THE UNIVERSITY OF CHICAGO

# PRICING ON TWO-SIDED MARKET OF RIDE-SHARING PLATFORM

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

KENNETH C. GRIFFIN DEPARTMENT OF ECONOMICS

BY
CHIEN-YU LAI

CHICAGO, ILLINOIS
AUGUST 2021

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# ACKNOWLEDGMENTS

Here, I want to thank many people who give me supports through past few years as a graduate student in the University of Chicago.

First of all, my deepest gratitude goes to professor John List, my advisor. His comments always provide clear intuition and new angle for my research topics. This truly helps me strengthen my way of thinking and skills for research.

Secondly, I would like to express my thank to all professors in the Department of Economics. I truly learned a lot from the courses and workshops here. Professor Ali Hortaçsu and Professor Steve Levitt give lots valuable comments and advice to my thesis.

I also appreciate the assistance from the team in Lyft, especially Thomas Yu and Ian Muir. Thomas' excellent skill and patience help me derive the empirical results of my thesis. Ian's understanding of Lyft's experiments and coordination makes the project of my thesis possible.

Finally, I want to thank my family and all my friends. They always give me warm encouragement and helpful discussion which are necessary for me to continue my research.

# ABSTRACT

Ride-sharing platforms match riders and drivers by setting two prices. One price is what riders pay, the other is what drivers receive. There is no bargaining process between riders and drivers and there is also no direct transaction between these two sides. Thus, we may think of this market as two-sided. We model equilibrium in this two-sided market by endogenizing the number of riders and drivers and the two prices. In this paper, we study how (rider's) price affects the equilibrium outcome. From Lyft's market-level experiment on changing rider's price, we found higher rider's price decreases both rider's demand and waiting time, and it also decreases driver's supply and utilization in the equilibrium. We also compare the demand elasticities from market-level price treatment and individual-level price treatment, and we found that the former is smaller than the later in absolute value terms. This helps us do simple welfare analysis for rider's price change. From Lyft's driver side price treatments, we detect scaling problem that higher driver price might harm drivers by increased competition between drivers.

# CHAPTER 1

# INTRODUCTION

A ride-sharing platform matches riders and drivers. In this market, the platform sets two prices. One price is what riders pay, and the other is what drivers receive. If not only the gap between these two prices but the levels of these two prices affect the amount of trade, then this is a two-sided market. This paper focuses on how these two prices affect the outcome of this market. According to the characteristics of a ride-sharing platform, we construct a simple model for this market.

On a ride-sharing platform, riders care about the price they pay and eta, and drivers care about their (expected) hourly earnings. First, eta depends on the numbers of riders and drivers on the platform. Second, drivers' hourly earnings can be constructed from the driver's price and the average rate of ride requests they get, and the latter also depends on the numbers of riders and drivers on the platform. Therefore, in this market, a lower rider's price will attract more riders, and such an increase in riders will also attract more drivers to join if the driver's price remains the same because it makes it easier for drivers to get a rider. A similar argument could be made for the effect of a change in the driver's price. These can be interpreted as externalities in this market. In section 3, we show how to include such externalities in building a model characterizing the equilibrium of a ride-sharing platform.

Section 4 use experiment from Lyft to estimate the effect of prices on the platform. In the last section, we provide the application and extension of the structure in this paper.

# CHAPTER 2

# LITERATURE

Rochet and Tirole (2006) define a two-sided market. If the volume of transactions on a platform is affected by the allocation of total price between different ended users (for example, buyer and seller), this platform is a two-sided market. In this paper, they also provide some simple models to characterize two-sided markets and discuss the conditions for optimal pricing. Rochet and Tirole (2003) discuss platform competition with two-sided markets, and they show how governance structure affects price allocation and surplus between two sides of end-users. Weyl (2010) constructs a general model for multi-sided platforms. In this model, he turns the platform's problem into the choice of allocations, and he shows how user heterogeneity causes the properties and comparative statics of two-sided markets.

From literature of empirical work on two-sided markets, the structure depends on the characteristics of the market it focuses on. Wilbur (2008) discusses television advertising and viewing markets. Viewers care about how long the advertisement is when watching TV. Advertiser demand is affected by audience size, and other characteristics of the program. He uses IVs to estimate the parameters in the model. Landsman and Stremersch (2011) focus on multihoming effects in the video game console industry. They set up three equations for the whole system. They use prices of the platform in other countries and exchange rates as IVs to estimate the model. Rysman (2004) discusses competition in the market for yellow pages. The key property in its model is that advertisers may want less advertisement in a book and more users of the book, while the user may want more advertisement in a book. This paper uses some IVs to estimate parameters of the structural model. From these three examples, we should understand that models of two-sided markets and the factors we want to focus on can vary a lot.

My model is similar to Rochet and Tirole (2006) and Weyl (2010). Rochet and Tirole (2006) assume end-users care about the size of end-users from the other side. Thus, the aggregate number of end users of one side is a function of the size of the other side. However, this might not be appropriate for a ride-sharing platform. On the platform, riders might care about eta, and drivers might care about the average rate of ride requests they could get. In the equilibrium, Eta and average ride request rate should be generated by the numbers of both riders and drivers. I provide some micro-foundations to derive functions of aggregate demand and aggregate supply under this situation, and the result is similar to Rochet and Tirole (2006).

# CHAPTER 3

# SIMPLE MODEL

In this section, we construct a simple model describing the market of a ride-sharing platform. In this model, the ride-sharing platform has only one market, and it sets one price for riders and the other price for drivers. The platform earns the gap between these two prices. Both drivers and riders are price takers. After they see prices, they decide whether to stay on the platform or not. The equilibrium of the model is defined as how many drivers and riders stay on the platform given the prices, so the equilibrium is a concept of quantity. Followings are the notations for the model.

$j = S, D$. $S$ for supply (drivers), $D$ for demand (riders).

$N_j$: Size of $j$.

$P_j$: Price of $j$.

$M(N_D, N_S)$: Amount of trade generated by two sides.

$N_D(P_D, N_s), N_S(P_S, N_D)$: Demand function and supply function.

$N_j^*(P_D, P_S)$: Equilibrium size of $j$ ( $N_D(P_D, N_S(P_S, N_D^*)) = N_D^*$)

$P_D - P_S = P$: Gap of prices

Before entering the model, let's describe drivers' and riders' decisions about using the ride-sharing platform. This helps illustrate which behaviors of both sides are included in the model and how the equilibrium is constructed. Then, we derive the (aggregate) demand and supply functions and construct the equilibrium of the model.

## 3.1 Driver's and Rider's decisions

When using the ride-sharing platform, users may need to make two decisions. The first is whether they should open the app or not. If they open the app, the second decision is whether they should stay on the platform (rider's sending request and driver's receiving request) after they see some information from the platform (like estimated time of arrival, prices). The first decision is based on priors on situations users might face when opening the app. We separate users' decisions as a two-stage problem. Rider's two-stage behavior is as following:

$$\text{Stage 1:} \max\{u_{i0}, \int_{v_{i0}>u_i(E)-P_D} v_{i0} dF(P_D, E) + \int_{v_{i0}\leq u_i(E)-P_D} u_i(E) - P_D dF(P_D, E)\}$$

$$\text{Stage 2:} \max\{v_{i0}, u_i(E) - P_D\}$$

In stage 2, the rider already sees the eta (E, estimated time of arrival) and price ( $P_D$), and he could decide to get a ride on the platform or not. If he gets a ride on the platform, his utility is $u_i(E) - P_D$ where $u_i(E)$ is the utility from arriving the destination in time $E$. Otherwise, he gets $v_{i0}$. In stage 1, the rider decides to open the app or not based on priors on price and eta ( $F$). The second part of stage 1 problem is the expected utility of opening the app. Similarly, we also use a two-stage problem to describe drivers' decisions.

$$\text{Stage 1:} \max\{w_{j0}, \int_{\text{stop best}} r_{j0} dG(t_A, t_B, P_A S, P_B S, \epsilon, e) +$$
$$\int_{\text{A best}} [P_{AS} \cdot t_A + \epsilon] dG(t_A, t_B, \epsilon, e) + \int_{\text{B best}} [P_{BS} \cdot t_B + e] dG(t_A, t_B, \epsilon, e)\}$$

$$\text{Stage 2:} \max\{r_{j0}, P_{AS} \cdot t_A + \epsilon, P_{BS} \cdot t_B + e\}$$

In stage 2, after the driver opens the app, he could decide to close the app or choose a location to wait to get a potential request (here, we are using location A or B), and $t_A$ is the average ride request rate he could get at location A. In the simple one-market model, only location A will be remained in the model. Here, we just show how the model would work if there are multiple sub-markets for the drivers and the platform. In stage 1, the driver decides to open the app based on prior $G$.

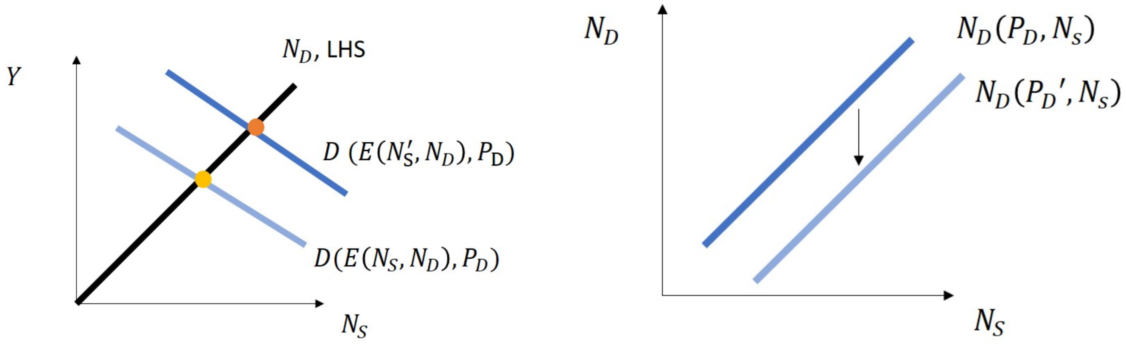
In this paper, we mainly focus on users' second stage. Therefore, the model is based on those who have already opened the app, and it links how prices and eta affect riders' decisions to send requests and drivers' decisions to stay on the platform. Therefore, the analysis in this paper can be interpreted as the short-run response to a price change. If the price change is made in the long run, users' priors will be changed as well. Then, stage 1 behavior should be considered in analyzing the market outcome.

## 3.2 Deriving the demand function and supply function

According to the previous section, riders care about the price they pay and eta, and drivers care about the price they receive and the rate of rides they could get. However, in the model, aggregate demand is set as a function of demand price and number of drivers, $N_D(P_D, N_s)$, and aggregate supply is set as a function of supply price and number of riders, $N_S(P_S, N_D)$. In this section, we will construct the aggregate demand and aggregate supply from what users of the platform care about and the concept of equilibrium.

In the ride-sharing market, the number of riders using the app is decided by demand price ( $P_D$) and eta ( $E$). Therefore, from the rider's viewpoint, the aggregate demand can be written as $N_D = D(E, P_D)$. Lower eta and lower demand price will attract more riders to use the platform. Thus, such an aggregate demand function should be characterized by $\frac{\partial D}{\partial P_D} < 0$ and $\frac{\partial D}{\partial E} < 0$. On the other hand, eta is the equilibrium result from the size of both the demand side and supply side. More drivers and fewer riders should generate lower eta, so $E$ can be expressed as $E(N_S, N_D)$ with $\frac{\partial E}{\partial N_S} > 0$ and $\frac{\partial E}{\partial N_D} < 0$. Then, we inject this to the aggregate demand above, and it becomes:

$$N_D = D(E(N_S, N_D), P_D) \tag{3.1}$$



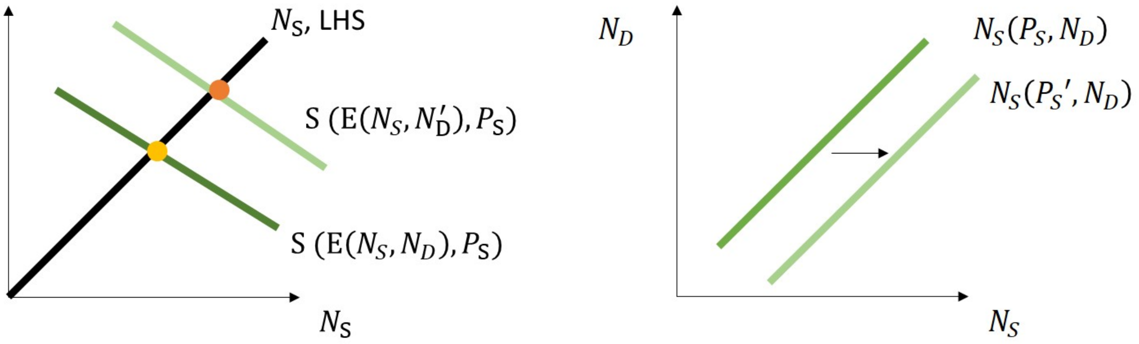
(a) (b)

Figure 3.1: Relationship between two sides from aggregate demand

The right-hand side of equation (3.1) is increasing in $N_S$ and decreasing in $N_D$. If we draw the left-hand side and the right-hand side both on a diagram given a fixed $P_D$, it shows the relationship between $N_D$ and $N_S$. It looks like Figure 3.1(a). Larger $N_S$ ( $N_S^{'} > N_S$) will generate lower $E$ and thus moves up $D(E(N_S, N_D), P_D)$, and the intersection is moving to the right as well. By collecting all these intersections, we could construct the relationship between $N_D$ and $N_S$ for aggregate demand as Figure 3.1(b). Therefore, the aggregate demand can be expressed as a function of demand price and number of drivers ( $N_D(P_D, N_S)$). In addition, higher demand price ( $P_D^{'} > P_D$) will move down the $N_D(P_D, N_s)$ line on Figure 3.1(b) from rule of demand.

Similar steps can be applied to aggregate supply as well. Drivers care about supply price ( $P_S$) and average rides per hour ( $T$) they could get from the platform. Thus, aggregate supply can be written as $N_S = S(T, P_S)$ with $\frac{\partial S}{\partial P_S} > 0$ and $\frac{\partial S}{\partial T} > 0$ because higher supply price and more ride requests make drivers earn more for staying on the platform. $T$ is an equilibrium outcome from size of supply and demand, so it can be expressed as $T(N_S, N_D)$ with $\frac{\partial T}{\partial N_S} < 0$ and $\frac{\partial T}{\partial N_D} > 0$. Replace this to the aggregate supply function from above, the equation becomes:

$$N_S = S(T(N_S, N_D), P_S) \tag{3.2}$$



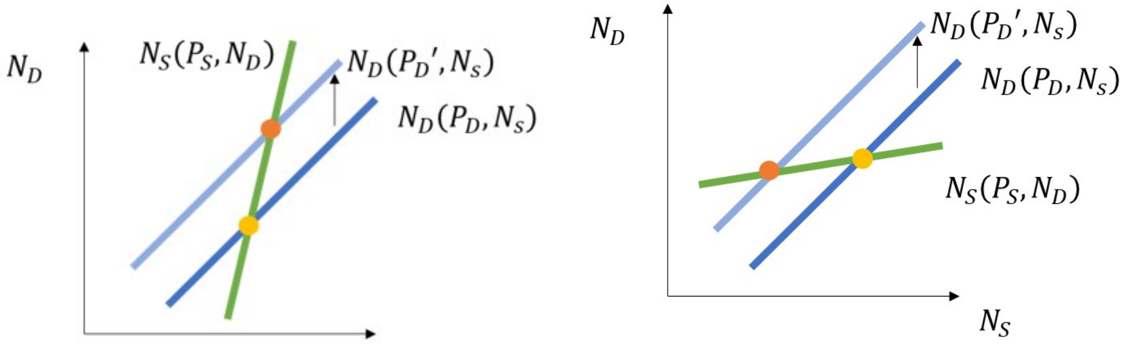
(a) (b)

Figure 3.2: Relationship between two sides from aggregate supply

The right-hand side of equation (3.2) is decreasing in $N_S$ and increasing in $N_D$. Then, we again draw the left-hand side and the right-hand side of equation (3.2) on the same diagram. We could construct the relationship between $N_D$ and $N_S$ for aggregate supply as Figure 3.2(b). Figure 3.2 ( $N_D^{'} > N_D$, $P_S^{'} > P_S$) shows that the aggregate supply can be re-written as a function of supply price and number of riders on the platform $N_S(P_S, N_D)$ with $\frac{\partial N_S}{\partial P_S} > 0$ and $\frac{\partial N_S}{\partial N_D} > 0$.

## 3.3 Equilibrium

Now, we have aggregate demand function $N_D(P_D, N_s)$ with $\frac{\partial N_D}{\partial P_D} < 0$, $\frac{\partial N_D}{\partial N_S} > 0$ and aggregate supply function $N_S(P_S, N_D)$ with $\frac{\partial N_S}{\partial P_S} > 0$ and $\frac{\partial N_S}{\partial N_D} > 0$. $\frac{\partial N_D}{\partial P_D} < 0$ and $\frac{\partial N_S}{\partial P_S} > 0$ come from the law of demand and supply. $\frac{\partial N_D}{\partial N_S} > 0$ and $\frac{\partial N_S}{\partial N_D} > 0$ implicitly come from how rider's eta and driver's average ride request rate connect to the number of riders and number of drivers in the equilibrium. In this model, we assume both riders and drivers are price takers. The equilibrium is defined as how many drivers and riders are staying on the platform given supply price and demand price. The equilibrium quantity is reached when $N_S^* = N_S(P_S, N_D(P_D, N_S^*))$ and $N_D^* = N_D(P_D, N_S(P_S, N_D^*))$. On the graph, the equilibrium sizes of both sides $(N_S^*, N_D^*)$ are the intersection of $N_D(P_D, N_S)$ and $N_S(P_S, N_D)$. The relationships between $N_D$ and $N_S$ are both positively correlated for



(a) Case 1 (b) Case 2

Figure 3.3: Possible cases for equilibrium with price comparative statics

the aggregate demand function and aggregate supply function. Therefore, the intersection of $N_D(P_D, N_s)$ and $N_S(P_S, N_D)$ doesn't necessarily exist. Or even if it exists, it may not be unique. If we first assume the intersection (equilibrium) exists, there are two possible case according to the relative slopes between $N_D(P_D, N_S)$ and $N_S(P_S, N_D)$. These two cases are shown in Figure 3.3.

In Figure 3.3, we suppose demand price decreases, $P_D^{'} < P_D$. From previous discussion, this price change will make $N_D(P_D, N_s)$ line move up. Then, in case 1, equilibrium driver size and equilibrium rider size both become larger. The change of equilibrium sizes in case 2 is opposite. The relative slopes of $N_D(P_D, N_s)$ and $N_S(P_S, N_D)$ is the reason for the difference between two cases. These two cases bring different comparative statics as following:

$$\text{Case 1:} \frac{\partial N_D^*}{\partial P_D} < 0, \frac{\partial N_S^*}{\partial P_D} < 0, \frac{\partial T^*}{\partial P_D} < 0 \tag{3.3}$$

$$\text{Case 2:} \frac{\partial N_D^*}{\partial P_D} > 0, \frac{\partial N_S^*}{\partial P_D} > 0, \frac{\partial T^*}{\partial P_D} > 0 \tag{3.4}$$

The sign of $\frac{\partial T^*}{\partial P_D}$ can be determined because in the thought experiment of Figure 3.3, we change only $P_D$ while keep $P_S$ the same. Thus, in the model, driver's average ride request rate in the equilibrium, $T^*$, would be the only factor for changing the equilibrium size of drivers. In next section, we could test these comparative statics through the result of experiments to distinguish which case matches to the real outcome of the ride sharing platform.

# CHAPTER 4

# EMPIRICS

## 4.1 Data description

In this paper, we use three experiments conducted on the ride-sharing platform Lyft: individual-level passenger price change, market-level passenger price change, and individual-level driver price change.

### *4.1.1 Individual-level passenger price change*

From December 20, 2018 to February 28, 2019, Lyft ran a 10-week experiment varying passenger prices at individual level in 19 regions (ATL, BWI, CLE, DCA, PHI, PHX, SFO, SJC, AUS, BNA, BOS, CHI, DEN, DFW, LAX, MIA, NYC, OCX, SEA). In this experiment, 16% of passengers on the platform were assigned one of 5 price multipliers, 1 (control), 0.9, 0.95, 1.05, and 1.1. Such assigned price multipliers would be applied to all of the passengers' rider prices for the entire 10-week period. The price multiplier was assigned right after the first time the passenger opened the Lyft app during the 10-week period (In theory, they need to log in to see the treatment/price), so we only look at data of a passenger starting from the week after his first exposure to treatment.

### *4.1.2 Market-level passenger price change*

In Lyft, PiscO (Price Optimizer) is a price adjusting system applying price multipliers to each passenger session to try to maximize consumer surplus. On June 20, 2019, Lyft increased overall price multipliers through PiscO in 11 regions (ATL, BWI, CLE, DCA, LAS, PHI, PHX, SBD, SFO, SJC, SMF). For these treated regions, we look 5 weeks before and 10 weeks after June 20, 2019 as the pre-treated period and post-treated period. During these 15 weeks, these treated regions had no other major price adjustments. Also, we chose some

Figure 4.1: Price multiplier change over time for treated regions

other top market regions with no major price adjustments in this period as control regions. The control regions are AUS, BNA, BOS, CHI, DEN, DFW, LAX, MIA, NYC, OCX, SEA.

### *4.1.3 Driver price change*

On November 22, 2019, Lyft rolled out an experiment related to driver pay at the individual level. Drivers were assigned to either treatment group or control group. The experiment was rolled out in stages, starting with 1% of drivers being assigned, and ramping up to 100% over time. A side effect (bug) of the experiment was that treated drivers experienced many of their rides paying out with a 20% commission rate instead of the usual 25% (the probability of a ride being affected conditioned on treated drivers was 86%). This bug was unannounced and not directly visible anywhere. It was spotted after some drivers noticed different commission rates between rides. This bug was fixed on January 6, 2020. We look at the 6 weeks between the start of the experiment and the bug fix. For each driver, we look at data starting from the week after their first exposure to treatment, up to the 3rd

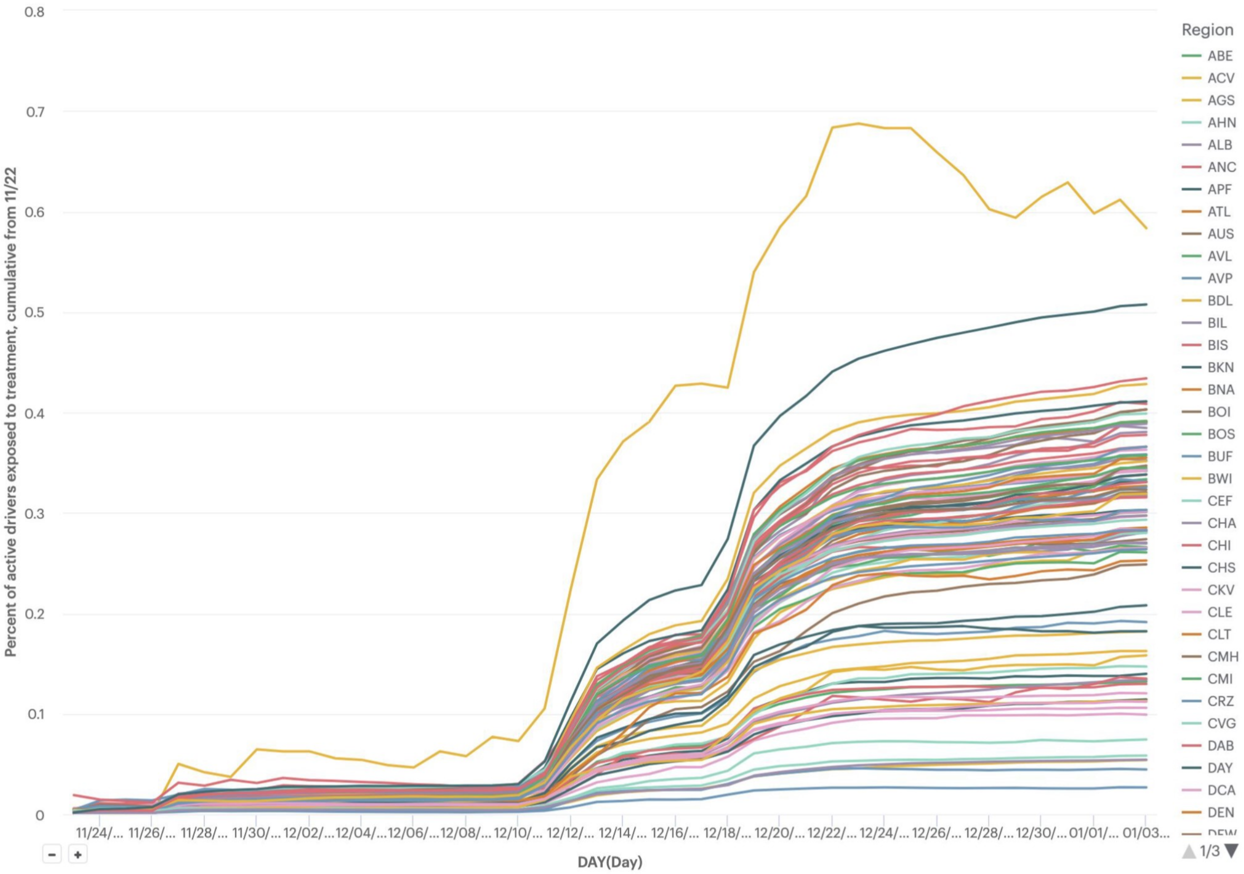


Figure 4.2: Percent of drivers affected by the bug across regions

week after their first exposure. To focus on active regions, we look at drivers from regions where, every day in the 6-week period, at least one treated driver is active – there are 97 such regions. Figure 4.2 shows percent of drivers affected by the bug across regions the six weeks.

## 4.2 Empirical result for rider side price change

### *4.2.1 Individual level*

In the individual level rider price change experiment, we look at data at the passenger-week level and run the following two regressions to derive the overall treatment effects and treatment effects over time.

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot Multiplier_i + \epsilon_{it}$$

$$Y_{it} = \alpha_i + \gamma_t + \Sigma_{w=0}^{9} \beta_w \cdot [1(Week_{it} = w) \cdot Multiplier_i] + \epsilon_{it}$$

In these equations, $Y$ is the outcome variable of interest, such as log of average rider's price, number of rider sessions, number of requested rides, number of completed rides, log of rider's average wait time, and log of average ride distance. $\alpha_i$is the rider's individual fixed effect. $\gamma_t$ is weekly fixed effect. $\beta$ and $\beta_w$ are treatment effects. $Multiplier_i$ is the price multiplier which rider $i$ got assigned to. $1(Week_{it} = w)$ is an indicator for the wth week after the passenger's first exposure to the price change.

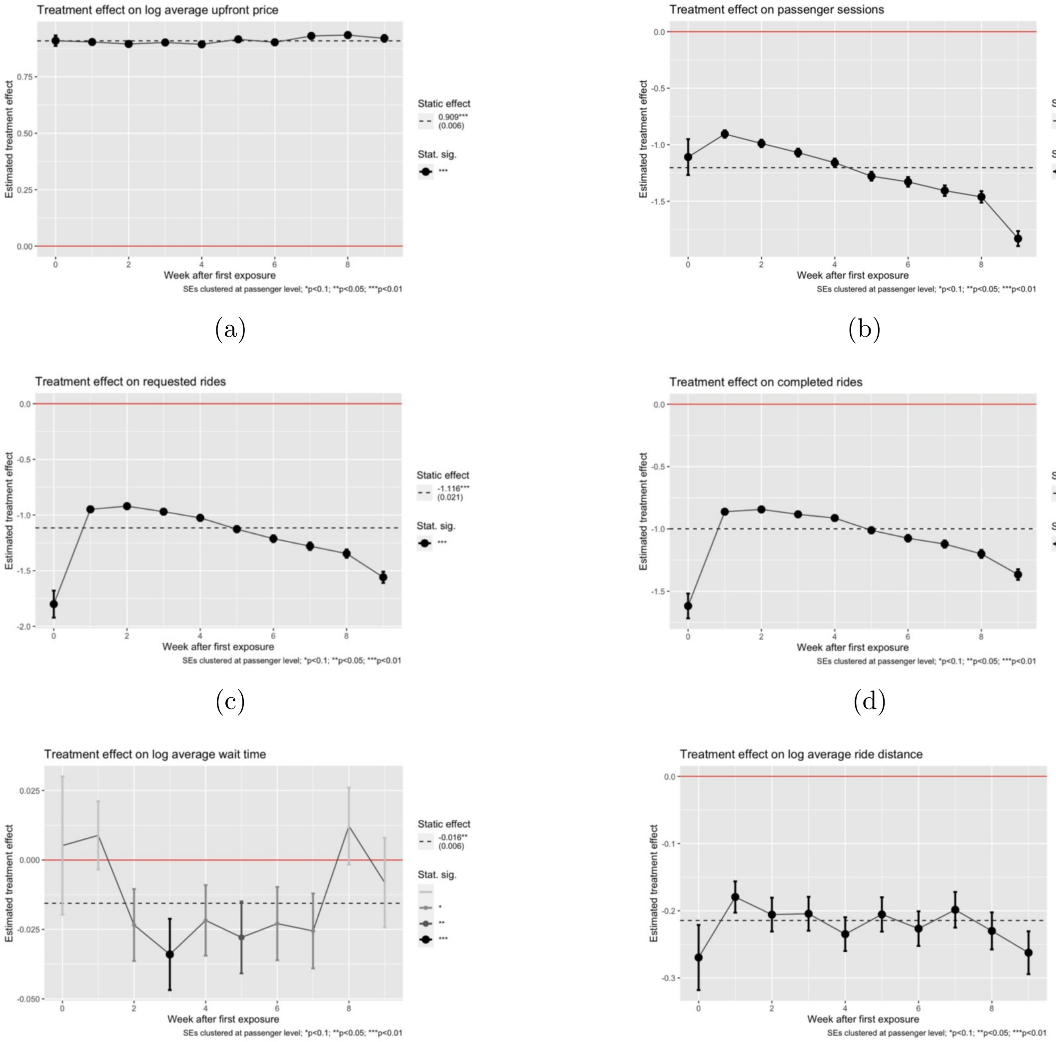
Figure 4.3 shows the treatment effect of an individual level price change. Overall, a one-unit increase on rider price multiplier generates a 91% increase on average in rider's price, 1.203 decrease in rider's weekly sessions, 1.116 decrease in rider's weekly requested rides, 0.999 decrease in rider's weekly completed rides, 1.6% decrease in rider's average wait time, and 21.4% decrease in average ride distance. These show that a higher rider price decreases rider demand, including rider's frequency of checking price (rider sessions), and the demand decrease is on rides with longer distance and longer wait time.

To derive the effect of price, we run modified 2SLS regressions as following:

$$LogPrice_{it} = \alpha_i + \gamma_t + \delta \cdot Multiplier_i + \epsilon_{it}$$

$$Y_{it} = \eta_i + \theta_t + \beta \cdot Log\hat{P}rice_{it} + \mu_{it}$$

Where $LogPrice_{it}$ is the log of average rider price. In the first stage, we use only passenger-weeks with price data because we don't have price data for passenger-weeks with no sessions (for example, we don't observe price for passengers who didn't log in in a whole



(a)

(b)

(c)

(d)

(e)

(f)

Figure 4.3: Treatment effect for individual level rider price change

week). In the second stage, we use the full sample. Table 4.1 is the result for 2SLS regressions.

Table 4.1: 2SLS results on rider side variables for market level rider price change

| | Rider sessions | Requested rides | Completed rides | Log waiting time | Log average ride distance |
|---|---|---|---|---|---|
| Predicted log rider's price | $-1.323^{***}$ | $-1.227^{***}$ | $-1.099^{***}$ | $-0.014^{**}$ | $-0.236^{***}$ |
| | (0.034) | (0.024) | (0.020) | (0.007) | (0.015) |
| Regional fixed effect | | | Yes | | |
| Week fixed effect | | | Yes | | |
| Observations | 10571211 | 10571211 | 10571211 | 4263170 | 2648158 |
| Adjusted $R^2$ | 0.019 | 0.008 | 0.009 | 0.069 | 0.030 |

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Note: Robust standard errors in parentheses

### *4.2.2 Market level*

For the market level price change experiment, we aggregate data at the region-day level. We run following OLS regressions to derive overall treatment effect and treatment effect over time. To derive the effect of price, we run modified 2SLS regressions as following:

$$LogPrice_{it} = \alpha_i + \gamma_t + \delta \cdot Multiplier_i + \epsilon_{it}$$

$$Y_{it} = \eta_i + \theta_t + \beta \cdot Log\hat{Price}_{it} + \mu_{it}$$

Where $LogPrice_{it}$ is the log of average rider price. In the first stage, we use only passenger-weeks with price data because we don't have price data for passenger-weeks with no sessions (for example, we don't observe price for passengers who didn't log in in a whole week). In the second stage, we use the full sample. Table 4.1 is the result for 2SLS regressions.

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot Post_t \cdot Treated_i + \epsilon_{it}$$

$$Y_{it} = \alpha_i + \gamma_t + \Sigma_{w=1}^{10}\beta_w \cdot [1(Week_t = w) \cdot Treated_i] + \epsilon_{it}$$

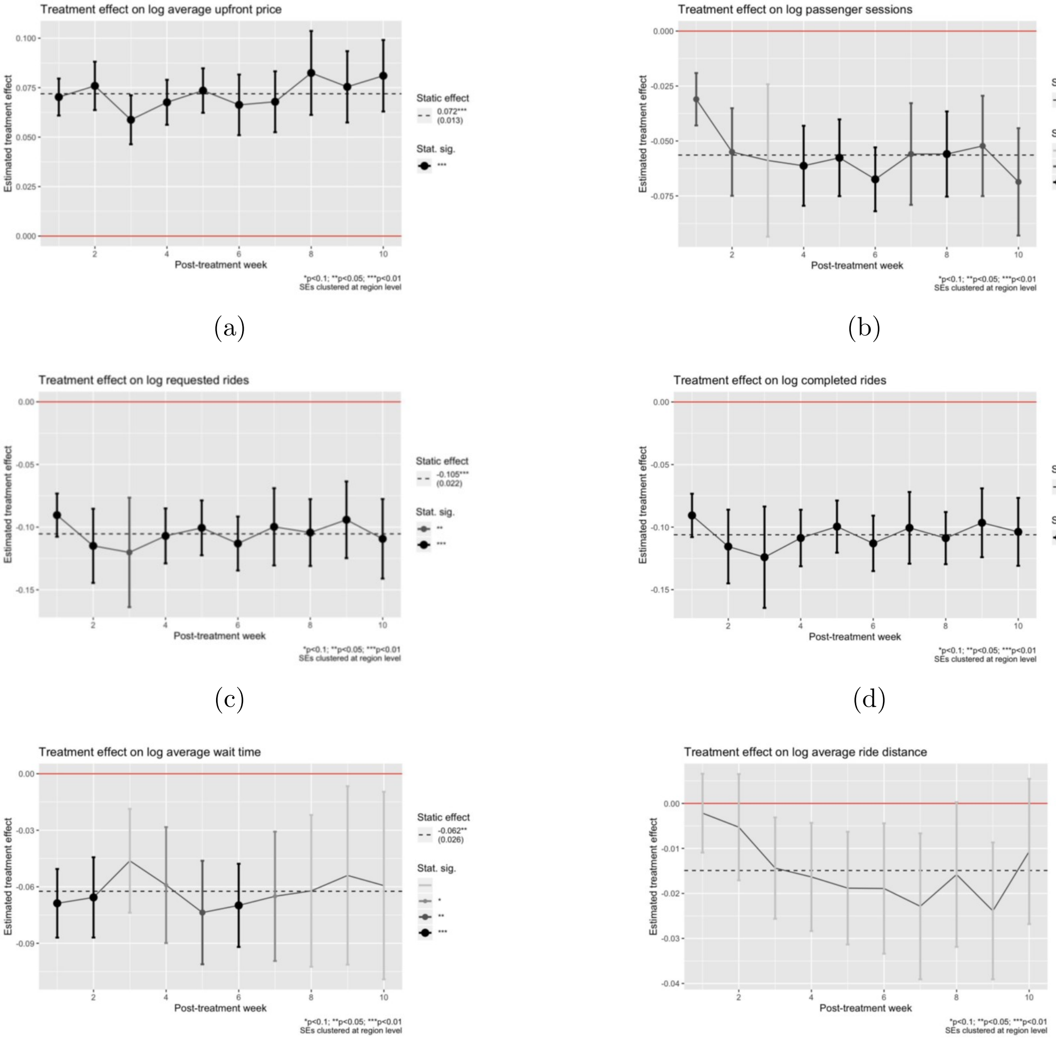
$Y$ is the outcome variable of interest which contains both rider side and driver side. Rider side outcome variables are log of average rider's price, log of rider sessions, log of requested rides, log of completed rides, log of rider's average wait time, and log of average ride distance. Driver side outcome variables are log of driver's utilization, log of driver's hourly earnings, log of market aggregate working hour, log of active drivers, log of working hours per active driver, and log of driver sessions per working hour. $\alpha_i$ is a regional fixed effect. $\gamma_t$ is a daily fixed effect. $\beta$ and $\beta_w$ are treatment effects. $Post_t$ is an indicator for days after 6/20 (start day of price change), and $Treated_i$ is an indicator for treated regions. $1(Week_t = w)$ is an indicator for days in wth week after 6/20.

Figure 4.4, 4.5 and 4.6 shows treatment effect on rider side outcome variables, unconditional driver side outcome variables, and driver side outcome variables conditioned on drivers existing on the platform before 6/20. On the rider side, this experiment increases rider's price by 7.2% and significantly decreases rider's demand on rider sessions (5.6%), requested rides (10.5%), and completed rides (10.6%). It also brings a significant decrease in rider's waiting time (6.2%), but there's no significant change in ride distance.

For the driver side, the experiment significantly decreases driver's utilization, driver's hourly earnings, and working hour per active driver both on all drivers and pre-existing drivers. We could see that most of the decrease in driver supply is reflected on the intensive margin because there's no significant decrease in number of active drivers and working hours per active driver significantly decreases by about 3%.

Similar to the above, we run 2SLS to derive price elasticities.

$$LogPrice_{it} = \alpha_i + \gamma_t + \delta \cdot Post_t \cdot Treated_i + \epsilon_{it}$$

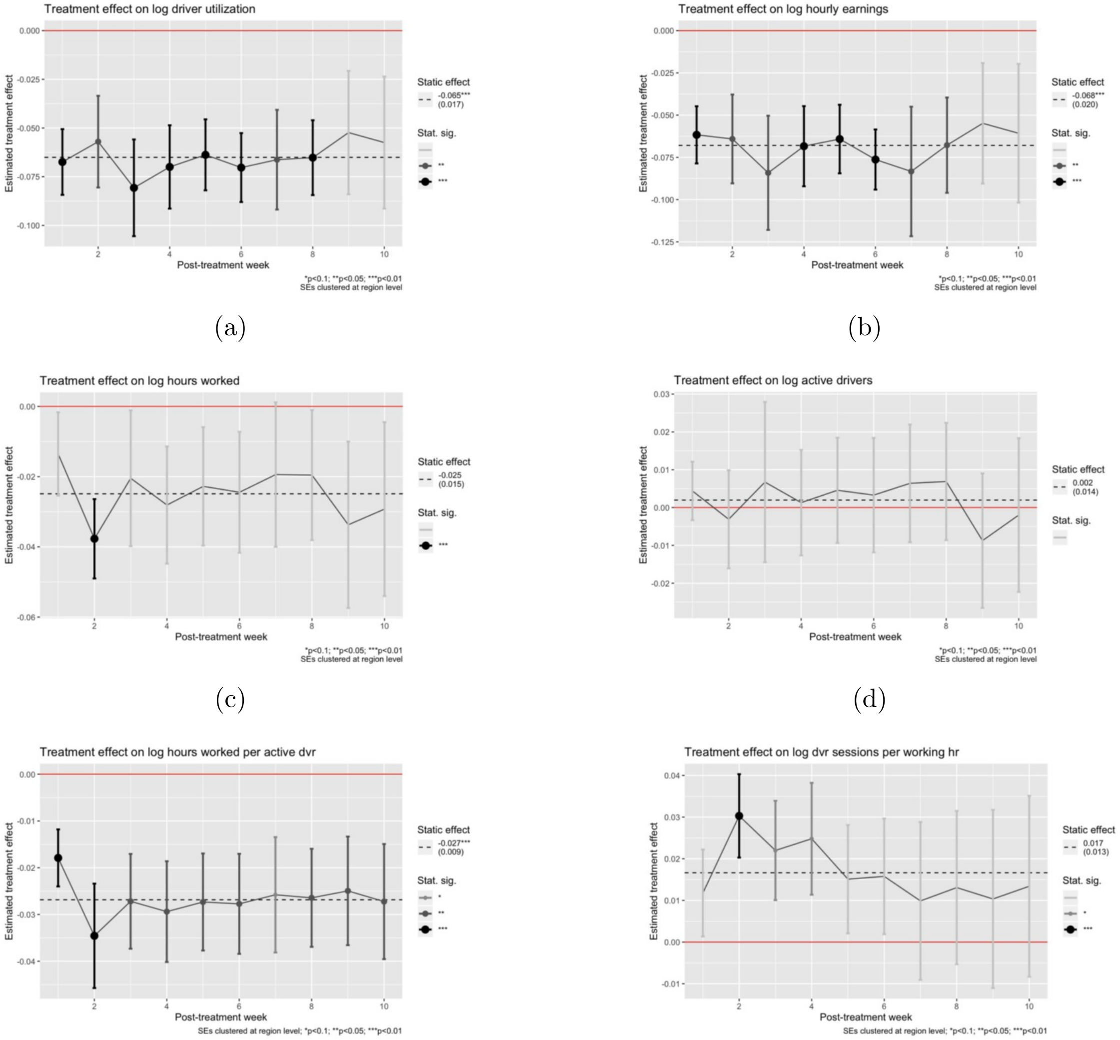


(a) (b)

(c) (d)

(e) (f)

Figure 4.4: Treatment effect on rider side variables for market level rider price change

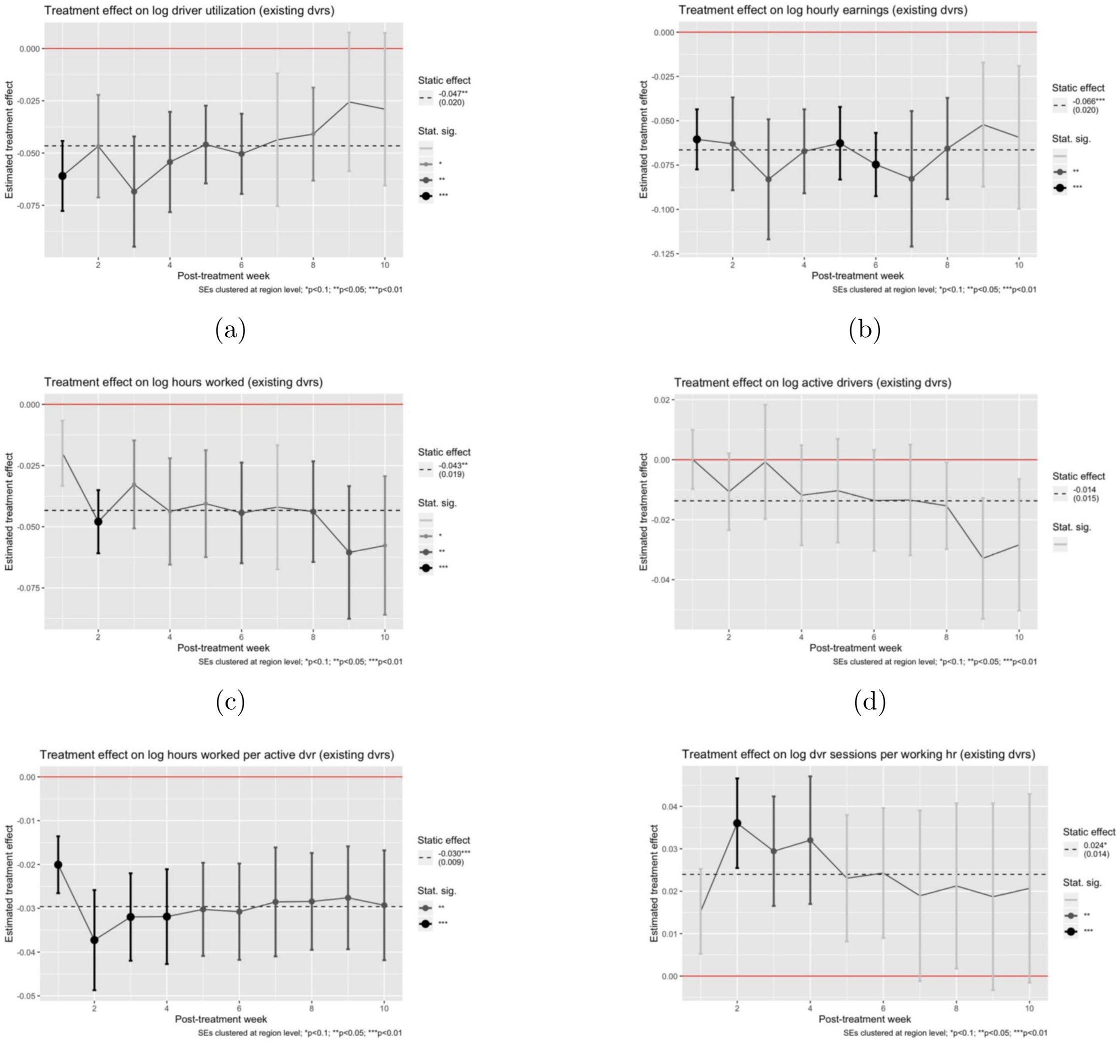


(a) (b)

(c) (d)

(e) (f)

Figure 4.5: Treatment effect on driver side variables for market level rider price change



(a) (b)

(c) (d)

(e) (f)

Figure 4.6: Treatment effect on driver side variables conditioned on pre-existing drivers for market level rider price change

$$Y_{it} = \eta_i + \theta_t + \beta \cdot Log\hat{Price}_{it} + \mu_{it}$$

$LogPrice_{it}$ is the log of average rider price at region-day level. Table 4.2, 4.3, and 4.4 are results for 2SLS regressions. Table 4.2 is for rider side variables, and Table 4.3 is for driver side variables. Table 4.4 is also for driver side variables, but it's conditioned on pre-existing drivers. We could see that both demand (rider sessions, requested ride, and completed rides) and supply (hour and hour per driver) decrease from rider price increase in the equilibrium. This fits the comparative statics of case 1 equilibrium (see Figure 3.3 and equation (3.3)). From the driver side, we don't see a significant effect on number of active drivers. This implies that the supply decrease mainly comes from the intensive margin. The decrease in utilization and earnings per hour also meet case 1 equilibrium because if driver's price is fixed, decrease in earnings and request rate should be the main reason for the decrease of supply in the model. In addition, an increase in rider's price decreases rider's waiting time in the equilibrium.

Table 4.2: 2SLS results on rider side variables for market level rider price change

| | Log rider sessions | Log requested rides | Log completed rides | Log waiting time | Log average ride distance |
|---|---|---|---|---|---|
| Predicted log average ride price | $-0.785^{***}$ | $-1.465^{***}$ | $-1.476^{***}$ | $-0.868^{*}$ | $-0.207$ |
| | (0.239) | (0.359) | (0.320) | (0.447) | (0.155) |
| Region fixed effect | | | Yes | | |
| Day fixed effect | | | Yes | | |
| Observations | | | 1995 | | |
| Adjusted $R^2$ | 0.981 | 0.964 | 0.972 | 0.785 | 0.949 |

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$
Note: Robust standard errors in parentheses

Table 4.3: 2SLS results on driver side variables for market level rider price change

| | Log utilization | Log earning per hour | Log earning per ride | Log hour |
|---|---|---|---|---|
| Predicted log average ride price | $-0.905^{***}$ | $-0.938^{**}$ | 0.219 | $-0.346^{*}$ |
| | (0.283) | (0.368) | (0.176) | (0.189) |
| Region fixed effect | | Yes | | |
| Day fixed effect | | Yes | | |
| Observations | | 1995 | | |
| Adjusted $R^2$ | 0.853 | 0.885 | 0.977 | 0.992 |
| | **Log driver** | **Log hour per driver** | **Log session per hour** | |
| Predicted log average ride price | 0.028 | $-0.373^{**}$ | 0.231 | |
| | (0.194) | (0.143) | (0.162) | |
| Region fixed effect | | Yes | | |
| Day fixed effect | | Yes | | |
| Observations | | 1995 | | |
| Adjusted $R^2$ | 0.996 | 0.957 | 0.918 | |

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Note: Robust standard errors in parentheses

### *4.2.3 Comparison and welfare analysis*

From above, we derive price elasticities from individual and market level experiments. Table 4.5 compares demand elasticities, and the difference between individual and market level is not significant. However, if we only look at the size, demand elasticities in absolute value are smaller in the market level experiment than in the individual level experiment. This might be explained by the change in rider's waiting time. From table 4.2, rider price elasticity on rider's waiting time is -0.868. This implies that riders suffer directly from the rider price increase but they get compensated by the waiting time decrease in the equilibrium. This outcome is similar to putting a tax on the demand side in a one-sided market. Table 4.5 lists rider price elasticities derived from individual an market level experiments

Table 4.4: 2SLS results on driver side variables conditioned on pre-existing drivers for market level rider price change

|  | Log utilization | Log earning per hour | Log earning per ride | Log hour |
|---|---|---|---|---|
| Predicted log average ride price | $-0.884^{***}$ | $-0.917^{**}$ | 0.191 | $-0.603^{**}$ |
|  | (0.280) | (0.364) | (0.171) | (0.229) |
| Region fixed effect | Yes |  |  |  |
| Day fixed effect | Yes |  |  |  |
| Observations | 1995 |  |  |  |
| Adjusted $R^2$ | 0.858 | 0.887 | 0.976 | 0.991 |
|  | Log driver | Log hour per driver | Log session per hour |  |
| Predicted log average ride price | $-0.191$ | $-0.412^{**}$ | $0.333^{*}$ |  |
|  | (0.191) | (0.145) | (0.174) |  |
| Region fixed effect | Yes |  |  |  |
| Day fixed effect | Yes |  |  |  |
| Observations | 1995 |  |  |  |
| Adjusted $R^2$ | 0.996 | 0.957 | 0.918 |  |

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Note: Robust standard errors in parentheses

According to these numbers, we could calculate the welfare change (change of surplus) from a rider price change by adding some assumptions on demand and supply function. Figure 4.7 show how we do such a welfare analysis.

For the demand side, we assume that demand curves are linear, and their slopes are independent of wait time. In Figure 4.7(a), when rider price increases from $P_D^*$ to $(1+\delta)P_D^*$, the quantity of demand in the equilibrium moves from the red dot to the blue dot which belong to different demand curves. The higher demand curve containing the blue dot corresponds to lower driver's waiting time after the price change. The black dot is the quantity of demand from the rider price increase given that riders waiting time stays the same. The difference between the red dot and black dot is calculated based on the estimation

Table 4.5: Comparison on price elasticities

| | Rider sessions | Requested rides | Completed rides |
|---|---|---|---|
| Market level | $-0.785^{***}$ | $-1.465^{***}$ | $-1.476^{***}$ |
| | (0.2390) | (0.3590) | (0.3200) |
| Individual level | $-0.931^{***}$ | $-1.645^{***}$ | $-1.664^{***}$ |
| | (0.0004) | (0.0024) | (0.0023) |

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$
Note: Robust standard errors in parentheses

from individual level experiment. The demand change on the rider side in the equilibrium can be separated into the movement from the red dot to the black dot and the movement from the black dot to the blue dot. Therefore, the change in consumer surplus can be shown as minus the pink area plus the blue area in the figure. The pink area is the welfare loss on the rider side directly from the rider price increase, and the blue area is the welfare gain on the rider side from the waiting time decrease. In the Figure 4.7(a), $\delta$ is the increase in rider price in percent, $\eta_i$ is the demand elasticity from individual level experiment, $\eta_m$ is the demand elasticity from market level experiment, and $N_D^*$ is the equilibrium demand quantity before the rider price change.

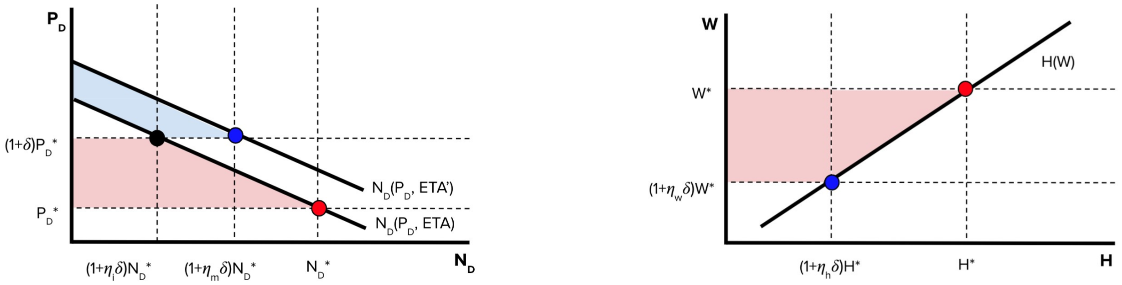
For the supply side, we assume the supply curve is linear and the unit of supply is driver's working hour. Thus, drivers care about hourly earnings ( $W$) in the supply function. In Figure 4.7(b), $\eta_w$ is the ride price elasticity on hourly earnings, and $\eta_h$ is the ride price elasticity on supply (working hour). Both $\eta_w$ and $\eta_h$ are derived from the market level experiment. The welfare change for the supply side would be minus the pink area for the ride price increase.

From Figure 4.7 and discussion above, welfare (surplus) change on the demand side and supply side for rider price increase can be calculated as:

$$\Delta DW = -\frac{(\eta_m - \eta_i)(1+\eta_i\delta)\delta P_D^* N_D^*}{\eta_i} - \frac{(\eta_m - \eta_i)^2 \delta^2 P_D^* N_D^*}{2\eta_i} - \delta(1+\eta_m\delta)P_D^* N_D^* + \frac{\eta_i \delta^2 P_D^* N_D^*}{2}$$

$$\Delta SW = \eta_w \delta (1 + \eta_h \delta) W^* H^* - \frac{\eta_w \eta_h \delta^2 W^* H^*}{2}$$

In this calculation, we take completed rides as the unit of demand because this is the amount related with the actual monetary transfer and enjoying ETA for the trade. Thus, we have $\eta_i = -1.664, \eta_m = -1.476, \eta_w = -0.938, \eta_h = -0.346$ from estimation. From 19 regions included in the market level experiment, we take average from May 17, 2019 to June 20, 2019 as the equilibrium variables before price change. $P_D = 16.586$, average ride price; $N_D^* = 899138.4$, completed rides per day; $W^* = 17.292$, hourly earnings; $H^* = 731746.017$, hours worked per day. If we consider a 1% increase in rider's price ( $\delta = 0.01$), we would get welfare (surplus) change as -\$131586.27 per day on rider side (0.88% of total revenue from riders) and -\$118483.11 per day on driver side (0.94% of driver's earning).



(a) Demand/Rider side

(b) Supply/Driver side

Figure 4.7: Change of surplus from rider price change

## 4.3 Empirical result for driver side price change

### 4.3.1 Individual level

In the driver side price change, we first analyze the data at the driver-week level. We run the following two regressions to derive overall and dynamic treatment effects.

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot Treated_i + \epsilon_{it}$$

$$Y_{it} = \alpha_i + \gamma_t + \Sigma_{w=0}^{3}[\beta_w \cdot 1(Week_{it} = w) \cdot Treated_i] + \epsilon_{it}$$

$Y$ is the outcome variable of interest, such as log earnings per ride, log hourly earnings, hour worked, log driver utilization, log driver sessions per working hour, and log percent of hours worked at peak time. $\alpha_i$ is the rider's individual fixed effect. $\gamma_t$ is weekly fixed effect. $\beta$ and $\beta_w$ are treatment effects. $Treated_i$ is the indicator for whether the driver was treated. $1(Week_{it} = w)$ is an indicator for the wth week after the passenger's first exposure to the treatment.

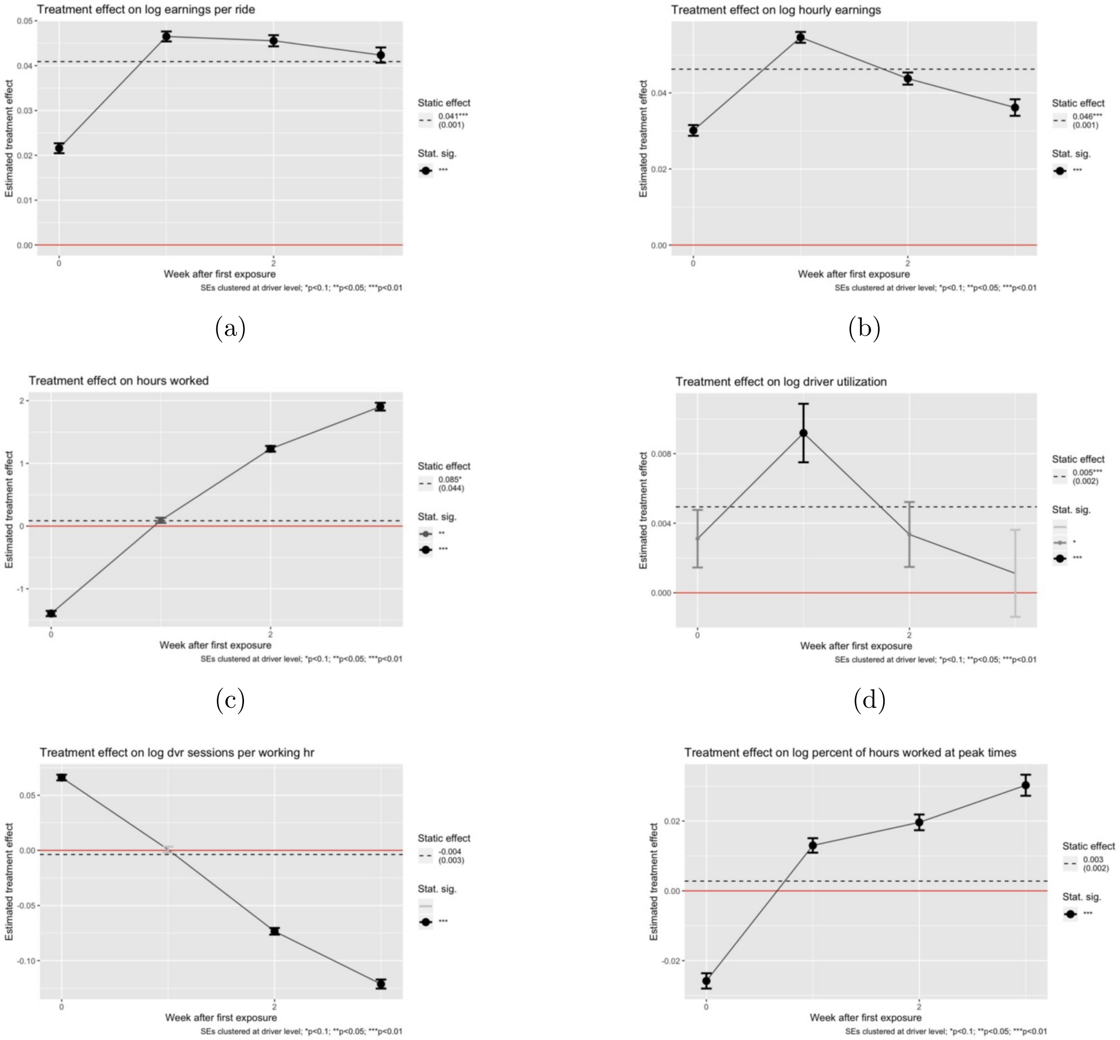
Figure 4.8 shows the treatment effect of an individual level price change. Overall, being treated generates a 4.1% increase on driver's earnings per ride, a 4.6% increase on driver's hourly earnings, 0.085 hour increase on working hour, a 0.5% increase on driver's utilization. Lower commission rate makes drivers have higher earnings, higher working hour, and higher utilization.

To derive the effect of price, we run modified 2SLS regressions as following:

$$LogRideEarnings_{it} = \alpha_i + \gamma_t + \delta \cdot Treated_i + \epsilon_{it}$$

$$Y_{it} = \eta_i + \theta_t + \beta \cdot LogRide\hat{E}arnings_{it} + \mu_{it}$$

Where $LogRideEarnings_{it}$ is the log of earnings per ride, $\alpha$is a region fixed effect, $\gamma$is week fixed effect, and $Treated_i$ is an indicator for a treated driver. In the first stage, we use only driver-weeks with completed rides because we don't have data for driver-weeks with no rides. In the second stage, we use the full sample. Table 4.6 is the result for 2SLS regressions. If $Y$ is logged, then $\beta$is the elasticity. Otherwise, the elasticity is given by the following equation:



(a) (b)

(c) (d)

(e) (f)

Figure 4.8: Treatment effect for individual level driver price change

$$Elasticity = \frac{\beta}{\bar{Y}}$$

where $\bar{Y}$ is the mean from the full sample.

Table 4.6: 2SLS result for individual level driver price change

| | Log earnings per hour | Hours | Is active | Hours if active |
|---|---|---|---|---|
| Predicted log average ride price | $1.143^{***}$ | $2.028^{*}$ | $0.071^{***}$ | 1.117 |
| | (0.036) | (1.091) | (0.024) | (1.159) |
| Region fixed effect | Yes | | | |
| Day fixed effect | Yes | | | |
| Observations | 820979 | 1019984 | 1019984 | 867600 |
| Adjusted $R^2$ | 0.242 | 0.025 | 0.015 | 0.02 |
| Elasticity | | 0.153 | 0.083 | 0.072 |

| | Rides | Peak hours | Log session per hour |
|---|---|---|---|
| Predicted log average ride price | $5.339^{***}$ | $1.467^{**}$ | $-0.087$ |
| | (1.958) | (0.602) | (0.070) |
| Region fixed effect | Yes | | |
| Day fixed effect | Yes | | |
| Observations | 1019984 | 1019984 | 846967 |
| Adjusted $R^2$ | 0.046 | 0.029 | 0.016 |
| Elasticity | 0.042 | | |

${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$
Note: Robust standard errors in parentheses

### *4.3.2 Use variation across regions to derive market level price change on driver side*

From Figure 4.2, we see the percent of drivers affected by the bug varies across regions. Thus, we try to use this variation to derive the effect of market level price change on driver side. We look at region-day level to run the following regression with the same 6-week period and 97 regions.

$$Y_{it} = \alpha_i + \gamma_t + \beta \cdot LogPctTreated_{it} + \mu_{it}$$

where $Y$ is the outcome variable of interest, $\alpha$is the region fixed effect, $\gamma$is the day fixed effect, $LogPctTreated$ is the logged percent of active drivers (cumulative from 11/22) who have been exposed to treatment. $\beta$is the treatment effect.

Table 4.7 shows that higher percent of drivers affected by bug leads to higher earnings per ride and higher work hours in the region, but it has no effect on earnings per hour. Table 4.8 shows that higher percent of drivers affected by bug has no effect on rider's wait time and demand.

We also run 2SLS to derive ride earning elasticities for market level change. The regression is run at region-day level.

$$LogRideEarnings_{it} = \alpha_i + \gamma_t + \theta \cdot LogPct\hat{T}reated_{it} + \mu_{it}$$

$$Y_{it} = \eta_i + \theta_t + \beta \cdot LogRide\hat{E}arnings_{it} + \mu_{it}$$

where $LogRideEarnings$ is logged earnings per ride, $\alpha$is a region fixed effect, $\gamma$is a day fixed effect, $LogPctTreated$ is the logged percent of active drivers (cumulative from 11/22) who have been exposed to treatment, $Y$ is the (logged) outcome variable of interest, $\eta$is a region fixed effect, $\theta$is a day fixed effect, $LogRide\hat{E}arnings_{it}$ is the predicted value from the 1st stage. $\beta$is the elasticity.

Table 4.9 and Table 4.10 are 2SLS results for market level driver price change. According to Table 4.9, there's no significant effect on driver-side variables. However, Table 4.10 shows that higher driver price leads to higher rider sessions and average ride price.

Table 4.7: OLS result for market level driver price change on driver side variables

| | Log earnings per ride | Log earnings per hour | Log hours | Log drivers |
|---|---|---|---|---|
| Log percent of drivers treated, cum | $0.009^{*}$ | 0.004 | $0.021^{**}$ | $0.012^{*}$ |
| | (0.005) | (0.010) | (0.008) | (0.006) |
| Region fixed effect | | Yes | | |
| Day fixed effect | | Yes | | |
| Observations | | 4065 | | |
| Adjusted $R^2$ | 0.952 | 0.893 | 0.995 | 0.997 |

| | Log hours per driver | Log session per hour | Log lapse per request | |
|---|---|---|---|---|
| Log percent of drivers treated, cum | $0.009^{*}$ | $-0.003$ | 0.011 | |
| | (0.005) | (0.006) | (0.017) | |
| Region fixed effect | | Yes | | |
| Day fixed effect | | Yes | | |
| Observations | | 4065 | | |
| Adjusted $R^2$ | 0.941 | 0.851 | 0.778 | |

$^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$
Note: Robust standard errors in parentheses

### *4.3.3 Scaling problem in the driver side price change*

The driver side price change experiment shows scaling problem similar to Crépon et al (2013). Table 4.11 and 4.12 are results of OLS regressions (individual-day level) restricted to regions with highest (top 20%) and lowest (bottom 20%) percent of drivers treated by bug. From the last column of table 4.11, we see that although treated drivers have higher hourly earnings, more drivers treated decreases hourly earnings for all drivers in the same region. Thus, lower commission rate can harm drivers through scaling. Such result may come from competition between drivers. From last column of table 4.12, percent of drivers treated increases working hour for all drivers in a region. Such increased competition (working hour) may cause the decrease on driver's hourly earning if the demand side doesn't change much.

Table 4.8: OLS result for market level driver price change on rider side variables

| | Log average wait time | Log sessions | Log requested rides |
|---|---|---|---|
| Log percent of drivers treated, cum | 0.0001 | $0.026^{**}$ | 0.001 |
| | (0.010) | (0.011) | (0.012) |
| Region fixed effect | | Yes | |
| Day fixed effect | | Yes | |
| Observations | | 4065 | |
| Adjusted $R^2$ | 0.847 | 0.993 | 0.989 |
| | **Log completed rides** | **Log average price** | **Log average distance** |
| Log percent of drivers treated, cum | 0.016 | $0.013^{**}$ | $-0.013^{***}$ |
| | (0.011) | (0.005) | (0.005) |
| Region fixed effect | | Yes | |
| Day fixed effect | | Yes | |
| Observations | | 4065 | |
| Adjusted $R^2$ | 0.992 | 0.885 | 0.917 |

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$
Note: Robust standard errors in parentheses

Table 4.9: 2SLS result for market level driver price change on driver side variables

| | Log earnings per hour | Log hour | Log driver |
|---|---|---|---|
| Predicted log earnings per ride | 0.406 | 2.264 | 1.318 |
| | (1.007) | (1.489) | (0.973) |
| Region fixed effect | Yes | | |
| Day fixed effect | Yes | | |
| Observations | 4065 | | |
| Adjusted $R^2$ | 0.910 | 0.987 | 0.994 |
| | Log hours per driver | Log session per hour | Log lapse per request |
| Predicted log earnings per ride | 0.947 | $-0.305$ | 1.180 |
| | (0.696) | (0.667) | (1.769) |
| Region fixed effect | Yes | | |
| Day fixed effect | Yes | | |
| Observations | 4065 | | |
| Adjusted $R^2$ | 0.890 | 0.839 | 0.788 |

${}^{*}p < 0.10$, ${}^{**}p < 0.05$, ${}^{***}p < 0.01$

Note: Robust standard errors in parentheses

Table 4.10: 2SLS result for market level driver price change on rider side variables

| | Log average wait time | Log sessions | Log requested rides |
|---|---|---|---|
| Predicted log earnings per ride | 0.016 | $2.842^{*}$ | 0.068 |
| | (1.080) | (1.657) | (1.322) |
| Region fixed effect | | Yes | |
| Day fixed effect | | Yes | |
| Observations | | 4065 | |
| Adjusted $R^2$ | 0.848 | 0.985 | 0.989 |
| | **Log completed rides** | **Log average price** | **Log average distance** |
| Predicted log earnings per ride | 1.670 | $1.431^{**}$ | $-1.446$ |
| | (1.484) | (0.619) | (1.220) |
| Region fixed effect | | Yes | |
| Day fixed effect | | Yes | |
| Observations | | 4065 | |
| Adjusted $R^2$ | 0.987 | 0.862 | 0.592 |

$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$
Note: Robust standard errors in parentheses

Table 4.11: OLS results on log hourly earnings for scaling problem

| | Log hourly earnings | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Treated | $0.051^{***}$<br>(0.003) | | $0.045^{***}$<br>(0.005) | | $0.040^{***}$<br>(0.006) |
| In high regions | | $-0.056^{***}$<br>(0.003) | $-0.062^{***}$<br>(0.005) | | |
| Treated × In high regions | | | 0.009<br>(0.006) | | |
| Percent of cumulative drivers treated | | | | $-0.584^{***}$<br>(0.012) | $-0.613^{***}$<br>(0.018) |
| Treated × Percent of cumulative drivers treated | | | | | $0.044^{**}$<br>(0.020) |
| Day fixed effect | | | Yes | | |
| Observations | | | 263276 | | |
| Adjusted $R^2$ | 0.008 | 0.008 | 0.010 | 0.022 | 0.024 |

${}^{*}\,p < 0.10$, ${}^{**}\,p < 0.05$, ${}^{***}\,p < 0.01$
Note: Robust standard errors in parentheses

Table 4.12: OLS results on sum hours worked for scaling problem

| | Sum hours worked | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Treated | $0.132^{*}$ | | 0.200 | | $-0.018$ |
| | (0.075) | | (0.128) | | (0.143) |
| In high regions | | $1.544^{***}$ | $1.608^{***}$ | | |
| | | (0.074) | (0.129) | | |
| Treated × In high regions | | | $-0.097$ | | |
| | | | (0.158) | | |
| Percent of cumulative drivers treated | | | | $7.269^{***}$ | $6.849^{***}$ |
| | | | | (0.283) | (0.440) |
| Treated × Percent of cumulative drivers treated | | | | | 0.637 |
| | | | | | (0.495) |
| Day fixed effect | | | Yes | | |
| Observations | | | 325484 | | |
| Adjusted $R^2$ | 0.007 | 0.010 | 0.010 | 0.011 | 0.011 |

$^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

Note: Robust standard errors in parentheses

# CHAPTER 5

# CONCLUSION, EXTENSION AND FUTURE WORK

In this paper, we construct a model characterizing the two-sided market of a ride-sharing platform. The model considers the externality of sizes between supply and demand. Then, we apply individual level and market level experiments on rider price changes to estimate the price effects on this two sided market, and we further do welfare analysis for rider price changes. We found out that demand price elasticity is smaller for the market level experiment than the individual level experiment, but such difference is not significant. Such a result may come from the fact that rider's waiting time is decreasing from a market level rider price increase. We also found out that the market level rider price increase causes driver's hourly earnings to decrease and then make supply decrease through the intensive margin.From driver side treatment, we found that higher earnings per ride increases driver's earnings per hour and working hour at individual level. However, we didn't find significant and remarkable effect at market level by using variation on treatment rate across regions. On the other hand, we detect scaling problem on driver price treatment by comparing regions with high and low treatment rates.

The structure of this paper can be applied to platform's optimization problem and tax analysis on the platform. Platform's optimization problem, such as maximizing profit, maximizing welfare, etc, can be separated into two parts. One is to decide the price gap between rider and driver. The other is to decide the price level of the two prices. The second part is needed if the platform is a two-sided market. The tax analysis is to calculate how tax arrangement on two sides affect the market outcome, like total amount of trade, profit, and welfare. If we are considering optimal tax arrangement, it goes back to the platform's optimization problem. The model can also be extended to describe cases with multiple platforms and multiple locations, like drivers are allowed to change between platforms and locations to wait for requests, by adding more assumptions and variables. The driver's two-stage problem in section 3.1 provides the potential of such extension.

# REFERENCES

Bruno Crépon, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. Do labor market policies have displacement effects? evidence from a clustered randomized experiment. *The Quarterly Journal of Economics*, 128(2):531–580, 2013.

Vardit Landsman and Stefan Stremersch. Multihoming in two-sided markets: An empirical inquiry in the video game console industry. *Journal of Marketing*, 75(6):39–54, 2011.

Jean-Charles Rochet and Jean Tirole. Platform competition in two-sided markets. *Journal of the European Economic Association*, 1(4):990–1029, 2003.

Jean-Charles Rochet and Jean Tirole. Two-sided markets: A progress report. *The RAND Journal of Economics*, 37(3):645–667, 2006.

Marc Rysman. Competition between networks: A study of the market for yellow pages. *The Review of Economic Studies*, 71(2):483–512, 2004.

E. Glen Weyl. A price theory of multi-sided platforms. *American Economic Review*, 100(3): 1642–1672, 2010.

Kenneth C. Wilbur. A two-sided, empirical model of television advertising and viewing markets. *Market Science*, 27(3):356–378, 2008.