Integrating Empirical Analysis into Analytical Framework: An Integrated Model Structure for On-Demand Transportation

Yuliu Su[1], Ying Xu[1(✉)], Costas Courcoubetis[1], and Shih-Fen Cheng[2]

[1] Engineering Systems and Design, Singapore University of Technology and Design,
Singapore 486830, Singapore
`xu_ying@sutd.edu.sg`

[2] School of Computing and Information Systems, Singapore Management University,
Singapore 178902, Singapore

Abstract. On-demand transportation services have been developing in an irresistible trend since their first launch in public. These services not only transform the urban mobility landscape, but also profoundly change individuals' travel behavior and demand for cars. In this paper, we propose an integrated model structure which integrates empirical analysis into a discrete choice based analytical framework to investigate a heterogenous population's choices on transportation mode and car ownership with the presence of ride-hailing. Distinguished from traditional discrete choice models where individuals' choices are only affected by exogenous variables and are independent of other individuals' choices, our model extends to capture the endogeneity of supply demand imbalance between ride-hailing service providers and users. Through equilibrium searching and counterfactual analysis, we further quantify the magnitude of impacts of platform operations and government policies on car demand, usage and traffic conditions. The structure of the model and managerial insights are explained in detail.

Keywords: Ride-hailing · Peer-to-peer sharing · On-demand platforms · Mobility as a service

1 Introduction

In the past decade, with the ubiquity of connectivity and proliferation of smart phones, we have witnessed the emergence of peer-to-peer ride-hailing platforms such as Uber, Grab and Go-Jek [1]. While this new form of peer-to-peer service has disruptively changed the on-demand mobility sector, discussions about its impacts on our society and environment have been inconclusive. On the one hand, this new form of peer-to-peer sharing has brought improvement on social welfare, in terms of creating more employment opportunities, enabling individuals to earn additional income [2, 3], encouraging shared rides,

eliminating wasteful driving, and thus reducing traffic and vehicle emissions [4–6]. On the other hand, criticism has also arisen for these platforms' quest for profits and market shares by diverting non-driving trips to driving mode [7], which leads to unnecessary driving demand and excess supply. For example, a survey in US shows that in the absence of ride-hailing, about 49% to 61% of ride-hailing trips would not have taken place at all or would have been made by other modes such as walking and public transportation [8]. Moreover, we find that the impacts of peer-to-peer ride-hailing platforms on car ownership are also two-fold. On the one hand, ride-hailing may have a cannibalization effect on demand for cars [9], reducing individuals' incentive to own a car. On the other hand, ride-hailing may lead to a value enhancement effect on car ownership [9], adding more valuations with purchasing a car [10, 11]. For example, a study on new vehicle registrations in China shows that Uber entry is significantly associated with an 8% increase in car ownership [12].

A key feature leading to the above unresolved debates is that, different from traditional taxi service, in the peer-to-peer ride-hailing sector the drivers are flexible individuals. As a result, ride-hailing platform could not fully control the capacity of their service supply. Moreover, since individuals could switch between the supply side and demand side (i.e., being a driver or a rider) [13], the interaction adds more challenges for the platforms to balance supply and demand, not to mention to optimize the utilization of the capacity. Hence, to study the impacts of ride-hailing, it is critical to understand how individuals' decisions on being a driver or a rider for the platforms are affected, and these decisions are also coupled with other decisions on travel model as well as car ownership and usage. Although there are a few attempts in investigating the mechanism of peer-to-peer markets, most of them focus on product sharing (e.g., car renting) and adopt an analytical framework (e.g., [9, 14] and [15]). As these analytical studies are usually formatted as an equilibrium searching problem through analytical calculation or simulation, they may be challenged by its simplifications and lack of validation by empirical data in practice.

To the best of our knowledge, empirical studies focusing on peer-to-peer on-demand service platforms are relatively limited. To enrich this research stream, this paper introduces an integrated model structure which connects theoretical models with empirical analysis to investigate a heterogenous population's choices on car ownership, usage and transportation mode with the presence of ride-hailing. Specifically, we first develop a discrete choice framework which captures heterogenous individuals' choices among different strategies facing ride-hailing. They decide whether to purchase a personal car, how to use the car if s/he owns one (either serve for the platform or solely use for personal need), and whether to use the platform service or not when needed. These strategies are affected by various factors such as wages, working hours, price sensitivity, perception of traveling in a private car compared to public transportation, and benefits and costs of owning a personal car, etc., which are all included in our model. Moreover, noticing that in the two-sided on-demand platform context, each individual's choice depends on the dynamics between supply and demand, which is a collective result of the whole population's decision. We also incorporate such endogeneity of the supply and demand imbalance and connect the two sides through a matching function, which is also embedded into the choice functions. We then empirically calibrate our model using publicly

disclosed data in the Singapore ride-hailing marketplace. Based on our model, we further conduct counterfactual analysis to quantify the magnitude of impacts of platform operations and government policies on ownership and traffic, which yields empirical assessment of the peer-to-peer transportation service platform.

The principal contributions of our paper are summarized as follow.

a. We introduce an integrated model structure which integrates empirical analysis into a discrete choice-based equilibrium framework. As in the two-sided on-demand platform context, each individual's choice depends on the dynamics between supply and demand, which is a collective result of the whole population's decision, endogeneity of the supply and demand balancing rises as a challenge (for example, more platform drivers lead to higher chance and shorter waiting time for riders to grab a car, and thus higher the utility for riders). Commonly used choice models are limited to capture such endogeneity as they consider the variables affecting individuals' choices as exogenous and the population's choices are independent of each other's. Distinguished from the traditional choice models, our framework overcomes this chal1lenge by incorporating the endogeneity of supply and demand imbalance through a matching function and formalizing individuals' choices as a game equilibrium as a function of others' choices. Moreover, although our model is developed with a focus of peer-to-peer ride-hailing, it can easily be generalized to other peer-to-peer markets including both product sharing and on-demand service platform.

b. Our integrated model demonstrates flexibility with data scarcity. A common challenge for empirical studies is the lack of access to data. For example, to develop a discrete choice model, large volume of individual-level choice data are usually required, which are commonly collected through large scale interview or discrete choice experiment survey. The data collection could be costly and time-consuming, and may also incur privacy issues (as the survey involves questions about individuals' demographic information such as wage, traveling needs and mode preference). Conversely, our model demonstrates the flexibility of parameters being estimated with sparse aggregate-level data. We use publicly disclosed data on distribution of car ownership and usage in the Singapore ride-hailing marketplace to calibrate our model parameters. Our approach sheds lights on applying scarce data on empirical settings.

c. Our analyses inform the debate on how ride-hailing may affect ownership, usage and traffic with the sophisticated consideration of supply demand matching propensity. The main findings suggest that ride-hailing may demonstrate both cannibalization effect and value enhancement effect on car ownership depending on different conditions. The cannibalization effect of ride-hailing on own car ownership is mainly due to its lower price compared to conventional taxis. In terms of individuals' decision of whether and how to participate in the peer-to-peer ride-hailing service, their choices are affected not only by ride-hailing price or driving cost but also the dynamics between supply and demand imbalance. For example, we observe that although the increase of service price increases platform drivers' earnings, it in fact leads to fewer platform drivers due to a diminishing user base, which lowers demand and matching propensity. We further investigate the relationship between platform revenue and

ride-hailing payment. Our analysis suggests that due to the supply demand dynamics, the platform's revenue does not increase monotonically with service charges; there exits an optimal price level which allows the platform to maximize its revenue.

The remainder of this paper is organized as follows. Section 2 reviews related literature in peer-to-peer sharing. In Sect. 3 we introduce the framework of our integrated model in detail. Section 4 specifies the data we used for our empirical setting and the calibration process of our model parameters. In Sect. 5 we present the equilibrium analysis under various counterfactual scenarios. Section 6 concludes the work with a discussion of the findings and future research directions.

2   Literature Review

Peer-to-peer sharing has attracted growing research interest in recent years. Studies related with peer-to-peer sharing can be broadly classified into two categories with respect to their study focuses. The first category focuses on peer-to-peer resource/product sharing, which allows owners to rent out their products/assets to non-owners for a short term (e.g., Airbnb for accommodation sharing and Zipcar for car renting). Most of the papers in this group study the impacts of product sharing on individuals' decisions on product ownership from different perspectives such as product manufacturer (e.g., [9, 16] and [14]) or the platform owner (e.g., [15]). The second category focuses on peer-to-peer on-demand service, which connects users requiring a time-sensitive service with independent service providers (e.g., Uber for transportation service and Deliveroo for food delivery service). Most papers in this category investigate how the interactions between the supply and demand in the two-sided market are affected by different factors, such as [17] by the factor of dynamic pricing, [18] by customers' sensitivity to delay and service providers' independence, [19] by payout ratio (i.e., the ratio of wage paid to service providers over price charged from consumers) and [20] by service providers' self-scheduling. In addition, a few other studies investigate the effectiveness of different contract designs for stakeholders of the on-demand service platform such as [21] and [22]. Some other studies adopt a game theoretic framework to examine the organization of the on-demand service such as [23] and [24]. The early works in peer-to-peer sharing are dominated by theoretical arguments that construct analytical framework to explain and investigate the mechanism of and agents' interactions in the innovative marketplace.

Although previous analytical studies provide the first insights about individuals' choices with the presence of sharing economy, the majority of them has focused on system equilibrium and/or welfare of service providers and customers, which are obtained through analytical calculation or simulation and are not validated by empirical data in practice [25]. The empirical studies on on-demand platforms are relatively limited, mostly focusing on using statistical regression models to study the impact of sharing economy on ownership or traffic congestion. For example, [7, 12] and [3] use longitudinal data and difference-in-difference models to investigate the influence of the entry of Uber on traffic congestion, new vehicle ownership and rates of entrepreneurial activities respectively. [26] structure a before-and-after assessment and employ a fixed-effects panel data regression model to analyze the impact of ride-hailing on traffic congestion

in a major city. [25] use a probit regression estimated by a modified Heckman two-stage method to study the impact of financial incentive on platform drivers' decision of whether to work and corresponding number of work hours. However, a key shortcoming of regression methods such as difference-in-difference is that their statistical results do not always imply an unbiased estimate of the causal effect as mentioned in [27].

It is worth noting that despite the growing abundance of literature in sharing economy, many controversial questions still remain and need to be answered. Given the fact that analytical work would be challenged by the simplifications and that existing empirical studies fail to effectively explain the causation and have limited prediction power, there are calls for research integrating analytical methodologies with empirical methodologies. Such action could improve the validity as well as the explanatory power and prediction capability of the model, which could provide in-depth managerial insights. So far, we have noticed one paper [28] attempting this integrated methodology. The authors first introduce a dynamic model and then use empirical data to calibrate the model. However, their focus is the peer-to-peer durable goods sharing market; and they do not consider the matching rates as a dynamic result of supply and demand but exogenously specified. As far as we know, no studies have yet attempted to analyze on-demand service platforms, in particular the ride-hailing platform, by using such integrated methodology. Motivated by the lack of relevant research, in this paper we propose a model structure that integrates empirical analysis into analytical framework with the consideration of endogenous supply demand matching to study the peer-to-peer on-demand service platform.

3   The Model

3.1   Agents' Strategies and Payoffs

Our study objects (agents) are a working age population consisting of individuals who are eligible for driving, below retirement age, and most importantly, have a regular travel need. To satisfy their travel needs, individuals may choose to own a personal car to satisfy the travel need. Once with a private car, with the presence of ride-hailing, the individual also has an opportunity to choose to work for the platform as a driver (either full-time or part-time). In contrast, without a personal car, the individual could decide between using the ride-hailing service or other transportation service such as public transportation. Accordingly, we construct a set of five strategies $\Sigma = \{N, P, F, U, A\}$ for the individuals to choose from.

- [N] Non-platform driver: An individual who owns a personal car but only uses the car to satisfy her own travel needs.
- [P] Part-time platform driver: An individual who owns a personal car and has a regular full-time job. She drives for the platform during after-work time to earn extra income.
- [F] Full-time platform driver: An individual who owns a personal car and forgoes the regular job with wage w to work full-time on the platform.
- [U] User: An individual who does not own a personal car but prefers to satisfy her travel needs by using the ride-hailing service.

- [A] Abstinent: An individual who does not own a personal car and whenever she needs a ride, she chooses public transportation.

Each individual chooses its strategy based on its expected long term payoffs, which are characterized as the average utility associated with each strategy per period. The individuals' travel needs is denoted by M_i per period. Due to the comfort and convenience, traveling in private cars generates positive payoffs compared with taking public transportation. We normalize the utility of taking public transportation to zero, and use γ_i to indicate individual i's relative utility gain of traveling in private cars per kilometer. Correspondingly, strategies N, P, F, U enjoy a utility gain of $\gamma_i M_i$ compared with strategy A. For strategies N, P, F, a fixed car ownership cost[1] k and a variable cost c (such as petrol and toll) will incur. Considering the convenience to access the car, strategies N, P, F also enjoy a convenience benefit of b_N, b_P and b_F respectively. Furthermore, for simplicity, we assume that all part-time and full-time platform drivers work the same amount of time t_P and t_F respectively to provide the ride-hailing service. More, all individuals but the full-time platform drivers (strategy F) are assumed to own a regular full-time job with wage w_i per period. Each individual is further characterized by its price sensitivity $\theta_i \geq 0$.

Before introducing the payoff functions, some characteristics of the ride-hailing market are worth mentioning. In the two-sided ride-hailing market, the matching between the drivers and riders could not be perfect due to spatial and temporal constraints. As a result, a platform driver could only be matched with a rider with a proportion of $\alpha \in (0, 1)$ out of the driver's total working time, while a platform user's request could only be successfully fulfilled with probability $\beta \in (0, 1)$. With the chance of $1 - \beta$, the platform user must resort to public transportation, which is assumed to be always available.

Next we specify the payoff functions for each strategy of individual i. The subscript i in the equations indicates that the corresponding variables are individual specific, variables with strategy subscript (such as b_N) are assumed to be homogeneous for individuals in the same strategy group ("N"), and variables without subscript are common across the whole population.

- [N] A non-platform driver has a regular job with wage w per period. She possesses a personal car and uses it to satisfy her own travel needs M_i and thus gain a payoff of $\gamma_i M_i$ from traveling in a private car. She also pays for the fixed ownership cost k and variable cost c per kilometer. Price sensitivity θ_i converts monetary value to utility. She also earns a convenience benefit b_N from owning a car. In this case the payoff is

$$V_{N,i} = \theta_i w_i + \gamma_i M_i - \theta_i (cM_i + k) + b_N.$$

- [P] A part-time platform driver possesses a car and has a regular job as a non-platform driver. However, she uses the car not only to satisfy her own travel need M_i but also provide ride-hailing service on the platform after this regular job. The term

---

[1] In order to participate in ride-hailing, platform drivers are required to possess a private car. They could either purchase a personal car or rent a car from the platform. In our model formulation we do not distinguish between these two ways as essentially they both represent a fixed ownership cost k.

$\theta_i t_P s(\alpha\rho - c) - t_P l$ represents the additional benefit from the part-time ride-hailing service. t_P is the time she works as a part-time ride-hailing driver while s is the average driving speed. The product of t_P and s converts the working hours to total kilometers supplied on the platform. Due to market friction and supply-demand imbalance, out of the $t_P s$ kilometers she supplies for the platform, only α fraction is effectively used by serving a passenger, generating an earning of $t_P s \alpha \rho$, where ρ is the per kilometer earnings from providing the service. We assume that while providing the service, the driver is always driving around with or without a traveler in her car as she may be driving towards higher demand areas to search for new customers. Therefore, she always incurs a variable driving cost c during time period t_P. As a result, the net profit from part-time ride-hailing service is $t_P s(\alpha\rho - c)$. Again, θ_i converts monetary profit to utility. Also, as the part-time platform driver works overtime with consideration of her regular job, she incurs a loss of leisure time utility l per time unit during her working time t_P on the platform. The above modeling assumptions imply the following payoff:

$$V_{P,i} = \theta_i w_i + \gamma_i M_i - \theta_i(cM_i + k) + b_P + \theta_i t_P s(\alpha\rho - c) - t_P l.$$

- [F] A full-time platform driver gives up her regular job and takes providing ride-hailing service as her full-time job. Therefore, she forgoes the wage w from her previous regular job (hence $\theta_i w_i \times 0$). Similar to a part-time platform driver, the net income from being a full-time platform driver is expressed by $t_F s(\alpha\rho - c)$. The average working time of the full-time platform driver's previous regular job is denoted by λ. The full-time platform driver may choose to work overtime on the platform. The term $(t_F - \lambda)_+$ is an indicator of whether the full-time platform driver incurs a utility loss of leisure time. If total working time on the platform t_F is larger than λ, it suggests that the full-time driver works overtime and incurs a leisure loss l per unit time. Hence, the payoff of a full-time platform driver is given by

$$V_{F,i} = \theta_i w_i \times 0 + \gamma_i M_i - \theta_i(cM_i + k) + b_F + \theta_i t_F s(\alpha\rho - c) - (t_F - \lambda)_+ l.$$

- [U] A user is by definition an individual that uses ride-hailing service to satisfy her transportation needs. She pays r dollars per kilometer for hailing a car. However, due to market frictions, only β percentage of times a user can be successfully matched with a car. If she fails to find a private-hire car, then with probability $1 - \beta$ the user opts for public transport. Thus, the corresponding payoff is given by

$$V_{U,i} = \theta_i w_i + \beta(\gamma_i M_i - \theta_i r M_i).$$

- [A] An abstinent individual chooses not to buy a car nor use ride-hailing service. Whenever she has a travel need, she opts for the outside option such as public transport or walking. We normalize the utility of the outside option to zero and the cost of the outside option is assumed to be negligible. Thus the payoff function of the abstinent only consists a term of her regular job wage. Therefore, the payoff of an abstinent is

$$V_{A,i} = \theta_i w_i.$$

3.2 Choice Model and Population Share

Among all the five agent types listed in Sect. 3.2, some individuals may be constrained by exogenous reasons to be of a certain type. For example, some may not have driving license or are not physically fit for driving. Therefore, we consider the population divided into two segments S_1, S_2. The first segment is eligible for driving and hence has access to all five strategies *N, P, F, U, A*. The other segment cannot act as a driver (at least in the short term where the decision model is defined) and hence has access to either the strategy of *U* or *A*, i.e., user or abstinent. We use discrete choice model to study the choices of the heterogeneous population.

Discrete choice models based on multinomial logit have been used extensively in the field of transportation research [29]. According to random utility theory (RUT) [30], in additional to the systematic part of the utility V_{ij} as specified in the payoff functions, there is a random part ε_{ij} which captures the unobservable factors. We assume ε_{ij} are independently and identically distributed type-I extreme random variables. Therefore, for the five strategies listed in Sect. 3.2, the probability P_{ij} of individual i choosing strategy $j \in \{N, P, F, U, A\}$ in the two population segments can be expressed as

$$S_1 : P_{ij} = \frac{e^{V_{ij}}}{e^{V_{i,N}} + e^{V_{i,P}} + e^{V_{i,F}} + e^{V_{i,U}} + e^{V_{i,A}}}, j = N, P, F, U, A,$$
$$S_2 : P_{ik} = \frac{e^{V_{ik}}}{e^{V_{i,U}} + e^{V_{i,A}}}, k = U, A.$$

As the heterogeneity of the population is characterized by individual-specific variables $(\theta_i, \gamma_i, w_i, M_i)$, the integral of above equations over the joint distribution of $f(\theta, \gamma, w, M)$ produces the choice shares of the five strategies within each segment. Specifically,

$$S_1 : S_{j,S1} = \int \frac{e^{V_{ij}}}{e^{V_{i,N}} + e^{V_{i,P}} + e^{V_{i,F}} + e^{V_{i,U}} + e^{V_{i,A}}} f(\theta, \gamma, w, M) d\theta d\gamma dw dM,$$
$$j = N, P, F, U, A,$$

$$S_2 : S_{k,S1} = \int \frac{e^{V_{ik}}}{e^{V_{i,U}} + e^{V_{i,A}}} f(\theta, \gamma, w, M) d\theta d\gamma dw dM, k = U, A.$$

Let us further assume that a fraction of the population belongs to S_1 and $1 - a$ fraction belongs to S_2, then the population shares of strategies *N, P, F* are

$$S_j = a \int \frac{e^{V_{ij}}}{e^{V_{i,N}} + e^{V_{i,P}} + e^{V_{i,F}} + e^{V_{i,U}} + e^{V_{i,A}}} f(\theta, \gamma, w, M) d\theta d\gamma dw dM, j = N, P, F,$$

$$\tag{1}$$

and the population shares of strategies *U* and *A* are

$$S_k = a \int \frac{e^{V_{ij}}}{e^{V_{i,N}} + e^{V_{i,P}} + e^{V_{i,F}} + e^{V_{i,U}} + e^{V_{i,A}}} f(\theta, \gamma, w, M) d\theta d\gamma dw dM$$
$$+ (1 - a) \int \frac{e^{V_{ik}}}{e^{V_{i,U}} + e^{V_{i,A}}} f(\theta, \gamma, w, M) d\theta d\gamma dw dM, k = U, A. \tag{2}$$

It is worth noting that Eqs. (1) and (2) have no closed form solution. However, we can approximate the integral through simulation. Specifically, we first draw R samples $(\theta^r, \gamma^r, w^r, M^r)$ from the joint distribution $f(\theta, \gamma, w, M)$, then use this as an empirical distribution to compute the average of the probability P_{ij} over individuals of the population instead of the integral. This average is the simulated population share of the five strategies:

$$
\begin{aligned}
\hat{S}_j &= a\left(\frac{1}{R} \sum_{r=1}^{R} \frac{e^{V_j^r}}{e^{V_N^r} + e^{V_P^r} + e^{V_F^r} + e^{V_U^r} + e^{V_A^r}} \right), j = N, P, F, \\
\hat{S}_k &= a\left(\frac{1}{R} \sum_{r=1}^{R} \frac{e^{V_k^r}}{e^{V_N^r} + e^{V_P^r} + e^{V_F^r} + e^{V_U^r} + e^{V_A^r}} \right) + (1-a)\left(\frac{1}{R} \sum_{r=1}^{R} \frac{e^{V_k^r}}{e^{V_U^r} + e^{V_A^r}} \right), j = U, A.
\end{aligned}
\tag{3}
$$

A challenge in estimating above population shares is the endogeneity of ride-hailing matching probabilities α and β, which are functions of the supply (population shares of platform drivers) and demand (population share of users). More precisely, controlling other factors, if α increases, a platform driver has higher chance to serve a customer and thus obtain higher income. Therefore, strategies P and F become more tempting and corresponding population shares S_P and S_F, i.e., the supply of service will increase. However, as the supply of the service increases, the matching rate α will conversely tend to decrease. User's matching probability β acts in a similar vein. This observation implies that there is a circular dependence between supply and demand and the matching rates α and β. To characterize this dependence, we incorporate below matching functions based on the work of [15] into our model:[2]

$$
\alpha = t \cdot \frac{D}{D+S}, \beta = t \cdot \frac{S}{D+S}.
\tag{4}
$$

Under our model specification, the supply and demand can be further expressed as the mileage supplied and requested:

$$
Supply = (S_P t_P + S_F t_F)S, Demand = S_U E(M).
\tag{5}
$$

In (4), the parameter $t \in (1, \min\{\frac{Demand+Supply}{Demand}, \frac{Demand+Supply}{Supply}\})$ characterizes the level of matching friction in the two-sided market: the higher t is, the lower the matching friction, hence the higher the matching propensity.

4   Data

Our empirical context will be the Singapore ride-hailing marketplace, which acts as an ideal research subject for the following reasons. Firstly, as an island country, the transportation system of Singapore is relatively closed with few foreign vehicles entering

---

[2] According to [15], Eq. (4) is derived based on deterministic fluid approximation of a multi-server loss queuing system with arrival rate D and total service rate S, in which, α corresponds to the system utilization and β corresponds to the probability at which a job request could find an idle server upon arrival.

from outside, which guarantees that supply of the peer-to-peer ride-hailing service is provided within the region. Given the fact that the Singapore government records the private hire/ride-hailing car population over time, this unique feature makes the number of platform drivers easily trackable, which is not the case in other cities. For example, in San Francisco, previous study finds that about 80% of the Uber drivers actually live outside of the city and bring their cars into the city every morning to provide for peer-to-peer service [5]. As the supply is from both internal and external, such an open system brings difficulty in modeling the dynamics between supply and demand. Secondly, due to the strict regulations of the Singapore government, the level of car ownership is being monitored over time. In order to own and use a vehicle (no matter for personal use or for ride-hailing service), car owners are required to obtain a quota licence called "Certificate of Entitlement (COE)" through bidding. Since the number of cars are strictly controlled by the government, the supply of private car is relatively stable over the years and the COE quotas made available for each year is limited. In view of the fact that the price of the car itself (determined by its open market value) is relatively stable, the bidding price of COE can be viewed as a proxy for demand of cars as the total number of cars allowed is relatively constant.

In the following sections, we elaborate on how we calibrate our model based on the Singapore ride-hailing marketplace. We focus on the population with age between 20 and 64 for those who are eligible to drive and below retirement age. Specifically, we use publicly disclosed data in year 2017 to calibrate our model and data in year 2016, 2018 and 2019 to conduct robustness check and validity test. Values of our model parameters are determined through two ways. First, some parameters such as travel demand and wage distribution do not depend on our model specification and can be *estimated* from available data directly. However, other parameters such as price sensitivity, perception of traveling in private cars, convenience benefit of car ownership and utility loss of working overtime cannot be obtained through the data directly. Therefore, we *calibrate* these parameters by searching for the values that make our model captures the reality i.e., generates the calibrated population shares as close as the actual ones.

4.1 Estimated Parameters

Table 1 summarizes the values of the parameters estimated directly from data for year 2017. The travel demand is estimated by fitting a lognormal distribution based on average monthly mileage of private cars and distribution of taxi trip distance (including both street hailing and requests from ride-hailing platforms). Wage distribution of the population is obtained by fitting a lognormal distribution to the household income data. As mentioned previously, the ownership cost of a private car consists of car retail price and the cost of COE. Therefore, we approximate the ownership cost as the sum of both car price and COE cost. Usually, ride-hailing platforms take a commission fee from drivers for each trip. They also provide various incentive schemes to attract more supply. Thus platform drivers' earning ρ is estimated as the sum of after-commission trip fare and incentive rewarded.[3] The value of matching friction level t is determined by setting the

---

[3] See example of a major ride-hailing company's incentive scheme on 2019: https://www.gojek.com/sg/blog/sg-driver-incentives-gojek-singapore/.

Table 1. Estimated parameter values for year 2017

Parameter	Symbol	Estimate
Travel demand	M	lognormal $(3.75, 0.38^2)$ km/day
Regular job wage	w	lognormal $(7.89, 0.86^2)$ \$/mth
Ownership cost of car	k	60 \$/day
Driving cost	c	0.13 \$/km
Average driving speed	s	40 km/hr
Regular working hours	λ	45 h/week
Part-time platform driver working hours	t_P	7 h/week
Full-time platform driver working hours	t_F	49 h/week
User payment for ride-hailing	r	0.99 \$/km
Driver earning from ride-hailing	ρ	1.00 \$/km
Market friction level	t	1.3

matching rates α and β reflecting the Singapore ride-haling market. As suggested by the taxi occupancy data (fraction of driving time a taxi is occupied with a passenger), the representative matching rate for platform drivers is around 0.5, hence we estimate t $= 1.3$ which leads to $\alpha = 0.505$ and $\beta = 0.895$. More details of the estimation refer to online Appendix.[4]

4.2 Calibrated Parameters

The remaining model parameters (sensitivity of money θ, convenience benefit of possessing a car b_N, b_P and b_F, utility difference between traveling in a private car and public transport γ, and loss of leisure utility l) cannot directly be estimated from the data. We assume that θ and γ independently follow a lognormal distribution with parameters $(\theta_\mu, \theta_\sigma^2)$ and $(\gamma_\mu, \gamma_\sigma^2)$ respectively. We calibrate $(\theta_\mu, \theta_\sigma^2, \gamma_\mu, \gamma_\sigma^2, b_N, b_P, b_F, l)$ by minimizing the sum of squared percentage differences between the calibrated and actual population shares by solving below optimization problem:

$$min \sum_j \left(\frac{\hat{S}_j - S_j}{S_j}\right)^2, j = N, P, F, U, A,$$

$$s.t. \, b_N, b_P, b_F, l \geq 0$$

(6)

where $\hat{S}_j$ are the calibrated population shares as in Eq. (3) and S_j are the actual population shares. The objective function value of our calibration is 9.98×10^{-7}, which suggests that the model matches the actual population shares very well. The values of the calibrated parameters are listed on online Appendix.

---

[4] Online Appendix available at: https://drive.google.com/file/d/1DQkbJSBWFoFIMmAEMxkf eo7qjJrT2CRI/view?usp=sharing.

4.3   Model Validation

We validate our calibrated model by testing its prediction power. Recall that we use only the 2017 data to calibrate the values of parameters $(\theta_\mu, \theta_\sigma^2, \gamma_\mu, \gamma_\sigma^2, b_N, b_P, b_F, l)$, thus we could apply the model to different years 2016, 2018 and 2019 to compare the predicted and actual population shares. Table 2 lists the prediction results. We can see that the predicted population shares are fairly close to the actual ones and capture the trend of its variation.

Table 2. Predicted vs. actual population shares

	Non-platform driver (S_N)				Platform drivers ($S_P + S_F$)			
	2016	2017	2018	2019	2016	2017	2018	2019
Predicted population share	18.80%	19.17%	20.70%	21.70%	2.91%	3.10%	2.97%	3.20%
Actual population share	19.32%	19.16%	19.37%	19.57%	2.81%	3.11%	2.90%	3.17%

5   Counterfactual Analysis

In this section, we use the calibrated model to conduct numerical experiments to examine how choice shares of the five strategies might change with new policies. In particular, we focus on the variations of driving cost and ride-hailing payment as they are the most commonly used means of market intervention and platform operations.

As mentioned in Sect. 4, due to the government's regulation, the population of private cars can be considered as a constant number. Explicitly, the number of non-platform drivers, part-time platform drivers and full-time platform drivers sums up to this constant number; and demand for cars (including both non-platform drivers and platform drivers) is captured by the bidding price of COE which is a component of car ownership cost k. Therefore, we have two equilibrium loops in our numerical experiments, one for car ownership constraint (total number of cars fixed) and the other for ride-hailing service supply-demand balancing (through dynamics between matching rates α and β). Below we present the main results of the counterfactual analyses.

5.1   Model Equilibria from Varying Driving Cost

Figure 1 illustrates the variation of population shares of platform drivers and users with respect to driving cost per kilometer at equilibria. As driving cost increases, the number of platform driver decreases whereas the population of users increases even with lower matching propensity. This may be because with rising driving cost, for those who prefer traveling in a private car, self-driving becomes too expensive and they would rather switch to ride-hailing services. The equilibrium results further show that the

increase of driving cost would lead to slight decrease in car demand and VKT, which suggests that actions affecting driving cost such as controlling road tax and fuel tax may help to mitigate traffic but with limited effectiveness. Therefore, other measures such as improving accessibility public transportation may worth considering (More results available on online Appendix).
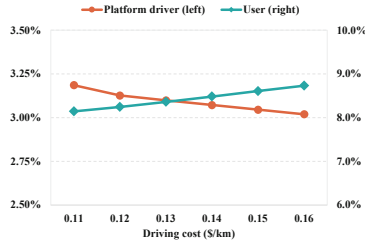

Fig. 1. Population share change w.r.t driving cost

5.2 Model Equilibria from Varying Ride-Hailing Payment

Figure 2 illustrates the variation of population shares of platform drivers and users w.r.t. ride-hailing payment. As expected, the proportion of ride-hailing users slashes with higher ride-hailing payment. Such a phenomenon brings a two-fold impact on platform drivers. On the one hand, the higher payment promises higher earnings for drivers as their income is proportionate of user charges. On the other hand, as there are fewer users, hence less demand, the matching rate for drivers decreases. In our analysis, the effect of the former is dominated by the latter. As ride-hailing payment increases, although the earnings from busy period increases for platform drivers, their total effective income instead decreases more significantly due to the lower matching propensity. Thus, we observe the diminishing population shares of both part-time and full-time platform drivers.
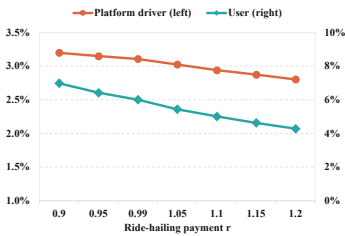

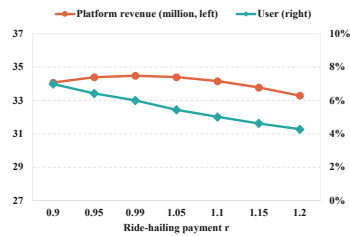**Fig. 2.** Population share change w.r.t ride-hailing payment

Fig. 3. Platform revenue w.r.t ride-hailing payment

The equilibrium results further show that, unexpectedly, the increment in riding price would lead to a slight increment in car demand. Since the total number of cars is a constant number, with the population of platform drivers dropping, non-platform

drivers increase. However, more non-platform drivers do not necessary lead to higher COE price. This observation suggests that as ride-hailing becomes more expensive, non-car owners who prefer to travel in a private car are more likely to buy a personal car rather than turning to public transportation. Moreover, as shown in Fig. 3, the platform's revenue does not increase monotonically with ride-hailing payment but peaks around 1.05 dollars per kilometer and then starts to diminish. This observation suggests that although higher charges bring higher commission fees for the platform, the enlarging loss of users counteracts such an incremental effect. Thus, increasing service charges to attain higher revenue is not sustainable and there exits an optimal payment level which allows the platform to maximize its revenue (More results available on online Appendix).

6 Conclusions

In this paper we introduce an integrated model structure that integrates empirical analysis into an analytical framework to assess the impact of on-demand platform on mobility choice, ownership and traffic congestion. By introducing matching functions, we further incorporate the endogeneity of supply demand imbalance of the two-sided on-demand service market into the model. Our analyses provide several managerial insights: to control traffic and curb VKT growth, measures through regulating driving cost such as increasing road tax and gasoline tax may have limited effectiveness. Other procedures such as improving accessibility public transportation may worth considering. For platform revenue management, our results suggest that there exits an optimal payment level which allows the platform to maximize its revenue. Although raising user charges within a certain range would increase the platform's profit, it is not sustainable as higher charges leads to enlarging loss of users, which brings a counteracting effect.

References

1. Tafreshian, A., Masoud, N., Yin, Y.: Frontiers in service science: ride matching for peer-to-peer ride sharing: a review and future directions. Serv. Sci. **12**(2–3), 44–60 (2020)
2. Hall, J.V., Krueger, A.B.: An analysis of the labor market for Uber's driver-partners in the United States. ILR Rev. **71**(3), 705–732 (2018)
3. Burtch, G., Carnahan, S., Greenwood, B.N.: Can you gig it? An empirical examination of the gig economy and entrepreneurial activity. Manage. Sci. **64**(12), 5497–5520 (2018)
4. Fellows, N., Pitfield, D.: An economic and operational evaluation of urban car-sharing. Transp. Res. Part D: Transp. Environ. **5**(1), 1–10 (2000)
5. Anderson, D.N.: "Not just a taxi"? For-profit ridesharing, driver strategies, and VMT. Transportation **41**(5), 1099–1117 (2014)
6. Alexander, L., Jiang, S., Murga, M., Gonzalez, M.C.: Origin–destination trips by purpose and time of day inferred from mobile phone data. Transp. Res. Part C: Emerging Technol. **58**, 240–250 (2015)
7. Li, Z., Hong, Y., Zhang, Z.: An empirical analysis of on-demand ride sharing and traffic congestion. In: Proceedings of the International Conference on Information Systems (2016)

8. Clewlow, R.R., Mishra, G.S.: Disruptive transportation: the adoption, utilization, and impacts of ride-hailing in the United States. Institute of Transportation Studies, University of California, Davis (2017)
9. Jiang, B., Tian, L.: Collaborative consumption: strategic and economic implications of product sharing. Manage. Sci. **64**(3), 1171–1188 (2016)
10. Weber, T.A.: The question of ownership in a sharing economy. In: 2015 48th Hawaii International Conference on System Sciences, pp. 4874–4883. IEEE (2015)
11. Filippas, A., Horton, J.J., Zeckhauser, R.J.: Owning, using, and renting: some simple economics of the 'sharing economy.' Manage. Sci. **66**(9), 4152–4172 (2020)
12. Gong, J., Greenwood, B.N., Song, Y.: Uber might buy me a Mercedes Benz: an empirical investigation of the sharing economy and durable goods purchase. Available at SSRN 2971072 (2017)
13. Benjaafar, S., Hu, M.: Operations management in the age of the sharing economy: what is old and what is new? Manuf. Serv. Oper. Manag. **22**(1), 93–101 (2020)
14. Weber, T.A.: Product pricing in a peer-to-peer economy. J. Manage. Inf. Syst. **33**(2), 573–596 (2016)
15. Benjaafar, S., Kong, G., Li, X., Courcoubetis, C.: Peer-to-peer product sharing: implications for ownership, usage, and social welfare in the sharing economy. Manage. Sci. **65**(2), 477–493 (2018)
16. Abhishek, V., Guajardo, J., Zhang, Z.: Business models in the sharing economy: manufacturing durable goods in the presence of peer-to-peer rental markets. Available at SSRN 2891908 (2020)
17. Banerjee, S., Riquelme, C., Johari, R.: Pricing in ride-share platforms: a queueing-theoretic approach. Available at SSRN 2568258 (2015)
18. Taylor, T.A.: On-demand service platforms. Manuf. Serv. Oper. Manage. **20**(4), 704–720 (2018)
19. Bai, J., So, K.C., Tang, C.S., Chen, X., Wang, H.: Coordinating supply and demand on an on-demand service platform with impatient customers. Manuf. Serv. Oper. Manage. **21**(3), 556–570 (2019)
20. Gurvich, I., Lariviere, M., Moreno, A.: Operations in the on-demand economy: staffing services with self-scheduling capacity. In: Hu, M. (ed.) Sharing economy. SSSCM, vol. 6, pp. 249–278. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-01863-4_12
21. Cachon, G.P., Daniels, K.M., Lobel, R.: The role of surge pricing on a service platform with self-scheduling capacity. Manuf. Serv. Oper. Manage. **19**(3), 368–384 (2017)
22. Hu, M., Zhou, Y.: Price, wage and fixed commission in on-demand matching. Available at SSRN2949513 (2019)
23. Benjaafar, S., Bernhard, H., Courcoubetis, C., Kanakakis, M., Papafragkos, S.: Drivers, riders, and service providers: the impact of the sharing economy on mobility. Manage. Sci. (2021)
24. Benjaafar, S., Ding, J.-Y., Kong, G., Taylor, T.: Labor welfare in on-demand service platforms. Available at SSRN (2020)
25. Sinchaisri, P., Allon, G., Cohen, M.: The impact of behavioral and economic drivers on Gig economy workers. In: Academy of Management Proceedings, vol. 2019, p. 10216. Academy of Management, Briarcliff Manor (2019)
26. Erhardt, G.D., Roy, S., Cooper, D., Sana, B., Chen, M., Castiglione, J.: Do transportation network companies decrease or increase congestion? Sci. Adv. **5**(5), eaau2670 (2019)
27. Fredriksson, A., de Oliveira, G.M.: Impact evaluation using Difference-in-Differences. RAUSP Manage. J. **54**, 519–532 (2019)
28. Fraiberger, S.P., Sundararajan, A., et al.: Peer-to-peer rental markets in the sharing economy. NYU Stern School of Business Research Paper, vol. 6 (2015)

29. Keyes, A.K., Crawford-Brown, D.: The changing influences on commuting mode choice in urban England under Peak Car: a discrete choice modelling approach. Transport. Res. F: Traffic Psychol. Behav. **58**, 167–176 (2018)
30. McFadden, D.: The measurement of urban travel demand. J. Public Econ. **3**(4), 303–328 (1974)