Pricing and operational planning of a fixed-route ride-sharing service

Wanqing Zhu, Xi Chen, Bo Wang, Steven Lawrence, Hao Zhou, Armagan Bayram

Pricing and operational planning of a fixed-route ride-sharing service

Wanqing Zhu and Xi Chen*

Department of Industrial and Manufacturing Systems Engineering,
University of Michigan-Dearborn,
Dearborn, Michigan 48128, USA
Email: wanqingz@umich.edu
Email: xichenxi@umich.edu
*Corresponding author

Bo Wang, Steven Lawrence and Hao Zhou

Ford Motor Company,
Dearborn, Michigan 48124, USA
Email: bwang55@ford.com
Email: SLAWRE62@ford.com
Email: hzhou35@ford.com

Armagan Bayram

Department of Industrial and Manufacturing Systems Engineering,
University of Michigan-Dearborn,
Dearborn, Michigan 48128, USA
Email: armagan@umich.edu

Abstract: Fixed-route ride-sharing services, e.g., RidePal, OurBus, Urbvan, are becoming increasing popular among metropolitan areas. Effective pricing and operational planning of these services are undeniably crucial in their profitability and survival. However, the effectiveness of existing approaches has been hindered by the accuracy in demand estimation. In this paper, we develop a data-driven demand model using the multinomial logit model. We also construct a nonlinear optimisation model based on this demand model to jointly optimise price and operational decisions. A case study based on a real world fixed-route ride-sharing service in New York City is presented to demonstrate how the proposed models are used to improve the profitability of the service. We also show how this model can apply in settings where only limited public data are available to obtain effective estimation of demand and profit.

Biographical notes: Wanqing Zhu is a Data Scientist in Global Data Insight & Analytics in Ford Motor Company. She was a graduate student assistant in University of Michigan – Dearborn majored in Industrial and System Engineering. Her work focuses on the sharing economy transportation and the impact on customer behaviours. She is now focusing on doing research about the fleet management and network development for autonomous vehicles.

Xi Chen is an Assistant Professor of Industrial and Manufacturing Systems Engineering at the University of Michigan-Dearborn. Her research interests include supply chain management, sustainable operations, smart mobility operations, and resilient systems. Her primary research methodologies include optimisation, game theory, discrete choice modelling, dynamic programming and data analytics. She received her PhD degree in Industrial and Systems Engineering from the University of Minnesota.

Bo Wang was a mobility analytics manager working at Global Data Insights & Analytics (GDIA) organisation at Ford Motor Company. He was leading a team of data scientists and developers building AI-driven solutions and products from inception to production. He led several mobility analytics projects at Ford, including micro-transit operation, connected vehicle, and smart city analytics. He received his BS, MS, and PhD in Transportation Engineering from Iowa State University.

Steven Lawrence is a Senior Data Scientist and Supervisor at The Ford Motor Company. He received his PhD in Consumer Behaviour (2013) from Wayne State University in Detroit, Michigan. He joined Ford's Smart Mobility Analytics team in 2016 and has since specialised in the business applications of autonomous vehicles. At Ford's AV LLC, he is the geospatial business analytics product owner.

Hao Zhou earned his PhD in Industrial Engineering from University of Michigan. He has worked in Smart Mobility research at Ford Motor Company since 2013. He joined Amazon in 2017 as research scientist working on last mile delivery routing optimisation. His research interest is in optimisation of transportation system.

Armagan Bayram is an Assistant Professor in the Department of Industrial and Manufacturing Systems Engineering at the University of Michigan-Dearborn. Her research interests include the development of stochastic models and solution methods for capacity and resource allocation problems. Her particular interests are stochastic optimisation and dynamic programming models that involve non-profit, healthcare, and mobility applications.

1   Introduction

The rapid growth of the sharing economy has been witnessed around the world in recent years. That is, the economy is undergoing a paradigm shift away from single ownership and towards shared ownership of goods and services. Successes have been seen, for example, in businesses that share habitation (e.g., Airbnb), financial services (e.g., CrowdFunding), vehicles (e.g., Car2go, ZipCar), and other mobility solutions (e.g., Uber, RidePal). Among these sharing services, ride-sharing is a particularly popular category, as evidenced by the popularity of Uber Pool and Lyft Line. Ride-sharing refers to the sharing of partial or whole trips among multiple riders using the same vehicle. By having more people using one vehicle, the travelling cost of each person can be reduced

while vehicle capacity utilisation can be significantly improved. Moreover, reduction in air pollution and traffic congestion may also result due to a reduction in the number of vehicles per trip demand. There are a number of ride-sharing companies operating different modes of sharing. Large cities have seen the most successful implementations of such ride-sharing services due to the immense opportunities of common trip segments.

Among ride-sharing services, the specific business models take several forms, including the door-to-door model (e.g., Uberpool and Lyft Line), the corner-to-corner model (e.g., Via) and the fixed-route model (e.g., OurBus, RidePal, urbvan). The focus of this work is on fixed-route ride-sharing services, in which shuttles operate on fixed routes with predetermined stations. Customers of the service send request in advance to reserve a seat and then walk to and wait at a station by the scheduled time for their pick-up. Customers are charged a lower price than door-to-door ride-sharing service, while incurring longer travel times due to walking and waiting. Fixed-route ride-sharing services predominantly aim at serving commuters for completing trips to and from work. On the other hand, efforts from the public transportation sector to provide smarter and more adaptive services have also been witnessed. Take the SmartBus in Melbourne Australia as an example, it offers more frequent services, extended operating time, updated timetable at bus stops and customer flow statistics at a fixed flat rate.

Traditionally, the main goal of public transit services is to maximise the welfare of the riders. As a result, they tend to charge very minimal fares and rely on government subsidy to stay in operation. However, cities nowadays are increasingly partnering with private services to solve their transportation problem. Arlington, Texas, for example, is replacing its downtown bus service with Via's ride-sharing service. With the growing instances of privately owned ride-sharing services, more focus needs to be placed on profitability. Pricing subsequently becomes a crucial decision for these services. In addition to price, how the service is operated can also have a significant effect on its profitability and market share. For example, for a given customer, an affordable service with stations within close proximity would be most attractive, whereas an expensive service whose stations are far away would unlikely be a good choice.

Despite the importance of pricing and operational decisions, they have not been adequately addressed primarily due to the difficulty to predict demand. This is due to several reasons. First, there is a lack of demand data of fixed-route ride-sharing and similar services. Most fixed-route ride-sharing services are relatively recent startups where systematic data collection have not been developed. Traditional transit services rarely vary their prices, resulting in particularly limited data points. In addition, customers may be reluctant to provide personal information due to privacy concerns. Second, as transportation systems become growing complex, there are many alternatives that compete for the demand of commuters. Third, the demand is affected by complex factors such as personal preference. For example, the attractiveness of an affordable service that requires significant amount of walking largely depends on the sensitivity of the customer towards cost versus time.

In this research, we develop a method for the optimal pricing and operational planning of a fixed-route ride-sharing service that addresses the above challenges. Our contributions are as follows. First, we develop a data-driven demand model that incorporates concerns about cost, time, customer heterogeneity, and competing transportation alternatives. Using publicly available data, we show that this demand model is able to effectively predict customer mode choice and hence demand. Second, we

develop a nonlinear optimisation model for jointly optimising profit, fleet size and shuttle frequency based on the proposed demand model. Third, a case study of a real world fixed-route ride-sharing service is provided to demonstrate how the proposed models are used to improve the profitability of this service. Fourth, using the case study, we derive several interesting insights pertaining to the fixed-route ride-sharing service in New York City (NYC). For example, we find that introducing a per-distance rate has limited effect on the profitability of the service, and that there is an opportunity to further increase adoption rate with little compromise on profitability by adjusting the flat rate.

The rest of the paper is organised as follows. Section 2 offers a review of related literature. Section 3 describes the model setup and formulates the joint pricing and operational planning problem as a nonlinear optimisation problem. Section 4 provides a case study based on a real world fixed-route ride-sharing service, in which NYC commuter mode-choice are fitted using real data and the joint optimisation problem proposed in Section 3 is solved. Section 5 offers concluding remarks.

2   Literature review

This work is related to the literature on the pricing of transportation services. As public transportation usually employs a flat rate, the majority of this literature has focused to the pricing of taxi services. Douglas (1972) develops an aggregate model with a constant fare per unit time and per unit travel distance to optimise the vacancy rate for the taxicab industry. However, Douglas (1972) does account for spatial effect on demand. This aggregate pricing model is widely used in later papers such as De Vany (1975), Arnott (1996), Chang and Chu (2009), He et al. (2018) and Zha et al. (2018). Arnott (1996) provides a dispatching model for taxis to reduce the subsidisation for the taxicab industry, where he considers a space (a two dimensional city) within which taxis are randomly and uniformly distributed for implementation. Chang and Chu (2009) solve for the welfare-maximising price for cruising taxi market, where the demand is assumed to be a log-linear function of price and average waiting time. Our work contributes to this literature by developing a data-driven approach for estimating demand and an optimisation program based on this demand model for jointly optimising price as well as operational policies. He et al. (2018) propose an equilibrium framework to depict the operations of a regulated taxi market and investigate the optimal pricing strategy. Zha et al. (2018) analyse the impact of spatial pricing on ride-sourcing market and examine the equilibrium under spatial pricing. Unlike the above literature, the demand estimation method in this work is based on real mode-choice decisions rather than a stylised demand function form. We also account for the effect of accessibility of the route in the demand function, which is not considered in taxi pricing since taxis provide door-to-door service. Moreover, our objective is to maximise the total operational profit, which differs from the typical objective of taxi fare optimisation of maximising the social willingness-to-pay.

The operational decisions considered in this paper include shuttle frequency, which has been studied in the literature on transit route configuration. This literature considers decisions include selection/improvement of routes as well as the optimal transit frequency (see for example Lampkin and Saalmans, 1967; Silman et al., 1974; Marwah et al., 1984; Soehodho and Nahry, 1998; Lee and Vuchic, 2005; Jha et al., 2019; Feng et al., 2019; Mahdavi Moghaddam et al., 2019). Within this literature, few have considered the joint optimisation of both price and frequency of transit services, with

Delle Site and Filippi (1998) and Chien and Spacovic (2001) being exceptions. However, they assume constant elasticity demand functions and do not provide methods for the calibration demand elasticity.

A key element of this work is the modelling of demand of the ride-sharing service based on travel mode-choice decisions of customers. The literature on travel mode choice is extensive and we review some of the most relevant ones below. Deneubourg et al. (1979) develop a dynamic model to study the effect of behavioural fluctuations on the competing modes of automobile and public transportation. However, they do not consider the cost or service region of the transportation modes, or model the specific choice process. Much of the later literature use multinomial logit (MNL) model to model the decision making process, (see for example Swait and Ben-Akiva, 1987; Cervero and Kockelman, 1997; Cervero, 2002; Miller et al., 2005; Frank et al., 2008; Ding et al., 2018; Hu et al., 2018; Guo et al., 2018). Cervero and Kockelman (1997) introduce 3Ds: density, diversity and the design into the MNL model, and found them to be significant in mode-choice decisions. Cervero (2002) performs a model comparison between the original model and an expanded model with land-use and socio-economic variables based on a dataset based in Montgomery County, Maryland. Koppelman and Bhat (2006) introduce mode-choice modelling using multinomial and nested logit models in a report to the US Department of Transportation. They also carry out a micro-simulation on the SF bay area to validate the mode-choice model. Using a public transportation survey in Chicago, Javanmardi et al. (2015) conclude that individual and household socio-demographic, transit availability and vehicle availability play an important role in the modelling process. Consistent with this literature, we use the MNL model for mode-choice decisions, and consider all factors mentioned in the above paper including land-use, socio-economy and demographic factors. Using NYC Regional House Hold Travel Survey and other data, we fit a MNL model that can be used to predict NYC commuters' mode-choice decisions using only information of their origins and destinations. In contrast to the above literature which focuses on fitting the mode-choice model, we utilise the fitted model to simulate the demand of a fixed-route ride-sharing service for inputting in pricing optimisation.

Finally, this research also contributes to the growing literature on operational decisions in the sharing economy. However, the focus of this literature has primarily been on car-sharing services such as Zipcar and Car2go, or on-demand services such as Uber and Lyft. For example, He et al. (2017) study the service region design problem for a one-way electric vehicle sharing system. They develop an adoption rate model and compute the profit using queueing theory. Qi et al. (2018) provide logistical planning models for shared last-mile delivery services. Lu et al. (2018) study the allocation of vehicle fleet to service zones when the demand is uncertain using a two-stage stochastic integer program. In a related paper, Chang et al. (2017) consider the location design and fleet rebalancing of a car-sharing service under emission constraints. Bellos et al. (2017) examine the interaction between car sharing decisions and an original equipment manufacturer's product design under the corporate average fuel economy standards. Agatz et al. (2012) study the dynamic matching between vehicles and customers for an on-demand ride-sharing service. Cachon et al. (2017) analyse the effect of several pricing schemes for on-demand services and find one that achieves near-optimal profit. He (2018) investigates the incentive design and spatial allocation of capacity in shared mobility systems. Benjaafar et al. (2018) study the effect of on-demand service platforms

on labour welfare. Kim and Lee (2017) evaluate the impact of reallocation in the optimisation of one-way carsharing systems. This work contributes to the above literature by studying the joint pricing and operational decisions of a fixed-route ride-sharing system.

3   The model

Without loss of generality, we consider a fixed-route ride-sharing service provider who operates a one-way route, which we denote as RS. However, this assumption can be easily extended to two-way service, which is omitted for notational brevity. The set of stations is denoted as $S = \{1, 2, \ldots, s\}$. For any $k \in \{1, 2, \ldots, s\}$, service zone centred at station k is the region in which each point is located no more than γ (in Manhattan distance) away from station k, denoted as $A(k)$. For example, the maximum distance that residents are willing to walk is considered as the radius of the service region of the service provider's stations. We denote the origin of a customer i as o_i and his/her destination as d_i. The origin station for customer i is denoted as O_i and the destination station is denoted as D_i. For simplicity, we discretise each service region into evenly spaced (in Manhattan distance) Q points, and assume that the origins and destinations of customers are uniformly distributed among these Q points. Figure 1 illustrates this discretisation in an example with $Q = 25$, where the star and dots represent the locations of the station and possible origin/destination of customers respectively and the grid represent the road directions. That is, if $o_i = k$ ($d_i = k$), the customer i's origin (destination) locates at any point within the service region $A(k)$ with equal probability of $\frac{1}{Q}$. In addition to origin and destination, other characteristics of customer i (e.g., demographics, income) is summarised in an additional variable x_i. We define *potential customers* of RS as travellers whose origin is covered by the service region of a station of RS and whose destination is covered by the service region of a subsequent station of RS. We denote the probability density function of potential customers in service region $A(k)$ as $f_k(\cdot)$ and the set of possible values of x_i in $A(k)$ as X_k.

Figure 1   Discretisation of possible origin/destination locations in a service region
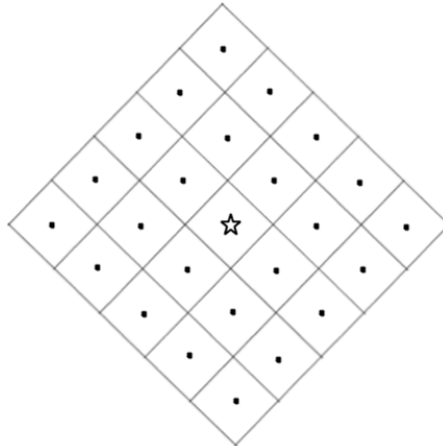
3.1 The raw adoption rate

The *raw adoption rate* of the ride-sharing service refers to the proportion of potential customers who prefer the service over other competing options. Note that the raw adoption rate differs from actual adoption rate of the service (see definition in Section 4.4) in that raw adoption rate does not factor in the capacity constraint of the service, and hence reflect solely customer preference. We estimate the raw adoption rate by modelling the travel mode-choice process for each potential customer whose origin and destination fall in the route's service regions using the classic MNL model. We define Φ as a set of all available travel modes and $RS \in \Phi$. The utility a customer i derives from choosing mode $\phi \in \Phi$.

$$U_\phi\left(x_i, c_i^\phi, t_i^\phi\right) = V_\phi\left(x_i, c_i^\phi, t_i^\phi\right) + \in_i^\phi \tag{1}$$

where c_i^ϕ is customer i's cost of travel associated with model ϕ, and t_i^ϕ is customer i's time of travel associated with model ϕ. $V_\phi(x_i, c_i^\phi, t_i^\phi)$ is the deterministic part of the customer's utility, which depends on cost of travel, time of travel, as well as customer i's personal characteristics. $\in_i^\phi$ represents the random component of the utility function and is assumed to follow an extreme value distribution. We note that c_i^ϕ is a known fixed number once the origin and destination of the customer is known. It is also important to highlight that t_i^ϕ consists of not only the time of travel spent on the vehicle (denoted as $t_i^{\phi,IV}$), but also walking time to and from the corresponding travel mode for its access, denoted as $t_i^{\phi,WT}$ and $t_i^{\phi,WF}$ respectively. That is,

$$t_i^\phi = t_i^{\phi,WT} + t_i^{\phi,IV} + t_i^{\phi,WF} \tag{2}$$

As a result, customer i chooses RS as his/her mode of travel if and only if $U_{RS}(x_i, c_i^{RS}, t_i^{RS}) \geq \max_{\phi \in \Phi/\{RS\}} \{U_\phi(x_i, c_i^\phi, t_i^\phi)\}$. Note that this formulation applies to scenarios where one or more travel modes that are unavailable to the customer, in which case high values may be assigned to c_i^ϕ or t_i^ϕ, or both. Let δ_i be a binary decision variable indicating whether customer i chooses to ride with RS (i.e., $\delta_i = 1$) or not (i.e., $\delta_i = 0$). Therefore, the probability of customer choosing RS as his/her mode of travel (given the origin, destination and personal characteristics of the customer) can be calculated using the MNL formula (Anas, 1983) and given by:

$$P(\delta_i = 1 | o_i, d_i, x_i) = \frac{\exp\left(V_{RS}\left(x_i, c_i^{RS}, t_i^{RS}\right)\right)}{\sum_{\phi \in \Phi} \exp\left(V_\phi\left(x_i, c_i^\phi, t_i^\phi\right)\right)} \tag{3}$$

where c_i^ϕ and $t_i^\phi (\phi \in \Phi)$ are known constants corresponding respectively to the costs and times determined by the origin and destination of the customer.

From the perspective of the service provider, x_i, o_i and d_i are often not directly observable. To address this issue, we propose the following steps for estimating the raw adoption rate. It is worthwhile to note that δ_i depends on the cost and time of travel associated with each mode, which is in turn affected by the locations of the customer's

origin and destination. For example, a customer is more likely to choose to ride with RS if the other travel modes between his/her origin and destination are costly, time-consuming, or inaccessible. Therefore, it is important to differentiate between the raw adoption rates between different origin-destination pairs. The probability of customer i who travels from region $A(k)$ to region $A(j)$ requesting a ride with RS can be derived as follows:

$$P\left(\delta_i = 1 \middle| O_i = k, D_i = j\right)$$
$$= \sum_{o_i \in A(k)} \sum_{d_i \in A(j)} \int_{X_k} P_i\left(\delta_i = 1 \middle| o_i, d_i, x_i\right) P\left(o_i \middle| O_i = k\right) P\left(d_i \middle| D_i = j\right) f_k\left(x_i\right) dx_i \tag{4}$$

$$= \sum_{o_i \in A(k)} \sum_{d_i \in A(j)} \int_{X_k} \frac{\exp\left(V_{RS}\left(x_i, c_i^{RS}, t_i^{RS}\right)\right)}{\sum_{\phi \in \Phi} \exp\left(V_\phi\left(x_i, c_i^\phi, t_i^\phi\right)\right)} \frac{1}{Q^2} f_k\left(x_i\right) dx_i \tag{5}$$

where $P(o_i \mid O_i = k)$ $(P(d_i \mid D_i = j))$ is the probability that customer i's trip starts from (ends at) o_i (d_i) given that the pick-up (drop-off) station is station $k(j)$, and recall that $f_k(\cdot)$ is the probability density function of potential customers in service region $A(k)$ and X_k is the set of possible values of x_i in $A(k)$. This value can be seen as the raw adoption rate of RS among customers who travel from region $A(k)$ to region $A(j)$. In what follows, we denote $AR_{kj} = P(\delta_i = 1 \mid O_i = k, D_i = j)$ for notational convenience.

Finally, let p_k^j denote the probability that the pick-up station of potential customer is station k and the drop-off station is station j, then the overall raw adoption rate of RS among all customers covered by its service regions can be calculated as:

$$P\left(\delta_i = 1\right) = \sum_{k=1}^{s-1} \sum_{j=k+1}^{s} P\left(\delta_i = 1 \middle| O_i = k, D_i = j\right) p_k^j \tag{6}$$

3.2 *Planning for price and operations*

A well-designed pricing and operating policy allows the service provider to balance the desire to increase demand and revenue with the associated cost. On the one hand, it benefits the customers and increases demand if the service provider either increases the shuttle departure frequency on each route or lowers the price. On the other hand, a higher frequency or a lower price may lead to lower profit margin. Therefore, it is important to assess whether increasing revenue or profit margin is more effective, and whether either objective should be achieved through changing price or operations, or both.

In this section, we develop an optimisation model for jointly optimising the price and operations, i.e., shuttle departure frequency, of the ride-sharing service. For ease of exposition, we consider a linear pricing rule:

$$c_i^{RS} = c_1 d_{O_i D_i} + c_2 \tag{7}$$

where c_1 is rate per unit of distance travelled with RS (d_{xy} represents the distance between x and y), and c_2 is the flat rate charged per ride. Linear pricing is common in practice among various transportation modes used for commuting, e.g., public transit, taxi.

The operating cost consists of two components: a fixed cost per day per vehicle, denoted as F_f, and a variable cost dependent on the number of times each vehicle drives

through the route, denoted as F_v. Examples of the fixed cost include vehicle rental, driver wage, insurance, parking, cleaning and maintenance and data fee. The variable cost includes for example fuel cost and hourly payments to drivers. Total operating time per day is denoted as T, and the total time for completing a trip and back to the starting station is T_R. The fleet size is n. Shuttles depart every β minutes and the capacity of each shuttle is N. Overall customer travel demand (including all transportation modes) from service region of station k to station j per unit time is denoted as D_{kj}. The walking and driving speed are assumed to be constant and denoted as s_w and s_d, respectively. d_{kj} denotes the distances between stations k and j for simplification.

The service provider RS simultaneously chooses c_1, c_2, β and n in order to maximise its profit. Note that not all customer travel requests are guaranteed to be satisfied by RS due to its capacity limitation. To capture this consideration, we also introduce an intermediate decision variable y_{kj} to represent the expected number of requests from station k to station j that are fulfilled by RS. Hence, the optimisation problem of RS can be formulated as:

$$\max_{c_1,c_2,\beta,y_{kj}} \frac{T}{\beta}\left(\sum_{k=1}^{s-1}\sum_{j=k+1}^{s}\left(c_1\times d_k^j + c_2\right)\times y_{kj} - F_v \right) - nF_f \tag{8}$$

s.t.

$$y_{kj} \leq \beta D_{kj}AR_{kj}, \quad \text{for} \quad k=1,...,s-1, \; j=k+1,...,s \tag{9}$$

$$\sum_{k\leq l, j>l} y_{kj} \leq N, \quad \text{for} \quad l=1,...,s-1 \tag{10}$$

$$AR_{kj} =$$

$$\sum_{o_i \in A(k)}\sum_{d_i \in A(j)} \int_{X_k} \frac{\exp\left(V_{RS}\left(x_i, c_1 d_{kj}+c_2, \left(d_{o,k}+d_j\right)/s_w + d_{kj}/s_d\right)\right)}{\sum_{\phi\in\Phi}\exp\left(V_\phi\left(x_i, c_i^\phi, t_i^\phi\right)\right)} \frac{1}{Q^2} f_k\left(x_i\right)dx_i \tag{11}$$

$$\text{for} \quad k=1,...,s-1, \; j=k+1,...,s$$

$$\beta \geq \frac{T_R}{n} \tag{12}$$

$$c_1, c_2 \geq 0, \; y_{kj} \geq 0, \quad \text{for} \quad k=1,...,s-1, \; j=k+1,...,s, \; n\in N \tag{13}$$

The objective function is equal to the profit generated from each route completion (revenue generated minus the variable cost per route completion) multiplied by the total number of times the route is completed, and minus the total fixed cost per day. The first constraint guarantees that the number of fulfilled trips between stations k and j does not exceed the total number of trips requested by customers. The second constraint ensures that the capacity constraint of each shuttle is not violated at each station. The third constraint is derived from the adoption rate model illustrated in Section 3.1, relating the price of service to the adoption rate of the service among potential customers travelling from $A(k)$ to $A(j)$. The fourth constraint assures that the shuttle departure interval is feasible given the shuttle fleet size, i.e., the shuttles have sufficient time to return to the starting station after completing the route in order to follow the schedule.

4   A case study

In this section, we present a case study based on a real world fixed-route ride-sharing service to demonstrate how the model from Section 3 can be applied. We continue to refer to this service as RS. The route analysed in the following operates in NYC and has 18 stations; see Figure 2 for an illustration. Customers can book a ride by specifying pick-up and drop-off stations 10 minutes before departure time during the operating hours of 7:00 AM–10:30 AM and 4:00 PM–7:30 PM on weekdays. Shuttles depart every 10–15 minutes on this route. The current pricing policy is a flat rate of $4/ride.

Figure 2   Route of RS (see online version for colours)
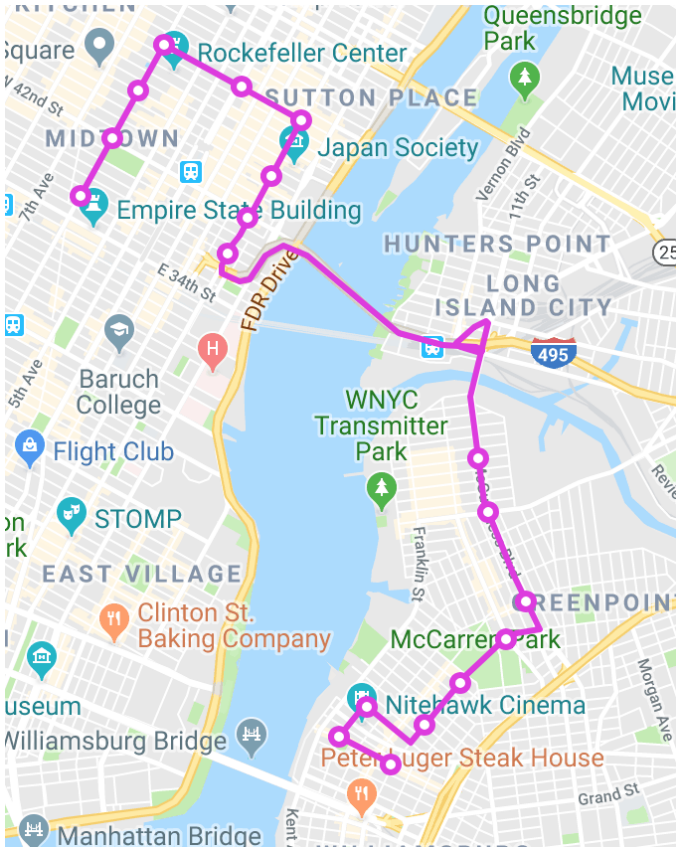
4.1   Datasets

Several datasets are used in this case study. First, the 2010/2011 Regional Household Travel Survey (RHTS) data are used to model the consumer mode-choice model. RHTS data is collected by the New York Metropolitan transportation council and provides travel statistics from fall of 2010 to fall of 2011 in New York, New Jersey and Connecticut. Nearly 19,000 households across 28 counties participated in the survey. Travel information including trip purpose, trip time and distance, and activities during the trip.

Other geographic information and demographic information are also recorded using self-reported data and GPS data. For the majority of our analysis (except for estimating average speed of transportation), RHTS data related to commuter trips within and between Manhattan and Brooklyn were selected because the target customers of RS are commuters fitting this description. In total, 3,020 records of such trip provided information on trip duration, distance, modes, activities, etc. Three travel modes considered include *auto*, *walk* and *transit*, while other modes, i.e., bike, taxi and auto passenger are excluded from consideration due to the scarcity of their records (we also found that the performance of the mode-choice model improves as a result of excluding these modes).

The CTPP 2006–2010 Census Tract Flows (CTF) data is used to estimate the total commuter trip travel demand. This data records total worker counts and its associated margins of error for all tract pairs countrywide including Puerto Rico. Moreover, the FIPS codes are also provided for residence and workplace state, county and census tract. The 2010 American Census Data and the MTA NYC Travel Survey Data are used to estimate the distribution of the demographic and socio-economic factors of potential customers. The research team also has been given the operational cost data of RS.

In addition, we have also collected from various other data sources to supplement the above datasets, such as using the Google Maps API, the details of which will be discussed in the following subsections.

4.2   Parameter estimation

In this section, we estimated the input parameters for the model described in Section 3 for optimising the price and operations of RS.

4.2.1   Travel cost and time

In order to estimate customer's adoption rate using the model described in Section 3.1, we need to estimate the cost and time associated with all modes considered. For the chosen mode of each trip, this information is readily available in the RHTS data. However, the cost and time associated with other (not chosen) modes have to be estimated. We explain the methods for their estimation below.

We first compute the average speed for the three travel modes considered, by directly dividing trip distance used by a given mode recorded in RHTS data by its corresponding trip duration (see Table 1). All recorded trips are used to ensure a sufficient sample size and result reliability.

Table 1   Average speed of travel modes

Alternatives	Transit	Auto	Walk
Avg. speed (mi/min)	0.098	0.117	0.0325

Walk time

Travel time using mode walk (walk time) is estimated directly using average walking speed:

$$Walk\ time = \frac{Trip\ distance}{0.0325\ \text{mi/min}}$$

*Auto time*

Travel time using the auto mode is estimated as:

$$Auto\ time = \frac{Trip\ distance}{0.117\ \text{mi/min}} + Out\ of\ vehicle\ travel\ time$$

where the *Out of vehicle travel time* consists of the walking time to the parking lot and to the workplace, and is assumed to average 5 minutes.

Transit time

Travel time using mode transit (transit time) is more variable and may be affected by a number of factors, such as travel distance, time of the day, accessibility of transit stations and transit frequency. We postulate that the population and income level (of both origin and destination) are key indicators of socio-economic characteristics of an area. For example, the higher the population, the more congestion there may be, while the more accessible transit may be. As a result, transit time may vary from census tract to census tract. To capture this feature, we use census tract as our lowest resolution for population and income, and then uniformly generate a large number of possible origins or destinations using the geographic package in R (points that fall in areas where residence is impossible, e.g., central park, river, are eliminated). We note that only locations in Manhattan are considered for the purpose of estimating transit time, due to the lack of transit time data for trips originating/terminating in Brooklyn. Figure 3 illustrates Manhattan census tracts and uniformly generated origins or destinations considered.

Figure 3   Census tracts and origins/destinations considered for Manhattan, (a) Manhattan census tracts (b) uniformly generated origins or destinations

(a)   (b)
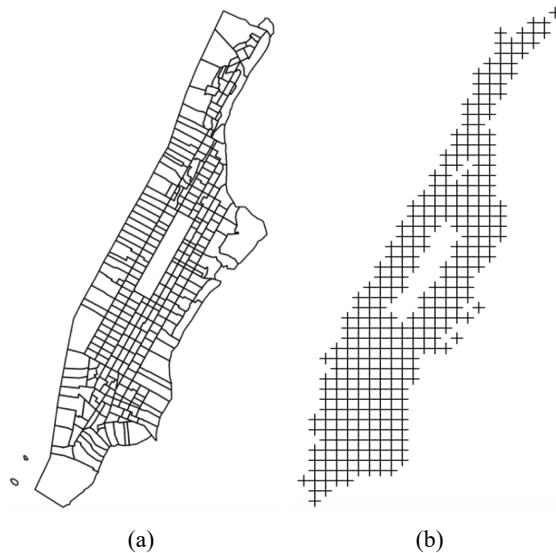
The Google Map direction API is used to obtain the real transit travel times between any two-points in Figure 3(b). We collected times for trips during both rush hours and non-rush hours on weekdays. Weekend trips are not considered because RS does not operate on weekends. We also find the average income and population for the corresponding census tract of each point in Figure 3(b).

Table 2   Regression summary of transit time estimation

Variable	Coefficient	t-value
Distance (mi)	0.0322***	696.5
Average income of origin census tract ($)	0.0017***	120.18
Average income of destination census tract ($)	0.0019***	133.83
Rush hour (Y/N)	0.0176***	63.98
Multiple R^2	0.909	
Number of observations	190,345	

Note: ***significant at the 0.01 level.

The transit travel time is fitted using multiple linear regression in R software. Table 2 summarises the regression output. We can see that the coefficients for all four independent variables, i.e., distance (trip distance), ini_inc (average income of the origin census tract), des_inc (average income of the destination census tract) and rush (a dummy variable indicating whether the trip is taken during rush hours), are significant. The R-squared value of the model is 0.909, which indicates a good fit.

Travel cost

The monetary cost of the each travel mode considered is estimated as follows. The cost of walking is zero. For transit, we estimate the cost using the fare for a subway or local bus ride, which is a fixed flat rate of $2.75/ride. For the auto mode, the cost estimate consists of three components: gas fee, parking fee, as well as a toll fee (source: NYC DOT) for passing the bridge if a customer travelling from Manhattan to Brooklyn or vice versa. The value of each auto mode cost component is provided in Table 3. For example, the cost of a trip from Manhattan to Brooklyn, is estimated as 0.15 × *distance* + 9.5 × *parking time* + 5.76. Here, the parking time is estimated as the full-time working hours per day, i.e., 8 hours.

Table 3   Auto mode travel cost

	Manhattan-Manhattan	Brooklyn-Brooklyn	Intra-Brorough
Gas fee ($/mi)	0.15	0.15	0.15
Parking fee ($/day)	10	9	9.5
Toll fee ($)	0	0	5.76

Note: ***significant at the 0.01 level.

4.2.2 Total travel demand

We first estimate the overall inter-census-tract travel demand based on the CTPP 2006–2010 CTF data of commuter trips. We use the flow between census tracts

provided by CTF data to represent the total travel demand along RS's route. For travel between census tract pairs missing in CTF data and intra-census-tract travel, we supplement the input with travel flow estimates using the gravity model (Anderson, 2011) as detailed below.

Based on the gravity model, we estimate the total travel demand between census tract *i* and census tract *j* to be:

$$T_{ij} = \beta_1 \frac{Pop_i Pop_j Inc_i^{\beta_2} Inc_j^{\beta_3}}{dist_{ij}^{\beta_4}} \qquad (14)$$

where Pop_i (Pop_j) is the population of census tract $i(j)$, Inc_i (Inc_j) is the average income in census tract $i(j)$, and $dist_{ij}$ is the distance between the geometric centres of census tracts i and j. We then fit the above gravity model using the CTF data.

Using the methods described above, we estimate the total potential travel demand for RS by assuming that demand is evenly split between either direction in both mornings and afternoons. We find that a significant portion of travel demand are for intra-census-tract travels, which is consistent with the observation of high proportion of walk mode in the RHTS survey.

We assume that potential origins and destinations are uniformly distributed among 25 points within the service region of each station (see Figure 1 for an illustration). The radius of the service region of each station is assumed to be the maximum distance NYC residents are willing to walk, that is, 0.25 miles (Yang and Diez-Roux, 2012). Origins and destinations in areas where two or more service regions overlap are assigned to the service region of their closest respective station. This is to ensure that no customer gets counted twice and that everyone takes the shuttle from the nearest station. Using the above method, we generate all potential origins and destinations for customers.

4.2.3   Other parameters

Based on the station coordinates, the Manhattan distances between pairs of stations are calculated to represent the distances of travel between RS's stations. The RS shuttle travel time are estimated using the Google Maps API by treating the automobile travel time from one station to another as the corresponding RS shuttle travel time. Finally, we also recognise that a large portion of NYC residence do not own a car (77% according to StatsBee), making the auto mode infeasible for them. To factor in this consideration, we randomly select 77% of the potential trips and assign a very large cost to the auto mode for them, effectively excluding the auto mode from the available modes for these trips.

4.3   The mode-choice model

We consider several demographic and socio-economic factors suggested by Javanmardi et al. (2015) in the initial model, which include age, income, gender, access to transit, vehicle ownership, employment, gross population density, land-use diversity. We also introduce dummy variables for whether the origin/destination is in Brooklyn or not. These factors are estimated based on the American Fact Finder Census 2010 (AFC) and the NYC Travel Survey Data (NTS). After initial model fit, we find that 6 out of 12 factors are insignificant and hence are dropped from the model. Table 4 summarises the above results.

Table 4 Characteristics selection

Demographics	Data source	Significant? (Y/N)
Age	NTS	Y
Income	NTS	Y
Gender	NTS	Y
Origin Brooklyn	AFC	Y
Destination Brooklyn	AFC	Y
Access to transit	AFC	N
Vehicle ownership	AFC	Y
Full-time employed	AFC	N
Gross density of origin census tract	AFC	N
Land-use diversity of origin census tract	AFC	N
Gross density of destination census tract	AFC	N
Land-use diversity of destination census tract	AFC	N

Table 5 Mode-choice model parameters for transit and walk modes, with auto as base mode

Variable	Transit		Walk	
	Coefficient	*t-value*	*Coefficient*	*t-value*
Constant	0.2715	0.7651	0.7844	1.8556
Cost	–0.2148***	–9.0320	–0.2148***	–9.0320
Time	–0.0371***	–21.0925	–0.0371***	–21.0925
Age	0.0020	0.3940	–0.0080	–1.3959
Gender	0.1090	0.9456	0.0189	0.1416
Income	4.0281e-06*	2.5677	9.0963e-07	0.5031
Origin from Brooklyn	–0.5520***	–3.5340	–0.5593*	–2.1360
Destination in Brooklyn	–0.7327***	–4.7394	–0.5328*	–2.0422
Vehicle ownership	–1.0943***	–15.2441	–0.7809***	–9.7836
Log-likelihood	–2,063.8			
McFadden R^2	0.2828			
Likelihood ratio test	Chisq = 1627.8 (p-value ≤ 2.22e-16)			
Number of observations	3,020			

Notes: *significant at the 0.10 level; **at the 0.05 level and ***at the 0.01 level.

After eliminating the insignificant characteristics, we consider trip-specific variables including time and cost, and personal characteristics including, age, income, gender, origin from Brooklyn, destination in Brooklyn and vehicle ownership when fitting the MNL model. Table 5 summarises the best-fitting MNL model. We find that cost and time are both highly significant in mode-choice decisions for all travel modes. As the cost and/or time of a mode increases, the probability of customers choosing the mode decreases. Number of vehicles in the household (VEHNO) is significant for both walk and transit. If the number of vehicles per household increases, a person in that household is more likely to choose auto rather than walk or transit. The origin and/or destination

being in Brooklyn has a significant impact on mode choice, however, it is more significant for choosing transit than choosing walk. Income is significant for choosing transit but not for choosing walk. Age and gender are not significant in all modes. The likelihood ratio test is significant at 99.9% confidence level, suggesting a reliable model.

We performed a ten-fold cross validation (randomly selecting 90% of the data as training data and using the rest 10% as the test data) to examine the prediction accuracy of the fitted mode-choice model, where prediction accuracy is calculated as (number of correct predictions) / (total number of predictions). Table 6 summarises the result. We can see that the prediction accuracy is reasonable overall (59.64–81.45%), and higher for transit (66.67–91.86%) and walk (65.79–80%) modes. Prediction accuracy for auto mode is less reliable, varying from 6.98% to 60.71%. This is a result of the small sample size of the auto mode (566 records out of 3,020 in total). However, the target customer group of RS is commuters in NYC, whose main competition comes from traditional or new forms of public transit and walking. Therefore, the prediction accuracy of auto mode is expected to have limited effect on the analysis.

Table 6   Ten-fold cross validation of mode-choice model

Predict accuracy	Transit	Auto	Walk	Overall
1	84.48%	28.26%	80.00%	74.18%
2	84.15%	33.33%	74.47%	74.18%
3	88.65%	13.16%	69.23%	74.55%
4	84.21%	6.98%	67.21%	68.36%
5	84.12%	20.93%	79.03%	73.09%
6	66.67%	39.13%	65.79%	59.64%
7	89.36%	44.23%	78.05%	77.45%
8	91.86%	42.11%	76.92%	81.45%
9	82.35%	38.78%	75.34%	72.73%
10	83.22%	60.71%	68.75%	73.82%

4.4   Optimal policy

In this section, we discuss the optimal pricing and operational policy for RS. Several competing travel modes are considered, including transit, auto, walk. In addition, we also consider another major competitor of RS – Via, which is another ride-sharing service provider offering corner-to-corner service for $5 flat fee in NYC. Note that because the RHTS data does not contain ride-sharing modes, the utility function of the transit mode is used for both RS and Via due to the similarity between their services. We study three cases:

1   when RS chooses optimal prices only given the current operation policy of 10-minute departure intervals and fleet size of 7 (i.e., the smallest fleet size that enables 10-minute departure interval)

2   when RS simultaneously chooses optimal prices and operational decisions

3   when RS simultaneously chooses optimal prices and operational decisions and prices are flat rates only.

For each case, we examine the performance of the optimal policy in terms of profit and average adoption rate (percentage of customers adopting RS's service among those who know of RS) under varying customer *knowledge levels*, i.e., percentage of people who know about RS, of 10%, 20%, 50% and 100%.

Table 7 Optimal pricing under current operational policy

Knowledge	c_1 ($/mi)	c_2 ($)	β (min)	n	Average adoption rate	Profit ($)
0.1	0.1661	5.2668	10	7	1.35%	−1,322.1
0.2	0.1661	5.2668	10	7	2.71%	−977.1184
0.5	0.1661	5.2668	10	7	6.77%	57.7013
1	0.1661	5.2668	10	7	13.54%	1,782.4

Table 8 Jointly optimal pricing (linear price) and operational policy

Knowledge	c_1 ($/mi)	c_2 ($)	β (min)	n	Average adoption rate	Profit ($)
0.1	0.1887	5.2589	194.1266	1	1.45%	$110.1669
0.2	0.1887	5.2589	97.0633	1	2.90%	$453.1939
0.5	1.5268	4.7659	65	1	5.72%	$1,382.2
1	1.5268	4.7659	32.5	2	11.45%	$2,764.5

Table 9 Jointly optimal pricing (flat rate only) and operational policy

Knowledge	c_1 ($/mi)	c_2 ($)	β (min)	n	Average adoption rate	Profit ($)
0.1	0	5.5707	187.8345	1	1.44%	$109.8949
0.2	0	5.5407	93.9173	1	2.89%	$452.6498
0.5	0	8.5292	65	1	4.11%	$1,269.3
1	0	6.3472	21.6667	3	11.34%	$2,689.4

Tables 7–9 present the optimal policy and performance for the three cases mentioned above. For example, in case 2, when the knowledge level is 5%, the optimal price is a flat rate of $4.7659 and a distance-based rate of $1.5268/mi under the current departure interval. We can also see that the optimal prices do not vary in case 1. This is because in case 1 the shuttles depart so frequently that there is always excessive capacity.

Figure 4 illustrates the comparison between optimal profits in the three cases considered under different knowledge rates. As expected, in each case, the profit increases with knowledge rate. The current operational policy can not profit unless the knowledge rate is higher than around 0.5. Both cases 2 and 3 significantly improve profitability of the service compared to the current operational policy, suggesting that the current operational policy provides excessive capacity. The improvement in profit from including distance-based price is relatively moderate, varying from less than 1% to 8.9%. This perhaps surprising finding is a result of existing competition, that is, the competing modes' pricing policy are predominantly flat-rate only. As a result, the benefit of charging a per-distance price is limited by the reduced competitive advantage. This is further reflected in the phenomenon that the optimal per-distance rate is increasing in the

knowledge level as increased knowledge level enhances RS's competitiveness. These results highlight the advantage of incorporating competing mode choices in the demand model compared to traditional demand function forms.
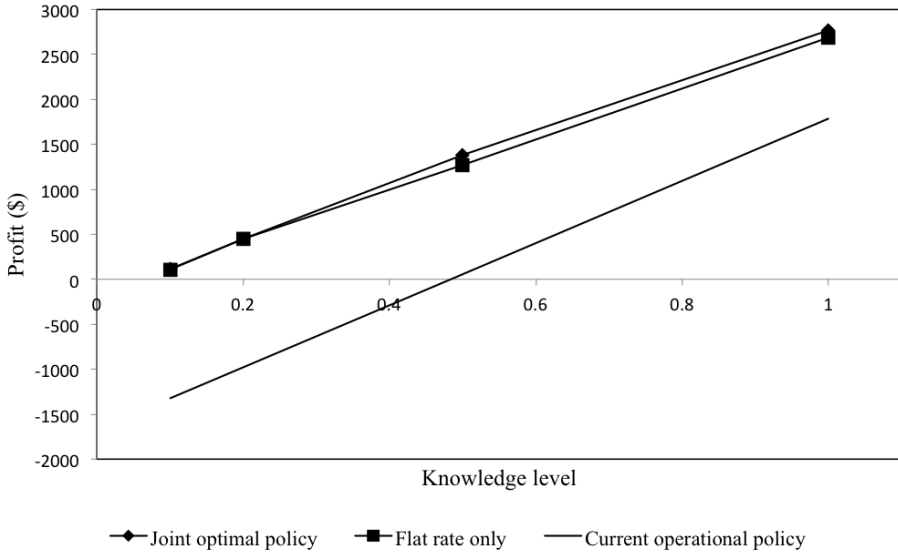
Figure 4   Comparing profits
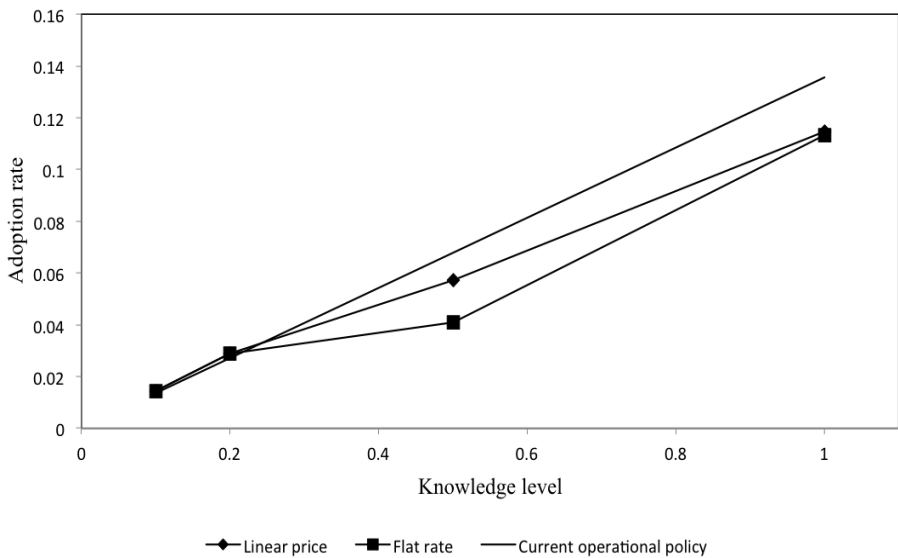

**Figure 5**   Comparing adoption rates

Figure 5 illustrates the comparison between *adoption rates* of RS in the three cases considered under different knowledge rates, where adoption rate of a travel mode is defined as the trips served by the mode divided by the total travel demand that can be served by this mode. Clearly, current policy obtains the highest adoption rate due to the

excessive capacity the service provides. The adoption rates under linear pricing policy are higher than those under flat-rate only policy, especially for intermediate knowledge levels.

Figure 6   Trade-off between profit and adoption rate (at knowledge level 1), (a) varying per-distance price c_1 (fixing c_2 at \$4.7659) (b) varying at rate c_2 (fixing c_1 at \$1.5268/mi)
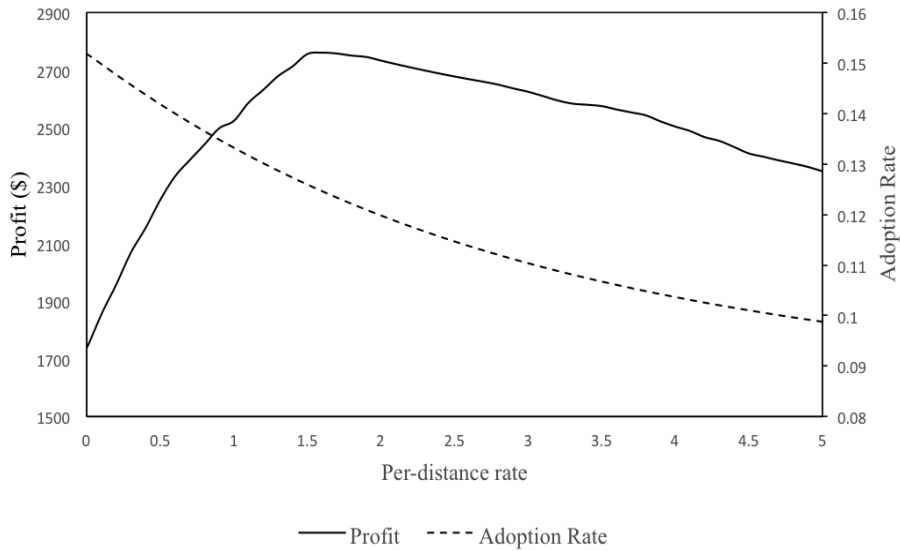
(a)
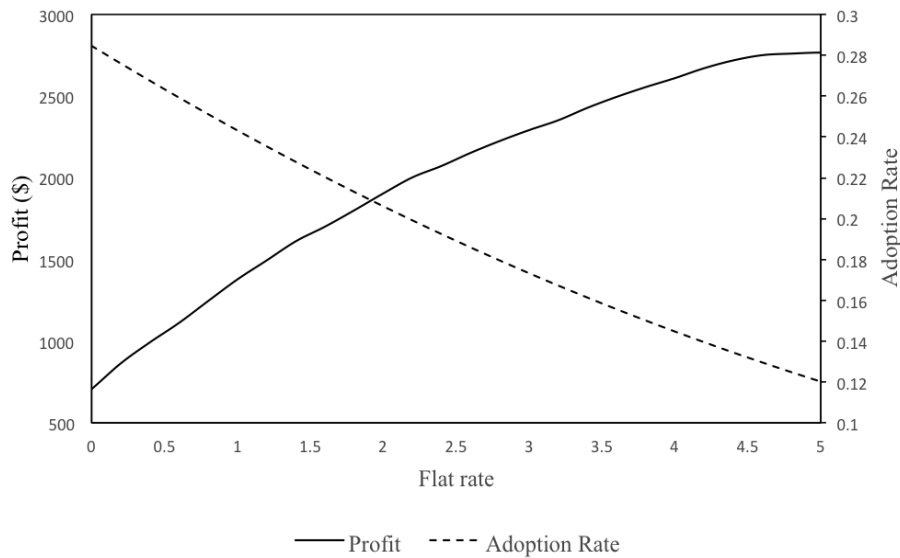
(b)

It is not difficult to see that there exist a trade-off between profit and adoption rate. Figures 6(a) and 6(b) illustrate this trade-off by varying the per-distance rate and flat rate, respectively. We can see that in reducing price from optimality, adoption rate is improved at the expense of profitability. However, interestingly, we observe that the profit decreases at a lower rate than the adoption rate increases in the neighbourhood of the optimal price. This effect is particularly strong when varying the flat rate. For example, reducing the flat rate from $4.8 to $4.6 leads to a 3.5% increase in adoption rate with only 0.4% decrease in profit. This effect suggest an opportunity to significantly increase adoption rate with little compromise on profit, which has important implications as customer adoption is critical to the success of a startup in a competitive environment.

5   Conclusions

In this paper, we investigated the joint pricing and operational policy design problem for a fixed-route ride-sharing service with the objective of maximising profit. To solve the challenge of estimating the demand of the service, we developed a data-driven model based on MNL mode-choice model. Using this model, we then constructed an optimisation model for the joint planning of price, fleet size and shuttle frequency.

Using publicly available travel survey data, we showed that our mode-choice model is effective in predicting customer mode-choices and therefore demand of the service. In a case study of a real-world fixed-route ride-sharing service RS in NYC, we calculated the optimal prices and operational policies for varying levels of commuter knowledge of RS. Our results suggest that jointly optimising price and operational policy significantly improves the profitability of RS's service compared to optimising only price under the current operational policy, and that the optimal departure interval is substantially larger than the current value. We also find that, for RS, having a distance-based price only moderately affect the profitability of the service, and its effect on the adoption rate may be more notable. Moreover, we illustrated the trade-off between RS's profit and adoption rate, and highlighted the opportunity for significantly increasing the adoption rate with little compromise of profit by reducing the flat rate from its optimal value.

References

Agatz, N., Erera, A., Savelsbergh, M. and Wang, X. (2012) 'Optimization for dynamic ride-sharing: a review', *European Journal of Operational Research*, Vol. 223, No. 2, pp.295–303.

Anas, A. (1983) 'Discrete choice theory, information theory and the multinomial logit and gravity models', *Transportation Research Part B: Methodological*, Vol. 17, No. 1, pp.13–23.

Anderson, J.E. (2011) 'The gravity model', *Annu. Rev. Econ.*, Vol. 3, No. 1, pp.133–160.

Arnott, R. (1996) 'Taxi travel should be subsidized', *Journal of Urban Economics*, Vol. 40, No. 3, pp.316–333.

Bellos, I., Ferguson, M. and Toktay, L.B. (2017) 'The car sharing economy: interaction of business model choice and product line design', *Manufacturing & Service Operations Management*, Vol. 19, No. 2, pp.185–201.

Benjaafar, S., Ding, J., Kong, G. and Taylor, T. (2018) *Labor Welfare in On-Demand Service Platforms*, Working Paper, University of Minnesota.

Cachon, J., Daniels, K.M. and Lobel, R. (2017) 'The role of surge pricing on a service platform with self-scheduling capacity', *Manufacturing & Service Operations Management*, Vol. 19, No. 3, pp.368–384.

Cervero, R. (2002) 'Built environments and mode choice: toward a normative framework', *Transportation Research Part D: Transport and Environment*, Vol. 7, No. 4, pp.265–284.

Cervero, R. and Kockelman, K. (1997) 'Travel demand and the 3Ds: density, diversity, and design', *Transportation Research Part D: Transport and Environment*, Vol. 2, No. 3, pp.199–219.

Chang, J., Yu, M., Shen, S. and Xu, M. (2017) 'Location design and relocation of a mixed car-sharing fleet with a CO2 emission constraint', *Service Science*, Vol. 9, No. 3, pp.205–218.

Chang, S. and Chu, C-H. (2009) 'Taxi vacancy rate, fare, and subsidy with maximum social willingness-to-pay under log-linear demand function', *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2111, No. 1, pp.90–99.

Chien, S. and Spacovic, L. (2001) 'Optimization of grid bus transit systems with elastic demand', *J. Adv. Transp.*, Vol. 36, No. 1, pp.63–91.

De Vany, A.S. (1975) 'Capacity utilization under alternative regulatory restraints: an analysis of taxi markets', *Journal of Political Economy*, Vol. 83, No. 1, pp.83–94.

Delle Site, P. and Filippi, F. (1998) 'Service optimization for bus corridors with short turn strategies and variable vehicle size', *Transp. Res. Part A*, Vol. 32, No. 1, pp.19–38.

Deneubourg, J., de Palma, A. and Kahn, D. (1979) 'Dynamic models of competition between transportation modes', *Environment and Planning A*, Vol. 11, No. 6, pp.665–673.

Ding, C., Wang, Y., Tang, T., Mishra, S. and Liu, C. (2018) 'Joint analysis of the spatial impacts of built environment on car ownership and travel mode choice', *Transportation Research Part D: Transport and Environment*, Vol. 60, pp.28–40.

Douglas, G.W. (1972) 'Price regulation and optimal service standards: the taxicab industry', *Journal of Transport Economics and Policy*, Vol. 6, No. 2, pp.116–127.

Feng, X., Zhu, X., Qian, X., Jie, Y., Ma, F. and Niu, X. (2019) 'A new transit network design study in consideration of transfer time composition', *Transportation Research Part D: Transport and Environment*, Vol. 66, pp.85–94.

Frank, L., Bradley, M., Kavage, S., Chapman, J. and Lawton, T.K. (2008) 'Urban form, travel time, and cost relationships with tour complexity and mode choice', *Transportation*, Vol. 35, No. 1, pp.37–54.

Guo, Y., Wang, J., Peeta, S. and Anastasopoulos, P.C. (2018) 'Impacts of internal migration, household registration system, and family planning policy on travel mode choice in China', *Travel Behaviour and Society*, Vol. 13, pp.128–143.

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

He, L., Mak, H-Y., Rong, Y. and Shen, Z-J.M. (2017) 'Service region design for urban electric vehicle sharing systems', *Manufacturing & Service Operations Management*, Vol. 19, No. 2, pp.309–327.

He, Q. (2018) *Incentive-Driven Two-Sided Matching and Capacity Control in Shared-Mobility Systems*, Working Paper, University of North Carolina, Charlotte, NC.

Hu, H., Xu, J., Shen, Q., Shi, F. and Chen, Y. (2018) 'Travel mode choices in small cities of China: a case study of Changting', *Transportation Research Part D: Transport and Environment*, Vol. 59, pp.361–374.

Javanmardi, M., Langerudi, M.F., Shabanpour, R. and Mohammadian, K. (2015) 'Mode choice modelling using personalized travel time and cost data', in *14th Conference of International Association on Travel Behavior Research (IATBR)*, Windsor, UK.

Jha, S.B., Jha, J. and Tiwari, M.K. (2019) 'A multi-objective meta-heuristic approach for transit network design and frequency setting problem in a bus transit system', *Computers & Industrial Engineering*, Vol. 130, pp.166–186.

Kim, K.H. and Lee, Y.H. (2017) 'Vehicle-relocation optimization for one-way carsharing', *International Journal of Industrial Engineering: Theory, Applications and Practice*, Vol. 24, No. 5.

Koppelman, F.S. and Bhat, C. (2006) *A Self Instructing Course in Mode Choice Modeling: Multinomial and Nested Logit Models* [online] https://www.caee.utexas.edu/prof/bhat/COURSES/LM_Draft_060131Final-060630.pdf.

Lampkin, W. and Saalmans, P.D. (1967) 'The design of routes, service frequencies and schedules for a municipal bus undertaking: a case the design of routes, service frequencies and schedules for a municipal bus undertaking: a case study', *Oper. Res. Q.*, Vol. 18, No. 4, pp.375–397.

Lee, Y-J. and Vuchic, V.R. (2005) 'Transit network design with variable demand', *J. Transp. Eng.*, Vol. 131, No. 1, pp.1–10.

Lu, M., Chen, Z. and Shen, S. (2018) 'Optimizing the profitability and quality of service in carshare systems under demand uncertainty', *Manufacturing & Service Operations Management*, Vol. 20, No. 2, pp.162–180.

Mahdavi Moghaddam, S.H., Rao, K.R., Tiwari, G. and Biyani, P. (2019) 'Simultaneous bus transit route network and frequency setting search algorithm', *Journal of Transportation Engineering, Part A: Systems*, Vol. 145, No. 4, p.04019011.

Marwah, B.R., Farokh, S., Umrigar, S. and Patnaik, S.B. (1984) 'Optimal design of bus routes and frequencies for Ahmedabad', *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 994, pp.41–47.

Miller, E.J., Roorda, M.J. and Carrasco, J.A. (2005) 'A tour-based model of travel mode choice', *Transportation*, Vol. 32, No. 4, pp.399–422.

Qi, W., Li, L., Liu, S. and Shen, Z-J.M. (2018) 'Shared mobility for last-mile delivery: design, operational prescriptions, and environmental impact', *Manufacturing & Service Operations Management* [online] https://doi.org/10.1287/msom.2017.0683.

Silman, L.A., Barzily, Z. and Passy, U. (1974) 'Planning the route system for urban buses', *Comput. Oper. Res.*, Vol. 1, No. 2, pp.201–211.

Soehodho, S. and Nahry, S.S. (1998) 'Optimal scheduling of public transport fleet at network fleet', *J. Adv. Transp.*, Vol. 34, No. 2, pp.297–323.

Swait, J. and Ben-Akiva, M. (1987) 'Empirical test of a constrained choice discrete model: mode choice in Sao Paulo, Brazil', *Transportation Research Part B: Methodological*, Vol. 21, No. 2, pp.103–115.

Yang, Y. and Diez-Roux, A.V. (2012) 'Walking distance by trip purpose and population subgroups', *American Journal of Preventive Medicine*, Vol. 43, No. 1, pp.11–19.

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