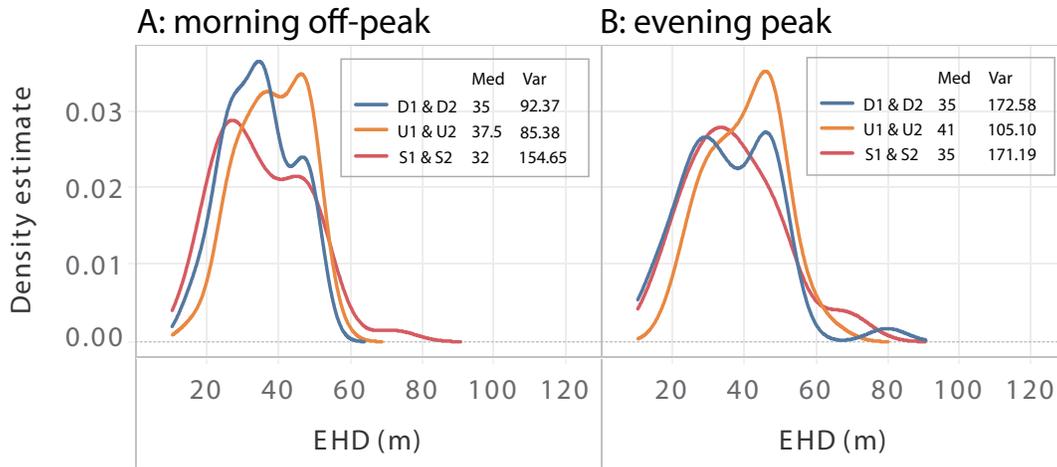
Figure 8.1. Estimated density function for calibrated EHD of s-hail in 2016
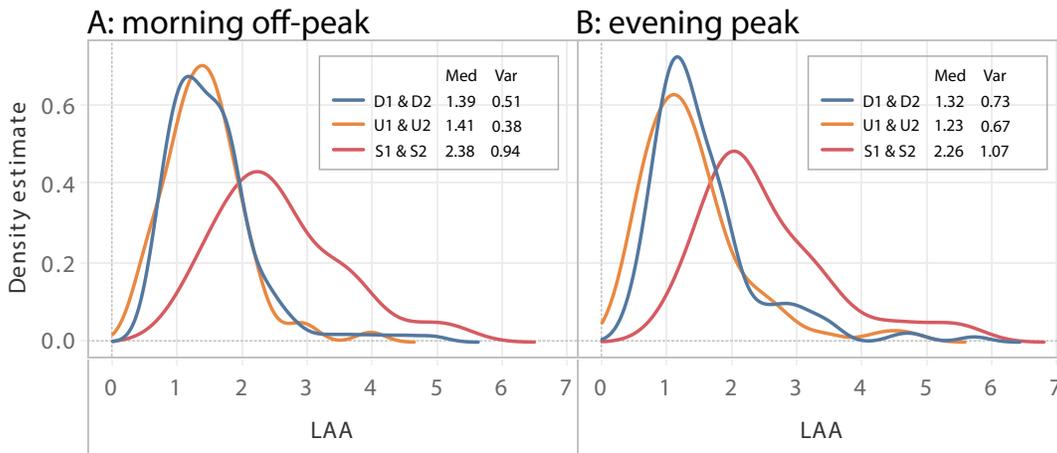
Figure 8.2. Estimated density function for calibrated LAA of s-hail in 2016

locations where ride-hail pickup events are concentrated. These location stand out precisely because they are activity hot spots that attract both passengers and vacant vehicles. In downtown and urban areas, there are more and closely spaced hot spots, and thus the competition reduces the average attractiveness of each individual hot spot. Moreover, while the PDFs of LAA in the two periods are largely similar, it

is worth noting, in all three areas, the median is lower in the peak period than the off-peak period. This trend is likely related to drivers' tendency to avoid traffic congestion.

Figure 8.3 shows the matching efficiency k in the off-peak period generally peaks around 0.35 in all three areas, though the variance is much greater in downtown areas. Unlike EHD and LAA, which seem relatively insensitive to the peaking effect, the distribution of k looks very different in the peak period. In urban and suburban areas, the distributions have a considerable rightward shift, almost doubling the medians. In downtown areas, the median slightly increases, accompanied with a quite significant drop in variance. Overall, the e-hail operator appears to perform better in the peak period than in the off-peak period, but it evidently has greater success outside the downtown area. We speculate that the greater efficiency observed from data may be attributed to various rewarding schemes designed to boost productivity in the peak period.
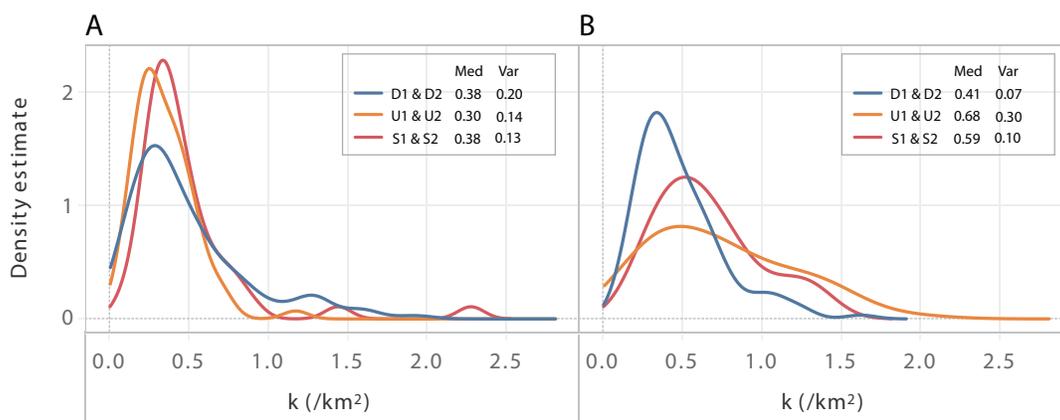
Figure 8.3. Estimated density function for calibrated matching efficiency of e-hail

Figure 8.4 compares the estimated and observed average passenger wait times for the purpose of validation. 8.4(A) is for s-hail and 8.4(B) for e-hail. Points on the diagonal dash line mean the model make correct predictions. The color of each point represents the sample rate. It shows that e-hail models tend to have better goodness-of-fit, likely because the data used in their calibration are of higher quality (they are actual observations of passenger wait time). As expected, a lower sample rate leads to worse goodness-of-fit; see Figure 8.4(B).
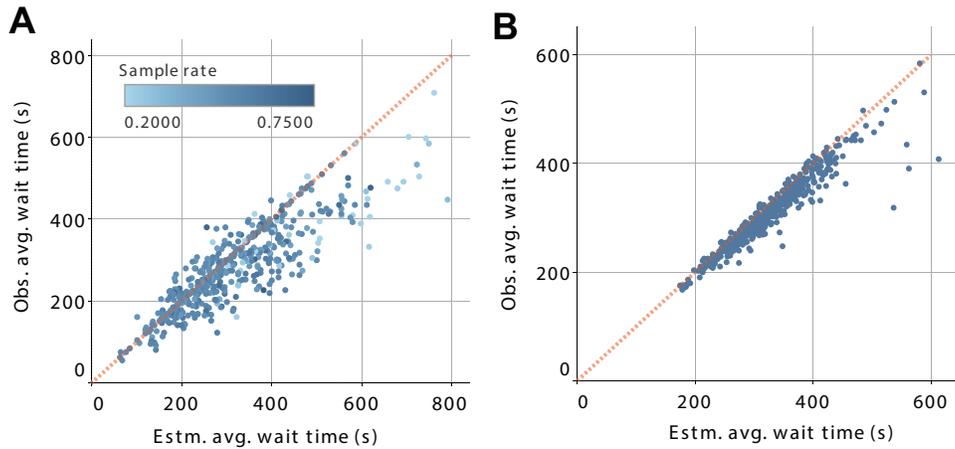
Figure 8.4. Simple validation of estimated average passenger wait time

CHAPTER 9

System performance evidences

Below, Section 9.1 reports the relative performance of the two services by district, and compares it against the qualitative prediction of the spatial matching model. In Section 9.2, a regression analysis is performed to find an empirical relationship between the pickup rate, vacant vehicle density and waiting passenger density for each service. The results are then compared with the analytical Cobb-Douglas functions derived in Section 4.2.

9.1. Service performance

Figure 9.1 reports the average pickup rate[1] and vacant vehicle density averaged over six different districts (see Section 5.2) during (A) morning off-peak and (B) evening peak. D-1 and D-2 are downtown districts, U-1 and U-2 are urban districts and S-1 and S-2 are suburban districts. In the two suburban districts, e-hail holds a clear advantage over s-hail, in terms of both the number of pickups and vacant vehicle density. Notably, a substantial portion of the vacant e-hail vehicles are matched—meaning they are no longer available to take orders—and more so in the peaking period than the off-peak period. In fact, the number of unmatched vacant e-hail vehicles is similar to that of vacant s-hail vehicles, even though the total supply

---

[1]The pickup rate associated with core area; see Eq. (5.1) in Section 5.3.

of vacant e-hail vehicles is significantly higher in these two districts. In the two down-town districts, where the densities of both passengers and vehicles are supposedly higher, e-hail still outperforms s-hail in terms of the total supply. Yet, the disadvantage of having to spend much of that supply on picking up passengers becomes an even greater burden for e-hail here, especially during the peak-time. In D-2 during the peak-time, for example, almost 80% of the total supply are stuck in the pickup phase. Consequently, the pickup rate by e-hail falls behind in both D-1 and D-2. Results in the urban districts, where the vehicle supply ranges between downtown and suburban districts, are mixed: e-hail serves more passengers in district U-1 but s-hail has the upper hand in U-2.

The above observations indicate the relative performance of the two services indeed varies substantially with market locality. Specifically, the density seems to hurt the productivity of e-hail but benefit that of s-hail. This finding is consistent with our analytical result that asserts e-hail has a lower returns to scale than s-hail. We shall see more direct empirical evidence on this point in the next section.

Figure 9.2(A) compares the average passenger wait time of the two services in the six districts. We can see that the district-level average wait time ranges between 4 and 6 minutes in most cases. A notable exception is U-2, in which the average wait time
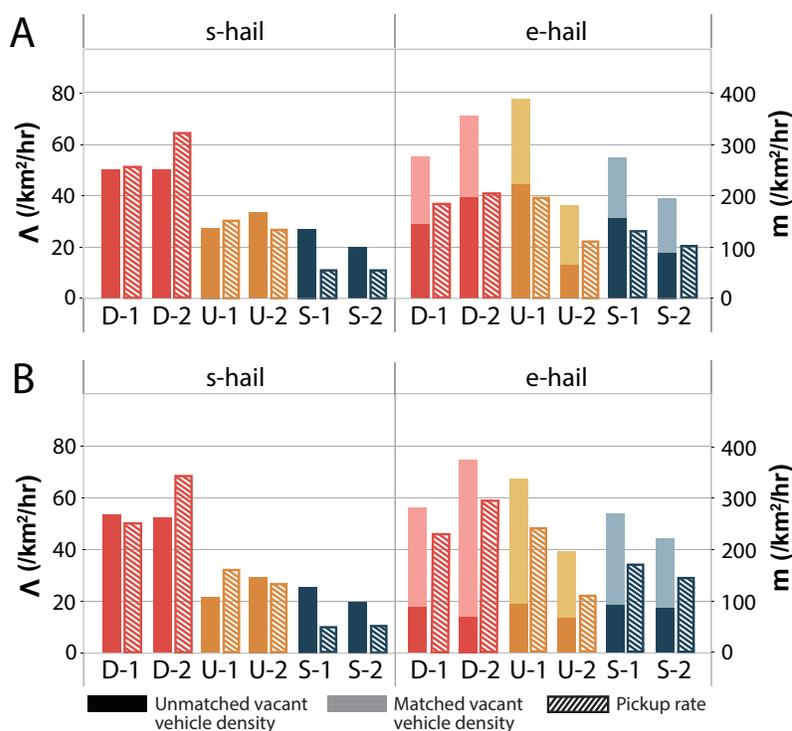
Figure 9.1. Average pickup rate and vacant vehicle density

for s-hail exceeds 8 minutes in both periods[2]. In the urban and suburban districts, e-hail has a small (with the exception of U-2) but consistent lead over s-hail in average wait time. Yet, it falls behind again in downtown districts in both periods, with a discrepancy up to 2 minutes, which amount to roughly a third or half of the total wait time. Another trend from Figure 9.2(A) is that wait time for e-hail is more sensitive to peaking than that for s-hail. In D-1, D-2 and U-1, in particular, e-hail experiences

---

[2]A closer look at the data indicates s-hail's struggle in this district is related to its substantially lower attractiveness to vacant vehicles. The median of LAA in U-2 is 1.09 for the off-peak and 0.70 for the peak. In contrast, in U-1, the median is around 1.4 in both periods. The local markets in U-2 are less attractive likely because it is fairly close to Shenzhen's Baoan International airport. As a result, not only are many vacant vehicles observed in the data more likely to move towards the airport direction (the northwestern corner, see Figure 5.1), they might be actually heading towards the airport in the first place.
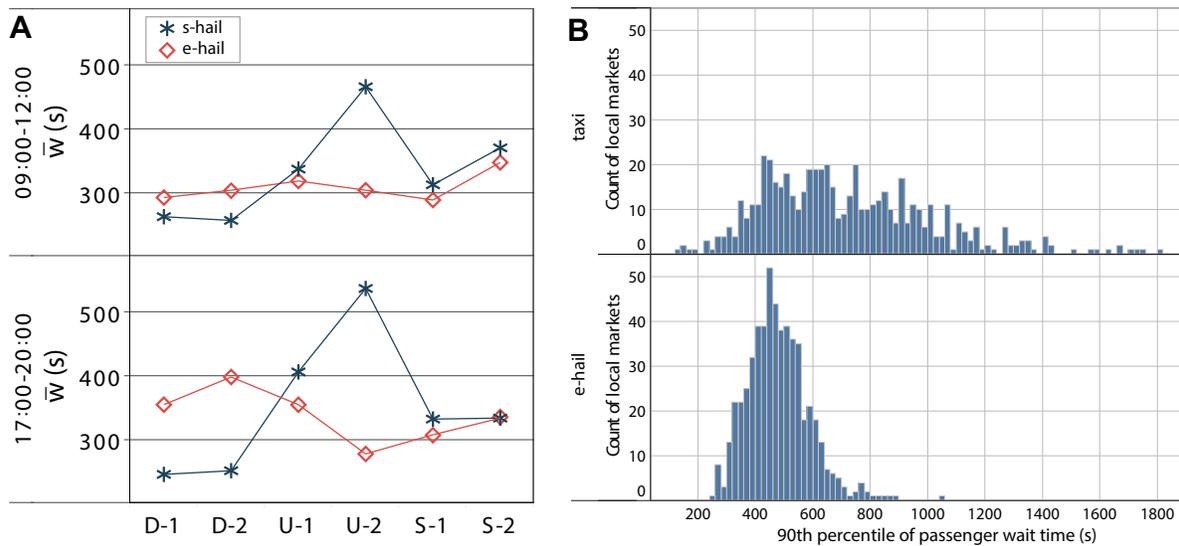
Figure 9.2. Passenger wait time of s-hail and e-hail in the six districts

a significantly higher wait time (up to 30%) in the peak period than in the off-peak period. In contrast, the discrepancy between the two periods is largely negligible for s-hail.

Again, the rising density seems to impact s-hail and e-hail in different manners. Whereas e-hail outperforms s-hail in almost all metrics in less dense areas, it struggles to keep up with s-hail in downtown areas where the density is the highest. In downtown districts, not only does s-hail produce more trips with less vacant vehicles than e-hail, it also lowers passengers' average wait time.

E-hail is often praised for providing a consistent level of service (e.g., Cramer and Krueger, 2016). That is, it largely avoids excessively long wait that taxi passengers often experience. Our empirical analysis uncovers evidence supporting this claim. As shown in Figure 9.2(A), at the aggregate level, s-hail has greater cross-district variations than e-hail. Recall that the distribution of passenger wait time for s-hail
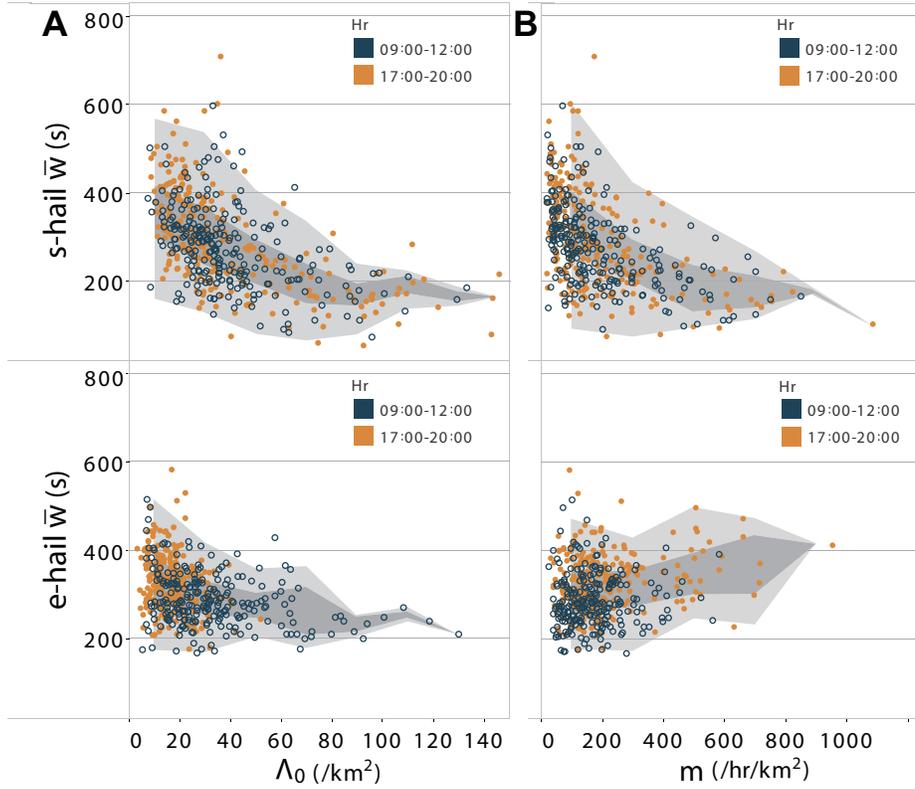
Figure 9.3. Passenger wait time against unmatched vacant vehicle density and pickup rate for s-hail and e-hail

has a longer and fatter tail than that for e-hail (see Eqs. (4.2) and (4.4)). This is also demonstrated in Figure 9.2(B), which plots the empirical distribution of 90th percentile passenger wait time in all local markets. Specifically, a maximum wait time of 10 minutes for e-hail is guaranteed in 88% of local markets with 90% confidence, while this is only ensured in 40% of local market for s-hail.

We further explore the relationships between average wait time $\bar{w}$, unmatched vacant vehicle density Λ_0, and pickup rate m at the local market level [3]. Figure 9.3 shows passenger wait time ($\bar{w}$) against (A) unmatched vacant vehicle density (Λ_0) and

---

[3]The pickup rate associated with the supply area; see Eq.(5.2) in Section 5.3.

(B) pickup rate (m) for s-hail (top) and e-hail (bottom). Each data point represents one local market in a period. The dark-gray areas denote 1st and 3rd quartiles, while the light-gray areas represent 1.5 IQR (interquartile range).

Figure 9.3(A) shows, as expected, an increase in Λ_0 tends to reduce average wait time for both s-hail and e-hail. The plot also confirms that s-hail wait time has a much wider range than e-hail does across different levels of Λ_0. It also reveals another trend: $\bar{w}$ seems to decrease more sharply for s-hail than for e-hail as unmatched vacant vehicle density increases. This finding agrees with the predictions of expected wait time formula, Eqs. (4.6) and (4.7), derived in Section 4.1. Specifically, Eq. (4.6) asserts that $\bar{w}$ is proportional to $1/\Lambda_0$ (recall $\Lambda_0 = \Lambda$ for s-hail) whereas Eq. (4.6) indicates $\bar{w}$ is proportional to $1/\sqrt{\Lambda_0}$ (recall $\Lambda_0/\Pi_0 = k\Lambda/\Pi$ from Eq. (3.10)).

The top plot of Figure 9.3(B) shows $\bar{w}$ for s-hail tends to decrease as the pickup rate increases. In other words, the service gets better as more people use it. This seemingly counter-intuitive phenomenon is known in mass transit systems as the Mohring effect (Mohring, 1972), which states that an increase in demand may shorten passenger wait time because it drives up service frequency. As explained in Section 4.2, if increasing demand attracts additional vacant vehicles, then it is bound to reduce s-hail wait time. Interestingly, e-hail demonstrates an almost opposite trend in this relationship: As the pickup rate grows, $\bar{w}$ becomes slightly longer; see the bottom panel of Figure 9.3(B). Again, our model anticipates this empirical finding. According to Eq. (4.11), as the pickup rate m rises, the vacant vehicle density Λ must increase in proportion in order to hold the wait time constant. If vacant vehicles fail to keep up with the pickup rate, the wait time may actually trend up, as revealed in the plot.

9.2. Returns to scale: verified tale of two markets

In this section we further examine the economies of scale for s-hail and e-hail. Our approach is to conduct regression on the Cobb-Douglas production function that describes the relationship between trip production (m) and inputs (Λ and Π). Our matching model (Eqs. (4.12) and (4.13)) suggests that the output elasticities of both Λ and Π are identical in both services (1 for s-hail and 0.5 for e-hail). Below, Section 9.2.1 keeps this structure in the regression. In Section 9.2.2, we relax this assumption and further control the impact of location (districts) and time (peak vs. off-peak).

9.2.1. Identical elasticity

Under the assumption that the output elasticities are identical for Λ and Π, the Cobb-Douglas function takes the following form

$$(9.1) \qquad\qquad m = \beta(\Pi\Lambda)^{\alpha},$$

where β is the total factor productivity (TFP), interpreted as the pickup rate when $\Lambda = \Pi = 1/\text{km}^2$; and α is the output elasticity of the two inputs. Note that the pickup rate used in this section is associated with the supply area of each local market; see Appendix 5.3 for more details.

To validate Eq. (9.1), we build a regression model of the following linear form:

$$(9.2) \qquad\qquad \log m = a_0 + a_1(\log \Lambda + \log \Pi) + \epsilon,$$

where ϵ is a statistical error term that follows normal distribution.

Eq. (9.2) is fitted using observed m, Λ and Π in all 490 local markets and two periods for both s-hail and e-hail. Figure 9.4 reports both the original data and the fitted linearized Cobb-Douglas functions. Each data point represents one local market at a given time period. The fitted linear equations are reported at the lower-right corner, along with their goodness-of-fit R^2. The intercept of a fitted line a_0 essentially measures the logarithm of TFP. For s-hail, $\hat{a}_0 = 0.266$ and for e-hail, it is 1.112. Hence, with unit densities of waiting passengers and vacant vehicles, the pickup rate in an e-hail market is 6 times of that in an s-hail market (i.e., $\beta^e / \beta^s \simeq 10^{\hat{a}_0^e - \hat{a}_0^s} = 7.015$). Thus, in areas with very low density, e-hail produces much more trips than s-hail, thanks to its technological superiority.

The slope of a fitted line represents the output elasticity. For s-hail, $\hat{a}_1 = 0.707$ whereas for e-hail, it is 0.461. The returns to scale of a production depends on the sum of the two elasticities. With the assumption of identical elasticity, the sum ($2\alpha \simeq 2\hat{a}_1$) is 1.413 and 0.922 for s-hail and e-hail respectively. According to the standard economic theory, therefore, s-hail displays strong increasing returns to scale ($2\alpha \gg 1$) and e-hail displays near-constant returns to scale ($2\alpha \simeq 1$). Therefore, while e-hail's technological superiority significantly boosts TFP (hence the performance in low density areas), it hurts economies of scale (hence the performance in high density areas).

We next compare the above finding with our analytical results. Eq. (4.12) suggests that s-hail display increasing returns to scale with $2\alpha = 2$ (compared to 1.413 obtained from regression), and Eq. (4.13) suggests that e-hail display constant returns to scale with $2\alpha = 1$ (compared to 0.922 obtained from regression). Thus, the theory not only broadly agrees with the underlying message that s-hail enjoys stronger economies
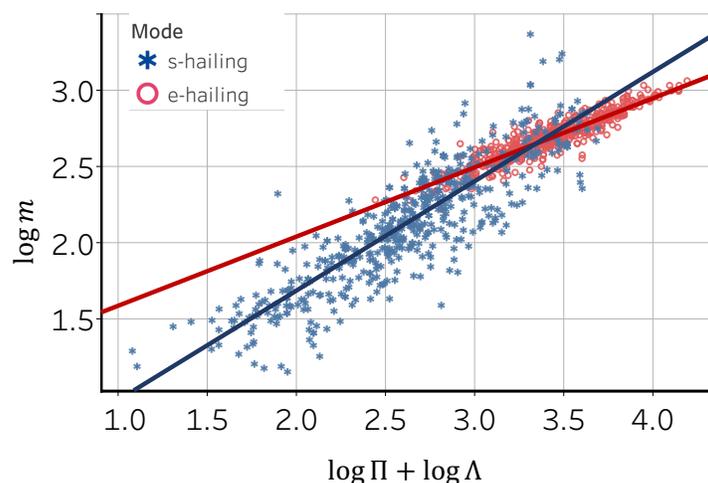
Figure 9.4. Regression results of the Cobb-Douglas function with the identical elasticity for both inputs

of scale than e-hail, it actually predicts the magnitude of returns to scale reasonably well, especially for e-hail (the error is within 10%). The model seems to overestimate s-hail's returns to scale by a relatively large margin (about 30%), likely because it completely ignores inter-passenger competition in s-hail. As to TFP, our model predicts that the value of TFP for e-hail is $2\sqrt{k}/(\sigma d)$ times of that for s-hail (see Eqs. (4.12) and (4.13)). Using the mean values reported in Chapter 8 for these parameters, We estimate $2\sqrt{k}/(\sigma d) \simeq 23.4$ ($d \simeq 0.036$ km, $\sigma \simeq 1.5$ and $k \simeq 0.4$). In other words, our model predicts that, with unit densities of vacant vehicles and waiting passengers, the pickup rate of e-hail is about 22 times greater than that of s-hail, compared to about 6 times obtained from regression. Considering this prediction is based on the three calibrated parameters (k, d and σ) that are themselves subject to substantial variances across local markets, this large discrepancy is not particularly surprising. However, the model prediction and the regression results still agree with each other that the

productivity of e-hail, as measured by TFP in the Cobb-Douglas function, is about an order of magnitude higher than that of s-hail.

9.2.2. Heterogeneous elasticity and fixed effects

In this section, we relax the assumption that Λ and Π share the same output elasticity and consider spatial and temporal impacts on the matching. Using the same data set, we fit the following model:

$$(9.3) \qquad \log m = a_0 + a_1 \log \Pi + a_2 \log \Lambda + \mathbf{x}^T \mathbf{b} + \epsilon,$$

where a_1 and a_2 are output elasticity corresponding to Π (waiting passenger density) and Λ (vacant vehicle density) and thus the returns to scale is computed as $a_1 + a_2$; $\mathbf{x}$ is a vector of dummy variables introduced to account for the time- and location-specific fixed effect (FE) with coefficients $\mathbf{b}$. Specifically, we consider six different districts and two time periods (peak and off-peak). In total we fit six models in this section, three for each service: a model with heterogeneous elasticity but without FE (model (1)), a model with both heterogeneous elasticity and FE (model (2)), and a model that considers the interactions between the time and location FE, in addition to heterogeneity (model (3)).

Table 9.1 reports the main results of the above six regressions models. For comparison, the results with identical elasticities are also reported (model (0)). All models agree with the main conclusion of the analytical result: e-hail has a substantially lower returns to scale, compared to s-hail. Models with highest adjusted R^2 values concludes taxi service yields $a_1 + a_2 = 1.314 > 1$ and e-hail yields $a_1 + a_2 = 0.957 \simeq 1$.

Table 9.1. Main results of regression

	Taxi				E-hail			
model	(0)	(1)	(2)	(3)	(0)	(1)	(2)	(3)
	$\log m$	$\log m$	$\log m$	$\log m$	$\log m$	$\log m$	$\log m$	$\log m$
a_0	0.266***	0.392***	0.488***	0.506***	1.112***	1.201***	1.114***	1.097***
	(0.048)	(0.054)	(0.070)	(0.065)	(0.030)	(0.024)	(0.027)	(0.027)
a_1	0.707***	0.821***	0.842***	0.851***	0.461***	0.676***	0.758***	0.779***
	(0.018)	(0.029)	(0.034)	(0.034)	(0.008)	(0.014)	(0.022)	(0.022)
a_2		0.541***	0.474***	0.463***		0.205***	0.185***	0.178***
		(0.038)	(0.043)	(0.044)		(0.016)	(0.021)	(0.021)
$a_1 + a_2$ or $2a_1$	1.414	1.362	1.316	1.314	0.922	0.881	0.943	0.957
Df. model	1	2	8	13	1	2	8	13
R^2	0.759	0.770	0.800	0.802	0.853	0.911	0.921	0.930
Adj. R^2	0.759	0.769	0.796	0.796	0.853	0.910	0.920	0.929
Time FE	No	No	Yes		No	No	Yes	
Location FE	No	No	Yes		No	No	Yes	
Interaction				Yes				Yes
	Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1							

Introducing FEs has minor influence on the estimates of elasticities but substantially changes the intercept of s-hail model. This finding implies the matching process of s-hail is more sensitive to spatial and temporal properties of the local market.

Notably, the estimates of a_1 and a_2 are quite different in all six models. For both s-hail and e-hail, the waiting passenger density seems to contribute more to the pickup rate, and this discrepancy is more pronounced in the case of e-hail (the ratio between the two elasticities is about 4:1, rather than 1:1 suggested by the model). While this disagreement between the model prediction and empirical results does not affect the main findings and insights, it does suggest that our parsimonious modeling effort may be unable to capture the physics of the matching process to its full extent. To close this section, it may be useful to speculate some of the reasons behind the inconsistency. On the one hand, precisely measuring the "production inputs" (i.e., Π and Λ) in each

local market is difficult because of the interactions beyond the arbitrarily selected market boundaries (see Chapter 5). Miscounting either of the inputs from data could distort the elasticities. On the other hand, some of the simplifying assumptions that are introduced for analytical tractability may also play a role. For s-hail, as we have mentioned above, the low elasticity on the vacant vehicle density may be caused by inter-passenger competition that is ignored in the model. For e-hail, a potential caveat is the assumption that the matching efficiency k is independent of Λ and Π. Given the complex relationship between matched and unmatched inputs in reality, this assumption may not always hold. However, to eliminate this assumption would require explicitly modeling the matching and pickup phases, which is beyond the scope of the dissertation.

CHAPTER 10

Conclusions

In this dissertation, we have proposed a general matching theory to describe the passenger–driver matching process in various ride-hail markets. It uncovers the underlying mechanism that governs the passenger–driver matching process and identifies the key parameters required to fully specify it. Novel calibration methods are developed to extract empirical information about passenger wait time from available data. Such information is critical for appropriate comparison of various ride-hail modes. Reliable estimates of key parameters and system performance metrics further helps better understand the relative performance of various ride-hail services.

Our results reveal, both analytically and empirically, why the ride-hail industry is such a *tale of two markets*. Essentially, there are two opposing forces at work. On the one hand, e-hail reduces the search friction by increasing the number of vacant vehicles that a passenger can reach. In low-density markets, this advantage helps dramatically improve the productivity (by as much as one order of magnitude) and lower the likelihood of unpleasantly long waits. On the other hand, the seemingly unlimited connectivity between passengers and drivers also leads to an unintended consequence. By enabling a large number of waiting passengers to compete for the same pool of vacant vehicles, e-hail induces a *congestion* effect that results in a much lower returns to scale than s-hail. The impact is most prominent in high-density markets, where e-hail holds no clear advantage over s-hail. Most importantly, unlike

s-hail, e-hail does not enjoy increasing returns to scale, a feature that would support the theory of "winner-take-all".

What lessons could we learn from the findings above? First, simply scaling up may not help an e-hail operator become more efficient, because the industry displays near-constant returns to scale. This finding may be good news for small operators, because scale is hardly a barrier to entry. Second, limiting connectivity may be beneficial sometimes. For example, an e-hail operator could intentionally limit the matching radius to avoid excessive competitions among passengers, or simply price some out ((known as surge pricing in practice, see e.g., Castillo et al., 2018).

Our analysis focuses on the relationship between the output (the pickup rates) and the inputs (the waiting passenger density and the vacant vehicle density) in the matching at some already-formed equilibrium. This approach allows us to obtain key insights about the market as a whole while staying agnostic on some factors (e.g., pricing, inter-mode competition). These factors, however, are integral to the market equilibrium and future research should definitely explore how they affect the formation of the equilibrium. We have also discussed some limitations of the proposed matching model in the previous section. Of these, the most noteworthy are the lack of demands-side congestion in s-hail and the constant matching efficiency in e-hail. To account for the congestion effect in s-hail, one must consider the possibility that the closest matchable vehicle of one passenger is "intercepted" by another passenger. Allowing the matching efficiency to vary is more complicated and a future study could pursue different paths to address the issue. One possibility, as mentioned before, is to separate the matching and the pickup phases. That is, instead of letting the wait

time depend only on Λ and Π, we could instead assume it is also affected by Λ_0 and Π_0. Alternatively, we could make exponents on Λ and Π as parameters, rather than fixing them at 1. Which approaches prevails likely depends on what data is available, and how the model is calibrated. Finally, our model is validated using data collected in one city. It would be interesting to test whether the main findings can be generalized to other cities in other countries. Given the limited availability of similar data elsewhere, this effort is also left to future research.

We close by noting ride-hail is unlikely the only industry prone to the impact of the efficiency paradox. After all, the key insight here is simple and universal: Unless the service itself can be rendered more efficiently (e.g., by capacity sharing), quick match may only bring competition that undermines the overall efficiency. Uberization is continuing to reshape our everyday life, from grocery shopping (e.g., InstaCart), fund raising (e.g., GoFundMe) to education (e.g., VipKid). Therefore, it is important to understand which industries the paradox is likely to affect, the extent to which they are affected, and in what manner. We hope this dissertation motivates others to examine this issue further, and that our theory finds applications beyond its domain.

References

, 2008. Zhenhua road closure for construction (in Chinese). `http://jt.sz.bendibao.com/news/2008519/69944.htm`, accessed: 2018-01-23.

, 2017. Introduction of shenzhen metro line 2 (in Chinese). `http://ditie.mapbar.com/shenzhen/news/124939.html`, accessed: 2018-01-23.

Afeche, P., Liu, Z., Maglaras, C., 2018. Ride-hailing networks with strategic drivers: The impact of platform control capabilities on performance. Columbia Business School Research Paper, 18–19.

Aitkin, M., Clayton, D., 1980. The fitting of exponential, weibull and extreme value distributions to complex censored survival data using glim. Applied Statistics, 156–163.

Arnott, R., 1996. Taxi travel should be subsidized. Journal of Urban Economics 40 (3), 316–333.

Banerjee, S., Riquelme, C., Johari, R., 2015. Pricing in ride-share platforms: A queueing-theoretic approach, available at: SSRN 2568258 (Accessed: 2018-11-12).

Beesley, M. E., Glaister, S., 1983. Information for regulating: the case of taxis. The Economic Journal 93 (371), 594–615.

Besbes, O., Castro, F., Lobel, I., 2018. Spatial capacity planning, available at SSRN 3292651 (Accessed: 2019-10-23).

Bimpikis, K., Candogan, O., Saban, D., 2019. Spatial pricing in ride-sharing networks. Operations Research.

Boscoe, F. P., Henry, K. A., Zdeb, M. S., 2012. A nationwide comparison of driving distance versus straight-line distance to hospitals. The Professional Geographer 64 (2), 188–196.

Buchholz, N., 2019. Spatial equilibrium, search frictions and efficient regulation in the taxi industry, available at `https://scholar.princeton.edu/sites/default/files/nbuchholz/files/taxi_draft.pdf` (Accessed: 2020-03-04).

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

Caldieraro, F., Zhang, J. Z., Cunha Jr., M., Shulman, J. D., 2018. Strategic information transmission in peer-to-peer lending markets. Journal of Marketing 82 (2), 42–63.

Castillo, J., Knoepfle, D. T., Weyl, E. G., 2018. Surge pricing solves the wild goose chase, available at SSRN 2890666 (Accessed: 2018-05-03).

Chen, L., Mislove, A., Wilson, C., 2015. Peeking beneath the hood of Uber. In: Proceedings of the 2015 Internet Measurement Conference. ACM, pp. 495–508.

Chen, M. K., Sheldon, M., 2016. Dynamic pricing in a labor market: Surge pricing and flexible work on the Uber platform. In: Ec. p. 455.

Coase, R. H., 1937. The nature of the firm. economica 4 (16), 386–405.

Cobb, C. W., Douglas, P. H., 1928. A theory of production. American Economic Review 18, 139–165.

Cramer, J., Krueger, A. B., 2016. Disruptive change in the taxi business: The case of uber. American Economic Review 106 (5), 177–82.

Daidj, N., 2018. Uberization (or uberification) of the economy. In: Advanced Methodologies and Technologies in Digital Marketing and Entrepreneurship. IGI Global, pp. 116–128.

Davis, G. F., 2015. What might replace the modern corporation: Uberization and the web page enterprise. Seattle UL Rev. 39, 501.

De Vany, A. S., 1975. Capacity utilization under alternative regulatory restraints: an analysis of taxi markets. The Journal of Political Economy, 83–94.

Douglas, G. W., 1972. Price regulation and optimal service standards: The taxicab industry. Journal of Transport Economics and Policy, 116–127.

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

Fairthorne, D., 1964. The distance between pairs of points in towns of simple geometric shape. In: Proceedings 2nd International Symposium on the Theory of Traffic Flow, Paris, 1964. OECD.

Feng, G., Kong, G., Wang, Z., 2017. We are on the way: Analysis of on-demand ride-hailing systems, available at SSRN 2960991 (2019-5-23).

Frechette, G. R., Lizzeri, A., Salz, T., 2019. Frictions in a competitive, regulated market: Evidence from taxis. American Economic Review 109 (8), 2954–92.

Halaburda, H., Jan Piskorski, M., Yıldırım, P., 2017. Competing by restricting choice: The case of matching platforms. Management Science 64 (8), 3574–3594.

Hall, J., Kendrick, C., Nosko, C., 2015. The effects of ubers surge pricing: A case study. The University of Chicago Booth School of Business.

Hall, J. V., Horton, J. J., Knoepfle, D. T., 2019. Pricing efficiently in designed markets: The case of ride-sharing, available at `http://john-joseph-horton.com/papers/uber_price.pdf` (Accessed: 2019-05-31).

He, F., Shen, Z.-J. M., 2015. Modeling taxi services with smartphone-based e-hailing applications. Transportation Research Part C: Emerging Technologies 58, 93–106.

He, F., Wang, X., Lin, X., Tang, X., 2018. Pricing and penalty/compensation strategies of a taxi-hailing platform. Transportation Research Part C: Emerging Technologies 86, 263–279.

Hotelling, H., 1938. The general welfare in relation to problems of taxation and of railway and utility rates. Econometrica: Journal of the Econometric Society, 242–269.

Hu, M., Zhou, Y., 2018. Dynamic type matching, available at SSRN 2592622 (Accessed: 2019-8-4).

Kay, R., 1977. Proportional hazard regression models and the analysis of censored survival data. Applied Statistics, 227–237.

Kim, K., Baek, C., Lee, J.-D., 2018. Creative destruction of the sharing economy in action: The case of uber. Transportation Research Part A: Policy and Practice 110, 118–127.

Lagos, R., 2000. An alternative approach to search frictions. Journal of Political Economy 108 (5), 851–873.

Lagos, R., 2003. An analysis of the market for taxicab rides in New York City. International Economic Review 44 (2), 423–434.

Laird, N., Olivier, D., 1981. Covariance analysis of censored survival data using log-linear analysis techniques. Journal of the American Statistical Association 76 (374), 231–240.

Little, J. D., 1961. A proof for the queuing formula: L= λ w. Operations research 9 (3), 383–387.

Mohring, H., 1972. Optimization and scale economies in urban bus transportation. The American Economic Review 62 (4), 591–604.

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.

Nourinejad, M., Ramezani, M., 2019. Ride-sourcing modeling and pricing in non-equilibrium two-sided markets. Transportation Research Part B: Methodological.

Özkan, E., Ward, A. R., 2020. Dynamic matching for real-time ride sharing. Stochastic Systems.

Rysman, M., 2009. The economics of two-sided markets. Journal of economic perspectives 23 (3), 125–43.

Schroeter, J. R., 1983. A model of taxi service under fare structure and fleet size regulation. The Bell Journal of Economics, 81–96.

Shapiro, M. H., 2018. Density of Demand and the Benefit of Uber, availabel at `http://www.shapiromh.com`. (Accessed: 2019-02-19).

Tucker, C., Zhang, J., 2010. Growing two-sided networks by advertising the user base: A field experiment. Marketing Science 29 (5), 805–814.

Wang, H., Yang, H., 2019. Ridesourcing systems: A framework and review. Transportation Research Part B: Methodological 129, 122–155.

Wang, X., He, F., Yang, H., Gao, H. O., 2016. Pricing strategies for a taxi-hailing platform. Transportation Research Part E: Logistics and Transportation Review 93, 212–231.

Xu, Z., Yin, Y., Ye, J., 2019. On the supply function of ride-hailing systems. Transportation Research Part C: Emerging Technologies 00.

Yan, C., Zhu, H., Korolko, N., Woodard, D., 2019. Dynamic pricing and matching in ride-hailing platforms. Naval Research Logistics (NRL).

Yang, H., Ke, J., Ye, J., 2018. A universal distribution law of network detour ratios. Transportation Research Part C: Emerging Technologies 96, 22–37.

Yang, H., Leung, C. W., Wong, S., Bell, M. G., 2010. Equilibria of bilateral taxi–customer searching and meeting on networks. Transportation Research Part B: Methodological 44 (8), 1067–1083.

Yang, H., Qin, X., Ke, J., Ye, J., 2020. Optimizing matching time interval and matching radius in on-demand ride-sourcing markets. Transportation Research Part B: Methodological 131, 84–105.

Yang, H., Yang, T., 2011. Equilibrium properties of taxi markets with search frictions. Transportation Research Part B: Methodological 45 (4), 696–713.

Yang, T., Yang, H., 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., Sze, N. N., 2014b. Returns to scale in the production of taxi services: an empirical analysis. Transportmetrica A: Transport Science 10 (9), 775–790.

Yao, S., Mela, C. F., 2008. Online auction demand. Marketing Science 27 (5), 861–885.

Zervas, G., Proserpio, D., Byers, J. W., 2017. The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry. Journal of marketing research 54 (5), 687–705.

Zha, L., Yin, Y., Xu, Z., 2018. Geometric matching and spatial pricing in ride-sourcing markets. Transportation Research Part C: Emerging Technologies 92, 58–75.

Zha, L., Yin, Y., Yang, H., 2016. Economic analysis of ride-sourcing markets. Transportation Research Part C: Emerging Technologies 71, 249–266.