2021-07-24 10:31:12-04:00

Information Design for Platform Drivers

by

Andreas Haupt

B.Sc., University of Bonn (2014)
M.Sc., University of Bonn (2017, 2018)
B.Sc., University of Frankfurt (2019)

Submitted to the Institute for Data, Systems, and Society
in partial fulfillment of the requirements for the degree of

Master of Science in Social and Engineering Systems

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author . . . .
Institute for Data, Systems, and Society
August 7, 2021

Certified by . . . .
Jinhua Zhao
Associate Professor
Thesis Supervisor

Accepted by . . . .
Fotini Christia
Program Head, Social and Engineering Systems

Information Design for Platform Drivers

by

Andreas Haupt

Submitted to the Institute for Data, Systems, and Society
on August 7, 2021, in partial fulfillment of the
requirements for the degree of
Master of Science in Social and Engineering Systems

Abstract

With the rise of app-based matching platforms, gig workers become important suppliers of labor to transportation and food delivery. An important dimension in the management of this labor supply are pricing decisions by platforms. This thesis makes a case for the importance of another dimension of platform design for platform drivers: Demand information given to drivers. Using a large-scale survey and platform-provided data, we give evidence for the relevance of demand information for drivers' labor supply decisions. We then start our investigation into how information should be designed, and derive policy recommendations for platforms and their regulators.

Thesis Supervisor: Jinhua Zhao
Title: Associate Professor

Acknowledgments

I thank my family (Helga and Katrin Haupt), my roommates (Isabell, Juan-Pablo, Sophia, Salome, Bernardo) and friends and mentors (Christina, Marieke, Richard) for the emotional support I needed in this time. I thank Stephen Zoepf (Lacuna.ai, Stanford University), Ryan Green (Gridwise), Dan Calacci (MIT Human Dynamics), Xiaotong Guo (MIT Joint Transportation Lab), and Hai Wang (Singapore Management University) for helpful conversations. I thank my reader John Horton (MIT Sloan) for his open feedback. Finally, I thank my supervisor Jinhua Zhao for his visions, and for being such a kind lab leader.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

> [Gig economy,] the collection of markets that match providers to consumers on a [job] basis in support of on-demand commerce. [...] Prospective clients request services through an Internet-based technological platform or smartphone application that allows them to search for providers or to specify jobs. Providers (gig workers) engaged by the on-demand company provide the requested service and are compensated for the jobs.
> – Congressional Information Service, TOCITE

1.1 Information Design for Platform Drivers

This thesis studies the relevance of demand information to platform drivers and how it should be designed.

By platform drivers we mean all food or grocery delivery drivers and ridesourcing drivers. While formally offering distinct services, in many urban spaces, platforms offering ridesourcing, so-called Transportation Network Companies (TNCs), also offer delivery of food or groceries. Examples of TNCs—which all offer also delivery services—are Grab and Gojek in Southeast Asian countries, Didi Yuching in China and several South American countries, as well as Uber in a large part of the world.

The Congressional Information Service's definition of gig work highlights the opportunity of search for consumers. However, it does not mention search, or any form of information acquisition by gig workers.

In most of gig work, and in particular in driving for platforms, gig workers need to make important decisions which require information or expectation on the demand for rides or deliveries: When to work, where to work, and for how long to work. Not only do these decisions have an impact on driver earning, but also on TNC operations.

Information that could affect decisions of gig workers to supply or not supply their labor often takes the form of demand information. For example, information on an area or a time period of high demand for rides can affect drivers' earnings.

This information becomes particularly crucial as platform drivers usually do not set prices for their services.

While information provision to platform driver influences decision-making of drivers, information provision becomes also important for other parts of the gig economy. For example, knowledge of demand and expectations have been demonstrated in the market for freelance software development TOCITEwhen the market environment changed, or in investments into AirBnB housing TOCITE.

This thesis focuses on the design of information to drivers. It highlights that merely considering *pricing*, i.e. the setting of prices to consumers and gig workers, might be beneficially complemented by considering *information provision*, both for TNC operations, a market for third-party information providers and regulation.

After positioning platform drivers in the spheres of urban transportation and the gig economy in the rest of this chapter, and reviewing important dimensions of platform design which interact with information design and information provision through third parties in chapter 2, we make three contributions.

The first part of our study, chapter 3 investigates which other factors besides expected earnings affect driver labor supply decisions. In a large survey of platform drivers in Jakarta, Indonesia, we find significant inconsistencies with earnings maximization of drivers in terms of reported earnings, and find cultural, as well as informational aspects in drivers' labor supply decisions. We also find that drivers often do not reposition themselves, but follow the algorithm whenever possible.

We supplement our observations on the relevance of non-pricing TNC design questions in chapter 4 using earnings differences in platform-provided data. We provide an approximate value of demand knowledge using earnings differences under optimal *repositioning*, i.e. movement of idle drivers to other areas in the city. We make this estimate by inferring real earnings for platform driver trips undertaken for TNC DiDi Yuching in November 2016 in Chengdu, China and comparing the earnings differences to average earnings of delivery drivers on the platform. We find that optimal repositioning, which can be seen as a very strong form of knowledge of demand, can lead to significant (at least two-fold) increase in earnings if a small group of drivers is perfectly informed about demand.

Third, we take first steps into the design of information for platform drivers in **??**. In a benchmark model we formalize properties of the informational environment. We first show that never providing the same demand information to all drivers is welfare-maximizing. We show that in our model, an informational Braess' paradox arises, which is distinct from existing notions of an informational Braess' paradox.[1] We expose that platform drivers might ignore information presented to them by the platform as a result of limited commitment of a TNC which does not allow them to credibly commit to not maximize their revenue with a choice of information. We show that third-party information aggregators can remedy this inefficiency, as can a change in the payment structure of gig drivers, and that aggregators can do so even with information provided to all subscribed drivers, as long as these are a small proportion of all drivers.

The final chapter 6 collects policy recommendations both for TNC operators and regulators. We first show that information design is of different significance for TNC operators and regulators. While the former will need to rely on third-party aggregators to commit to disclosing information in a way that drivers will not ignore, and are in this endeavour aligned with drivers, the latter should think more carefully about equity in the proposed models: Information needs to be provided asymmetrically. While asymmetrically, it still needs to require that drivers are equally treated. We

[1]Manxi Wu studies this in road networks; Manish Raghavan studies it for several classification technologies TOCITE.

conclude with engineering challenges arising from our observations.

While we provide several facets of information design for platform drivers, we largely abstract both pricing (who pays what) and matching (which driver gets to serve which rider) decisions and competition concerns away. Real-world platforms solve joint optimization problems of pricing and information in competition with other platforms. While this means that all of our analysis needs to be seen as having a narrow view of the design possibilities of platforms, we highlight, on the contrary, that viewing platform design purely as a *pricing* question might fall short of optimal mobility platform design.

Our analysis can also only be seen as mid-term, as information design to Autonomous Vehicles will depend on whether all vehicles are centrally controlled and cooperative or not. Throughout this thesis, we maintain that platform drivers are human.

1.2 Background

Platform drivers take an increasing role in urban mobility (subsection 1.2.1), and have special characteristics compared to other forms of gig work (subsection 1.2.2). Some of these characteristics have led to more prominent regulatory interventions into platform driving (subsection 1.2.3). Regulators have for the biggest part not intervened into information provision to drivers, which puts drivers into a status quo of an information environment given by their TNC apps and third parties aggregating information (subsection 1.2.4).

1.2.1 Platform Drivers and Urban Mobility

As of 2021, services provided by platform drivers now make up $^1/_3$ of the global taxi market TOCITE. With the rise of the first transportation network companies about 15 years ago, the rapid changes leaves many of the design dimensions of this market open.

An important difference to other travel modes such as public transit is the need for dynamic rebalancing, i.e. ensuring a distribution of drivers such that in each part of the urban area expected ride demand can be met at each point in time. While dynamic rebalancing is also an important question in public transit operations, schedules can be designed with much less uncertainty.

Part of the uncertainty making dynamic rebalancing hard comes from the need to design incentives that ensure participation. Even with accurate demand models for riders, platform drivers usually have the opportunity to log off their platform or reject (some of) the rides requested by a customer. If they are not incentivized to continue working on the platform, they will not.

Incentive designs for a supply side of mobility have historically been less prominent in transportation research TOCITE. This opens an opportunity for the introduction of tools used in competitive supply of labor.

1.2.2 Platform Drivers as Gig Workers

Platform drivers are also gig workers *qua* the definition introducing this chapter. Riders search on the platform and are matched to a driver for their ride or delivery. Nevertheless, platform drivers are extreme in some of their characteristics.

First, platform driving is, compared to other parts of the gig economy very flexible. This characteristic comes from the nature of gigs being short. Compared to other parts of the gig economy, e.g. lodging (e.g., on AirBnB, gigs usually at least one day), and services for care (e.g., care.com), technology (e.g., Andela), design (e.g., 99designs) and home services (e.g., Porch), typically require scheduling or completion within at least a day. This makes management of incentives compared to drivers

Second, the terms of the contract are fully determined by the platform. This is in contrast to other platforms such as in lodging, technology and design, where gig workers can set their own prices.

Third, reaching out to customers outside of the platform is much more challenging for platform drivers than for other gig workers. Compared to other gig work, giving a ride on the platform is short, and as both the complete contract and all payments are processed by the platform, it is hard for drivers to build gigs independently of the platform. This gives the platform additional bargaining power in negotiations with the driver.

Fourth, many of the drivers are full-time and work many hours a day. This both raises concerns over work safety, but also over wage bounds, some of which have been introduced as regulations.

1.2.3 Regulation of Platform Driver Work

Flexibility, limitations on contracting and communication, and the ubiquity non-wage controlled, full-time work led to policy interventions regarding minimum wages and the employment status of platform drivers.

The New York City Transport and Limousine Commission introduced an earnings standard which guaranteed a *proxy* minimum wage for ridesourcing drivers in New York City TOCITE. Drivers earn for rides an additional amount to increase the average as calibrated against historical demand data to the minimum wage level.

In 2020, California accepted via a referendum a special provision for ridesourcing drivers to be excluded from legal sufficient conditions for an employment relationship.

In contrast, the UK Supreme Court decided that ridesourcing drivers *are* employees of TNCs, with labor implications for drivers, with effects for payment and insurance of dirvers.

These regulations do not include provisions on the information that drivers get from TNCs, and other actors work in this sphere.

1.2.4 Platform Drivers' Payment and Information Environment

Most platforms drivers face an app on a smartphone, which offers them potential gigs, which they can accept or reject. At each point in time, drivers can select to log off the system, which drivers report to use strategically TOCITE. In the U.S. TOCITEand in our survey conducted in Jakarta (chapter 3), we find that a minority of drivers has several apps open at the same time. Hence, drivers are limited in their ability to observe demand on the other platforms.

This interacts also with payment design of platforms. Many platforms associate bonuses with the acceptance of rides, or punishments up to being blocked on the platform for not accepting ride requests. This makes being online in different platforms complicated for many drivers.

An important distinction will be three time periods, which we call, in accordance with nomenclature introduced by Uber TOCITE*phases 1, 2, and 3.*

Phase 1 corresponds to times where the driver is online but idle. In these times, the driver is not sent ride requests and does (except for bonus payments for repositioning) not earn money.

Phases 2 and 3 refer to the time between the acceptance of a ride request and the pickup of the passenger, and the ride, respectively. In these phases, riders are paid.

The goal of this thesis is to establish the relevance of providing drivers with demand information, and will only significantly affect driver behavior in phase 1. Drivers that are picking up or driving a rider have a clear task, and even a contract for their work. The main question we ask in our theoretical section will be when the platform can transmit any information that will not be ignored.

Chapter 2

Related Work

Our work at the intersection of the Labor Economics in the Gig Economy, Supply Management in Transportation, Information Design, and information gathering from platforms contributes to several literatures.

2.1 Estimating Labour Supply Through Surveys

Our contribution first contribution, a survey to platform drivers in Jakarta, Indonesia, uses a survey to study the labor supply decisions of drivers, and does so using a survey.

2.1.1 Models of Platform Driver Labour Supply

Labor supply choices of drivers, or more generally participation decisions in two-sided platforms, is assumed as exogenous, or rational expectations of drivers is assumed.

Papers showing the impact of multi-homing use either stylized models such as those inspired by Hotelling lines [**bryan2019theory**, **bakos2020platform**], assume exogenous models of [**castillo2017surge**] or do not model multi-homing decisions [**liu2017impact**].

Other papers assume that in moments where agents join a platform, they have exact knowledge of the surplus that is gained and the prices that both sides receive [**liu2019multihoming**].

[**Wang_Yang_2019**] offers a comprehensive review on the ride-hailing system including demand, supply and platform perspectives. Driver behavior is modeled differently in the ride-hailing literature. Regarding operations and optimizations for ride-hailing platforms, e.g., driver-customer matching, idle vehicle rebalancing and vehicle routing, drivers are assumed to follow the instructions of platforms ([**Bertsimas_Jaillet_Martin_20** **Alonso-Mora_2017**, **Braverman_Dai_Liu_Ying_2019**, **Wen_Zhao_Jaillet_2017**]).

Simple driver behavior models are proposed in the pricing literature of the ride-hailing system. [**Bai_So_Tang_Chen_Wang_2019**] propose a queuing model to optimize platforms' profit while considering price-sensitive customers and earning-sensitive drivers. Each driver has a reservation earning rate based on his outside option and the driver will provide services only if earning rate exceeds the reserved value. [**Taylor_2018**] evaluate the impact of uncertainty in passenger delay sensitivity and driver independence on platforms' optimal per-service price and wage. Each driver has an opportunity cost and the driver will participate in the platform when having non-negative expected utility.

To better understand driver behavior, several studies discuss the general reasons why drivers work in the ride-hailing system. An econometric framework with closed-form measures is proposed by [**Sun_Wang_Wan_2019**] to estimate both the participation elasticity and working-hour elasticity for drivers. In particular, this article assumes that there [**Chaudhari_Byers_Terzi_2018**] design an earnings-maximization strategy for drivers and show that strategic behavior about when and where to drive have significant impact on drivers' earnings. [**Jiang_Kong_Zhang_2021**] show that regret aversion and ignorance of suggestion are the two major behavioral factors which influence drivers' re-positioning decisions.

2.1.2 Labour Economics and Industrial Economics

Our analysis in chapter 3 pertains to

2.2 Investigations into Platform Pricing

To understand platform pricing, many platforms

2.2.1 Scraping-Based Studies

Some platforms scrape data from platforms, such as studies on platform competition in New York TOCITE(rosaia) and a consumer surplus comparison of taxi and ridesourcing TOCITE(shapiro). Similarly, TOCITE(peaking under the hood) uses large-scale simulation of apps.

2.2.2 Information from Platforms

2.2.3 Gathering Driver Data

2.3 Driver Repositioning

The repositioning of drivers, also known as Dynamic Rebalancing in the literature on platform design.

2.4 Cheap Talk and Information Design

Our third contribution contributes to a line of work on cheap talk and information design.

2.4.1 Cheap Talk

The standard model of Crawford-Sobel TOCITE, in which one agent, which has a piece of information, communicates with another agent, who can take an action affecting both agents' utility, too strong mismatch of incentives leads to no communication. A classical model of cheap talk, communication is costless, and does neither binds sender nor receiver to take a particular action in the future TOCITE(Farrell). In our first model of information provision, we model the strategic interaction between the platform and drivers as a cheap talk game, and show that if the incentives of the platform and drivers diverge sufficiently, that, in equilibrium, no information is transmitted, in terms of the cheap talk literature, the platform *babbles*.

2.4.2 Bayesian Persuasion and Information Design

In contrast to cheap talk, Information Design models allow the sender to commit to rules for which information they reveal TOCITE(kamenica gentzkow, also annual reviews). This can have profound implications for the outcome o

2.4.3 Information and Mechanism Co-Design

While

Chapter 3

Non-Monetary Shapers of Supply Decisions

> Since Uber's introduction in 2009, ridesharing platforms like Uber, Didi, Grab, and Lyft have radically transformed the taxi and limo industry. These services, which allow consumers to order a car to their location via a smartphone application, now control roughly 1/3 of the international taxi market. In other words, a ridesharing firm acts as a platform matching drivers to riders and setting the pricing terms between them. —TOCITE

This chapter presents significant frictions in a market for urban mobility in Jakarta, Indonesia. We use quantitative results to present significant deviation from driver earnings maximization, and code free-text answers to gather evidence for reasons outside of earnings maximization.

3.1 Survey

We analyze data from a survey conducted in spring 2021 in Jakarta, Indonesia.

3.1.1 Sample

Participants were asked to complete an online survey hosted on Qualtrics and received 846 complete survey responses after being distributed and circulated to drivers through known social media (WhatsApp groups) used by drivers.

Figure 3-2a shows the age distribution of participants, where around 50% of drivers are within the age group of 30 to 39. Figure 3-2b reveals that over 97% of drivers who participated in the survey are male and around 97% are motorbike drivers.

There exist three major ride-hailing companies in Jakarta: Grab, GoJek and Shopee. While Grab and Gojek are incumbents in this market, Shoppee entered the market only in 2021. Figure 3-1 shows the percentage of drivers working for each of the platforms, double-counting multi-homers. Around 94% of drivers work for the incumbents Grab or GoJek. In our sample of drivers, only 12% of drivers questioned in the survey are multi-homers.

We are not aware of demographics of TNC drivers in Jakarta and hence unable to verify representativity of our sample.

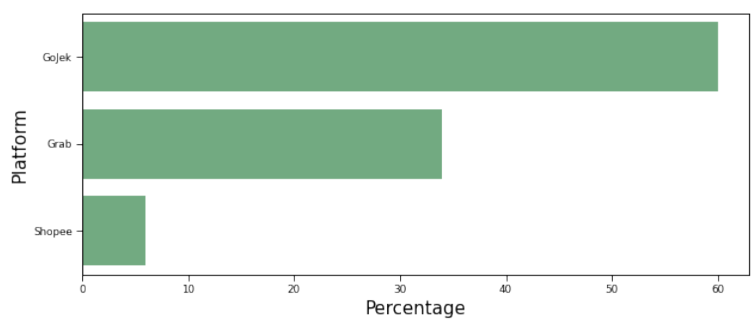
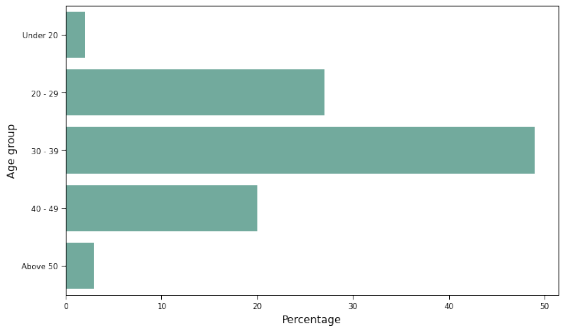


Figure 3-1: Platform distribution of drivers. (GoJek: 60%, Grab: 34%, Shopee: 6%)

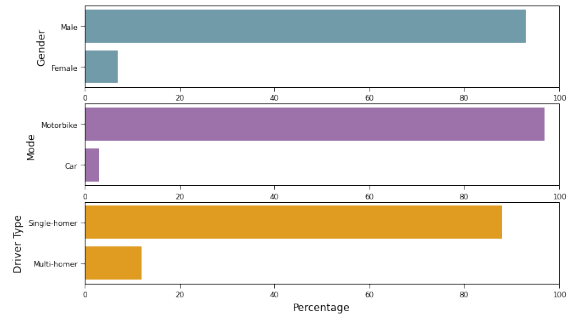
3.1.2 Questions

The survey consists of multiple questions blocks and is given in its entirety in Appendix **??**.

In a first block, participants were asked to check all TNCs they have been working for in the past 30 days. A survey respondent is defined as a *multi-homer* if they have



(a) Age distribution of drivers. (Below 20: 2%, 20 to 29: 27%, 30 to 39: 49%, 40 to 49: 20%, above 50: 3%)



(b) Gender, mode and driver type distributions of drivers. (Male: 97%, female: 3%; motorbike: 97%, car: 3%; single-homer: 88%, multi-homer: 12%)

Figure 3-2: Sociodemographic information and driving behaviors of survey participants.

been working for at least two platforms in the past 30 days. In this first block, basic information and driving behavior are asked as well, e.g., vehicle type used, previous occupation, behaviors while waiting for a next ride/order.

The second block depends on whether respondents were identified in the first block as multi-homers or not. Multi-homers are asked how they switch between different platforms, allocations of their working hours and reasons for multi-homing. For non-multi-homers, reasons for non-multi-homing and factors that would make them multi-home are asked.

In a third block, for each ride-hailing platform survey participants have worked for in the last 30 days, we asked a number of questions about driving activities during the last two weeks. These questions included whether drivers mostly worked on delivering food or transporting passengers, their most frequent driving area, days and hours to work, total driving distance, and daily salaries.

A final block elicited socio-demographic characteristics including age, gender, educational status, income level, living areas, and weekly expenses.

Table 3.1: Results of regression models predicting multi-homing behaviors of ride-hailing drivers in Jakarta based on sociodemographic information

Variable	coef	std err	t	P> |t|	[0.025	0.975]
Intercept	0.3168***	0.052	6.040	0.000	0.214	0.420
Age (30 - 39)	-0.0455*	0.025	-1.786	0.075	-0.095	0.005
Age (above 50)	-0.1189*	0.069	-1.721	0.086	-0.255	0.017
Male	-0.1213***	0.042	-2.911	0.004	-0.203	-0.040
Degree (bachelor's or higher)	0.1258***	0.044	2.840	0.005	0.039	0.213
Degree (junior high school)	0.0327	0.035	0.939	0.348	-0.036	0.101
Ride-hailing as main source of income	-0.0733*	0.038	-1.925	0.055	-0.148	0.001
Age (40 - 49) and ride-hailing as main source of income	-0.0480	0.033	-1.474	0.141	-0.112	0.016
R-squared:	0.033	**Adjusted R-squared:**			0.025	

Table 3.2: **Results of regression models predicting multi-homing behaviors of ride-hailing drivers in Jakarta based on previous occupation**

Variable	coef	std err	t	P> |t|	[0.025	0.975]
Intercept	0.1276***	0.012	10.735	0.000	0.104	0.151
Motorcycle taxi driver before	-0.0796**	0.031	-2.574	0.010	-0.140	-0.019
R-squared:	0.008	**Adjusted R-squared:**			0.007	

3.2 Findings

To understand the multi-homing behavior and who works for multiple ride-hailing platforms, we first constructed a Lasso regression model with the regularization coefficient 0.001 based on drivers' socio-demographic information, which were shown in Table 3.1.We find that male drivers whose age is between 30 to 39 and above 50 are less likely to multi-home in Jakarta market. Also, drivers are less likely to multi-home if working for ride-hailing platforms is the main source of their income. On the other hand, we find that drivers with higher degrees are significantly more likely to multi-home. Next, we observe the correlation between previous work as a motorcycle taxi driver before on multi-homing behaviors. The OLS regression model shown in Table 3.2 suggests that drivers who has been motorcycle taxi drivers before are less likely to multi-home in Jakarta market.

Table 3.3: **Results of regression models predicting multi-homing behaviors of ride-hailing drivers in Jakarta based on whether being part-time ride-hailing drivers**

Variable	**coef**	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
Intercept	0.1097***	0.012	9.258	0.000	0.086	0.133
Being part-time ride-hailing drivers	0.0441	0.032	1.384	0.167	-0.018	0.107
R-squared:	0.002	**Adjusted R-squared:**			0.001	

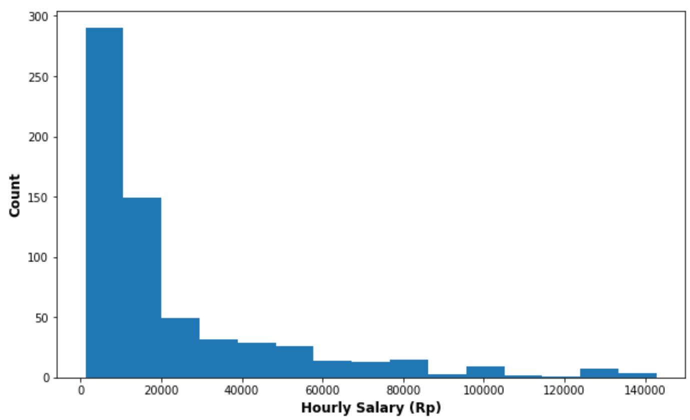


Figure 3-3: Hourly salary distribution of survey respondents (ride-hailing drivers in Jakarta)

Finally, we turn our attention to the results of our regression models exploring the impact of being a part-time ride-hailing driver on drivers' multihoming behaviors. From the model results shown in Table 3.3, we find that part-time ride-hailing drivers are more likely to multihome in our sample. However, it is a less significant predictor of multihoming behaviors compared to other variables above.

3.2.1 Earnings for Different Platforms

Then, we consider the Jakarta ride-hailing market and drivers' hourly salary and total working hours. Figure 3-3 displays the hourly salary distribution of all survey respondents in Jakarta, which fits into a negative binomial distribution. Most ride-hailing drivers earn less than 20,000 Rp (approximately 1.4 USD) per hour, which is significantly less than the average hourly salary 120,175 Rp (approximately 8.4 USD) in Jakarta [**Jakarta_hourly_salary**].

Table 3.4: **Results of regression models predicting hourly salary of ride-hailing drivers in Jakarta based on whether working for GoJek**

Variable	**coef**	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
Intercept	30710	2215.544	13.863	0.000	26400	35100
Only GoJek	-11000	2547.137	-4.319	0.000	-16000	-6009.224
R-squared:	0.038	**Adjusted R-squared:**			0.036	

Table 3.5: **Results of regression models predicting daily working hours of ride-hailing drivers in Jakarta based on whether working for GoJek**

Variable	**coef**	**std err**	**t**	**P> |t|**	**[0.025**	**0.975]**
Intercept	10.6642***	0.162	65.928	0.000	10.347	10.982
Only GoJek	1.0335***	0.210	4.911	0.000	0.620	1.447
R-squared:	0.036	**Adjusted R-squared:**			0.034	

To understand whether drivers working for different ride-hailing platforms are payed differently, we constructed a regression model on the impact of working for GoJek on hourly salary, shown in Table 3.4. In this model, we only consider motorcycle drivers who work exclusively for one platform. The model results suggest that GoJek drivers earn significantly less than drivers from other platforms, 11,000 Rp (approximately 0.77 USD) less per hour.

Furthermore, we built another regression model to investigate the difference of total working hours between ride-hailing platforms. The model results in Table 3.5 show that GoJek drivers work significantly longer than drivers from other platforms, 1.0335 hours more per day. Both regression results imply that there existing a large switching cost between GoJek and other platforms (Grab and Shopee) in Jakarta ride-hailing market. GoJek motorcycle drivers work longer hours but earn less compared to other drivers.

3.2.2 Free-Text

In this section, we use a topic model (Latent Dirichlet Analysis) to discover potential non-monetary reasons for the choice of TNC that drivers work for. All of the analysis is use answers to question 16 (compare Appendix **??**) "Anything else we should know about why you decided not to use multiple platforms?".

In a topic analysis, the word "focus" is highly descriptive of one of the clusters. Respondents indicate that "focus was needed for best results" and that they would not like to "confuse the apps".

Others refer to the character of one of the platforms as being an Indonesian incumbent Gojek that competes with an (in our sample) better-paying incumbent from Singapore. Respondents make reference to "children of the nation" to underline their choices.

We use survey evidence showing that in an a significant share of non-multi homing drivers in an oligopolistic market for ridehailing in Jakarta. Our findings hint at substantial frictions to multi-homing, non-monetary incentives in the decision what TNC drivers choose to drive for, and when to drive.

3.2.3 Discussion

As our evidence can only explain a small part of the variation in regression models, we propose other variables that could explain more of the variation in multi-homing and pay.

Experience Respondents mention *focus* prominently in reasons for multi-homing. To disambiguate whether this focus refers to learning of the platform or on permanent presence on the platform, joining date of the platform could be used as a proxy for experience.

Risk Preferences Another important explanation of variation in earnings could be given by the risk preferences of drivers. High variation in pay could stem from different behavior and different willingness to drive in times where getting a job from the platform are uncertain. Existing short surveys [**falk2016preference**] allow to estimate risk preferences.

Chapter 4

The Value of Information

After our investigation into potential multi-homing frictions in a market, we give evidence here on the value of information to drivers. This is aligned with claims by companies that offer information services to drivers TOCITE.

4.1 Description of Data

The DiDi KDD Challenge 2020, sponsored by DiDi Yuching, consisted of two challenge. An order dispatch challenge, and a driver repositioning challenge. We will work with data and submissions of the latter.

4.1.1 Annotated Order Data

We are given for November 2016 about 10 Million orders. An order consists of an origin-destination pair, a timestamp for beginning and end, an identifier for the driver that took the order, and a number of *reward units*. We argue below that reward units are a constant factor and exemplify this with a good fit of our model in **??**.

4.1.2 Reinforcement Learning Challenge Leaderboard

The KDD 2020 reinforcement learning challenge solves an optimal repositioning problem together with an optimal order dispatch problem. The problem is a Markov

Dep. Variable:	reward_units	**R-squared:**	0.887
Model:	OLS	**Adj. R-squared:**	0.887
Method:	Least Squares	**F-statistic:**	1380.
Date:	Sun, 18 Jul 2021	**Prob (F-statistic):**	0.00
Time:	14:45:52	**Log-Likelihood:**	-2998.9
No. Observations:	2000	**AIC:**	6004.
Df Residuals:	1997	**BIC:**	6021.
Df Model:	2		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
Intercept	-0.1664	0.077	-2.155	0.031	-0.318	-0.015
distance	0.3191	0.021	15.449	0.000	0.279	0.360
duration	0.0635	0.007	9.645	0.000	0.051	0.076

Decision Process, which is partially observable for the repositioning challenge. In Reinforcement Learning language, which we review in **??**

We also use the leaderboard's reinforcement challenge leaderboard. It is reproduced from

4.1.3 Earnings Table

4.2 Estimation

4.2.1 Model

We assume that there is a constant c such that

$$\text{rewardunits} = c\,\text{RMB}\,.$$

The reason for this lies in the following proposition.

4.2.2 Regression

4.2.3 Checks

4.3 Discussion

	coef	std err	z	P> |z|	[0.025	0.975]
Intercept	-0.2569	0.262	-0.981	0.326	-0.770	0.256
rush_hour[T.True]	0.0095	0.320	0.030	0.976	-0.618	0.637
distance	0.3318	0.047	7.113	0.000	0.240	0.423
rush_hour[T.True]:distance	-0.0292	0.048	-0.602	0.547	-0.124	0.066
duration	0.0653	0.011	6.170	0.000	0.045	0.086
rush_hour[T.True]:duration	0.0028	0.014	0.200	0.841	-0.024	0.030

Chapter 5

Information Design

> The information design problem has a literal interpretation: there really is an information designer (or mediator, or sender) who can commit to provide extra information to players to serve her own interests. While the commitment assumption may be problematic in many settings, it provides a useful benchmark.–TOCITE

In this section, we study pure information design for platform drivers. We refer to it as *pure* as we abstract from its interaction with pay.

5.1 Model

In our stylized model, two agents simultaneously decide to drive to a part of a city or to do something else. Before each of the drivers makes her decision on whether to drive, the system can give them a message m on whether there is demand on the platform.

We assume two ways in which the platform can communicate the information to the drivers.

		Driver 2	
		go	not go
Driver 1	go	$\frac{1}{2}$, $\frac{1}{2}$	1, 0
	not go	0, 1	0, 0

Figure 5-1: Payoff Matrix for the two-driver game.

5.1.1 Cheap Talk

In cheap talk, the platform discloses some information to a driver, which, given the update, updates their behavior. More concretely, there is an abstract set of *messages* $m \in \mathcal{M}$ that the platform can send. The players receive the message and play optimally. In particular, the timeline of the game is:

1. The demand at the area is realized.

2. The platform decides to send a message to the drivers.

3. The drivers decide to drive to the area or not.

5.1.2 Information Design

In information design, the drivers can commit to a driver

5.1.3 Public and Private Messages

5.2 Inefficiencies

In this environment, two sources of in

5.2.1 Inefficiencies through lack of commitment

5.2.2 Inefficiencies through public information

5.3 Remedies

Given the inefficiencies observed in the last section, some remedies for the market should be considered. We propose that

5.3.1 Commitment via third parties

5.3.2 Commitment via reputation

5.3.3 Approximate Efficiency of Public Information to Few Drivers

Chapter 6

Conclusion and Policy Implications

This thesis studied the relevance and optimal design of information provided to ridesourcing drivers. chapter 3 showed that besides earnings maximization, drivers also take into account cultural and infromational considerations in their labor supply decisions. chapter 4 showed that there are high potential earnings gains from additional information. We then went on to study commitment and privacy of messages as two challenges. In this last chapter, we consider three stakeholder groups, platforms, regulators and transportation engineers with policy recommendations and additional complications arising out of this work.

6.1 The Commitment Challenge for Platforms

We observed in chapter 5 that the limited commitment of the platform to providing different platform can lead to not information being transmitted and we proposed reputation

6.2 The Fairness Challenge for Transportation Regulators

6.3 Challenges for Transportation Engineers

The above two challenges give interesting challenges for transportation engineers.

6.3.1 Allowing for Commitment without Harming Businesses

6.3.2 Managing Fairness in Information Provision