Pricing and Efficiency in a Decentralized Ride-Hailing Platform*

Renata Gaineddenova†

January 16, 2022

Click here for the most recent version

Abstract

Asymmetric information about market participants' valuations and costs plays a crucial role in the efficiency of a platform's design. Using novel data from a ride-hailing platform called inDriver, I examine whether decentralizing the pricing mechanism improves market efficiency. Unlike its competitors, inDriver requires riders to offer a price for their requested trips, and allows drivers to either agree to the offer, ask for a higher price, or ignore the request. Under this mechanism, a rider with a high willingness to pay for a trip can offer a higher price to increase her chances of being matched. At the same time, under decentralized pricing riders might not truthfully reveal their valuations, which can result in lower average prices on the platform. To understand welfare implications of decentralized pricing for riders and drivers, I develop an equilibrium model of a decentralized ride-hailing market and estimate its parameters using user-level data on the universe of ride requests in a single city. I then use the obtained estimates to compare welfare under a decentralized mechanism to an alternative mechanism in which prices are chosen by the platform. I find that decentralized pricing significantly improves efficiency in the studied market.

*I would like to thank Jean-François Houde, Alan Sorensen, and Ken Hendricks for their invaluable advice and guidance. I am indebted to Arsen Tomsky and the inDriver team for providing the data, answering my questions, and making this project possible. I am also grateful to Nicolas Badaracco, Emilio Calvano, John Kennan, Lorenzo Magnolfi, Paolo Martellini, Arpita Patnaik, Andrea Pozzi, Dan Quint, Sandra Spirovska, Chris Sullivan, Christopher Taber, Joanna Venator, and participants of the University of Wisconsin-Madison IO Seminar for their useful comments and suggestions. All errors are my own. This research was supported by the Caves Dissertation Fellowship, Christensen Graduate Fellowship in Empirical Economics and Juli Plant Grainger Summer Research Scholarship.

†Department of Economics, University of Wisconsin-Madison, USA (e-mail: gaineddenova@wisc.edu).

1 Introduction

Efficient market design is one of the cornerstone problems in economics. Modern online matching intermediaries (platforms) use a myriad of mechanisms to connect buyers and sellers. While some of them choose to act as market planners, deciding which prices to charge and which agents to match, others allow market participants to reach an agreement on their own terms. The efficiency of an employed mechanism greatly depends on the presence of asymmetric information and the platform's ability to elicit match-relevant private information (Akerlof, 1978). If the matching intermediary is always perfectly informed, an efficient allocation can be reached simply through personalized posted prices. However, it would be unrealistic to assume that any platform could have complete information about all buyers' willingness to pay and sellers' willingness to sell at all points in time.

In this paper, I explore the importance of information disclosure through prices and study the trade-offs associated with centralized pricing and decentralized pricing in the context of a ride-hailing platform. Researchers tend to consider ride-hailing markets to be very homogeneous and argue that there is no need for price discovery (Einav, Farronato, Levin, and Sundaresan, 2018). Yet, there are still vast differences in how ride-hailing platforms choose to operate. While Uber-like centralized market forms have been dominant in the United States, other decentralized market organization forms have emerged around the globe.[1] Allowing agents to offer their own prices to signal private information can eliminate information gaps and improve match quality (Spence 1978, Tadelis and Zettelmeyer 2015). However, by delegating pricing decisions to market participants, an intermediary loses control over the prices. As a result, participants might not truthfully reveal (shade) their valuations; this behavior can lead to lower equilibrium prices, which can harm the other side of the market.

To study how decentralization in pricing affects efficiency in a ride-hailing market, I use rich consumer-choice data from a unique platform — inDriver, which provides on-demand trips. InDriver originated in 2013 in Yakutsk, Russia, reached a unicorn status in 2021[2] and now operates in 37 countries with 100 million registered users. The platform states that it provides "freedom of choice," allowing riders and drivers to bargain over the prices with a well-defined set of rules and a timeline. The platform is unique in its price-setting mechanism: the rider offers a price, and drivers

[1] The dominant U.S. players with a centralized model are Uber and Lyft. Examples of decentralized platforms in other parts of the world include inDriver, Liftago, and Drife.

[2] In finance, the term "unicorn" denotes a privately held startup company valued at over $1 billion.

then decide whether to accept, counter with a higher price, or decline the request altogether.

With data on the universe of attempted trips made in a single city, I document several important market patterns, which I later incorporate into the model. First, I show that riders have heterogeneous willingness to pay: different riders offer different prices for trips with exactly the same characteristics, including distance, creation time, origin, and destination, as well as idiosyncratic market conditions that each rider faces, such as the number of nearby drivers. Second, I document the existence of a trade-off between higher offered price and better chances of finding a quick match. On one hand, offering a higher price increases a rider's chances of getting a ride quickly; on the other hand, it increases his expected payment. Third, I document two important facts about drivers' behavior on the platform: (1) I show that drivers are selective, preferring shorter, more expensive trips with a shorter pickup distance, and (2) I present evidence that drivers behave strategically, participating less frequently when incentives to wait for a better match are higher.

To study the welfare effects due to decentralization in pricing, I model the decisions of both sides of the market. In the demand-side model, riders decide to take a trip prior to arriving at the platform and differ in their trip valuation. They try to get matched only once and choose to participate only if the trip generates a positive expected surplus ex ante.[3] Upon arrival, each rider observes market conditions (state) and forms beliefs about the respective probabilities of getting matched at two different prices, low and high.[4] He then chooses the price that maximizes his expected utility. Similarly to the setup in Kawai, Onishi, and Uetake (2021), the heterogeneity in willingness to pay and a trade-off between offered price and matching rates result in a single-crossing property. Riders with low valuations choose a low price, while high-valuation riders prefer to offer more to ensure higher matching rates (signaling).

In the supply-side model, I formulate an idle driver's decision about how to respond to a request. I model drivers' decisions as a simultaneous-move game with incomplete information. Idle drivers observe incoming requests, which vary by type, then choose whether to accept the request, make a counteroffer, or ignore the request and remain idle. This decision is dynamic: drivers can get matched now or wait and be matched with a potentially better request in the future. Accepting

[3] In the data, around 80% of riders are matched within 20 seconds. Among the remaining riders, a third increase their price eventually, while the majority keep the same price and end up waiting longer or dropping their requests entirely. These numbers vary significantly across different hours of the day. In the estimation, I use data from times of day when the one-shot assumption is closest to being satisfied.

[4] A peculiar (but convenient) aspect of the market I study is that the distribution of offered prices is discrete: during most times of day, just two price points account for over 93% of requests. This feature probably stems from the fact that transactions are cash-based.

an offered price or making a counteroffer (entering the matching stage) does not guarantee that the driver will get the ride, because other competing drivers may have also entered the matching stage. For each request type, a driver has to form beliefs about the expected payoffs from winning or losing the matching stage. If a driver wins, he forms a match and will (for a duration) no longer be searching for a trip. If a driver loses at the matching stage, he loses some time and remains idle. Participating drivers in the matching stage, therefore, always bear an implicit cost. While trying to get matched, a driver cannot observe other, potentially better, requests that arrive during a matching stage in which he participates.

Bringing the supply and demand sides of the market together, I define the equilibrium in the platform. I use a large market equilibrium concept, where the equilibrium is a set that consists of drivers' and riders' beliefs such that both riders and drivers behave optimally given those beliefs, and the beliefs are consistent with the played equilibrium. This concept is similar to the oblivious equilibrium concept of Weintraub, Benkard, and Van Roy (2008). I assume that all market participants make their decisions based only on their own state and the long-run average market state.

I then take the model to the data, estimating the demand-side and supply-side models separately. To estimate both models, I use a two-step approach. First, I estimate participants' equilibrium beliefs directly off the data. Second, I recover model primitives, assuming that participants behave optimally given these beliefs. The main aim of the demand-side estimation is to recover the distribution of trip valuations. For the demand-side model, my estimation approach is similar to that of Agarwal and Somaini (2018). To recover parameters of interest, I use plausibly exogenous variation in the choice environment that riders face when making their pricing decision. For the supply-side model, my estimation is similar to the one developed for dynamic models of imperfect information (Bajari, Benkard, and Levin 2007, Aguirregabiria and Mira 2007). I use the supply-side model to estimate the parameters of a drivers' static utility and a measure of impatience, as well as a variance of unobserved shocks that influence a driver's decision. The main identification assumption is that drivers cannot influence the type of requests they observe (they can, however, ignore them).[5] The fact that I observe drivers who participated (those who agreed or who made a counteroffer) as well as drivers who ignored the requests allows me to leverage the drivers' responses to various types of requests and identify parameters of interest. With the estimates of drivers' beliefs in hand, I estimate the parameters of interest using the nested fixed point procedure

[5]I consider a compact city, and the data support the fact that drivers do not significantly change their location between the end and the beginning of two consecutive trips.

developed by Rust (1987).

Given the described estimates and others I obtained directly from the data, I then turn to the counterfactual analysis to evaluate the welfare consequences of pricing decentralization for drivers and riders. I compare the existing mechanism to a counterfactual alternative in which the platform chooses either a high or a low price for each rider. I let the platform observe everything that a rider observes at the decision time: the number of idle drivers and their distances to a rider. The platform, however, does not observe a rider's valuation for a trip. I analyze how the equilibrium evolves between various centralized pricing regimes (i.e., how often the platform chooses a high price based on market observables) and what each of them implies for participants' welfare. I find that pricing centralization decreases drivers' welfare by at least 10% (in the regime that corresponds to drivers' optimum) and riders' welfare by at least 4% (in the regime that corresponds to riders' optimum) relative to the existing decentralized mechanism.

To better understand these empirical results and the effects of decentralized pricing (DP), I discuss how market participants' welfare changes when a centralized pricing (CP) is implemented. Riders differ in their attitudes toward centralized pricing. High-valuation riders always lose from CP: the centralized platform takes away their ability to shade (hide) and signal their valuations, and their matching rates are now flattened. Under CP, these riders prefer regimes in which the platform charges higher prices more often — such a regime guarantees them higher matching rates. On the contrary, low-valuation riders will potentially benefit from CP if the platform does not charge high prices too often. Signaling from high-valuation riders in a DP imposes a negative externality on low-valuation riders by making drivers more selective. Nevertheless, these riders strongly dislike CP with high prices, since they will be priced out of the market.

Driver's optimum welfare under CP corresponds to a regime with third-degree price discrimination (not all prices are low or high). This optimum value of welfare, however, still falls short of the welfare achieved under DP. Under CP, drivers benefit from higher average prices, but an increase in prices does not fully compensate for the share of lost riders who are priced out of the market. The results for drivers can vary by the market environment. Markets where riders have more incentives to shade their valuations (for instance, markets with more drivers) could theoretically show higher welfare for drivers under CP.

Related Literature

My paper contributes to four strands of literature. First, I study how different pricing mechanisms encourage voluntary information disclosure, extending past literature on private information and market efficiency to a new setting. Starting with the seminal work of Akerlof (1978), multiple studies have highlighted the importance of asymmetric information for various industries. In a recent theoretical paper, Akbarpour, Li, and Gharan (2020) show that efficiency of a centralized matching algorithm without prices depends on information provision: if the platform knows which agents are terminal, it can achieve a more efficient allocation.[6] Closer to my empirical setting, Tadelis and Zettelmeyer (2015) empirically show that information disclosure leads to improved efficiency in an auction market. In this paper, I argue that delegating pricing decisions to market participants facilitates voluntary information disclosure through signaling and can improve market efficiency.

Second, my paper provides a unique empirical framework for understanding the signaling effect, which has primarily been studied in a theoretical framework. While there exists an extensive theoretical branch on signaling in various settings (Milgrom and Roberts 1986, Milde and Riley 1988, Jullien and Mariotti 2006, Cai, Riley, and Ye 2007), the empirical evidence is still nascent (Backus, Blake, and Tadelis 2019, Sahni and Nair 2020, Kawai et al. 2021).

Third, my paper develops a novel equilibrium model of a ride-hailing market with decentralized pricing. Transportation markets have received much attention, starting with Lagos (2000). The change in how the industry operates, along with the availability of new datasets, has given researchers an opportunity to study various aspects of this market: labor-market outcomes (Angrist, Caldwell, and Hall 2017, Hall and Krueger 2018, Chen, Rossi, Chevalier, and Oehlsen 2019), gender gap in earnings (Cook, Diamond, Hall, List, and Oyer, 2018), and platforms' competition (Rosaia, 2020). Using a similar decentralized setting but in which prices are offered by drivers, Buchholz, Doval, Kastl, Matejka, and Salz (2019) study the time valuation of the riders. Several papers have focused on efficiency implications in particular: the effect of regulation on welfare in a taxicab market (Frechette, Lizzeri, and Salz, 2019), the effect of inefficient pricing tariffs (Buchholz, 2018), inefficiencies in decentralized transportation markets with search frictions (Brancaccio, Kalouptsidi, and Papageorgiou, 2020), and surge pricing in a centralized ride-hailing market (Castillo, 2019). In the model, I drop a bilateral meeting assumption and show how the interaction

[6] In an empirical application, using data from a small peer-to-peer segment of Didi, Liu, Wan, and Yang (2019) show that if participants are not too heterogeneous, the centralization in matching rule (without any pricing considerations) achieves better market outcomes.

between riders' and drivers' decisions forms the "market" through which agents meet each other. I then explore how this decentralized mechanism affects the welfare of market participants in a two-sided market.

Fourth, this project is closely tied to the recent empirical literature in industrial organization that studies various pricing-related questions in the online platforms (Wei and Lin 2017, Einav et al. 2018, Dubé and Misra 2019). In this paper, I study a new market-driven pricing mechanism that implements self-regulated price-discrimination and can simultaneously increase the welfare of both buyers and sellers.

The rest of the paper is organized as follows. Section 2 describes the empirical environment and presents some stylized facts about the market functioning, which I later incorporate into the model. Section 3 introduces the decentralized equilibrium model. Section 4 discusses the identification of model primitives. Section 5 presents the estimation results. Section 6 describes the counterfactual scenario, in which the platform (not the rider) sets the prices, and presents the welfare analysis. Section 7 concludes.

2 Empirical Environment and Data

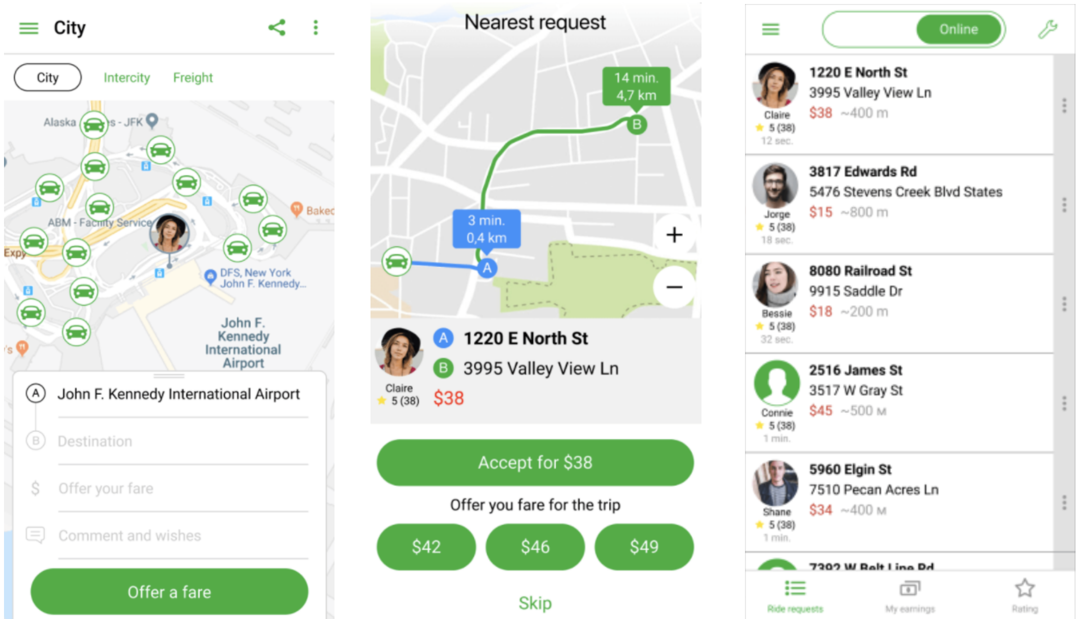
2.1 Platform: inDriver

To study a decentralized ride-hailing market, I use a proprietary, user-level dataset from the platform inDriver. This platform launched in 2013 in Yakutia, Russia, the region known for having the lowest temperature in the Northern Hemisphere. In January 2013, the temperature in Yakutsk (Yakutia's administrative center) dropped below -45C°, and the local taxicab companies quickly jacked up their prices. City dwellers reacted by starting a group on VK (the Russian analog of Facebook), where potential riders posted their requests for riders, and independent drivers could contact them and offer rides. This endeavor was an instant success: in six months, more than 60,000 people joined the group. Later that year, the administrative rights to the group were purchased by the company Sinet, which turned the group into the mobile app inDriver.[7] The company now operates in 37 countries, has more than 100 million registered users; in 2021, it became a unicorn. The app has gone through many transformations since 2013, but the main feature has remained unchanged: the request starts with a rider offering a price.

inDriver operates several separate platforms: within-city travel, between-city travel, freight,

[7] www.indriver.com

Figure 1: inDriver Screens



Notes: This figure shows various screens seen by app users. The left figure corresponds to the rider's screen when he opens an app. The central figure corresponds to a screen that pops up on the driver's screen if the platform has identified him as one of the closest drivers and directly notifies him. The right figure shows a driver's feed. The driver can press on a request to see more detailed information. In that case, he will see a screen similar to the one in the center.

and delivery segments. I study the within-city segment, which has the same purpose as Uber and Lyft: to instantly connect riders looking for a trip with drivers who are operating on the spot. The request starts with a rider specifying his current location, desired destination, and the price he is offering for a ride. When a rider opens the app, he sees the portion of the map surrounding his current location and nearby active drivers.

The left image of Figure 1 depicts a rider's screen. Once the rider requests a trip, inDriver identifies the closest available drivers and directly notifies them about the incoming request. However, close-by drivers are not the only ones who can see the request: idle drivers looking for a trip see the request in the feed and can also try to get matched with the rider. The middle image of Figure 1 depicts a screen of a directly notified driver. The right image of Figure 1 depicts the feed from which not-notified drivers can see the request. Drivers who tap on the request will be taken to a screen containing more detailed information about the request (very similar to a screen seen by a notified driver).

The fact that inDriver delegates the price choice to riders is not the only feature that differentiates it from centralized Uber-like centralized platforms. In fact, the name "inDriver" is a combination of two words, **in**dependent **driver**, highlighting that the drivers also make their own decisions on the platform. Since inDriver does not explicitly assign drivers to requests, drivers have to choose what to do when a request arrives on the platform. There are three options. A driver can agree to the posted price, $\{A\}$, make a counteroffer, $\{C\}$ (there are three discrete options for counteroffer price precalculated by the platform based on the rider's initial offer), or ignore the request, $\{I\}$. If the driver agrees or counteroffers, he enters a matching stage.

Once inDriver receives the drivers' responses, the platform proposes a match to a rider. During the period I study, inDriver prioritized drivers who agreed with the rider's price and allocated the requests among them using a proprietary algorithm based on the pickup distance.[8] If the platform does not observe any driver who has agreed, it shows the rider a menu of drivers who have made counteroffers and lets a rider pick. [9] If the rider faces a menu, he is given 20 seconds to make a final decision. In what follows, I focus on the first 20 seconds from the time when the request appears.

Participation in the platform has always been free of charge for riders. During the time span I study in this paper, the drivers had to purchase access to the platform to participate. Each interested driver could have purchased access for either 1 hour, 4 hours, or 8 hours. There are no additional charges for the drivers at the time.

2.2 Data Description

I use the universe of inDriver trip data (both attempted and completed) from a single city. The dataset spans the period from November 1st, 2018, to April 3rd, 2019. For each request originated on the platform, I observe the time stamp when the request appeared, a rider's ID (a randomly generated number that cannot be linked to rider's real identity), and the details of the request: the initially offered price, the origin and destination (both defined as string values), and the coordinates of the rider at the time of the request. I directly observe which drivers are identified as the closest ones to the rider and their respective coordinates. In addition, I see a set of drivers that chose to enter a matching stage (by either agreeing or counteroffering). If a driver won the request and was matched to the rider, the platform tracked the request completion, providing additional information about the driver's location. For each driver present on the platform, I know the time when he became active (purchased access) and when his system time was set to expire.

[8] Although the platform does not reveal the exact details of the algorithm to the drivers, they are aware that *both* factors affect their chances of getting the request.

[9] In late 2019, the platform decided to give full freedom to the rider; it not longer allocates drivers in any way, letting riders pick from a menu.

I study a relatively isolated, compact city with a well-defined grid structure and clear boundaries.[10] inDriver allows idle drivers in the system to observe requests within a certain radius, even if they are not directly notified about the request. This radius is rarely binding for the location I study. Drivers know where requests appear and do not have to search for riders physically. Drivers do not significantly change their locations after completing a request and before getting matched with the next rider. These two facts suggest that spatial differences do not play a significant role in the location I study. As a result, I do not distinguish between city locations in the model, and I do not model drivers' strategic location choices within the city.

The data required additional cleaning. First, riders record their desired destination as a string value. While this feature appears convenient for both riders and drivers (most people know the streets and addresses in the city well), it means that the full coordinates of the intended trips are not directly present. To overcome this issue, I use the following procedure. The city has K distinct administrative districts.[11] For each district, I collected the street names and building numbers. Then, using textual analysis, I successfully matched the destination address with the district identifier for more than 80% of the sample. None of the evidence I have suggests that the rest of the sample is significantly different. There are multiple ways one can write down the same address (and make typing errors), which makes it hard to match all of them precisely. A brief examination of the unmatched requests suggests that these observations could be successfully matched if necessary. To further classify trip distances, I focus on the successfully matched requests, as they provide additional information for the trip distances. The platform asks the driver to press the "Complete" button once he finishes the trip, then records his location at that time. For each pair origin - destination district, I calculate a distribution of trip distances between districts and classify the pair to represent either a short or a long trip. If the median value for a distance between the two districts is below 1.6 km (mean value in a sample), I treat all requests between these districts as short trips. Otherwise, I classify a pair to represent a long trip. I do not calculate the exact driving distance for every single request.

Second, I needed to construct a list of idle drivers for each request (potential matches). The original dataset contains the coordinates and IDs of all drivers who were directly notified about each request. However, a significant share of drivers participate in the matching stage even if

[10] Per the inDriver's request, I cannot disclose the city name. The city's population is between 100,000 and 200,000 people.

[11] I do not report the exact number K so as not to disclose the studied location. The reader can think of K as a number greater than 30.

they are not directly notified. Instead, they observe a request in their feed (around a quarter of drivers who enter a matching stage within the first 20 seconds are not notified). In addition, some drivers were not notified and eventually decided not to participate in a matching stage but were still considering a request. To define the complete set of idle active drivers for each request, I have identified the drivers who (1) had access to the system, (2) were actively participating (trying to match with any rider within one minute of the request's arrival),[12] and (3) were not busy with any other requests at the time. Since requests arrive frequently, some of them inevitably overlap, and drivers miss some requests while participating in the matching stage of another request. In addition, for drivers identified as idle and active, I have successfully reconstructed the coordinates at the time of the request for more than 91% of the drivers in my sample. I assume that the driver's location does not change significantly within 30 seconds. Because drivers' movements in the city are controlled by speed limits, I believe this approach approximates drivers' locations with a reasonable degree of accuracy.

2.3 Data Summary

Table 1 presents daily statistics for activity on the platform for the initial sample. The platform is in very active use both on weekdays (Monday through Friday) and the weekends. On average, there are more than 10,000 ride requests and 800 participating drivers, but the table suggests there is plenty of variation between days.

Table 1: Daily inDriver Statistics

	Observations	10th %tile	Mean	90th %tile	SD
Monday–Friday					
— Number of requests	110	8,632.1	11,272.8	15,209.1	2,866.0
— Number of unique drivers	110	750.8	855.4	956.1	84.0
Saturday–Sunday					
— Number of requests	44	9,739.3	11,914.6	14,453.7	2,006.1
— Number of unique drivers	44	756.3	854.5	941.1	81.6

Notes: This table shows summary statistics for the number of daily requests and the number of unique drivers participating on the platform for the period I study.

[12] Some other factors might make drivers look active, while in fact, they are not available. For instance, some drivers forget to turn off their app after completing a ride and going off the clock. Since the system does not force drivers to be present for the whole duration of their purchased shift, they can choose to stop driving at any time. The system eventually recognizes such drivers as inactive, but not instantly.

Figures A1 and A2 depict the number of daily requests and unique drivers over time. Overall, the platform is stable, with some notable exceptions. A close examination of days-outliers reveals that they correspond either to holidays or to days with extreme weather conditions (e.g., with weather temperature below –20C°). I drop such days from my main sample, along with some other requests: those with rare prices offered for a trip or unmatched coordinates, etc. For additional information on data cleaning and criteria for exclusion, see Table A1.[13]

In addition to the variation across days, there exists significant variability across hours within a given day. Figure 2 (upper panel) shows the hourly distribution of requests on weekdays and weekends. On weekdays, the number of requests in the morning hours spikes dramatically around 7 am. The demand levels off by midmorning, is stable in the afternoon, and spikes again in the evening. The arrival of riders seeking rides looks noticeably different on the weekends. The increase in the number of riders in the morning hours is modest, stays more or less level throughout the day, then peaks slightly in the evening.

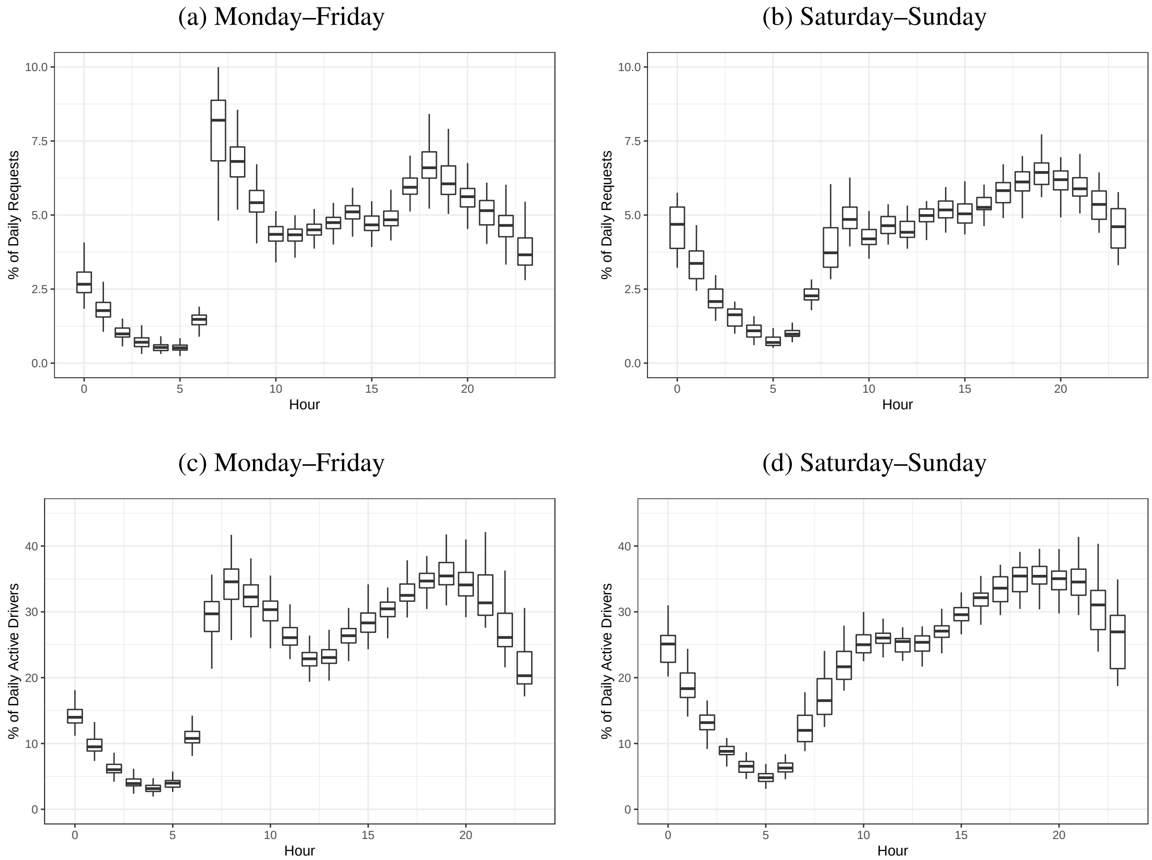
Supply closely follows the demand patterns closely. Figure 2 (lower panel) shows the hourly distribution of active drivers on weekdays and weekends. There are more active drivers (around 40% of all daily active drivers) in the morning and evening hours during the week. On the weekends, like demand, supply slowly increases in the morning, levels off, then continues to rise in the afternoon and into the early evening, before subsiding.

Though the platform did not make any pricing recommendations to the riders at the time of the sample, three prices emerged as the most frequent (defined in units of local currency): 300 (53.3%), 350 (34.1%) and 400 (6.8%).[14] Riders offered these three prices in more than 94% of requests. The discreteness in the observed prices was probably driven by the fact that all transactions were cash-based at the time: it was more convenient to pay with specific bills. There are potentially many reasons why a rider might have started his request with a certain price. Requests differ in many dimensions: some appear in rush hours, some have a longer trip distance, some originate or end in locations not favored by drivers. Later in this section, I present evidence that controlling for many observable factors does not fully explain the offered prices, suggesting that plenty of unobserved heterogeneity exists among riders.

[13] I scale arrival probabilities in the estimation to account for the fact that some observations were dropped in the cleaning process because I did not match all coordinates.

[14] I do not specify the currency so as not to disclose the city.

Figure 2: Hourly Distribution of Requests and Active Drivers



Notes: The left panel shows the distributions of the percentage of daily requests and the percentage of daily active drivers present on the platform in a particular hour on weekdays. The right panel depicts distribution for the same variables on weekends. The x-axis depicts military time.

2.4 Stylized Facts

Before moving to the model, I document several important stylized facts that characterize inDriver's decentralized ride-hailing system. I later incorporate these facts into the model: (1) the multiplicity of offered prices for similar trips (rider heterogeneity), (2) the existence of a trade-off between offered price and matching rates, (3) the fact that riders mainly use the platform to get an instant match, and (4) evidence that drivers are selective and (5) forward-looking in their decision on how to respond to a request.

Fact 1: Riders offer different prices for similar trips. As mentioned earlier, three different prices were offered on the platform in more than 94% of requests: 300, 350, and 400. However, it is possible that prices are uniform once conditioned on *trip characteristics* (e.g., day of the week, trip distance, origin and destination districts) and *changing market conditions* (e.g., how many drivers a rider sees once he opens the app). For example, all riders could offer high prices (350 or 400) at particular times of the day, such as morning rush hours. Alternatively, price variation could have been fully explained by the differences in the origin area (such as the city center or an airport). Nevertheless, the share of high prices never falls below 32% for any market.

To show that a significant share of choices in the sample cannot be simply explained by the observed variables (both stable and changing), I characterize the probability of offering a high price (either 350 or 400) using a logit regression where I include various controls that could potentially rationalize variation in prices offered by riders. Table 2 reports the results. Typically, R^2 is used to show how much of the variation in a dependent variable can be explained under the proposed model. However, binary outcome models do not allow using standard R^2. I employ a different measure to evaluate the proportion of variation explained by a proposed model widely used for the logit model: pseudo-R^2 developed by McKelvey and Zavoina (1975).[15] In addition, I calculate the mean absolute error for each model to assess the percentage of misclassified observations.

Table 2: Riders' Responses

Dependent Variable:	$\mathbb{1}$(Rider offers high price)				
	(1)	(2)	(3)	(4)	(5)
Long trip	✓	✓	✓		
Hour FE	✓	✓	✓	✓	
Weekday/Weekend	✓				
Origin FE			✓	✓	
Destination FE				✓	✓
Origin-Hour FE					✓
Date FE		✓	✓	✓	✓
Controls for changing market conditions	✓	✓	✓	✓	✓
Number of observations	763,005	763,005	763,005	763,005	763,005
McKelvey and Zavoina's pseudo-R^2	0.078	0.089	0.107	0.095	0.099
Mean Absolute Error	0.443	0.439	0.432	0.436	0.435

Notes: This table reports the results of the logit regressions of the likelihood of a rider offering a high price (either 350 or 400) on various controls. The controls included in the model are marked. I use the city-district identifier for the origin and destination FE. The number of idle drivers within a 0.5-km radius is included as the control for changing market conditions. I calculate McKelvey and Zavoina's pseudo-R^2 to assess the proportion of explained variation under each model.

The results show that even the most comprehensive fixed effects do not fully explain riders'

[15] The most common substitute for R^2 is McFadden's pseudo-R^2. However, it is commonly used for assessing goodness-of-fit rather than fo showing the proportion of explained variance. It tends to be low, even when the underlying regression is a perfect fit.

choices. Pseudo-R^2 is never above 0.11, and the mean absolute error never falls below 0.44. These findings suggest that other factors significantly contribute to the rider's decision beyond what can be explained by observables. This result is similar to the result in empirical auction literature that highlights that significant heterogeneity is present on the bidder's side (e.g., Krasnokutskaya, 2011). I treat this result as an indication that riders value trips differently and offer varying prices even when facing the same choice environment.

Fact 2: Offering a higher price is associated with better chances of facing an immediate match, all other things equal. I have previously established that riders offer varying prices for trips that look homogeneous based on observables. Thus, a trade-off must exist for such behavior to emerge. To study the trade-off involved in the rider's decision about which price to offer, I explore whether offering high or low prices results in different outcomes for a rider after controlling for observables.

Specifically, I focus on the probability with which a rider finds a match based on the price he offers. When calculating the total probability to find a match, I sum up the probabilities (1) that a request is accepted at the initial price and (2) that a rider receives a counteroffer, even if he refuses it afterwards. I run a linear probability model to explore whether offering a higher price is associated with a higher probability to find a match within 20 seconds. Table 3 reports the results.

The results suggest that offering a higher price is indeed associated with better chances to find an immediate match both for total probability (by more than 7 percentage points if there are no drivers within a 500-meter radius) and a probability of acceptance (by 18 percentage points in the same circumstances). This gap narrows for a higher number of idle drivers, to 4.2 percentage points (total probability) and 15.4 percentage points (acceptance probability) for the mean number of visible drivers, which is 3.6.

Fact 3: Most riders try to get matched only once. An assumption that riders try to find a match only once significantly simplifies the model. Table 4 presents empirical outcomes for the requests in the sample. A vast majority of requests, more than 75%, were matched within the first 20 seconds. Among unmatched riders, the majority (around 65%) did not increase the price before dropping out or being matched. Those that did not increase their price and stay incurred waiting costs. On average, these riders spent around 1.5 minutes on the platform before they get matched to a driver. Although the presented empirical evidence suggests that, on average, riders indeed made one attempt to get matched, the numbers differ significantly across markets. The weekday morning rush hour is characterized by much lower probabilities of an instant match, and more

Table 3: Probability of Facing a Quick Match (within 20 seconds)

	Total prob.	Acceptance prob.
High price	0.071***	0.183***
	(0.002)	(0.002)
N. of idle drivers within 500 meters	0.044***	0.046***
	(0.0003)	(0.0003)
High price $\times$ N. of idle drivers within 500 meters	-0.008***	-0.008***
	(0.0005)	(0.0005)
Short trip	0.027***	0.085***
	(0.001)	(0.001)
Constant	0.263***	0.072***
	(0.007)	(0.007)
Hour FE	✓	✓
Date FE	✓	✓
Observations	763,005	763,005
R^2	0.104	0.130

Notes: This table presents the results for the analysis of a trade-off that a rider faces while making a pricing decision. The left column corresponds to the regression analysis results, where the dependent variable is the total probability that a rider faces a match (either agreement or counteroffer). The right column shows the regression analysis results where the dependent variable is the probability that a rider faces an agreement. Standard errors are in parentheses; *p<0.1; **p<0.05; ***p<0.01.

riders eventually offered higher prices. Therefore, the assumption that riders tried to get matched once is more justified for some markets than others. For this reason, I exclude markets where riders often made several attempts from the estimation.

Table 4: One-Shot Actions

	N	% of Total	% of Unmatched (within 20 seconds)
Total Observations:	763,005		
· Matched within 20 seconds	581,191	76.17	
· Not matched (within 20 seconds):	181,814		
Matched after 20 seconds at requested price	74,164	9.72	40.79
Cancelled at requested price	30,950	4.06	17.02
Canceled at higher price	13,205	1.73	7.26
Matched at higher price (rider increased)	49,358	6.47	27.15
Matched at higher price (driver increased)	13,793	1.81	7.59
Others	265	0.03	0.15

Notes: This table presents final matching outcomes for all requests in the sample. More than 75% of riders were matched within 20 seconds. The remainder did not increase their offered price often.

Fact 4: Drivers are selective. I explore the main determinants of a driver's decision to participate in a matching stage. To this end, I characterize the probability that a driver participated in a matching stage using a linear probability model. Since the effects for some variables can vary with others, I include in the controls the pickup distance (in km), an indicator for a high offered price (I keep observations only for prices 300 and 350), an indicator for a short trip, and their combinations. I explore different participation decisions: a decision to agree to the offered price and a decision to make a counteroffer. Table 5 reports the results.

I find that drivers are selective. Specifically, the regression coefficients suggest that drivers tended to agree more often to the trips with higher prices, lower trip distance, and trips with lower distances to a rider. Drivers more often counteroffered long trips with lower rider bids.

Fact 5: Drivers are forward-looking. Market participants in a two-sided market with a high frequency of arrivals may have dynamic incentives: one might decide not to get matched even if a transaction generates a positive surplus in a static sense. This happens because market participants anticipate a better transaction in the future. Several studies have previously considered the skimming effect that can arise in a decentralized setting (Liu et al., 2019, Romanyuk and Smolin, 2019). I show that drivers indeed behaved strategically on the platform and became more selective when incentives to wait were high.

Consider the following thought experiment. Suppose there are two possible environments. In one, requests appear with a higher frequency, riders offer high prices more often, and competition from fellow drivers is not fierce. An example of such a market is a weekday rush hour. In another environment, riders arrive with a lower frequency and offer lower prices more often. In addition, competition is more intense relative to the previous environment. Such a market would correspond to a weekday late-morning hour. Between these two environments, a strategic driver has more incentives to ignore the lower-price requests when arrival rates and offered prices are higher (in a rush hour). In contrast, the decisions of a static driver should be equivalent in both environments.

To see whether the described relationship holds in the data, I compare two markets: a morning rush hour (requests that appear weekdays between 7 a.m. and 8 a.m.) and a late-morning hour (requests that appear weekdays between 10 a.m. and 11 a.m. during the working days). I run two linear probability models. In the first one, the dependent variable is the probability that a driver agrees to the offered price. In the second one, the dependent variable is the probability that a driver participates in the matching stage (either agrees or makes a counteroffer). Independent variables include an indicator for the high offered price (350), the pickup distance (in meters), an

Table 5: Drivers' Decisions

	Prob. to agree	Prob. to counter
Pickup distance (in km)	-1.020^{***}	-0.326^{***}
	(0.005)	(0.003)
High price	0.255^{***}	-0.218^{***}
	(0.003)	(0.002)
Short trip	0.190^{***}	-0.130^{***}
	(0.002)	(0.001)
Distance to a rider (in km) × High price	-0.204^{***}	0.285^{***}
	(0.009)	(0.005)
High price × Short trip	-0.128^{***}	0.108^{***}
	(0.005)	(0.003)
Short trip × Pickup distane (in km)	-0.301^{***}	0.178^{***}
	(0.007)	(0.004)
High price × Short trip × Pickup distance (in km)	0.174^{***}	-0.156^{***}
	(0.014)	(0.008)
Constant	0.542^{***}	0.327^{***}
	(0.003)	(0.002)
Hour FE	✓	✓
Weekday/Weekend FE	✓	✓
Observations	1,429,437	1,429,437
R^2	0.130	0.051

Notes: This regression reports the coefficients for regressions of the likelihood that a driver chose an action (either agree or counteroffer) on trip characteristics: pickup distance (in km), an indicator for high offered price, and an indicator for short trip. In addition, I include the interaction terms between mentioned characteristics to capture that the effects might vary with other trip characteristics. I include only responses of drivers located at a distance lower than 500 meters from a rider. A driver whose distance exceeded this threshold rarely entered the matching stage (less than 2% of all drivers located at a longer distance did in my sample). Standard errors in parentheses; $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

indicator for trip distance, an indicator for a rush-hour market, and an interaction term between the high-offered-price and the rush-hour market indicators. If drivers do not behave strategically and respond to the same type of requests similarly between the two markets, I would expect coefficient in front of the interaction term to be small and insignificant.

Table 6 reports the results. The magnitude of the coefficient of interest is large, and the effect is statistically significant for both dependent variables. This implies that the drivers' preferred actions toward the same request (trip distance and pickup distance) made at a low price in the rush

Table 6: Drivers' Dynamic Behavior

	Prob. to agree	Prob. to participate
Low price	-0.103^{***}	-0.032^{***}
	(0.004)	(0.004)
Pickup distance (in km)	-1.206^{***}	-1.327^{***}
	(0.011)	(0.011)
Short trip	0.066^{***}	0.020^{***}
	(0.003)	(0.003)
Rush hour	-0.074^{***}	-0.029^{***}
	(0.005)	(0.005)
Low price × Rush hour	-0.108^{***}	-0.065^{***}
	(0.006)	(0.006)
Constant	0.916^{***}	0.976^{***}
	(0.005)	(0.005)
Observations	118,805	118,805
R^2	0.138	0.127

Notes: This table reports the coefficients for regression of the likelihood of a driver's decision (to agree or not ignore the request) on an indicator for rush hour interacted with an indicator for a low offered price. Only requests that appeared weekdays between 7 a.m. and 8 a.m. (rush hour) and 10 a.m. and 11 a.m. (non–rush hour) are included in the regression. Similar to the previous exercise, I include only observations for which the pickup distance does not exceed 0.5 km. Standard errors in parentheses; $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

hour and non–rush hour differ. Drivers tend to reject low prices more often in the morning hours when compared to later in the day.

3 Model of a Decentralized Ride-Hailing Market

3.1 Overview

To analyze the welfare effects of a pricing decentralization, I develop an equilibrium model of the ride-hailing market based on the empirical patterns presented above. There are two types of agents on the market: riders and drivers. Riders are short-lived and try to get matched only once. Drivers behave dynamically and have a discount factor β. In the model, time is discrete with an infinite horizon. All time intervals are equal to one second. Each period there is a probability

$(1-\pi)$ that it will be the driver's last period at the platform.[16] A driver's measure of impatience can be expressed as $\delta = \beta \times \pi$.

The model has demand side, supply side and a notion of equilibrium. Riders (demand side) decide which price to offer based on the state of the market they observe and their valuation of their trip. Drivers (supply side) decide on how to respond to an incoming request. The interaction of these two decisions forms an equilibrium where matching probabilities are defined.

All ex ante observable characteristics of the environment form a market definition m (e.g., time of the day, working day/weekend status). There is a known exogenous probability $\lambda_m(\bar{d})$ that a rider will show up on the platform at the beginning of a time interval to a search for a trip with distance $\bar{d}$. This probability captures between-market variation in the number of riders that search for a trip. If a rider appears (at most one per period), the following sequence of events takes place:

1. The rider observes a personalized state of the market, s (the number of idle drivers and their respective distances to a rider), and forms beliefs about the probability to get matched at different prices: η.

2. The rider chooses to offer a price b that maximizes his expected utility from participating.

3. Platform shows the request to all idle drivers. They observe an offered price b and a trip distance $\bar{d}$, common for all of them. Each driver i observes his own pickup distance, d_i, and a random shock associated with the request, ϵ_i, which he receives only if he ends up being matched to the rider and completes the trip. Drivers do not know how many competitors they face or how far are they from the rider.

4. Each driver with the state $(\bar{d}, b, d_i, \epsilon_i)$ decides whether to participate in the matching stage by agreeing or making a counteroffer. Each action is associated with the probability of winning the request: $\mu_m(A, \bar{d}, b, d_i)$ if a driver agrees, and $\mu_m(C, \bar{d}, b, d_i)$ if he makes a counteroffer.[17]

5. The platform collects idle drivers' responses and allocates the request based on its algorithm.[18] The algorithm gives priority to the drivers who have agreed to the offered price.

[16] Using an infinite-horizon framework allows me to abstract away from drivers' entry decision. However, in reality, drivers are not present on the platform forever. I assume that they become terminal due to exogenous reasons. Another heterogeneity on the driver's side (how far the driver is from the end of the shift) would complicate equilibirum beliefs computation.

[17] Note that A corresponds to driver's decision to agree to an offered price and C corresponds to his decision to counteroffer.

[18] The platform makes a decision based on a driver's distance and ranking.

6. The rider then faces a match, which he can refuse or accept. If he accepts the price, the match is formed, and the driver becomes occupied. If not, the rider leaves the platform and the driver remains idle.

3.2 Demand Model

Riders appear on the platform sequentially at the beginning of a time interval. A rider in a market m intends to take a trip with distance $\bar{d}$ that takes one of two values: $\{S, L\}$ (short or long). This rider appears with probability $\lambda_m(\bar{d})$. Each rider makes only one attempt to get matched and drops out if that attempt fails. I assume that the attempt takes 20 seconds.

Prior to opening the app, a rider draws his valuation for being matched within 20 seconds, v, from the distribution with density $f(v|\psi_m(\bar{d}))$. Upon arrival on the platform, the rider observes a state of the market — s, which consists of the number of idle drivers and their respective distances to the rider. The state of the market s determines the rider's beliefs about the corresponding probabilities to get matched at various prices: $\eta_m(A|b, \bar{d}, s)$, a probability that a rider's request will be matched at the offered price b, and $\eta_m(C|b, \bar{d}, s)$, probability that rider's request will be counteroffered at the offered price and a price $b + \Delta$ will be requested. In the model, riders choose between two prices: b_L or b_H, such that $b_H = b_L + \Delta$. The value of nonparticipation is normalized to zero.

A rider with valuation v solves the following optimization problem:

$$\max_{b \in \{b_L, b_H\}} u_m(b, v, \bar{d}, s),$$

where:

$$u_m(b, v, \bar{d}, s) = \underbrace{\eta_m(A|b, \bar{d}, s)}_{\text{Pr.that request is accepted at } b} \max[v - b, 0] + \underbrace{\eta_m(C|b, \bar{d}, s)}_{\text{Pr.that request is counteroffered at } b} \max[v - b - \Delta, 0]$$

Under these two assumptions, the problem has a threshold solution:

- $\eta_m(A|b_L, \bar{d}, s) < \eta_m(A|b_H, \bar{d}, s)$: the probability of acceptance at a low price is below a probability of acceptance at a high price.
- $\eta_m(A|b_L, \bar{d}, s) + \eta_m(C|b_L, \bar{d}, s) < \eta_m(A|b_H, \bar{d}, s) + \eta_m(C|b_H, \bar{d}, s)$: the total probability of getting matched at a low price is lower than the total probability of getting matched at a high price.

The threshold solution takes the following form:

$$b^*(v, \bar{d}, s) = \begin{cases} \emptyset & v < b_L \\ b_L & b_L < v < \hat{v}(\bar{d}, s) \\ b_H & v > \hat{v}(\bar{d}, s) \end{cases}$$

Threshold $\hat{v}_m(\bar{d}, s)$ can be calculated analytically (Appendix B). A rider never accepts any offer that exceeds his valuation. In his choice between a low price and a high price, the rider faces the following trade-off: on one hand, offering a high price increases his chances to be matched; on the other hand, offering a higher price is associated with a higher expected payment.

The probability that the rider offers b_H for a trip with distance $\bar{d}$ given the state s in the market m is:

$$\omega_m(b_H|\bar{d}, s) = \int_{\hat{v}_m(\bar{d},s)}^{\infty} f(v|\psi_m(\bar{d}))\,\mathrm{d}v \tag{1}$$

The chosen price plays two concurrent roles. First, it acts as a signal of a rider's valuation. Two riders who face the same market state s would offer different prices only if their valuations are different. The rider with a lower valuation would never choose a higher price than a rider with a higher valuation facing the same choice environment. Second, price acts as a signal about potential competitors that a driver faces. Riders who face tougher market conditions will be more likely to offer higher prices.

3.3 Supply Model

The next part of the model focuses on an idle driver's decision on how to respond to an incoming request. At most, one request can appear at the beginning of a time interval. For each observed request, a driver can agree to the offered price (A), make a counteroffer (C), or ignore the request (I).

An idle driver observes requests that appear on the platform over time. All idle drivers simultaneously observe an offered price b and a trip distance $\bar{d}$, but each driver i independently learns how far he is from a rider d_i. From a driver's point of view, requests vary in several dimensions: by offered price ($b \in \{b_L, b_H\}$), by trip distance ($\bar{d} \in \{S, L\}$), and by pickup distance (d

$\in \{d_1, ..., d_X\}$).[19] All idle drivers independently draw their distances to the rider at the moment when a request arrives to the platform from a commonly known discrete distribution: $\mathbb{P}(d_i = d_x)$. This distribution is the same across all markets and depends only on the city structure.

In addition, each driver i learns his own shock, ϵ_i, associated with the request at the time when the request appears. This shock represents the driver's idiosyncratic cost component that is unobservable to the platform and the econometrician. The driver receives the value of the shock only if he is matched to the rider. The shocks are drawn from a common normal distribution with a density function $f_\epsilon(\cdot)$. I assume that this distribution has a zero mean and a standard deviation σ_ϵ (drivers do not think ex ante that a particular request will be associated with positive or negative factors). The distribution of the shocks does not vary across markets or different types of trips, i.e., ϵ represents only truly idiosyncratic aspects of a driver's payoff for a given ride. Since I do not model the differences between destinations, in this context it can represent the value of heading to a particular area that can be different for idle drivers. Therefore, the same request will have a different appeal to two idle drivers due to differences in their distances to a rider and the idiosyncratic cost component ϵ.

Drivers do not observe their competitors directly. I consider the steady state of the industry, following the tradition of Hopenhayn (1992). In particular, drivers are aware only of the long-run distributions of the number of idle drivers and their distances in each market but do not know the exact realization for each given request. This known distribution of the number of possible competitors is summarized in the long-run equilibrium beliefs, and I assume that drivers do not learn between requests. Any success or failure is attributed to the steady-state probabilities, not to a specific shock to the number of idle drivers. This assumption is rather strong, given short time intervals in the model. However, the platform has high transaction frequency, and there are many reasons why two neighboring periods might look very different.[20] Learning on driver's side can be incorporated into the model, and will change how participants form the beliefs. In the steady-state that I consider, if a driver decides to enter the matching stage (agrees or makes a counteroffer), he faces probabilities of winning the request that are determined by a long-run equilibrium: $\mu_m(A, \bar{d}, b, d)$ if he chooses to agree and $\mu_m(C, \bar{d}, b, d)$ if he chooses to make a counteroffer. Drivers are also aware of the probabilities with which each type of request can appear, $\alpha(\bar{d}, b, d)$.

I start by describing how driver i makes a decision to respond to a request of type $(\bar{d}, b, d_i, \epsilon_i)$

[19] In the estimation, these values correspond to $\{100, 200, ..., 6000\}$ meters.

[20] Drivers finish their previous trips, drivers decide to enter and exit, drivers have been blocked in a matching stage of a different request, etc.

and then show how the value of being idle is formed. The driver chooses between agreeing to the offered price $\{A\}$, making a counteroffer $\{C\}$, or ignoring $\{I\}$.[21] Making a counteroffer is equivalent to asking for an increase in the price equal to Δ.

The corresponding utilities from each option:

$$
\begin{aligned}
\text{A}: \tilde{U}_m(A, \bar{d}, b, d_i, \epsilon_i) &= \mu_m(A, \bar{d}, b, d_i) \times [\bar{u}_m(\bar{d}, b, d_i) + \epsilon_i + \delta^{g(d_i+\bar{d})} V_m] + (1 - \mu_m(A, \bar{d}, b, d_i)) \times \delta^{20} V_m \\
\text{C}: \tilde{U}_m(C, \bar{d}, b, d_i, \epsilon_i) &= \mu_m(C, \bar{d}, b, d_i) \times [\bar{u}_m(\bar{d}, b + \Delta, d_i) + \epsilon_i + \delta^{g(d_i+\bar{d})} V_m] + (1 - \mu_m(C, \bar{d}, b, d_i)) \times \delta^{20} V_m \\
\text{I}: \tilde{U}_m(I, \bar{d}, b, d_i, \epsilon_i) &= \delta^5 V_m
\end{aligned}
$$

Utilities $\tilde{U}_m(A, \bar{d}, b, d_i, \epsilon_i)$ and $\tilde{U}_m(C, \bar{d}, b, d_i, \epsilon_i)$ have similar structures. Neither agreeing nor counteroffering guarantees that the driver will be matched. With some probability, $\mu_m(A, \bar{d}, b, d_i)$ or $\mu_m(C, \bar{d}, b, d_i)$, a rider wins the request and receives the sum of static utility and a shock, $\bar{u}_m(\bar{d}, b, d_i) + \epsilon_i$ or $\bar{u}_m(\bar{d}, b + \Delta, d_i) + \epsilon_i$, respectively, as well as a discounted continuation value once the request is completed, $\delta^{g(d_i+\bar{d})} V_m$. Here, $g(d_i + \bar{d})$ is a function that defines the time it takes to complete a trip with total distance $d_i + \bar{d}$, which is estimated directly off the data. Moreover, any participation comes at a cost. If the driver loses at the matching stage, he returns to the pool of idle drivers only after 20 seconds.

If the driver chooses to ignore the request, he still spends five seconds inspecting the details of the request and returns to the pool of idle drivers afterward. Therefore, participation is costly, for two reasons. First, if the driver is matched, he misses an opportunity to be matched with a potentially better rider (who is closer, offers a higher price, and/or wants to take a short trip). Second, if the driver decides to participate and loses, he misses an additional fifteen seconds, during which a better request could have appeared.

The driver's problem then takes the following form:

$$
W_m(\bar{d}, b, d_i, \epsilon_i) = \max_{D \in \{A, C, I\}} \tilde{U}_m(D, \bar{d}, b, d_i, \epsilon_i),
$$

where D is driver's decision whether to agree (A), make a counteroffer (C), or ignore (I) the request. $W_m(\bar{d}, b, d_i, \epsilon_i)$ is the value of the best response to the request of type $(\bar{d}, b, d_i, \epsilon_i)$.

Since both $\tilde{U}_m(A, \bar{d}, b, d_i, \epsilon_i)$ and $\tilde{U}_m(C, \bar{d}, b, d_i, \epsilon_i)$ are linear in ϵ_i, $W_m(\bar{d}, b, d_i, \epsilon_i)$ has the fol-

[21] Although the driver can choose from several discrete options on the platform, in the data, most drivers in the studied location ask for an increase equal to $\Delta = 50$.

lowing piece-wise form:[22]

$$W_m(\bar{d}, b, d_i, \epsilon_i) = \begin{cases} \delta^5 V_m & \epsilon_i \leq \bar{\epsilon}_m(\bar{d}, b, d_i) \\ \mu_m(C, \bar{d}, b, d_i)[\bar{u}_m(\bar{d}, b+\Delta, d_i) + \epsilon_i + \delta^{g(d_i+\bar{d})} V_m - \delta^{20} V_m] + \delta^{20} V_m & \bar{\epsilon}_m(\bar{d}, b, d_i) \leq \epsilon_i \leq \hat{\epsilon}_m(\bar{d}, b, d_i) \\ \mu_m(A, \bar{d}, b, d_i)[\bar{u}_m(\bar{d}, b, d_i) + \epsilon_i + \delta^{g(d_i+\bar{d})} V_m - \delta^{20} V_m] + \delta^{20} V_m & \hat{\epsilon}_m(\bar{d}, b, d_i) \leq \epsilon_i \end{cases}$$

where:

$$\bar{\epsilon}_m(\bar{d}, b, d_i) = \frac{\delta^5 V_m - \delta^{20} V_m}{\mu_m(C, \bar{d}, b, d_i)} - \bar{u}_m(\bar{d}, b+\Delta, d_i) + \delta^{20} V_m - \delta^{g(d_i+\bar{d})} V_m \tag{2}$$

$$\hat{\epsilon}_m(\bar{d}, b, d_i) = \frac{\mu_m(C, \bar{d}, b, d_i) \times \bar{u}_m(\bar{d}, b+\Delta, d) - \mu_m(A, \bar{d}, b, d_i) \times \bar{u}_m(\bar{d}, b, d)}{\mu_m(A, \bar{d}, b, d_i) - \mu_m(C, \bar{d}, b, d_i)} + \delta^{20} V_m - \delta^{g(d_i+\bar{d})} V_m \tag{3}$$

Intuitively, the full support of ϵ values can be split into three intervals. The left interval (with large negative values) is where the driver decides to ignore the request: even if he wins at the matching stage, his payoff will be lower than the value he receives by simply ignoring a request and remaining idle. The right interval (with large positive values) is where the driver decides to accept the request without counteroffering (which is associated with lower chances of winning). It is important to note that $\mu_m(A, \bar{d}, b, d_i) \geq \mu_m(C, \bar{d}, b, d_i)\ \forall\ \bar{d}, b, d, m$ by design (the system gives priority to drivers who agree over those who make a counteroffer). Therefore, if the driver receives a large-enough positive shock, he will not take chances by counteroffering it. The driver finds it optimal to make a counteroffer for the interval in the middle: the shock is not large enough to agree right away; however, this option is still better than ignoring the request.

Given equations 2 and 3, ex ante probabilities (without observing a shock) of each action in

[22] The condition under which all three options have a chance to be chosen is as follows:
$\delta^5 V_m - \delta^{20} V_m < \frac{\mu_m(A,\bar{d},b,d_i)\mu_m(C,\bar{d},b,d_i)[\bar{u}_m(\bar{d},b+\Delta,d_i) - \bar{u}_m(\bar{d},b,d_i)]}{\mu_m(A,\bar{d},b,d_i) - \mu_m(C,\bar{d},b,d_i)}$
If this condition is not satisfied, a driver would never choose to counteroffer a request of such type. For these cases, $W_m(\bar{d}, d, d_i, \epsilon)$ consists of two linear parts instead of three. There is one threshold, which looks similar to $\bar{\epsilon}(\bar{d}, b, d_i)$. In that case, in the equation 2, $\mu_m(C, \bar{d}, b, d_i)$ is replaced with $\mu_m(A, \bar{d}, b, d_i)$ and $\bar{u}(\bar{d}, b+\Delta, d_i)$ is replaced with $\bar{u}(\bar{d}, b, d_i)$.

response to a request of type $(\bar{d}, b, d_i)$ in a market m are given by:[23]

$$\rho_m(A|\bar{d}, b, d_i) = \int_{\hat{\epsilon}_m(\bar{d},b,d_i)}^{\infty} f(\epsilon|\sigma_\epsilon)d\epsilon \tag{4}$$

$$\rho_m(C|\bar{d}, b, d_i) = \int_{\bar{\epsilon}_m(\bar{d},b,d_i)}^{\hat{\epsilon}_m(\bar{d},b,d_i)} f(\epsilon|\sigma_\epsilon)d\epsilon \tag{5}$$

$$\rho_m(I|\bar{d}, b, d_i) = \int_{-\infty}^{\bar{\epsilon}_m(\bar{d},b,d_i)} f(\epsilon|\sigma_\epsilon)d\epsilon \tag{6}$$

The decision to take any action, however, depends on the value of being idle. The driver does not observe ϵ_i until a request appears and forms an expectation over the support of ϵ-shocks to obtain an ex-ante value of a request of type $(\bar{d}, b, d_i, \epsilon_i)$:

$$E_\epsilon[W_m(\bar{d}, b, d_i, \epsilon_i)] = \int_{-\infty}^{\infty} W_m(\bar{d}, b, d_i, \epsilon_i) f(\epsilon_i|\sigma_\epsilon)d\epsilon_i$$

Using the ex ante value for each type of request and ϵ's independence from all other request characteristics, I define driver's ex ante value of being idle in a market m as follows:

$$V_m = \sum_{\substack{\bar{d}\in\{S,L\}\times \\ b\in\{b_L,b_H\}\times \\ d_i\in\{d_1,\ldots,d_X\}}} \alpha_m(\bar{d}, b, d_i) \times E_\epsilon\Big[W_m(\bar{d}, b, d_i, \epsilon_i)\Big] + \Big(1 - \sum_{\substack{\bar{d}\in\{S,L\}\times \\ b\in\{b_L,b_H\}\times \\ d_i\in\{d_1,\ldots,d_X\}}} \alpha_m(\bar{d}, b, d_i)\Big) \times \delta V_m \tag{7}$$

This expression can be interpreted in the following way. With some probability, $\alpha_m(\bar{d}, b, d_i)$, a request with specified characteristics appears at the beginning of a period in market m, and a driver chooses an optimal action given the type of the request and the value of the shock, ϵ, he observes. When no request appears, the driver remains idle and transitions into the next period.

3.4 Equilibrium

The interaction between riders' and drivers' decisions makes the platform two-sided. The distribution of riders' valuations, the distribution of states they observe upon arrival, and the arrival rates influence the probabilities with which drivers see the requests. While a rider observes the

[23] These correspond to the case when all three options are chosen. For two options, probabilities can be calculated very similarly and the probability that a request will be counteroffered is zero.

full state of the market, s, which consists of a driver's i pickup distance, d_i, and the pickup distances of driver i's competitors, d_{-i}, driver i observes only his own pickup distance. However, the offered price is suggestive of the state of the market, s, that the rider faces and affects a driver's equilibrium beliefs about the probability of winning μ. The probabilities of the driver's actions (agree, counteroffer or ignore the request) in turn form the rider's beliefs about the corresponding probabilities to be matched (η). The beliefs summarize all information relevant to the market so that players make decisions based solely on their beliefs and their own state. A detailed description of beliefs formation can be found in Appendix B. I assume that all players have rational beliefs.

In the model, I consider a steady state in the market. Drivers do not update their beliefs about the competitors between requests: the perceived value of being idle now and in five seconds is the same (without discounting). Naturally, there is plenty of variation in the data: significant differences arise across hours and weekends, weather changes, etc. However, I assume that the idiosyncratic shocks to the market should not play a huge role in agents' decisions, conditional on all observable characteristics. For each request, drivers draw new distances to a rider and ϵ-shocks, significantly limiting dependency across periods. The only information that could potentially link two periods is the number of idle drivers on the market. However, in a large market where riders arrive frequently and drivers make decisions often, two neighboring periods should look almost independent. I use a large-market equilibrium concept, where the equilibrium is a set that consists of drivers' and riders' beliefs, such that riders and drivers alike behave optimally given those beliefs, and the beliefs are consistent with the played equilibrium.[24]

Definition: A rational-expectations equilibrium in a market m is a tuple: $\{\eta_m, \mu_m\}$, such that:

1. Riders maximize their expected utilities, given their beliefs about the corresponding probabilities to be matched, η_m, in the state they face, s.
2. Actions of riders influence probabilities of arrival for different types of requests, α_m.
3. Vacant drivers follow policy functions that maximize their expected utility based on their beliefs about winning probabilities, μ_m, and about the value of remaining unmatched, V_m, and the shock ϵ they receive.
4. V is a fixed point of:

$$V_m = \sum_{\substack{\bar{d}\in\{S,L\}\times \\ b\in\{b_L,b_H\}\times \\ d_i\in\{d_1,\ldots,d_X\}}} \alpha_m(\bar{d}, b, d_i) \times E_\epsilon\Big[W_m(\bar{d}, b, d_i, \epsilon_i)\Big] + \Big(1 - \sum_{\substack{\bar{d}\in\{S,L\}\times \\ b\in\{b_L,b_H\}\times \\ d_i\in\{d_1,\ldots,d_X\}}} \alpha_m(\bar{d}, b, d_i)\Big) \times \delta V_m$$

[24]This idea is close to the oblivious equilibrium concept of Weintraub et al. (2008).

5. Market participants have rational expectations, so that beliefs $\{\eta_m, \mu_m\}$ are self-fulfilling, given optimizing behavior.

4 Identification

There are two sets of model primitives: demand-side primitives and supply-side primitives. On the demand side, I estimate the probabilities of riders' arrival, $\lambda_m(\bar{d})$, and the distribution of riders' valuations, $f(v|\psi_m(\bar{d}))$. On the supply side, I estimate static utilities $\bar{u}(\bar{d}, b, d)$, the measure of drivers' impatience δ, the variance of ϵ-shocks that represent drivers' private information, the distribution of idle drivers $f_m(N)$, and the distribution of distances to a rider $\mathbb{P}(d_x)\ \forall x$. Some primitives are estimated directly off the data: $\lambda_m(\bar{d})$, $f_m(N)$ and $\mathbb{P}(d_x)\ \forall x$. Other primitives — such as $f(v|\psi_m(\bar{d}))$, $\bar{u}(\bar{d}, b, d)$, δ, and σ — are estimated using a proposed model, separately for supply and demand sides. In this section, I discuss a nonparametric identification of the parameters using the proposed model. In the next section, I estimate a parametric version of the model.

The demand-side dataset contains each rider's decision about which price to offer (b_L or b_H), information about the market state s that a rider faced at the decision time, and observable characteristics of the request: detailed information when request appeared (which I translate into a market definition m for the estimation) and trip distance $\bar{d}$. For each request I observe an outcome: if the rider received an agreement, faced a counteroffer, or was ignored. The outcome variables allow me to estimate the rider's beliefs η directly off the data. In this section, I treat them as given, and I explain how the beliefs can help to identify the distribution $f(v|\psi_m(\bar{d}))$.

The supply-side dataset contains idle drivers' decisions on how to respond to all types of requests. For each idle driver, I observe a price offered by a rider b, trip distance $\bar{d}$, and the pickup distance d_i. I also observe if a driver won the request, given request type, and a chosen action. This allows me to estimate drivers' beliefs directly off the data. For the identification argument, I consider them as given, and I discuss how variation in the beliefs and actions identifies the rest of the parameters.

4.1 Identification of Demand-Side Primitives

In the demand model, I aim to identify the distribution of riders' valuations, $f(v|\psi_m(\bar{d}))$, which determines the prices offered and, from a driver's point of view, the probabilities that requests will arrive. To recover the underlying parameters, I use variation in the choice environment. The argument for identification is similar in spirit to the one used in Agarwal and Somaini (2018). Every time a rider decides to take a trip, he faces a spatial distribution of idle drivers currently present on the platform. In the demand model, I specified a variable, s, that can be interpreted as an aggregate description of the picture that the rider sees in the app. Let us assume that there are two states in which we observe a rider: s and $\tilde{s}$. Without loss of generality, I will assume that $\tilde{s}$ is a more optimistic state (e.g., there are more drivers on the screen and their pickup distances are lower). Two different states, $\tilde{s}$ and s, are characterized by two different sets of beliefs, $\{\eta_m(A|b, \bar{d}, \tilde{s}), \eta_m(C|b, \bar{d}, \tilde{s})\}$ and $\{\eta_m(A|b, \bar{d}, s), \eta_m(C|b, \bar{d}, s)\}$, which then imply two different thresholds for choosing a high price, $\hat{v}_m(\bar{d}, \tilde{s})$ and $\hat{v}_m(\bar{d}, s)$. It follows that the threshold above which the rider offers a high price should be higher for $\tilde{s}$: $\hat{v}_m(\bar{d}, \tilde{s}) > \hat{v}_m(\bar{d}, s)$. That is, if the state is more optimistic, the rider must have a higher valuation to offer a high price. To show how variation in choice environment helps to identify the distribution of valuations, I use the following equation:

$$\underbrace{\omega_m(b_H|\bar{d}, s)}_{\substack{\text{Pr.that a rider offers } b_H \\ \text{when facing a state } s}} - \underbrace{\omega_m(b_H|\bar{d}, \tilde{s})}_{\substack{\text{Pr.that a rider offers } b_H \\ \text{when facing a state } \tilde{s}}} = \int_{\hat{v}_m(\bar{d},s)}^{\infty} f(v|\psi_m(\bar{d}))\,\mathrm{d}v - \int_{\hat{v}_m(\bar{d},\tilde{s})}^{\infty} f(v|\psi_m(\bar{d}))\,\mathrm{d}v$$

$$= \int_{\hat{v}_m(\bar{d},s)}^{\hat{v}_m(\bar{d},\tilde{s})} f(v|\psi_m(\bar{d}))\,\mathrm{d}v$$

The difference in probabilities that a rider chooses a high price between the two states traces the part of the distribution of v that lies between $\hat{v}_m(\bar{d}, \tilde{s})$ and $\hat{v}_m(\bar{d}, s)$.

4.2 Identification of Supply-Side Primitives

I intend to identify several parameters using the supply model: (1) static utilities from being matched with the requests of different types that vary across markets, $\bar{u}_m(\bar{d}, b, d)$; (2) a measure of impatience, δ; and (3) the variance of ϵ-shocks, σ_ϵ. The supply model is nonlinear, since it requires the computation of an inner fixed-point; however, below, I will briefly outline how the parameters are mainly identified.

Two unusual features of the platform—that the same request can be observed under both a low and a high price and that a driver can counteroffer—turn out to be extremely helpful for identifying the participation costs ($\delta^5 V_m - \delta^{20} V_m$) and the variance of ϵ shocks. In addition, I leverage that in the described setting, an econometrician observes drivers' responses to various types of requests (including drivers' decisions not to participate at all).

Since the shock thresholds define the probabilities with which a driver chooses a particular action (ignore, counteroffer, or agree to the request), one can use the observed probabilities for each type of request to find the values of the shocks (scaled to the variance). For instance, one can use equation 6 to invert a probability to ignore a request to find the threshold $\bar{\epsilon}_m(d, \bar{b}, d_i)$:

$$\bar{\epsilon}_m(d, \bar{b}, d_i) = \sigma_\epsilon \Phi^{-1}\Big(\rho_m(I|\bar{d}, b, d_i)\Big)$$

Similarly, one can find $\hat{\epsilon}_m(d, \bar{b}, d_i)$ using equation 4:

$$\hat{\epsilon}_m(d, \bar{b}, d_i) = \sigma_\epsilon \Phi^{-1}\Big(1 - \rho_m(A|\bar{d}, b, d_i)\Big)$$

The difference between the two thresholds above does not depend on the dynamic part of the utility function for any type of request and can be written as a linear combination of two terms: the participation costs ($\frac{\delta^5 V_m - \delta^{20} V_m}{\sigma_\epsilon}$) and the difference in static utility if the request is counteroffered ($\frac{u(\bar{d}, b+\Delta, d) - u(\bar{d}, b, d)}{\sigma_\epsilon}$, equal to $\frac{\Delta}{\sigma_\epsilon}$ in the case of linear utility). Therefore, the observed probabilities of drivers' responses to different types of requests, with varying probabilities to win, identify participation costs and variance of the shocks (σ_ϵ).

Once these two variables are pinned down, equations 2 and 3 allow me to find $u(\bar{d}, b, d) + \delta^{20} V_m - \delta^{g(d+\bar{d}) V_m}$ and $u(\bar{d}, b + \Delta, d) + \delta^{20} V_m - \delta^{g(d+\bar{d}) V_m}$. Taking these values back to equation 7 allows me to obtain another equation that depends on δ and V_m. Combined with the pinned-down participation costs $\delta^5 V_m - \delta^{20} V_m$, both δ and V_m can be recovered. Using these values, one can find the values of $u(\bar{d}, b + \Delta, d)$ and $u(\bar{d}, b, d)$.

5 Model Estimation and Results

I estimate the model primitives in several steps. The detailed nature of the dataset allows me to estimate some primitives directly off the data: (1) the requests' arrival probabilities $\lambda_m(\bar{d})$, (2) a distribution of idle drivers $f_m(N)$, and (3) a distribution of distances to a rider $\mathbb{P}(d_x)$. I estimate the rest of the parameters using the proposed demand and supply models.

Using identification arguments outlined in the previous section, both models can be estimated using variation in participants' beliefs. I estimate both models in two steps. First, I estimate participants' beliefs directly off the data: η (riders' beliefs about getting matched at different prices), μ (drivers' beliefs about winning a request), and α (probability for drivers that a request appears on the platform) for drivers. With the estimates for these beliefs in hand, I use a maximum likelihood approach to recover the primitives of interest.

5.1 Market Definition

Before moving to the details of the estimation, I first define the market m. After eliminating obvious outliers (such as national holidays and days with extreme weather conditions), I still observe significant variation across different hours of the day. In addition, meaningful differences exist between weekends and weekdays. I do not model the transitions between different hours of the day; I treat each hour as a separate market.

In the estimation, I focus only on the markets that represent even hours of weekdays. I have several reasons for this exclusion. The choice of weekdays is driven by the fact that I have more weekdays in the sample, which allows for a more precise estimation of the beliefs. In addition, most weekday hours (with the exception of the 7 a.m. market) satisfy assumption that riders make only one attempt to get matched, which I used in the model. Third, my supply-model estimation is computationally intensive, and the burden increases with the number of parameters to be estimated.

Figure 3 illustrates requests' arrival probabilities and the distribution of the number of idle drivers across different hours on weekdays. Markets that I use in the estimation are marked with blue vertical lines.

5.2 Estimation of Demand-Side Primitives

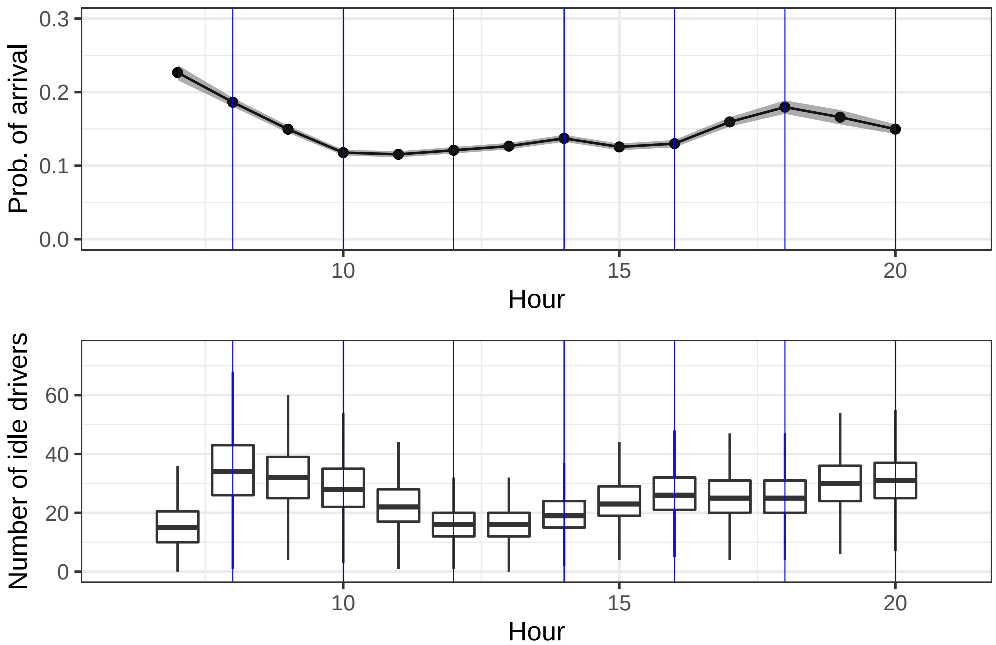
The main aim of my demand-model estimation is to obtain the distribution of valuations for each market m. For purposes of estimation, I parametrize $f(v|\psi_m(\bar{d}))$ to be a gamma distribution truncated from below, where the lower limit is determined by the lowest price seen on the platform.

$$f(v|\psi_m(\bar{d})) \sim G[\underbrace{k_m(\bar{d})}_{\text{Shape parameter}}, \underbrace{\theta_m(\bar{d})}_{\text{Scale parameter}}, \underbrace{b_L}_{\text{Lower limit}}]$$

I let all demand-side model primitives vary across markets. I observe meaningfully different probabilities of requests' arrivals across different hours: more requests arrive in the morning and in the early evening. I let the distribution of valuations change across markets as well, since the composition of travelers is likely to change: it could be that more business-associated travel happens in the morning, and more leisure-associated travel happens in the afternoon and evening.

I estimate the demand model in two steps. First, I estimate the rider's beliefs off the data. As discussed previously, to identify the parameters of the distribution $f(v|\psi_m(\bar{d}))$, I need an exoge-

Figure 3: Requests' Arrival Probabilities and Number of Idle Drivers: Monday–Friday



Notes: The upper panel of this figure depicts the probability (with 95% CI) that a request appears each second during different hours of the day. The lower panel shows the distribution of the number of idle drivers by different hours. Blue lines mark the markets used in the estimation. Time is in military time.

nous variation in the state that a rider faces, s, that shifts rider's beliefs but does not affect the riders' valuations. In the estimation, I exploit the variation in the number of idle drivers in the vicinity (R-meter radius)[25] of a rider at the time when the request appears on the platform. I chose this proxy for two reasons. First, the R-meter radius is roughly consistent with the area that a rider sees by default once he opens the app.[26] Second, conditioning on market definition m allows me to use the residual variation in the state that should not be correlated with aggregate demand shocks and is driven by many random factors, including where and how many drivers have just started their shifts, how many drivers have just completed their shift, and where and how many drivers have just tried to get matched with another rider but did not succeed. The random nature of requests' appearance and their exact location makes it unlikely that the variation in the number of idle drivers in a relatively small radius depends on the aggregate demand shocks.

Once I estimate riders' beliefs η, I form a maximum likelihood estimator for riders' observed actions to obtain the parameters that characterize the distribution of valuations for each market and

[25] A reader can think of R as being between 200 and 400 meters. The exact value depends on the initial map scale of the city.

[26] Nothing prevents a rider from zooming out to see a larger area; however, it is unknown whether they do so frequently.

trip distance.

Stage 1: Estimation of beliefs—probabilities of getting matched (η)

Recall that there are three possible outcomes each rider can face once he places a request: his request can be accepted (he faces a price b), his request can be counteroffered (he faces a price $b + \Delta$), or his request can be ignored. I estimate the probabilities of these outcomes using an ordered probit model. Beliefs are estimated separately for each market m. The independent variable is the outcome. The dependent variables include trip distance ($\bar{d}$), a rider's personalized state of the market (s), and a price that a rider offers (b). The results of beliefs' estimation are presented in Table C1. Consistent with what would have been expected, offering a high price, taking a short trip, and seeing more drivers nearby are associated with better chances that a request is accepted. Using the obtained estimates, I then predict probabilities of being matched under a low price and a high price for each observation in my sample, given trip distance $\bar{d}$ and market state s. Using these predicted probabilities, I calculate a threshold $\hat{v}_m(\bar{d}, s)$ for each choice situation in the sample.

Stage 2: Estimation and results

Once I have calculated the choice thresholds, I use a maximum likelihood approach to estimate the parameters of the distribution—$\psi_m(\bar{d})$. I perform estimation separately for each market and for different trip distances. Equation 1 computes a predicted probability that a rider in a market m offers b_H for a trip with distance $\bar{d}$ when facing a market state s.

Full estimates of the primitives (shape and scale coefficients of gamma distribution) are reported in Table C2. Table 7 presents partial results for two important statistics that characterize the estimated distributions: $P(v < b_H)$, a share of low-valuation riders whose valuations are below a high price, b_H, and a valuation of a median rider who wants to take a trip with distance $\bar{d}$ in a market m (reported in local currency). I find significant variation across different hours of the day and considerable heterogeneity across riders within the same market. The share of riders who would decide not to participate under a high price significantly fluctuates between 9.4% and 47% for short trips and between 3.9% and 28.8% for long trips. This statistic is important for the counterfactual analysis, as it shows the share of riders who would have been priced out of the market if the platform had mandated a high price for a trip.

Table 7: Demand-Side Estimates

	Short Trips		Long Trips	
Market	$P(v < b_H)$	Median valuation	$P(v < b_H)$	Median valuation
8 a.m.	0.302	533.447	0.159	818.703
10 a.m.	0.094	1020.139	0.039	1357.894
noon	0.445	378.715	0.266	545.238
2 p.m.	0.406	403.776	0.265	539.723
4 p.m.	0.231	586.821	0.122	817.696
6 p.m.	0.466	367.862	0.288	566.807
8 p.m.	0.256	550.638	0.138	773.144

Notes: This table reports the (1) estimated shares of riders who would have been priced out of the market if b_H is charged for a trip and (2) the valuation of a median rider in the market. The left side of the table presents the results for the short trips and the right side presents the same statistics for the long trips.

5.3 Estimation of Supply-Side Primitives

The main aim of my supply-side estimation is to obtain the estimates for the static utilities $\bar{u}_m(\bar{d}, b, d)$, measure of drivers' impatience δ, and variance of the idiosyncratic cost component σ_ϵ. I parametrize $\bar{u}_m(\bar{d}, b, d)$ to be:

$$\bar{u}_m(\bar{d}, b, d) = b - c \times d - \alpha_{0,m} - \alpha_1 \mathbb{1}(\bar{d} = L) \tag{8}$$

The static utility from a request of type $(\bar{d}, b, d)$, conditional on winning, is defined as a difference between the offered price, b, and the costs associated with taking a trip with distance $\bar{d}$ that is d meters away from a driver. The differences in response rates to requests with different pickup distance d and trip distances $\bar{d}$ allow me to identify the cost-per-meter parameter c, and the difference in the fixed costs between a long and a short trip, α_1. $\alpha_{0,m}$ represents the fixed costs of taking any trip in a market m. I assume that the three parameters do not vary across markets: the additional costs associated with taking a long trip α_1, cost-per-kilometer parameter c, and measure of drivers' impatience δ.

Similar to the demand model, I estimate the supply model in two steps. First, I estimate drivers' beliefs about probabilities of winning a request, μ, and probabilities of requests' arrival, α. Second, I use a maximum likelihood approach to obtain the estimates of the parameters using drivers' observed actions.

Stage 1: Estimation of beliefs—winning probabilities (μ) and arrival probabilities (α)

To proceed with the main supply-side estimation, I first need to estimate drivers' beliefs about winning probabilities (μ) and their beliefs about probabilities that certain types of requests arrive (α).

I use a logit model to characterize the probability of winning a request with a price b, trip distance $\bar{d}$, and pickup distance d. The dependent variables include an indicator for a chosen action (agree or counteroffer), an indicator for a trip type (a combination of offered price and trip distance), pickup distance, an indicator for a market m in which request appeared, and interaction terms between the above-mentioned above variables and market indicators. Table C3 reports the results of the estimation. Consistent with my expectations, the probability of winning a request is significantly higher if the driver agrees to an offered price relative to a situation in which he counteroffers. The probability of winning the request decreases with pickup distance. Among trip types, the driver has a higher probability of winning a request for a long trip with a low price; the requests for short trips with a high price are the most competitive. There is an additional variation in the winning probabilities across markets as well.

I estimate α using empirical probabilities that a request of a certain type appears. For each market, I compute the conditional probability that a request of type $(\bar{d}, b, d)$ appears, then I multiply these probabilities by the probability that any request appears in a market m. Figure C1 shows an example of conditional probabilities of different request types for a single market.

Stage 2: Estimation and results

With the estimates of the beliefs, I use a maximum likelihood approach to recover the parameters of interest. Some parameters do not vary across markets, so I pool all markets together in the estimation.

The estimation of the outlined supply model is computationally demanding, because the estimation procedure has a nested structure. With fixed values of parameters (δ, c, σ_ϵ, $\alpha_{0,m}$ and α_1), one needs to find V_m first (a fixed point of 8) in an inner loop, and the outer loop needs to find parameters that maximize the probabilities of the observed choices defined in equations 5, 6 and 7. This algorithm (NFXP) was first introduced by Rust (1987) and had since been applied to various applications.

Although it is convenient to think about the coefficients and variance of the shocks in money terms, in the estimation, for the computational purposes, I normalize σ_ϵ to be one and estimate a coefficient γ in front of the price b instead (it is implicitly normalized to one in equation 8). Table 8 reports the estimates for the supply-side primitives. I present the estimation results for normalized coefficients and transform them back into money terms in the last column of Table 8. Below I discuss each of the estimated parameters.

Table 8: Supply-Side estimates

Parameter	Estimate	(S.E.)	Transformed (money terms)
γ	0.009	(0.000)	1.000
c (per-km)	3.288	(0.006)	365.333
$\alpha_{0,m}$: 8 a.m.	1.636	(0.011)	181.778
$\alpha_{0,m}$: 10 a.m.	1.458	(0.011)	162.000
$\alpha_{0,m}$: noon	1.287	(0.010)	143.000
$\alpha_{0,m}$: 2 p.m.	1.465	(0.010)	162.778
$\alpha_{0,m}$: 4 p.m.	1.527	(0.010)	169.667
$\alpha_{0,m}$: 6 p.m.	1.701	(0.010)	189.000
$\alpha_{0,m}$: 8 p.m.	1.634	(0.010)	181.556
α_1	0.105	(0.003)	11.667
δ	0.969	(0.001)	0.969
σ_ϵ	1.000	-	111.111
Num.obs.	1,346,804		
Log-likelihood	-707,348.476		

Notes: This table reports the results of the supply-side estimation. I use choices of drivers that are within 1 km of a rider. This exclusion is driven by the fact that drivers who are farther than 1 km from a rider do not often participate.

Per-km cost parameter *c*. A 100-meter increase in the pickup distance generates additional costs equal to 36.5 in local currency. To put this number into context, it roughly corresponds to 12% of the low price. This estimated cost parameter is relatively high and is driven by data patterns. To better understand the reasons, one can use Table 5 to see how drivers respond to requests of different types. In the data, drivers sharply dislike trips that are farther away from them. The results for the linear probability model in Table 5 imply that the probability of accepting a request falls by more than 1 when the pickup distance increases by 1 km. The model rationalizes drivers' behavior through high values of the cost parameter c, since this is the only parameter that captures the differences between trips with different distances to a rider. Another possible reason why drivers might "dislike" trips with higher distances could be associated with their beliefs. Drivers might think their chances of winning fall more sharply than they actually do as the pickup distance increases.

Fixed costs $\alpha_{0,m}$. In the model, I let fixed costs vary across markets. I assume that this parameter captures not only the actual physical costs associated with taking a short trip (which I would expect to stay the same across different markets) but also some differences in the perceived value of not responding to a request, which can change across markets. The estimates indeed suggest that $\alpha_{0,m}$ varies across markets. The market with the lowest costs is the noon market; the market with the highest costs is the 6 p.m. market. These parameters significantly affect the value function of being idle: the higher $\alpha_{0,m}$ is, the lower the value of being idle. The value of being idle is highest for the noon market. A driver who participates in a noon market faces the lowest number of competitors relative to all other markets. These estimates also imply that drivers have relatively low margins: the lowest fixed cost corresponds to almost 48% of the low price observed on the platform.

Additional costs of taking a long trip $\alpha_{1,m}$. The additional costs of taking a long trip are not large. The difference is approximately 4% of the low price. In the data, action probabilities for short and long trips vary, but not drastically. Drivers might not pay significant attention to the distance of the trip, for two reasons. First, the considered city is not large, and drivers might consider the differences in trip distances to be insignificant. Second, the platform did not calculate the exact distance for them during the span of the sample.

Drivers' impatience parameter δ. The estimate of δ implies that drivers are impatient. One should not interpret this parameter as a discount factor. Recall that δ is a product between a discount factor β and a probability π that the driver will become terminal next period. The estimate implies that drivers are eager to match quickly and do not appear to wait long. Some drivers will become terminal because their time in the system expires, some will decide to quit for external reasons, and some might have their cars break down, etc.

Variance of idiosyncratic cost component σ_ϵ. The variance of the idiosyncratic cost component is large: one standard deviation in ϵ-shock accounts for more than 37% of the low price on the platform. This estimate suggests that the driver's private information about the costs plays an important role in his decision.

6 Counterfactual Analysis

In the counterfactual analysis, I compare market outcomes and the welfare of riders and drivers between two scenarios. In the first one, I consider the existing platform: riders choose prices, and drivers can make counteroffers.[27] I call this scenario decentralized pricing (DP). In the second scenario, the platform chooses the prices: riders do not make offers, and drivers are not allowed to counteroffer.[28] I let the platform choose a personalized price for each rider conditional on individual market conditions that the platform observes. In this scenario, the platform has the exact same information about market state s (the number of idle drivers and their distance) as a rider except for one piece: the rider's private valuation. Similar to the decentralized platform, only drivers know their idiosyncratic cost components. I refer to this scenario as centralized pricing (CP).

The main aim of the counterfactual analysis is to explore the welfare trade-offs and measure the effects associated with pricing decentralization. Different economic outcomes, such as matching rates, equilibrium prices, and welfare measures for riders and drivers can be compared across price-setting mechanisms. I do not consider total welfare and do not specify the platform's optimal objective function.[29] Instead, I show how equilibrium evolves across various centralized pricing rules and explore implications for participants' welfare. I start by highlighting the main trade-off that a price-setting platform faces and discuss how participants' welfare will be affected by the immediate change in pricing. I then discuss the equilibrium effect and present the results of the welfare analysis.

The Platform's Trade-off

Figure 4 illustrates the trade-off between arrival rates and average prices for a price-setting platform. Here, a rider faces a state of the market s: a collection of idle drivers and their distances to a rider. This particular choice situation under a decentralized equilibrium is characterized by a threshold $\hat{v}(s)$. A rider with valuation v would have chosen a high price, b_H, if his valuation exceeds this threshold: $v > \hat{v}(s)$. If his valuation is below the threshold, $v < \hat{v}(s)$, a rider would

[27] I limit drivers' ability to counteroffer and let them ask for a high price only if the low price is offered.

[28] I do not consider a fully centralized platform; I still allow drivers to make their own decisions. Therefore, I specifically focus on the effects of asymmetric information in the pricing decision. Efficiency consideration for a fully centralized platform is beyond the scope of this paper.

[29] During the span of my sample, the platform was making profits charging drivers an upfront fee to enter the system. There were no per-transaction fees. Since I do not model the driver's entry decision, I do not consider the platform's objective function.

have chosen a low price, b_L. Nevertheless, the platform does not possess information about the rider's valuation and only observes the state of the market s to decide which price to charge. Recall from the demand model that low values of $\hat{v}(s)$ correspond to the situation that a rider would mark as a "bad market": for instance, when the number of idle drivers is low and they are far away.

The left panel of Figure 4 demonstrates the case when the platform charges a low price. This price ensures that riders of all types stay on the platform and the arrival probabilities are not affected when compared to the decentralized platform. In this scenario, riders of all types pay a low price under CP. However, under DP, some types of riders ($v > \hat{v}(s)$) would have chosen to offer a high price to signal their valuations to ensure better matching rates. The graph with a CDF shows the shares of riders for whom the price stays the same and for whom the price goes down under CP with a low price.

The right panel presents the same choice situation, but now the platform charges a high price. Since the platform does not know the exact valuation of a rider, it will occasionally price riders out of the market ($v < b_H$). However, shading riders ($b_H < v < \hat{v}(s)$) will be forced to pay b_H instead of b_L, but will still stay on the platform. The graph with the CDF shows the share of riders (1) priced out of the market ($v < b_L$), (2) not being able to shade their valuations anymore ($b_H < v < \hat{v}(s)$), and (3) not experiencing an immediate change ($v > \hat{v}(s)$).

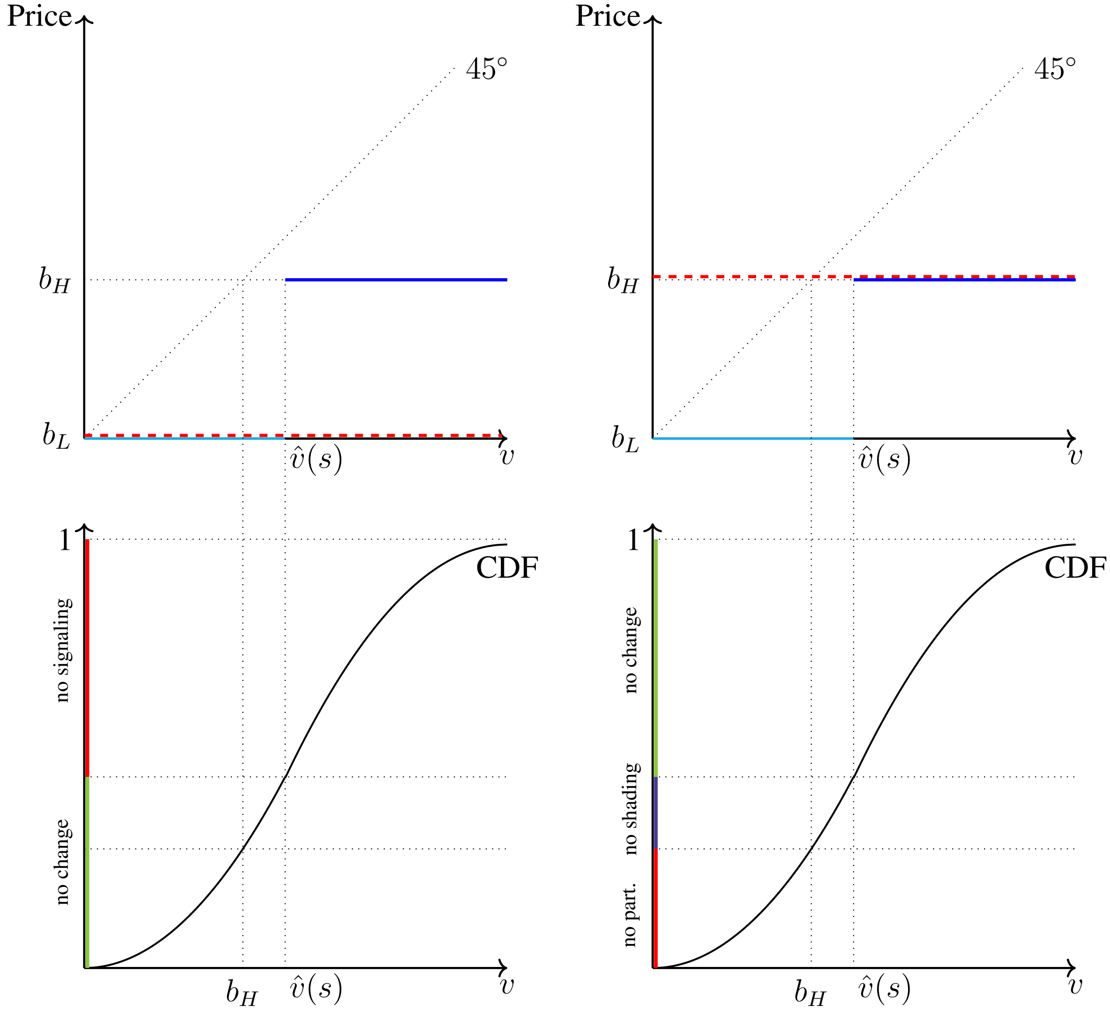
Therefore, the platform faces the following trade-off when choosing a price: by choosing a high price in CP more often, the platform prices some riders out of the market but eliminates shading. By choosing a low price in CP, the platform shuts down the signaling channel for high-valuation riders, thus lowering prices but ensuring that all riders stay on the platform.

Immediate Effect

With fixed matching rates for drivers' and riders', we can discuss the direct effect of the change in pricing on participants' welfare, which I call an *immediate effect*. I start by discussing welfare implications for riders. When the platform charges a low price, the welfare of low-valuation riders ($v < \hat{v}(s)$) does not change immediately and the welfare of high-valuation riders ($v > \hat{v}(s)$) inevitably decreases. When given a choice, high-valuation riders maximized their expected utility and chose a high price. When the platform charges a high price, a share of riders ($v < b_H$) is priced out of the market and experiences a welfare loss. The shading riders ($b_H < v < \hat{v}(s)$) now face a higher price than before, meaning that they also experience a welfare loss, since the high price was not optimal for them initially. Therefore, no riders would strictly benefit from an immediate effect

Figure 4: The Platform's Trade-Off

(a) Platform charges low price (b) Platform charges high price



Notes: The left panel shows the case when a platform charges a low price. The right panel demonstrates the case when a platform charges a high price.

of pricing centralization.

The assessment of the drivers' welfare due to an immediate effect is less straightforward. I consider two extreme cases first: when a platform charges a low price for all trips and when a platform charges a high price for all trips. Clearly, if the platform charges a uniform low price for all requests (all possible s), drivers experience a negative immediate effect: with the same arrival rates, all prices are now low. The impact of a uniform high price is less obvious. On one hand, it will lead to lower probabilities of seeing a request. On the other hand, all prices would be high. For any platform's decision between these two extreme cases, there exists a trade-off for driver welfare between the price and arrival rates: the more trips the platform prices high, the lower the share of riders who stay. If the platform does not lose too many riders due to higher prices but successfully shuts down shading behavior, drivers could benefit from CP. However, if the platform charges high prices more often and ends up losing a significant share of riders, CP will decrease driver welfare.

These assessments, however, are made under a rather strong assumption that the matching rates stay the same. After the platform switches to CP, drivers will adjust their behavior, resulting in new equilibrium matching rates that will affect both riders and drivers.

Equilibrium Effects and Pricing

Equilibrium effects. With the initial change in probabilities and prices, drivers will experience an *immediate effect* on their value of being idle. However, after the initial change, the new equilibrium will be reached through adjustment in drivers' actions and beliefs. I refer to the welfare change due to this adjustment as an *equilibrium effect*. The exact equilibrium and matching rates, however, depend on how the platform chooses prices.

Pricing. As mentioned previously, when a rider chooses a price herself, the optimal price plays a double role: it reflects the rider's own valuation, and it signals the market conditions, s, a rider faces. In the counterfactual scenario, I entirely shut down the signaling associated with the rider's valuation, but I preserve the signaling associated with market conditions by allowing a platform to vary prices based on s.

Recall that $\hat{v}(s)$ summarizes a state s in which a rider finds herself under a decentralized platform. Lower values of $\hat{v}(s)$ are associated with a "tough market" (bad s) from a rider's point of view. The higher $\hat{v}(s)$ is, the more optimistic the rider should be (good s). In the counterfactual, I assume that the platform tries to mimic the rider's actions and charges a high price when the market is "tough" and a low price when the market is "good" based on s. However, one needs to specify which markets are considered "good" and "bad," a definition that also changes with the pricing regime. In the counterfactual, the platform prices trips using a cutoff rule. When $\hat{v}(s) > \tilde{v}$ (chosen cutoff), the platform offers a low price to a rider, and when $\hat{v}(s) < \tilde{v}$, the platform offers a high price to a rider.[30]

As mentioned above, I do not aim to find the optimal pricing rule from the platform's point of view. Instead, I calculate the driver's and rider's welfare as well as other equilibrium characteristics under various pricing rules. I explore how equilibrium changes with respect to a chosen cutoff, $\tilde{v}$, and report the results. By performing this exercise, I intend to analyze trade-offs associated with centralization at different levels of price discrimination under CP. I show which pricing rules are favored by drivers and which would be preferred by riders. I evaluate all of them based on how they perform relative to a decentralized platform.[31]

Results

I present the results for one particular market: the noon hour of a weekday. Later I discuss how the results could change with the market environment. I compute the equilibrium for each specific pricing rule ($\tilde{v}$). Since I have short and long trips in the model, I need to specify two different $\tilde{v}$'s: one for short trips and one for long trips. I focus on cases where $\tilde{v}_{Long} > \tilde{v}_{Short}$, i.e., when the platform charges a high price for long trips more often than for short trips. However, showing results for various $\tilde{v}$ is not very intuitive, and I utilize the fact that the choice of $\tilde{v}$ results in different shares of high-priced trips. The relationship between $\tilde{v}$ and a share of high-priced trips is monotonic: higher $\tilde{v}$ leads to a larger share of high-priced trips. Therefore, all of the results here are presented for the shares of the trips that are priced high. The original results for the thresholds appear in Appendix D.

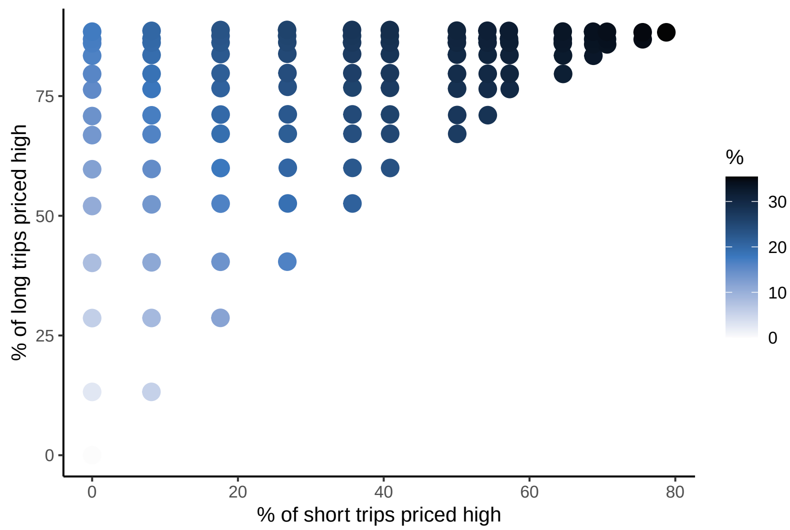
Prices and arrival rates. I compute an equilibrium for pairs of $\tilde{v}$'s (one for short trips and one for long trips) and calculate the corresponding shares of short and long trips that are priced high. With higher average prices on the platform, some riders will decide not to participate. Figure 5

[30] $\hat{v}$ here is different from the one under a decentralized platform. It is an equilibrium outcome under a CP regime.

[31] Association between market conditions that a rider faces and a chosen price introduces a constraint in the computation of both the equilibrium where prices are chosen by riders and a case where the platform sets the prices. Given the discrete nature of distances and number of competitors, one could have potentially generated all possible states, s, in which a rider could have found himself. However, with a large number of drivers and a large number of possible distances, this exercise quickly becomes unfeasible. However, the equilibrium computation is still possible through simulations. Details of the computation appear in Appendix B.

presents the percentage of riders who will be priced out of the market if a particular pricing rule is implemented. The lower-left corner corresponds to the case in which the platform sets a uniform low price for all trip types and, not surprisingly, all riders stay on the platform. As the platform chooses the high price more often (we move away from the lower-left corner to the upper-right corner of the graph), more and more riders are priced out of the platform.

Figure 5: Percentage of Priced-Out Riders



Notes: This figure depicts the percentage of riders who would be priced out under various pricing rules implemented by the platform.

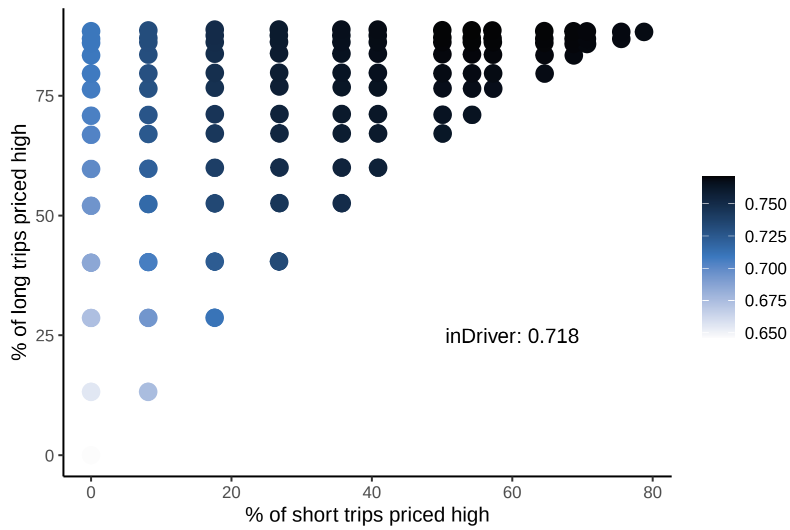
Matching rates. By pricing some riders out of the market, the platform can improve the average matching rates for riders who stay on the platform. Figure 6 displays the average matching rates conditional on participation. As the platform charges higher prices more often (moving away from the lower-left corner toward the upper-right corner), conditional matching rates increase.

Drivers. To assess the driver welfare, I look at the changes in the value of being idle (V). I calculate the value function of an idle driver under each pricing regime in CP. I then compare the value function of an idle driver under DP (V^{DP}) to the value function of an idle driver under CP (V^{CP}), and I compute a percentage change for each pricing regime:

$$\%\Delta V = 100\% \times \frac{V^{DP} - V^{CP}}{V^{DP}}$$

Figure 7 presents the percentage loss in driver welfare under CP relative to a decentralized platform. The highest loss corresponds to a case where the uniform price is low for all trips:

Figure 6: Probability of Getting Matched



Notes: This figure shows the average probabilities of getting matched for a participating rider given a proposed price, under various pricing rules implemented by the platform.

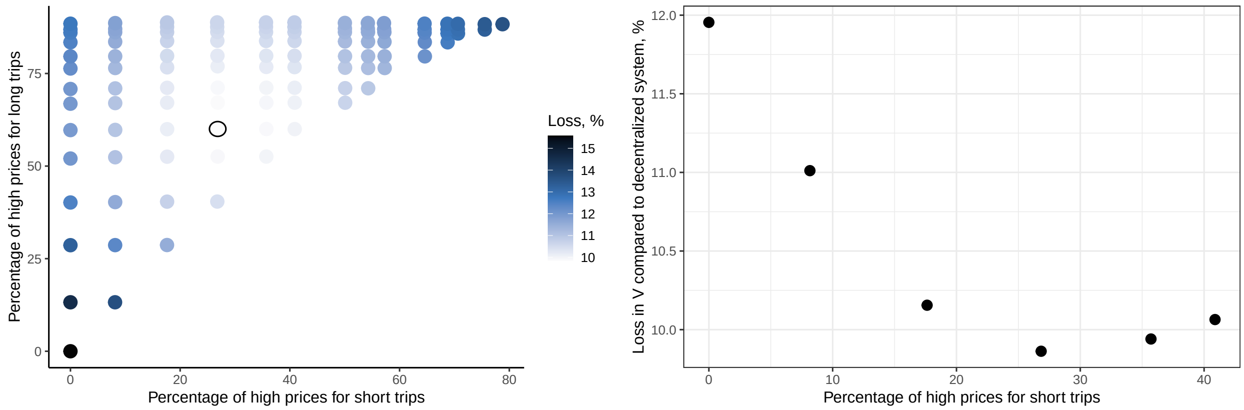
drivers strictly prefer some level of price discrimination. The lowest loss relative to a decentralized platform is achieved when a fair share of trips is priced high. But even under a pricing regime that corresponds to the highest driver welfare under CP, drivers on a centralized platform would still lose 10% of the welfare that could have been generated under a decentralized platform. The higher average prices cannot fully compensate for a share of priced-out riders, which is slightly more than 20% under that pricing regime. As the platform increases prices even more (moving to the upper-right corner), it loses more and more riders and driver welfare continues to decline.

To provide a simplified graphic representation of driver's loss, I take a horizontal slice of the left panel of Figure 7 and show how the loss varies only with the percentage of short trips that are priced high in the right panel of Figure 7. The percentage of high-priced long trips is fixed at the level that corresponds to a minimum of the loss in the left panel.

Riders. The *immediate effect* of the centralization will never improve rider welfare. However, after drivers adjust their behavior to a chosen pricing scheme, some riders could benefit from the updated matching rates. The benefits/losses depend on a rider's valuation and a chosen pricing regime.

Low-valuation riders: Depending on the pricing regime the platform implements, low-valuation riders can benefit or lose from centralization. If the platform does not choose high prices too often, low-valuation riders benefit from CP. The presence of signaling from high-valuation riders in a decentralized platform creates a negative externality for low-valuation riders by decreasing their

Figure 7: Loss in a Driver's Value Function Relative to a Decentralized Platform



Notes: The left panel presents a loss in the driver's value function relative to a decentralized platform under various pricing rules implemented by the platform. The minimum loss is marked with a circle. The right panel presents the driver's loss relative to a decentralized system with fixed $\tilde{v}$ for long trips.

matching rates. When the platform flattens matching rates under low prices, these riders gain from pricing centralization.

High-valuation riders: Riders with higher valuations do not benefit from CP. The pricing regime that corresponds to the lowest loss still cannot guarantee them higher welfare when compared to DP. The same high-valuation rider might signal her valuation in some situations (bad s) and shade her valuation in other situations (good s) under a decentralized platform. Under CP, she loses the ability to efficiently distinguish herself during bad market conditions (when the platform chooses low prices often) and loses the ability to shade her valuation (when the platform chooses high prices often).

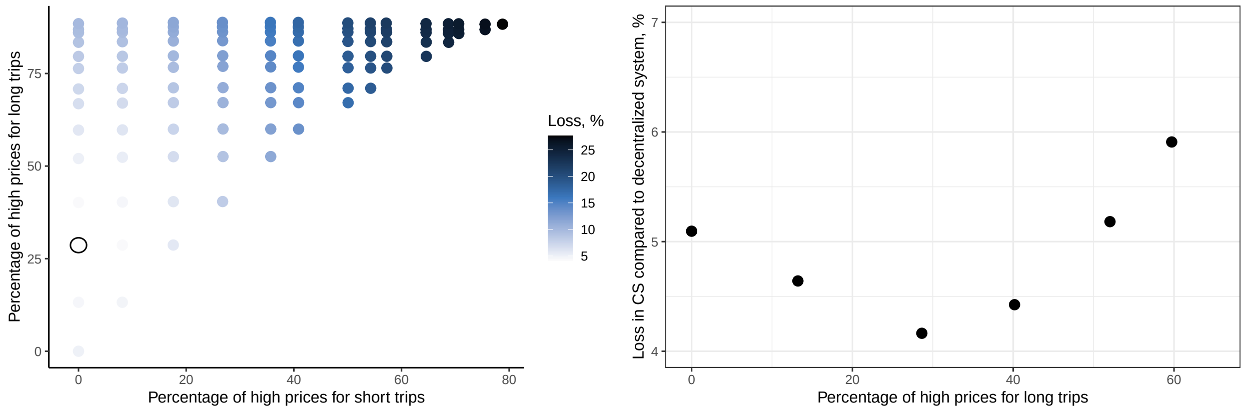
The exact benefit/loss, however, depends on each state s that a rider with a valuation v faces. For each s and v, one can find a price that a rider would have offered under the decentralized platform: $b^*(v,s) \in \{b_L, b_H\}$. A rider's expected surplus under DP is:

$$CS^{DP}(v,s) = \underbrace{\eta\big(A|b^*(v,s),s\big)}_{\text{Pr.req.accepted at } b^*(v,s)} [v - b^*(v,s)] + \underbrace{\eta\big(C|b^*(v,s),s\big)}_{\text{Pr.req.counteroffered at } b^*(v,s)} \max[v - b^*(v,s) - \Delta, 0]$$

Under CP, a rider with valuation v in state s faces a price $\hat{b}(s) \in \{b_L, b_H\}$ chosen by the platform. Her expected surplus is:

$$CS^{CP}(v,s) = \underbrace{\mathbb{1}(v > \hat{b}(s))}_{\text{Participation condition}} \times \underbrace{\tilde{\eta}\big(A|\hat{b},s\big)}_{\text{Pr.req.accepted at } \hat{b}(s)} [v - \hat{b}(s)]$$

Figure 8: Loss in Rider Surplus Relative to a Decentralized Platform



Notes: This left panel presents a loss in the rider's expected surplus relative to a decentralized platform under various pricing rules implemented by the platform. The minimum loss is marked with a circle. The right panel presents the rider's loss relative to a decentralized system with fixed $\tilde{v}$ for short trips.

The total percentage change in rider welfare between DP and CP can be calculated as:

$$\%\Delta CS = \int_v \int_s \left[100\% \times \frac{CS^{DP}(v,s) - CS^{CP}(v,s)}{CS^{DP}(v,s)}\right] f(v) f(s)\, \mathrm{d}v\, \mathrm{d}s$$

To approximate $\%\Delta CS$, I take 1,000,000 random draws for s and compute $\frac{CS^{DP}(v,s)-CS^{CP}(v,s)}{CS^{DP}(v,s)}$ for a grid of v. I then calculate $\%\Delta CS$ using various upper bounds for the distribution $f(v)$.[32]

The left panel of Figure 8 presents $\%\Delta CS$ under various pricing rules. I find that the total consumer surplus under CP always falls short of the rider's welfare under a decentralized system. The minimum loss is around 4% relative to DP. To provide a simplified representation of a rider's loss, I take a vertical slice of the left panel of Figure 8 and show how the loss varies only with the percentage of long trips that are priced high in the right panel of Figure 8. The percentage of high-priced short trips is fixed at the level that corresponds to a minimum of the loss on the right panel.

To show which riders win and lose under CP, I present how $\%\Delta CS(v)$[33] changes for low-valuation and high-valuation riders separately across different regimes of CP.[34] Figure 9 shows that CP is favored by low-type riders when prices in the platform are low. This effect emerges because by implementing CP the platform shuts down signaling from high-valuation riders and flattens

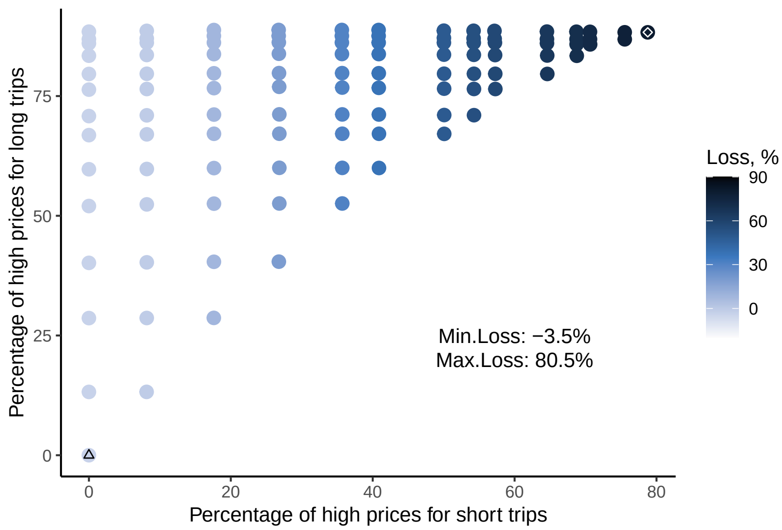
[32]The results do not change with various values of the upper bounds. The results reported here correspond to the upper bound, which is 17 times more than the lowest price in the platform.

[33]$\%\Delta CS(v) = \int_s 100\% \times \frac{CS^{DP}(v,s)-CS^{CP}(v,s)}{CS^{DP}(v,s)} f(s)\,\mathrm{d}s$

[34]I consider a rider with a valuation at the 10th percentile to be a low-valution rider and one at the 90th percentile to be a high-valuation rider.

matching rates for all riders, i.e., eliminates an externality imposed by high-valuation riders on low-valuation riders. The minimum loss, denoted with a triangle, corresponds to a 3.5% increase in CS relative to a DP. However, under higher prices in CP, low-valuation riders would prefer a decentralized system. This is not at all surprising, given that they are priced out of the market more often under higher prices. The maximum loss in CS for a low-valuation rider among pricing regimes presented on the graph, marked with a diamond, and is more than 90%. Clearly, under an equilibrium when all trips are priced high, it will become 100%.

Figure 9: Welfare Loss of Low-Valuation rider



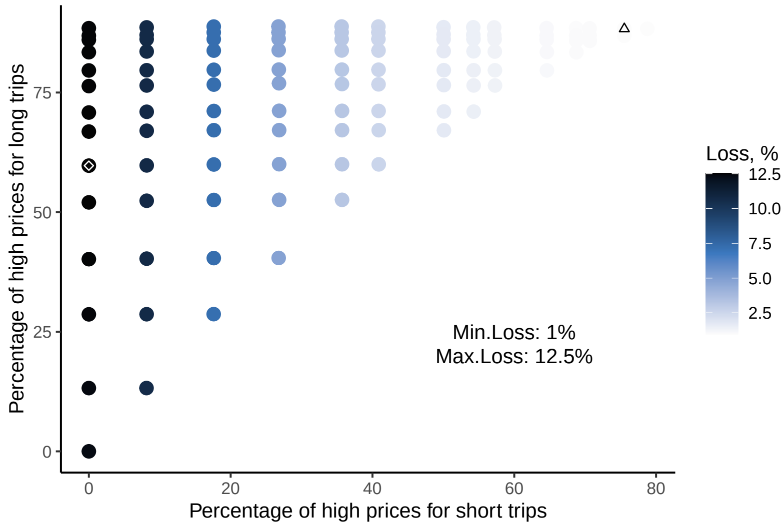
Notes: This figure presents welfare loss for a low-valuation rider. The triangle marks the lowest loss, and the diamond marks the highest loss.

Figure 10 presents welfare loss for a high-type rider. Such riders lose under all pricing regimes in CP. The loss is significantly lower for CP with high prices (upper right corner).

Discussion

The analysis above sheds light on welfare gains associated with pricing decentralization and explores which market participants could benefit from it. The exact magnitudes of the benefits and losses, however, could vary across different market environments and depend on several parameters, including the number of idle drivers, the shape of the distribution of riders' valuations, and riders' arrival rates, etc. Under some specific cases, pricing out low-valuation riders can improve driver welfare through higher average prices achieved under CP. This, however, will inevitably hurt riders.

Figure 10: Welfare Loss of High-Valuation Rider



Notes: This figure presents welfare loss for a high-valuation rider. The triangle marks the lowest loss, and the diamond marks the highest loss.

The presented counterfactual analysis intends to capture effects associated primarily with asymmetric information. The model, however, abstracts away from the equilibrium effects associated with the change in the number of idle drivers. Two effects arise when the platform switches from DP to CP. First, the number of idle drivers will change immediately. Clearly, as the platform decides to charge higher prices more often, the number of requests goes down, and the number of idle drivers increases. This can have only a negative effect on drivers: intensified competition will further lower their value of being idle. The effect for low-valuation riders will not be reversed: they will still often be priced out of the market. In theory, high-valuation riders could benefit from such change, but their new matching rates should increase dramatically to guarantee them higher welfare under CP. Second, the number of idle drivers will change in the long run as well. Driven by new expected earnings, drivers will resolve their entry problem. Future extensions of this model should incorporate drivers' entry decisions to provide the exact numerical values of the benefits and/or losses associated with CP.

7 Conclusion

This paper studies the effects of pricing decentralization (enabling privately informed agents to offer prices) in a ride-hailing market. To explore and quantify the welfare implications of asymmetric information on market outcomes, I build an equilibrium model of a decentralized ride-hailing market and present a framework in which effects of pricing decentralization can be analyzed. The findings suggest that private information is important for a two-sided market's efficiency. I find that both riders and drivers benefit from decentralized pricing. When prices are set by the platform, the welfare is redistributed between riders and drivers. By charging higher average prices, the platform shuts down riders' shading behavior and guarantees that drivers see higher average prices. However, higher prices that are chosen by a centralized platform inevitably hurt riders: some of them are priced out of the market, and some can no longer shade/signal their valuations. Decentralized pricing does not fully entail this redistribution trade-off and might offer higher welfare to both sides of the market when shading incentives are not very high. I show that decentralization of pricing decisions, which are nowadays often delegated to an intermediary, plays a crucial role in market efficiency, and could lead to higher benefits for both sides of the market.

The question "Who should choose prices on a two-sided market with asymmetric information?" is also relevant for regulatory purposes. With the pressure that arises when platforms become big players, it is important to understand which pricing mechanisms could be implemented and the effects that such changes would bring to the market.[35] The implications of the counterfactual analysis can also be extended to other two-sided markets with asymmetric information.

[35] For instance, regulatory pressure in the state of California in 2020 has incentivized Uber to switch from its usual matching mechanism to the one in which prices were set up by the drivers. However, in 2021, Uber switched back (see https://www.uber.com/blog/california/upcoming-changes-to-the-driver-app/), claiming that drivers asked for more too often and riders canceled the trips more frequently. Anecdotal evidence suggests that the drivers are not happy about the rollback (https://www.washingtonpost.com/technology/2021/06/09/uber-lyft-drivers-price-hike/), although this change is too recent to allow observers to gauge the reaction of all market participants.

References

Nikhil Agarwal and Paulo Somaini. Demand analysis using strategic reports: An application to a school choice mechanism. *Econometrica*, 86(2):391–444, 2018.

Victor Aguirregabiria and Pedro Mira. Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53, 2007.

Mohammad Akbarpour, Shengwu Li, and Shayan Oveis Gharan. Thickness and information in dynamic matching markets. *Journal of Political Economy*, 128(3):783–815, 2020.

George A Akerlof. The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics*, pages 235–251. Elsevier, 1978.

Joshua D Angrist, Sydnee Caldwell, and Jonathan V Hall. Uber vs. taxi: A driver's eye view. Technical report, National Bureau of Economic Research, 2017.

Matthew Backus, Thomas Blake, and Steven Tadelis. On the empirical content of cheap-talk signaling: An application to bargaining. *Journal of Political Economy*, 127(4):1599–1628, 2019.

Patrick Bajari, C Lanier Benkard, and Jonathan Levin. Estimating dynamic models of imperfect competition. *Econometrica*, 75(5):1331–1370, 2007.

Giulia Brancaccio, Myrto Kalouptsidi, and Theodore Papageorgiou. Geography, transportation, and endogenous trade costs. *Econometrica*, 88(2):657–691, 2020.

Nicholas Buchholz. Spatial equilibrium, search frictions and dynamic efficiency in the taxi industry. *Working Paper*, 2018.

Nicholas Buchholz, Laura Doval, Jakub Kastl, Filip Matejka, and Tobias Salz. The value of time: A high-frequency analysis of ridehail auctions. Technical report, 2019.

Hongbin Cai, John Riley, and Lixin Ye. Reserve price signaling. *Journal of Economic Theory*, 135 (1):253–268, 2007.

Juan Camilo Castillo. Who benefits from surge pricing? Technical report, 2019.

M Keith Chen, Peter E Rossi, Judith A Chevalier, and Emily Oehlsen. The value of flexible work: Evidence from uber drivers. *Journal of political economy*, 127(6):2735–2794, 2019.

Cody Cook, Rebecca Diamond, Jonathan Hall, John A List, and Paul Oyer. The gender earnings gap in the gig economy: Evidence from over a million rideshare drivers. Technical report, 2018.

Jean-Pierre Dubé and Sanjog Misra. Personalized pricing and customer welfare. *Available at*

SSRN 2992257, 2019.

Liran Einav, Chiara Farronato, Jonathan Levin, and Neel Sundaresan. Auctions versus posted prices in online markets. *Journal of Political Economy*, 126(1):178–215, 2018.

Guillaume R Frechette, Alessandro Lizzeri, and Tobias Salz. Frictions in a competitive, regulated market: Evidence from taxis. *American Economic Review*, 109(8):2954–92, 2019.

Jonathan V Hall and Alan B Krueger. An analysis of the labor market for uber's driver-partners in the united states. *Ilr Review*, 71(3):705–732, 2018.

Hugo A Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150, 1992.

Bruno Jullien and Thomas Mariotti. Auction and the informed seller problem. *Games and Economic Behavior*, 56(2):225–258, 2006.

Kei Kawai, Ken Onishi, and Kosuke Uetake. Signaling in online credit markets. Technical report, National Bureau of Economic Research, 2021.

Elena Krasnokutskaya. Identification and estimation of auction models with unobserved heterogeneity. *The Review of Economic Studies*, 78(1):293–327, 2011.

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

Tracy Liu, Zhixi Wan, and Chenyu Yang. The efficiency of a dynamic decentralized two-sided matching market. *Available at SSRN 3339394*, 2019.

Richard D McKelvey and William Zavoina. A statistical model for the analysis of ordinal level dependent variables. *Journal of mathematical sociology*, 4(1):103–120, 1975.

Hellmuth Milde and John G Riley. Signaling in credit markets. *The Quarterly Journal of Economics*, 103(1):101–129, 1988.

Paul Milgrom and John Roberts. Price and advertising signals of product quality. *Journal of political economy*, 94(4):796–821, 1986.

Gleb Romanyuk and Alex Smolin. Cream skimming and information design in matching markets. *American Economic Journal: Microeconomics*, 11(2):250–76, 2019.

Nicola Rosaia. Competing platforms and transport equilibrium: Evidence from new york city. Technical report, mimeo, Harvard University, 2020.

John Rust. Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, pages 999–1033, 1987.

Navdeep S Sahni and Harikesh S Nair. Does advertising serve as a signal? evidence from a field experiment in mobile search. *The Review of Economic Studies*, 87(3):1529–1564, 2020.

Michael Spence. Job market signaling. In *Uncertainty in economics*, pages 281–306. Elsevier, 1978.

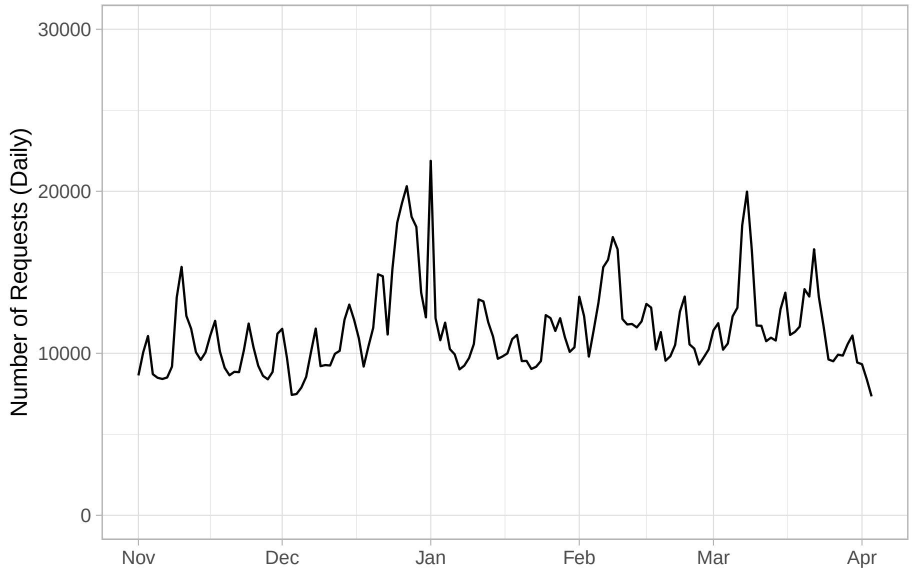
Steven Tadelis and Florian Zettelmeyer. Information disclosure as a matching mechanism: Theory and evidence from a field experiment. *American Economic Review*, 105(2):886–905, 2015.

Zaiyan Wei and Mingfeng Lin. Market mechanisms in online peer-to-peer lending. *Management Science*, 63(12):4236–4257, 2017.

Gabriel Y Weintraub, C Lanier Benkard, and Benjamin Van Roy. Markov perfect industry dynamics with many firms. *Econometrica*, 76(6):1375–1411, 2008.

Appendix A: Figures and Tables

Figure A1: Requests



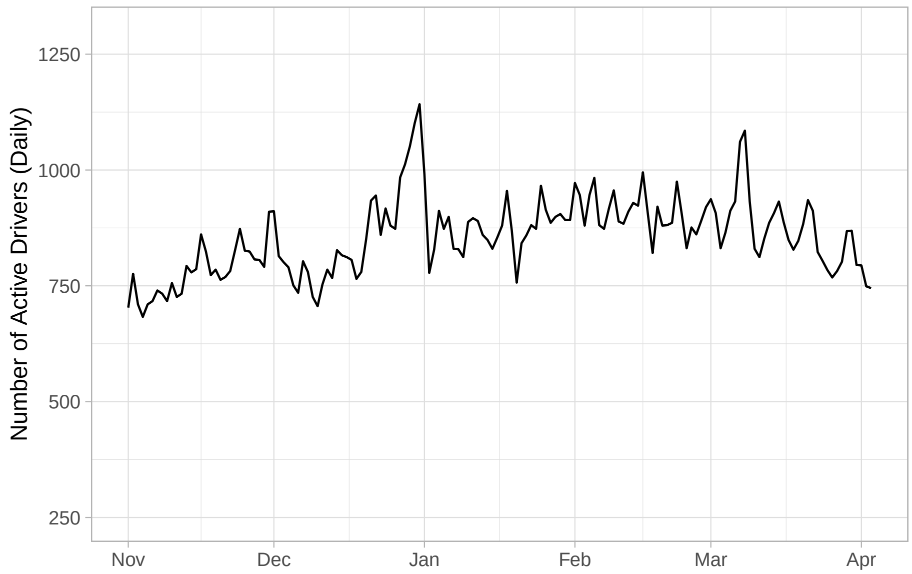
Note: This figure shows the daily number of requests originated on the platform during the study period.

Table A1: Data Cleaning

	N
Unique requests	1,764,250
-Holidays	-96,121
-Extreme weather	-234,369
-Rare prices	-45,724
-Outliers in coordinates	-14,773
-Night Hours(9PM-6AM)	-443,854
-Not matched address	-166,404
Total requests	763,005

Note: This table shows the criteria for sample selection.

Figure A2: Unique Drivers



Note: This figure shows the number of unique drivers participating on the platform during the study period.

Appendix B: Computational Details

Computing Existing Equilibrium

The actions of the riders (choosing between b_L and b_H) combined with exogenous arrival probabilities define the probabilities that riders see request of different types - $\alpha(\bar{d}, b, d)$. Although drivers do not observe their competitors at the time when the request arrives, I assume they know the long-run probabilities of various states that describe possible distances and the number of idle competitors on the platform. A part of the rider's state s not observed by a driver i is a vector $\boldsymbol{d_{-i}} \in D_{-i}$. Driver i does not explicitly know what is the state s that a rider observes. However, the price the rider offers signals both the rider's valuation v and the state s he faces.

The rider forms her beliefs based on observed state s, which can be written as a vector $[(d_1, k_1), ..., (d_X, k_X)]$, where d_1 is the lowest binned distance and d_X is the longest one, and k represents the number of drivers in each bin.

The rider then forms her beliefs η for each b:

$$\eta_m(A|b, \bar{d}, s) = \eta_m(A|b, \bar{d}, \underbrace{(d_1, k_1), ..., (d_X, k_X)}_{s}) = 1 - \prod_{j=1}^{X} \left(1 - \rho_m(A|\bar{d}, b, d_j)\right)^{k_j}$$

$$\eta_m(C|b, \bar{d}, s) = \eta_m(C|b, \bar{d}, \underbrace{(d_1, k_1), ..., (d_X, k_X)}_{s}) = \prod_{j=1}^{X} \left(1 - \rho_m(A|\bar{d}, b, d_j)\right)^{k_j} - \prod_{j=1}^{X} \rho_m(I|\bar{d}, b, d_j)^{k_j}$$

With the beliefs η, a threshold for offering a high price can be calculated. Under the condition $2\eta_m(A|b_L, \bar{d}, s) + 2\eta_m(C|b_L, \bar{d}, s) < \eta_m(A|b_H, \bar{d}, s)$, it is equal to:

$$\hat{v}_m(\bar{d}, s) = \frac{\overbrace{[\eta_m(A|b_L, \bar{d}, s)b_L + \eta_m(C|b_L, \bar{d}, s)(b_L + \Delta)]}^{\text{expected payoff if } b_L \text{ is offered}} - \overbrace{[\eta_m(A|b_H, \bar{d}, s)b_H]}^{\text{expected payoff if } b_H \text{ is offered}}}{\underbrace{[\eta_m(A|b_L, \bar{d}, s) + \eta_m(C|b_L, \bar{d}, s)]}_{\text{Pr.to get matched if } b_L \text{ is offered}} - \underbrace{[\eta_m(A|b_H, \bar{d}, s)]}_{\text{Pr.to get matched if } b_H \text{ is offered}}}$$

Alternatively, if $2\eta_m(A|b_L, \bar{d}, s) + 2\eta_m(C|b_L, \bar{d}, s) > \eta_m(A|b_H, \bar{d}, s)$, it is equal to:

$$\hat{v}_m(\bar{d},s) = \frac{\overbrace{[\eta_m(A|b_L,\bar{d},s)b_L + \eta_m(C|b_L,\bar{d},s)(b_L+\Delta)]}^{\text{expected payoff if } b_L \text{ is offered}} - \overbrace{[\eta_m(A|b_H,\bar{d},s)b_H + \eta_m(C|b_H,\bar{d},s)(b_H+\Delta)]}^{\text{expected payoff if } b_H \text{ is offered}}}{\underbrace{[\eta_m(A|b_L,\bar{d},s) + \eta_m(C|b_L,\bar{d},s)]}_{\text{Pr.to get matched if } b_L \text{ is offered}} - \underbrace{[\eta_m(C|b_H,\bar{d},s) + \eta_m(C|b_H,\bar{d},s)]}_{\text{Pr.to get matched if } b_H \text{ is offered}}}$$

With a particular order of actions on the platform (the rider draws a valuation prior to observing a state of the market $\rightarrow$ observes a state and forms beliefs about probabilities to get matched $\rightarrow$ offers an optimal price given valuation and beliefs $\rightarrow$ drivers observe the request and their own pickup distance), a probability that a driver observes a request of type $(\bar{d}, b, d_i)$ is:

$$\begin{aligned}
\alpha_m(\bar{d},b,d_i) &= \sum_{\boldsymbol{d_{-i}} \in \boldsymbol{D_{-i}}} \mathbb{P}_m(\bar{d},b,d_i,\boldsymbol{d_{-i}}) \\
&= \sum_{\boldsymbol{d_{-i}} \in \boldsymbol{D_{-i}}} \mathbb{P}_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\bar{d},d_i,\boldsymbol{d_{-i}}) \\
&= \sum_{\boldsymbol{d_{-i}} \in \boldsymbol{D_{-i}}} \mathbb{P}_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\bar{d})\mathbb{P}(d_i)\mathbb{P}_m(\boldsymbol{d_{-i}}) \\
&= \lambda_m(\bar{d})\mathbb{P}(d_i) \times \sum_{\boldsymbol{d_{-i}} \in \boldsymbol{D_{-i}}} \mathbb{P}_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}}) \\
&= \lambda_m(\bar{d})\mathbb{P}(d_i) \times \sum_{\boldsymbol{d_{-i}} \in \boldsymbol{D_{-i}}} \omega_m(b|\bar{d},\underbrace{d_i,\boldsymbol{d_{-i}}}_{s})\mathbb{P}_m(\boldsymbol{d_{-i}})
\end{aligned}$$

The probabilities of winning, conditional on agreeing or making a counteroffer for a request of type $(\bar{d}, b, d_i)$ are defined by the exact matching algorithm that was developed by the system. The algorithm is proprietary (the exact way how the system breaks the ties); however, the probabilities can be written in a generalized way:

$$\begin{aligned}
\mu_m(A,\bar{d},b,d_i) &= \sum_{\boldsymbol{d_{-i}} \in \boldsymbol{D_{-i}}} \mathbb{P}_m(\boldsymbol{d_{-i}}|\bar{d},b,d_i)\mathbb{P}_m(WM|A,\bar{d},b,d_i,\boldsymbol{d_{-i}}) \\
\mu_m(C,\bar{d},b,d_i) &= \sum_{\boldsymbol{d_{-i}} \in \boldsymbol{D_{-i}}} \mathbb{P}_m(\boldsymbol{d_{-i}}|\bar{d},b,d_i)\mathbb{P}_m(WM|C,\bar{d},b,d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(v > b+\Delta \mid b)
\end{aligned}$$

The two equations above are equivalent to:

$$
\begin{aligned}
\mu_m(A,\bar{d},b,d_i) &= \sum_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \frac{\mathbb{P}_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})}{\sum\limits_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \mathbb{P}_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})} \mathbb{P}_m(WM|A,\bar{d},b,d_i,\boldsymbol{d_{-i}}) \\
&= \sum_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \frac{\omega_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})}{\sum\limits_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \omega_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})} \mathbb{P}_m(WM|A,\bar{d},b,d_i,\boldsymbol{d_{-i}}) \\
\mu_m(C,\bar{d},b,d_i) &= \sum_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \frac{\mathbb{P}_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})}{\sum\limits_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \mathbb{P}_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})} \mathbb{P}_m(WM|C,\bar{d},b,d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(v > b+\Delta \mid b) \\
&= \sum_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \frac{\omega_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})}{\sum\limits_{\boldsymbol{d_{-i}}\in \boldsymbol{D_{-i}}} \omega_m(b|\bar{d},d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(\boldsymbol{d_{-i}})} \mathbb{P}_m(WM|C,\bar{d},b,d_i,\boldsymbol{d_{-i}})\mathbb{P}_m(v > b+\Delta \mid b)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}_m(WM|A,\bar{d},b,d_i,\boldsymbol{d_{-i}}) &= \mathbb{P}_m(WM|A,\bar{d},b,d_i,\underbrace{[(d_1,n_1),}_{n_1 \text{ competitors with distance } d_1} \quad \dots \quad \underbrace{,(d_X,n_X)]}_{n_X \text{ competitors with distance } d_X} \\
&= \prod_{j=1}^{X}[\rho_m(A|\bar{d},b,d_j)\times \mathbb{P}(win|d_i,d_j) + (1-\rho_m(A|\bar{d},b,d_j))\times 1]^{n_j} \\
\mathbb{P}_m(WM|C,\bar{d},b,d_i,\boldsymbol{d_{-i}}) &= \mathbb{P}_m(WM|C,\bar{d},b,d_i,\underbrace{[(d_1,n_1),}_{n_1 \text{ competitors with distance } d_1} \quad \dots \quad \underbrace{,(d_X,n_X)]}_{n_X \text{ competitors with distance } d_X} \\
&= \prod_{j=1}^{X}[\rho_m(C|\bar{d},b,d_j)\times \mathbb{P}(win|d_i,d_j) + (1-\rho_m(A|\bar{d},b,d_j)-\rho_m(C|\bar{d},b,d_j))\times 1]^{n_j}
\end{aligned}
$$

Algorithm to compute existing equilibrium:

0.0 I take 1,000 random draws of $\boldsymbol{d_{-i}}$ for each d_i and keep them fixed in all iterations.

0.1 Initialize beliefs μ_m^{iter}

0.2 Calculate η_m^{iter}

0.3 Solve rider's problem and find ω_m^{iter}

0.4 Compute α_m^{iter}

1. Compute V_m^{iter}
2. Derive predicted ρ_m^{iter}
3. Compute new $\eta_m^{iter'}$
4. Solve rider's problem and find $\omega_m^{iter'}$
5. Compute α_m^{iter+1}
6. Compute $\mu_m^{iter'}$
7. Update drivers' beliefs for $iter + 1$: $\mu_m^{iter+1} = a \times \mu_m^{iter} + (1 - a)\mu_m^{iter'}$
8. Repeat steps 1-7 until beliefs do not change and fixed point is found.

Appendix C: Estimation

Table C1: Riders' Beliefs (η) Estimation: Ordered Probit

	8 a.m.	10 a.m.	noon	2 p.m.	4 p.m.	6 p.m.	8 p.m.
High price	-0.434	-0.346	-0.519	-0.580	-0.539	-0.612	-0.597
	(0.026)	(0.040)	(0.029)	(0.028)	(0.033)	(0.022)	(0.027)
Number of visible drivers	-0.359	-0.213	-0.304	-0.232	-0.177	-0.262	-0.224
	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)	(0.006)	(0.007)
Short trip	-0.408	-0.342	-0.315	-0.451	-0.338	-0.319	-0.355
	(0.025)	(0.037)	(0.026)	(0.025)	(0.030)	(0.021)	(0.026)
Ordered probit thresholds							
Limit 1	0.301	1.067	-0.029	0.024	0.706	-0.290	0.328
	(0.021)	(0.032)	(0.023)	(0.022)	(0.027)	(0.020)	(0.024)
Limit 2	0.692	1.463	0.351	0.509	1.180	0.190	0.863
	(0.019)	(0.029)	(0.021)	(0.018)	(0.022)	(0.015)	(0.018)
Number of observations	42,656	26,881	28,664	31,707	30,388	43,711	37,263
Log-likelihood	-26,925.284	-13,012.190	-22,705.330	-25,076.019	-18,424.472	-36,443.109	-25,418.540

Notes: This table reports the results for the riders' beliefs estimation using an ordered probit model. Beliefs are estimated separately for each market. The dependent variable can take three values: 1 (Request is accepted), 2 (Request is counteroffered), and 3 (Request is ignored). The independent variables are (1) the indicator for a high price offered for a request, (2) the number of drivers within R-meter radius, and (3) an indicator for a short trip.

Table C2: Demand-Side Estimates

	Short Trip				Long Trip			
	Num. obs.	Shape (k_m)	Scale (θ_m)	Log-likelihood	Num. obs.	Shape(k_m)	Scale (θ_m)	Log-likelihood
8 a.m.	18,520	0.340	2323.226	-10,463	24,136	0.521	2111.673	-17,154
		(0.008)	(58.871)			(0.009)	(49.614)	
10 a.m.	12,435	0.680	1851.909	-6,843	14,446	0.926	1700.558	-10,127
		(0.023)	(58.024)			(0.023)	(50.984)	
noon	13,205	0.268	1474.325	-7,070	15,459	0.419	1528.473	-10,484
		(0.013)	(99.470)			(0.016)	(97.734)	
2 p.m.	14,226	0.295	1485.884	-7,707	17,481	0.431	1411.844	-11,808
		(0.016)	(104.020)			(0.020)	(108.124)	
4 p.m.	13,684	0.471	1413.321	-7,382	16,704	0.655	1413.504	-11,353
		(0.026)	(85.382)			(0.028)	(83.803)	
6 p.m.	18,671	0.237	1834.072	-10,610	25,040	0.343	2615.305	-17,249
		(0.012)	(153.526)			(0.014)	(229.934)	
8 p.m.	15,041	0.440	1413.002	-8,777	22,222	0.622	1411.199	-15,472
		(0.019)	(74.633)			(0.018)	(62.576)	

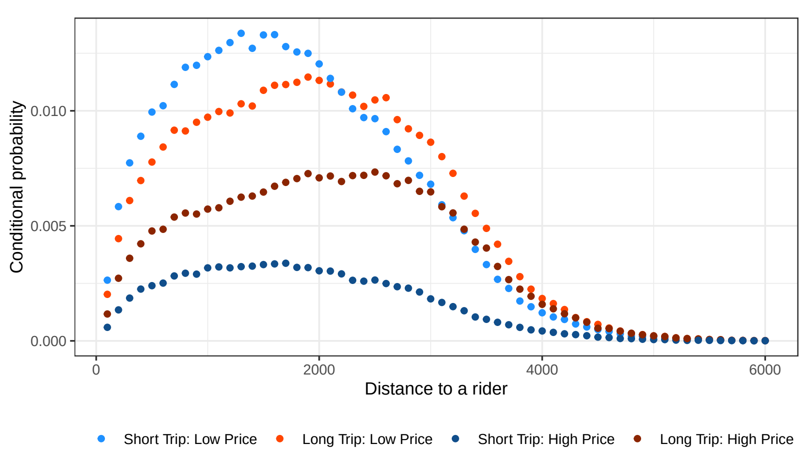
Notes: This table reports the estimates for the parameters of truncated gamma distribution obtained via maximum likelihood. The estimation is performed separately for long and short trips on each market.

Table C3: Drivers' Beliefs (μ) Estimation: Logit

	$\mathbb{1}$(Driver wins)
Agree	2.523 (0.039)
Pickup distance (d)	-0.647 (0.018)
Long trip $\times$ Low price	0.433 (0.027)
Long trip $\times$ High price	0.173 (0.028)
Short trip $\times$ Low price	0.133 (0.026)
10 a.m.	0.132 (0.076)
noon	0.331 (0.074)
2 p.m.	0.231 (0.068)
4 p.m.	0.156 (0.069)
6 p.m.	0.307 (0.063)
8 p.m.	0.156 (0.064)
Constant	-2.793 (0.046)
Interaction terms	✓
Number of observations	415,308
Log-likelihood	-252,394.400

Notes: This table reports the results for the drivers' beliefs estimation using a logit model. All interaction terms between market indicators and indicators for trip types, pickup distance, and a chosen action are omitted from the table for brevity.

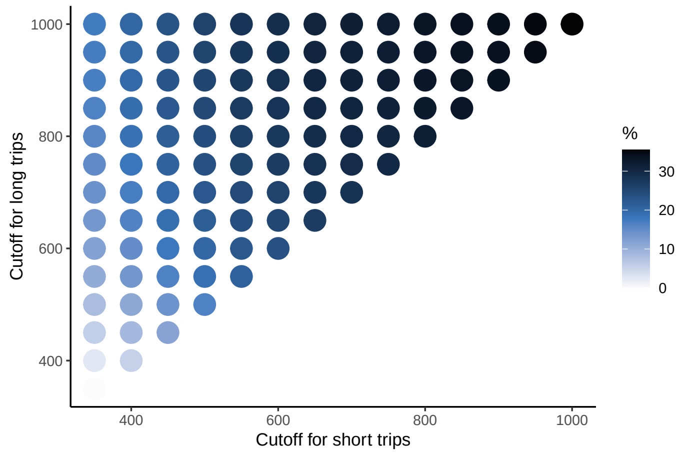
Figure C1: Conditional Probability of Request Arrivals



Note: This figure presents the conditional distribution of the requests of different types for a single market.

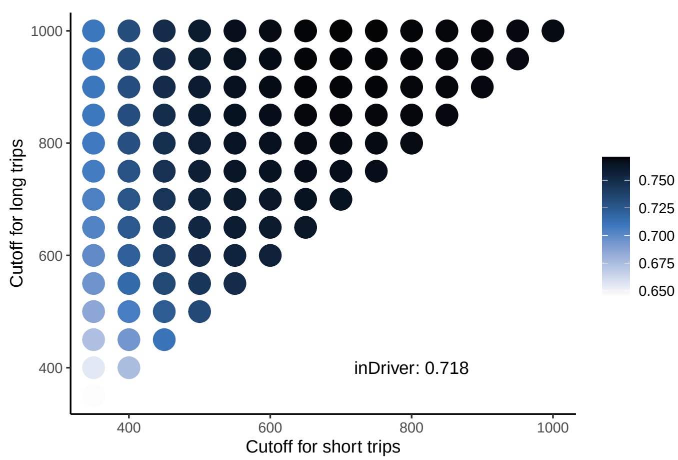
Appendix D: Counterfactual Analysis

Figure D1: Percentage of Priced-Out Riders



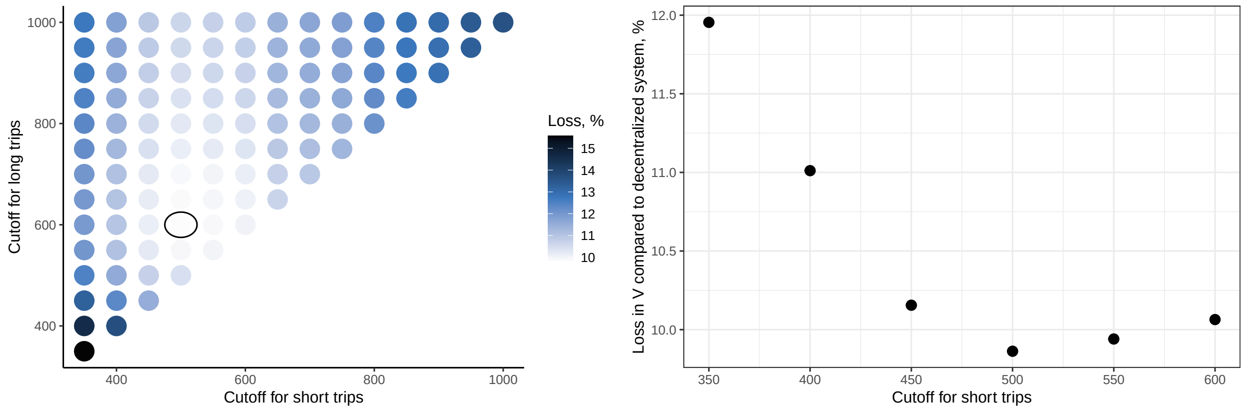
Note: This figure depicts the percentage of riders who would decide not to participate under a proposed price for various pricing rules implemented by the platform.

Figure D2: Probability of Getting Matched



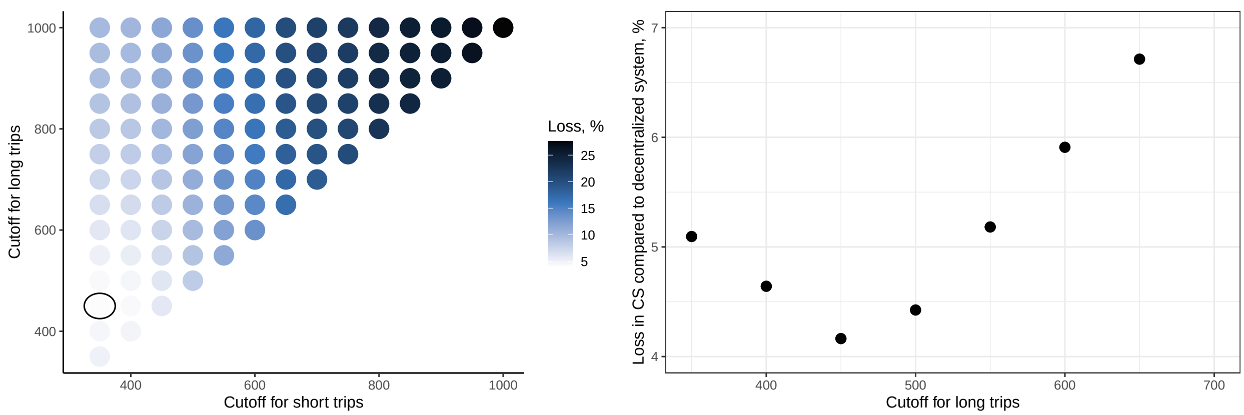
Note: This figure shows average probabilities of getting matched for a participating rider under a proposed price for various pricing rules implemented by the platform.

Figure D3: Loss in Driver's Value Function Relative to a Decentralized Platform



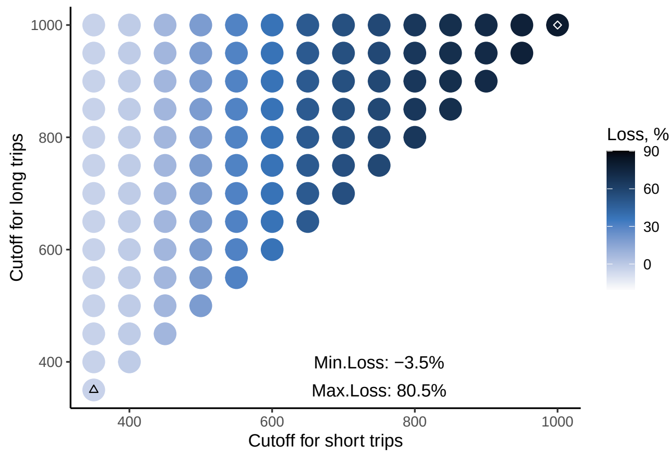
Notes: The left panel presents a loss in driver's value function relative to a decentralized platform under a proposed price for various pricing rules implemented by the platform. The minimum loss is marked with a circle. The right panel presents the driver's loss relative to a decentralized system with fixed $\tilde{v}$ for long trips.

Figure D4: Loss in Rider's Expected Surplus Relative to a Decentralized Platform



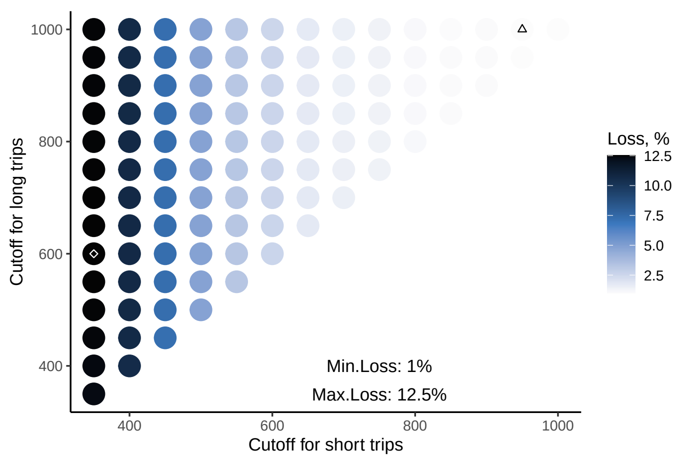
Notes: This left panel presents a loss in the rider's expected surplus relative to a decentralized platform under a proposed price for various pricing rules implemented by the platform. The minimum loss is marked with a circle. The right panel presents the rider's loss relative to a decentralized system with fixed $\tilde{v}$ for short trips.

Figure D5: Welfare Loss for Low-Valuation Riders



Notes: This figure presents the welfare loss for a low-valuation rider. The triangle marks the lowest loss, and the diamond marks the highest loss.

Figure D6: Welfare Loss for High-Valuation Riders



Notes: This figure presents the welfare loss for a high-valuation rider. The triangle marks the lowest loss, and the diamond marks the highest loss.