HOW DOES PRICE AFFECT THE UTILIZATION OF RIDESOURCING SERVICES: EVIDENCE FROM UBER JAPAN'S EXPERIMENTS

Krittanai SRIWONGPHANAWES [a], Daisuke FUKUDA [b]

[a] Department of Civil Engineering, The University of Tokyo
(Bunkyo, Hongo, Tokyo, 113-8656, Japan)
E-mail: krittanai-swpnw@g.ecc.u-tokyo.ac.jp
[b] Department of Civil Engineering, The University of Tokyo
(Bunkyo, Hongo, Tokyo, 113-8656, Japan)
E-mail: fukuda@civil.t.u-tokyo.ac.jp

Ridesourcing or ride-hailing has been introduced and has become popular in many major cities worldwide. The service often relies on the use of dynamic pricing, in which prices are adjusted in real time. Therefore, understanding the impact of price on demand is necessary for the operation of ridesourcing. The aim of this study is to empirically investigate the impact of price on demand through price elasticities using the session data of Uber taxis from Uber Japan's experiments in two cities: Nagoya and Kyoto. A mixed logit model with a flexible mixing distribution was estimated to capture the taste heterogeneity among riders, which increased the reliability of the result. The estimation results indicated that most riders were price inelastic, with an average price elasticity of approximately 0.1 to 0.2. These findings are useful for ridesourcing companies and policymakers because they provide valuable information to enable the maximization of profit or benefit to customers.

Key Words*: *Ridesourcing, price elasticity, revealed preference, choice model

1. INTRODUCTION

Ridesourcing, also known as ride-hailing, is a transportation service in which passengers book rides through a mobile app and connect directly with drivers (Rayle et al., 2016). Advances in technology have caused this service to grow rapidly and play a crucial role in modern transportation systems. Research has demonstrated both positive and negative impacts on transportation systems, with ridesourcing serving as either a substitute or complement to public transport (Cats et al., 2022; Kong et al., 2020; Hall, 2018; Rayle et al., 2016), potentially reducing car ownership (Tang et al., 2019), but also contributing to traffic congestion (Liang et al., 2022). Proper control and implementation are crucial for ridesourcing to benefit the transportation sector, but it is important to consider the unique laws and characteristics of each country because acceptance and use may vary. Further research in specific countries is required to fully understand the characteristics of ridesourcing.

The ridesourcing market involves a firm acting as a platform for two groups of consumers whose demands are interdependent, that is, a two-sided market (Wang and Yang, 2019; Filistrucchi et al., 2014). In numerous studies, researchers have focused on the supply side, that is, customer demand, which is influenced by factors such as service characteristics (price, travel, and waiting time), demographics (age and income), environment, demand management policies, and alternative options (Litman, 2022). Price is a key factor that affects demand and can be measured using price elasticity, which indicates the percentage change in demand that results from a 1% increase in price. Given that ridesourcing pricing is based on surge pricing, where prices change based on supply and demand (Guda and Subramanian, 2019), understanding price elasticity is crucial for understanding customer behavior. However, there have been limited studies on the price elasticity of ridesourcing, and most of them have relied on aggregated or stated preference data, which can led to possibly biased results. To achieve a more accurate and reliable understanding of price elasticity, disaggregated data that captures individual heterogeneity and reflects real behavior are required.

The aim of this study is to fill gaps in previous studies using revealed preference data from Uber's experiments in Nagoya and Kyoto, Japan. A discrete choice model is used to estimate customer behavior when an individual decides whether to request a ride. The study incorporates heterogeneity using a mixed logit model with a flexible semi-nonparametric distribution proposed by Fosgerau and Mabit (2013) to reveal the true population distribution of customer preferences.

Because of the data source, this study is limited to the Japanese context. First, Uber is a relatively new transportation mode in Japan: it has only been operating since 2020 and only in 15 cities. Second, ridesourcing is highly regulated. It is only available through taxi and premium services regulated by Japanese taxi laws, and surge pricing is currently not allowed. This indicates the ability of the government to control the price of ridesourcing. Therefore, the results of this study could be beneficial for countries in which ridesourcing operates in a highly regulated market, such as Japan.

The remainder of this paper is organized as follows: In Section 2, related literature is reviewed. In Section 3, the data and methodology used in this study are described. The estimation results for the choice models and price elasticities are presented in Section 4. Finally, in Section 5, a conclusion is provided, and future studies are proposed.

2. LITERATURE REVIEW

In this study, two main topics are considered: ridesourcing and price elasticity. Hence, papers related to these topics are summarized in this section.

(1) Ridesourcing

Ridesourcing is a popular alternative in urban transportation worldwide that relies on smartphone apps and wireless communication. The companies that provide this service call themselves transportation network companies (TNCs), for example, Uber, Lyft, and DiDi. To use the service, customers are required to specify the origin and destination, departure time, and vehicle type. The application then calculates and presents the customer with the estimated trip characteristics, such as the fare, waiting time, and travel time. Based on this information, the customer decides whether to request a ride. If the customer requests a ride, nearby drivers are dispatched. Once at least one driver has accepted the request, that driver picks up the passenger and drives the passenger to the destination. From the characteristics of the operation, the ridesourcing market is two-sided, that is, a firm acts as a platform for both riders and drivers, whose demands are interdependent (Wang and Yang, 2019; Filistrucchi et al., 2014).

Although ridesourcing is still a relatively new transportation mode, research devoted to the study of its users' characteristics and trip characteristics is increasing.

a) Users' characteristics

In most related studies, researchers have reported that most ridesourcing users tend to be young and well-educated (Chalermpong et al., 2022; Young and Farber, 2019; Alemi et al., 2018; Dias et al., 2017; Rayle et al., 2016). In some specific studies, such as those conducted by Alemi et al. (2018) and Dias et al. (2017), researchers also found that the adoption of ridesourcing is significantly related to the smartphone and tech savviness of a person. By contrast, Tang et al. (2019) found no significant differences in the educational backgrounds of ridesourcing users in China. This may indicate that the characteristics of ridesourcing users differ according to each country's environment and context.

Regarding the reasons for using ridesourcing, comfort and safety are the most valued attributes (Tirachini and del Río, 2019). Furthermore, the results from the study conducted by Young and Farber (2019) suggested that people use ridesourcing to avoid drunk driving.

b) Trip characteristics

Researchers have mostly observed the characteristics of ridesourcing trips using surveys. Some results are as follows: First, ridesourcing trips are flexible because they have no fixed schedules or routes. Second, the travel times are reported to be less than 30 minutes per trip (Tang et al., 2019). Third, people use ridesourcing mainly for social and recreational trips (Lavieri and Bhat, 2019; Clewlow and Mishra, 2017; Rayle et al., 2016). Tirachini and del Río (2019) found that in 40% of cases, people who use ridesourcing travel alone, whereas in 35% of the cases they travel with another person.

The waiting time for ridesourcing is worth explaining separately. Compared with trips using a similar mode, that is, conventional taxis, ridesourcing trips are more reliable in terms of waiting time: they are remarkedly shorter and more consistent (Zheng et al., 2019). It has been reported that ridesourcing users always experience a waiting time of 10 minutes or less (Chalermpong et al., 2022; Rayle et al., 2016)

(2) Transportation price elasticity

The main objective of this study is to understand how ridesourcing customers respond to price. Consumers of any goods or services are price sensitive, which means that price plays an important role in their consumption. Therefore, when the price is high, consumption is low, and *vice versa*. The demand for transportation products also follows this pattern.

The price elasticity of demand is a commonly used tool to study the price sensitivity of consumers (Litman, 2022). It measures the percentage change in demand caused by a 1% change in price. The magnitude of price elasticity indicates how sensitive a customer is to price changes. A value between -1 and 1 indicates that the customer is price inelastic, which means that price has a limited effect on consumption. A value outside this interval suggests that the customer is price elastic, which means that price has a significant impact on consumption.

Price elasticities are crucial for transportation planning, particularly pricing. Transportation planners and policymakers must understand how users will respond to various pricing scenarios for effective policy and goal achievement (Han and Li, 2009). For ridesourcing and conventional taxis, price elasticities are part of many analyses, such as estimating consumer surplus (Cohen et al., 2016), evaluating optimal prices (Gómez-Lobo et al., 2022), and evaluating spatial pricing (Ata et al., 2019). These examples show the significance of price elasticities and their applications.

a) Price elasticities of conventional taxis and ridesourcing

There is a small yet growing literature on the price elasticities of ridesourcing. Madanizadeh et al. (2022) used geographical regression discontinuity design to estimate the short-run and long-run price elasticities of Tapsi, a TNC in Iran. They found them to be -0.25 and -0.54, respectively. Gómez-Lobo et al. (2022), although not the main objective of their study, found that the demand for ridesourcing was elastic with a price elasticity of -2.22. Lam et al. (2021) used the discrete choice model to investigate demand elasticities and found that people are less sensitive to price during rush hours. Castillo (2018), using UberX data in Houston, found that the average price elasticity of the sample was 0.186. Cohen et al. (2016) used regression discontinuity design to estimate the price elasticities of UberX in four US. markets and found that the mean point estimate of the elasticity is -0.57. They also found that demand is most elastic during the day on weekends and least elastic during off-peak hours on weekdays.

The literature about the price elasticity of conventional taxis is also reviewed since both transportation modes have some common attributes. Ata et al. (2019) showed that the price elasticity of taxis in New York City was -0.4735. Rose and Hensher (2013) found that when the taxi fare increases by 10%, the price elasticity equals -1.042. Toner et al. (2010), using stated preference and transfer price experiments in four English cities, estimated the price elasticities to be between -0.8 and -1.4. Booz Allen Hamilton (2003) suggested that internationally average taxi price elasticities are most likely between -0.3 and -0.8.

This study differs from past literature by using a mixed logit model, a disaggregate approach, to estimate price elasticities instead of aggregate approaches. Unlike most studies, the model did not consider the effect of surge pricing, leading to potentially different results. The results were also expected to be between the price elasticity of conventional taxis and ridesourcing since the focus of this research was taxi-type ridesourcing.

3. DATA

(1) Outline of the data

Data for this study were obtained from Uber's price elasticity testing in Nagoya and Kyoto, Japan. Participants were randomly assigned to either a treatment group with discounts of 10%, 20%, or 40% or a control group with regular pricing. Customers in the treatment group were informed of their discount through in-app notifications, e-mails, and during the session.

The data contained information on all sessions of the Uber taxi service over the period February 15, 2021 to April 11, 2021. The estimated fare (maximum and minimum), travel time, and waiting time were available in these data, together with the date and time the sessions occurred, the location of the pick-up and drop-off point as a distance from the city center, and the trip identifier (id) corresponding to the session if it was requested. Additionally, the data contained all sessions created by each rider, which could be used as panel data. The original data contained 64,557 sessions and 8,513 unique riders for Nagoya, and 69,770 sessions and 10,065 unique riders for Kyoto.

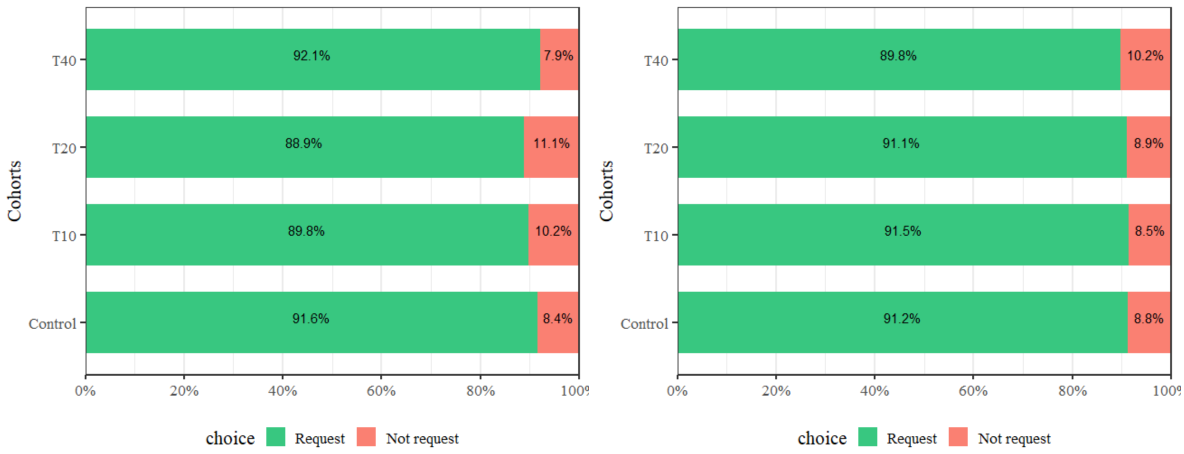
a) Data cleaning

The data were cleaned using the following process:

- Unusually long and short distance sessions were removed using an adjusted box plot for skewed distributions (Hubert and Vandervieren, 2008).
- Sessions that had a total online time, that is, the time spent on the application from opening the application to requesting a ride or closing the application, more than 40 minutes were removed.
- Sessions were removed that contained incomplete data, such as a successful session without the estimated fare or estimated waiting time.
- Requested sessions were removed if the riders

Table 1 Rider characteristics' descriptive statistics

Characteristic	Value	Sample frequency	
		Nagoya (N = 4,727)	Kyoto (N = 6,129)
Cohort	Control	33.56%	33.22%
	Treatment 10%	39.84%	40.62%
	Treatment 20%	18.28%	17.89%
	Treatment 40%	8.31%	8.27%
Frequency	Less than 10 times	71.24%	73.8%
	10 – 20	15.13%	15.71%
	20 – 30	6.01%	5.35%
	30 – 40	3.68%	2.33%
	40 – 50	1.65%	1.16%
	50 – 60	1.06%	0.91%
	60 – 70	0.53%	0.24%
	70 – 80	0.36%	0.18%
	80 – 90	0.11%	0.16%
	90 – 100	0.13%	0.03%
	more than 100 times	0.11%	0.11%



(a) (b)

Fig. 1 Aggregation of choices by cohort in (a) Nagoya and (b) Kyoto

canceled the request afterward because the reason for the cancellation was not observed.

- Sessions were removed for riders who had fewer than two sessions and fewer than one requested session. This was to ensure that the data were panel data and the effect of price on requesting a ride could be observed.

After the data were cleaned, there were 46,367 sessions and 4,726 unique riders for Nagoya, and 53,813 sessions and 6,129 unique riders for Kyoto.

b) Choice extraction

The choice of whether the session was requested was not stated directly in the data; however, it could be extracted using the corresponding trip id. The trip id is a unique identifier that is generated only when the session is requested, regardless of the success of the real trip, that is, trip cancellation from the rider or driver. By contrast, when the session is not requested, this trip id is blank. Using this fact, the choice of whether the session was requested can be extracted from the data.

(2) Descriptive analysis of the data

Table 1 summarizes the descriptive statistics of rider characteristics: cohort and frequency. In Nagoya and Kyoto, the treatment group with a 10% discount had the most percentage, followed by the control group, the treatment group with a 20% discount, and the treatment group with a 40% discount. During the experiment period, more than 70% of the riders created sessions fewer than ten times, whereas only around 0.1% created sessions more than 100 times. This might be because Uber and taxis were not commonly used in Japan.

The descriptive statistics of the session characteristics, such as day of the week, time of day, weather, estimated trip characteristics, and the result of the session (requested or not), are summarized in Table 2. Sessions were evenly distributed throughout the week, with the most frequent sessions occurring on Fridays and Saturdays, and they were created mostly from 8 AM to 10 AM and 4 PM to 10 PM. Regarding the weather, most sessions occurred when it was not raining. Regarding the estimated trip

Table 2 Session characteristics' descriptive statistics

Characteristics	Value	Sample frequency	
		Nagoya (N = 46,367)	Kyoto (N = 53,813)
Day of the week	Monday	13.45%	12.87%
	Tuesday	14.16%	13.93%
	Wednesday	13.98%	13.81%
	Thursday	14.32%	14.27%
	Friday	14.97%	15%
	Saturday	15.71%	15.71%
	Sunday	13.41%	14.42%
Time of day	6AM – 8AM	4.31%	5.48%
	8AM – 10AM	11.3%	12.73%
	10AM – 12PM	9.27%	9.88%
	12PM – 14PM	8.28%	8.76%
	14PM – 16PM	7.86%	8.12%
	16PM – 18PM	12.2%	12.2%
	18PM – 20PM	15.12%	12.99%
	20PM – 22PM	12.89%	12.14%
	22PM – 24PM	7.99%	7.73%
	Early morning	10.78%	9.96%
Weather	Raining	20.03%	25.28%
	Not raining	79.97%	74.72%
Estimated fare (JPY)	Less than 1,000	58.73%	53.95%
	1,000 – 2,000	30.30%	36.76%
	2,000 – 3,000	7.25%	7.86%
	3,000 – 4,000	2.99%	1.21%
	4,000 – 5,000	0.63%	0.21%
	More than 5,000	0.09%	0.01%
Estimated travel time (minute)	Less than 10	64.02%	46.60%
	10 – 20	32.35%	46.63%
	20 – 30	3.45%	6.23%
	More than 30	0.18%	0.54%
Estimated waiting time (minute)	1	0.68%	2.75%
	2	16.87%	24.22%
	3	20.68%	24.62%
	4	17.51%	18.7%
	5	13.54%	11.73%
	6	9.99%	7.52%
	7	7.71%	4.49%
	8	5.8%	3.28%
	9	4.14%	1.72%
	10	3.08%	0.97%
Estimated travel distance (kilometer)	Less than 2	40.17%	23.04%
	2 – 4	41.23%	40.93%
	4 – 6	10.01%	22.27%
	6 – 8	5.08%	9.29%
	8 – 10	2.61%	3.36%
	More than 10	0.90%	1.12%
Result	Requested	90.43%	91.18%
	Not requested	9.57%	8.82%

characteristics, the estimated fares were similar, but the estimated travel time and distance showed that riders in Kyoto made longer trips. Additionally, the estimated waiting time suggested that most riders in Kyoto experienced shorter waiting times. This suggests that there were more drivers in Kyoto than Nagoya. Regarding the result of the session, most of the sessions resulted in riders requesting a ride.

a) Aggregation of choices by cohort

Fig.1 shows the share of choices segmented by the cohorts of the riders to observe the effect of the experimental condition on the dependent variable. The results demonstrated that in Nagoya and Kyoto, the percentage of requested sessions for the treatment groups and control group were not significantly different, and the percentage for the control group was even larger than that for some of the treatment groups. This indicated that knowing that the prices were discounted did not have significant effect on the choices.

On the other hand, Table 3 shows the average number of sessions per rider during the experiment period segmented by cohort. It can be seen that the riders in the cohort with higher discount tended to create more sessions. This might be because while discounts encouraged riders to open the application

Table 3 Average number of sessions per rider segmented by cohort

Cohort	Average sessions per rider	
	Nagoya	Kyoto
Control	9.64	8.23
Treatment 10%	9.45	8.66
Treatment 20%	9.94	9.4
Treatment 40%	12.64	10.86

more frequently, the decision whether to request a ride depended on other factors.

4. METHODOLOGY

(1) Model specification

In this study, the effect of price and other factors on ridesourcing is investigated by examining riders' decision-making when using the service. A random coefficient binary mixed logit model is proposed to analyze the riders' choice of whether to request a ride based on estimated trip characteristics, weather conditions, and past experiences. The model accounts for heterogeneity in riders' preferences through a flexible nonparametric mixing distribution proposed by Fosgerau and Mabit (2013).

a) Mixed logit model

To investigate the behavior of riders, binary mixed logit models in which the unobserved inter-individual taste heterogeneity is observed through the random coefficients are estimated.

Suppose that individual n encounters a choice among two alternatives in each of T_n choice scenarios. The utility that individual n obtains from alternative j in choice scenario t is:

$$U_{njt} = \beta'_n x_{njt} + \varepsilon_{njt} \tag{1}$$

where,

x_{njt} : a vector of the observed explanatory variables,

β_n : a vector of the preference coefficients of rider n, and

ε_{njt} : the unobserved utility.

By specifying the unobservable factors to be distributed independent and identically distributed extreme value over time, individual, and alternatives, suppose β_n is known, the conditional choice probability of individual n choosing alternative j in situation t can be expressed as a logit probability:

$$L_{njt} = \frac{e^{V_{njt}}}{e^{V_{njt}} + 1} \tag{2}$$

However, β_n is unknown. The unconditional choice probabilities are the integrals of the conditional probabilities over the population density, $g(\beta|\theta)$, which depends on hyperparameter θ. Hence, the choice probability of choosing alternative j becomes:

$$P_{njt} = \int L_{njt}\, g(\beta|\theta) d\beta \tag{3}$$

As panel data, each individual makes a decision a total of T_n times. Suppose that individual n makes a sequence of choices, denoted by $y_n = \{y_{n1}, \dots, y_{nT_n}\}$. Conditioned on β_n, the probability that individual n makes a sequence of choices y_n is:

$$\boldsymbol{L}_{ny_n} = \prod_{t=1}^{T_n} \left(L_{nty_{nt}}\right) \tag{4}$$

Therefore, the unconditional probability is

$$P_{ny_n} = \int \boldsymbol{L}_{ny_n}\, g(\beta|\theta) d\beta \tag{5}$$

Based on this setting, the log-likelihood function is specified as

$$LL = \sum_{n=1}^{N} ln \int \prod_{t=1}^{T_n} \prod_{j=1}^{2} L_{njt}{}^{d_{njt}}\ g(\beta_n|\theta)\, d\beta \tag{6}$$

where,

d_{njt} equals 1 if individual n chooses j in choice scenario t and 0 otherwise.

Equation (6) does not have a closed form; therefore, it must be simulated. The simulated log-likelihood is expressed as:

$$SLL = \sum_{n=1}^{N} ln \left(\frac{1}{R} \sum_{r=1}^{R} \prod_{t=1}^{T_n} \prod_{j=1}^{2} L_{nit}{}^{d_{nit}} \right) \tag{7}$$

where,

R : the number of draws.

b) Utility function specification

The observable utility of choosing to request a ride is specified as a linear formulation. Because the main investigation considers the effect of price, the model is specified as an interaction model in which the price coefficient is affected by other variables. Additionally, the lagged dependent variable, that is, the choice in the previous session, is introduced into the model to represent the inertia of riders' behavior. The observable utility of rider n from choosing to request a ride in choice scenario t is:

$$V_{nt} = ASC + \beta_{F_n} F_{nt} + \beta_{TT} TT_{nt} + \beta_{WT_n} WT_{nt} + \gamma_{WE} WE_{nt} + \gamma_{PK} Peak_{nt} + \gamma_R Rain_{nt} + \gamma_y y_{nt-1} \quad (8)$$

where,

ASC : the alternative specific constant for choosing to request a ride,

F_{nt} : the average of the maximum and minimum estimated fare of choice scenario t presented to rider n (1000 JPY),

TT_{nt} : the estimated travel time of choice scenario t presented to rider n (minutes),

WT_{nt} : the estimated waiting time of choice scenario t presented to rider n (minutes),

TD_{nt} : the estimated travel distance of choice scenario t presented to rider n (kilometers),

WE_{nt} : the dummy variable equals 1 if choice scenario t of rider n occurs on weekends and 0 otherwise,

$Peak_{nt}$: the dummy variable equals 1 if choice scenario t of rider n occurs during public transportation's peak hours and 0 otherwise,

$Rain_{nt}$: the dummy variable equals 1 if it was raining when choice scenario t of rider n occurred and 0 otherwise,

y_{nt-1} : the dummy variable equals 1 if the previous session was requested and 0 otherwise; initial value y_{n0} is zero, and

β and γ : the vector of coefficients to be estimated.

Subscript n indicates that the coefficient is random.

c) Observed heterogeneity in price preferences

The observed heterogeneity in price preferences is captured by considering the interaction between price and other factors, that is, day of the week (weekends or weekdays), time of day (during peak hours or off-peak hours), and weather (raining or not raining), and time-distance ratio. As a result, the price coefficient of individual n, β_{F_n} is:

$$\beta_{F_n} = \bar{\beta}_{F_n} + \beta_{F,WE} WE_{nt} + \beta_{F,PK} Peak_{nt} + \beta_{F,R} Rain_{nt} + \beta_{F,TD} \left(\frac{TT}{TD}\right)_{nt} \quad (9)$$

d) Unobserved heterogeneity

To capture the unobserved heterogeneity in price and waiting time preference, their coefficients are assumed to follow some distributions. Initially, parametric distributions, such as normal and log-normal, were considered. However, the limitations of these distributions prompted the adoption of a more flexible distribution. Therefore, a semi-nonparametric distribution proposed by Fosgerau and Mabit (2013) is adopted. This method is easy to implement while maintaining performance when representing any distribution to any desired level. The coefficients can be computed by transforming draws from some distribution using a power series:

$$\beta = f(u|\alpha) = \sum_{k=0}^{K} \alpha_k u^k \quad (10)$$

where,

$\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_K)$: hyperparameters to be estimated.

However, the shape of the distribution obtained from Fosgerau and Mabit's approach relies heavily on the data, which allows the coefficients to be positive for price and waiting time. Therefore, a negative log-normal prior distribution controlled by the polynomial is also specified to ensure that the coefficients for price and waiting time are negative, as suggested by Lehmann et al. (2022):

$$\beta = -exp\left(\sum_{k=0}^{K} \alpha_k u^k\right) \quad (11)$$

(2) Price elasticity estimation

Generally, the mixed logit price elasticity of individual n when the individual encounters choice scenario t can be calculated using (11):

$$E_{x_{nt}} = F_{nt} \cdot \frac{\int \beta_{F_n} \cdot L_{nt,request}\left(1 - L_{nt,request}\right) g(\beta_n|\theta) d\beta}{\int L_{nt,request}\, g(\beta_n|\theta) d\beta} \quad (12)$$

The choices of each individual are observed as panel data, which allows the estimation of the conditional posterior distribution. The conditional posterior distribution is a distribution of taste among a group of individuals in the population who would make the same sequence of choices when encountering the same scenarios and provides better information than the population distribution. The conditional posterior distribution can be estimated by conditioning the population distribution on the sequence of choices of each individual. Using Bayes rule, the conditional posterior distribution of individual n is:

$$h(\beta_n|y_n, x_n, z_n, \theta) = \frac{P(y_n|x_n, z_n, \beta_n) g(\beta_n|\theta)}{P(y_n|x_n, z_n, \theta)} \quad (13)$$

where,

y_n : a vector of the observed sequence of choices of rider n,

x_n : a vector of the estimated trip characteristics

Table 4 Model performance (Nagoya)

	BL	MXL 1	MXL 2	MXL 3	MXL 4
Distribution of $\bar{\beta}_{F_n}$	Fixed	Normal	Log-normal	Polynomial	Log-polynomial
Distribution of β_{WT}	Fixed	Normal	Log-normal	Log-polynomial	Log-polynomial
No. of parameters	12	14	14	17	17
Log-likelihood	-13,828.27	-13,362.79	-13,421.79	-13,409.02	-13,396.13
Adjusted $\bar{\rho}^2$	0.5694	0.5838	0.5819	0.5823	0.5827
AIC	27,680.54	26,753.59	26,871.57	26,852.04	26,826.27
BIC	27,785.47	26,876.01	26,993.99	27,000.7	26,974.92

Table 5 Model performance (Kyoto)

	BL	MXL 1	MXL 2	MXL 3	MXL 4
Distribution of $\bar{\beta}_{F_n}$	Fixed	Normal	Log-normal	Polynomial	Log-polynomial
Distribution of β_{WT}	Fixed	Normal	Log-normal	Log-polynomial	Log-polynomial
No. of parameters	12	14	14	17	17
Log-likelihood	-15,391.58	-14,783.16	-14,895.32	-14,822.52	-14,822.13
Adjusted $\bar{\rho}^2$	0.587	0.6033	0.6003	0.6022	0.6022
AIC	30,807.16	29,594.33	29,818.65	29,679.04	29,678.27
BIC	30,913.88	29,718.83	29,943.15	29,830.23	29,829.45

of each choice scenario, and

z_n : a vector of other observed factors for each choice scenario.

Population density in (12) is substituted by conditional posterior distribution, which requires simulation for the estimation:

$$E_{x_{nt}} = x_{nt} \cdot \frac{\int \beta_{F_n} \cdot L_{nt,request}(1 - L_{nt,request}) h(\beta_n | y_n, x_n, z_n, \theta)\, d\beta}{\int L_{nt,request} h(\beta_n | y_n, x_n, z_n, \theta)\, d\beta} \tag{14}$$

With reference to the method in the book by Train (2009), the elasticities can be simulated as:

$$\widetilde{E}_{F_{nt}} = F_{nt} \cdot \frac{\sum_R w^r \cdot \beta_{F_n}{}^r \cdot L_{nt,request}{}^r (1 - L_{nt,request}{}^r)}{\sum_R w^r \cdot L_{nt,request}{}^r} \tag{15}$$

where,

R : the number of draws, and

w^r: the weight for draw $\beta_n{}^r$

$$= \frac{P(y_n | x_n, z_n, \beta_n{}^r)}{\sum_R P(y_n | x_n, z_n, \beta_n{}^r)}.$$

The aggregated direct elasticities can be calculated as follows using sample enumeration:

$$\overline{\widetilde{E}_{F_{nt}}} = \frac{\sum_{n=1}^{N} P_{nt} \widetilde{E}_{F_{nt}}}{\sum_{n=1}^{N} P_{nt}} \tag{16}$$

5. ESTIMATION RESULTS

In this section, the results of the analyses performed in this study are reported and discussed, which includes a descriptive analysis of the data, choice model estimation results, and price elasticity estimation results.

(1) Choice model estimation results

The software package Apollo 0.2.8 (Hess and Palma, 2019) was used for the maximum simulated likelihood estimator, with 1,000 Halton draws used in the simulation process. Furthermore, the random coefficients were assumed to be inter-individual coefficients.

Five models were estimated and compared. The first model was a standard binary logit model (BL) for which all the coefficients were fixed. To incorporate unobserved taste heterogeneity in the model, four mixed logit models were introduced, where the coefficient of the estimated fare and waiting time were assumed to be randomly distributed. The distributions of MXL 1 and MXL 2 were assumed to be normal and negative log-normal, respectively. For MXL 3 and MXL 4, Fosgerau and Mabit's polynomial-estimated distributions were introduced.

For both models, the distribution of the estimated waiting time coefficient was assumed to be a negative log-normal prior distribution, and Fosgerau and Mabit's (2013) polynomial was used to control the heavy tail property (Lehmann et al., 2022). The models differed in the distribution of the estimated fare coefficient. The distribution of MXL 3 was assumed to be the polynomial-estimated distribution directly, whereas the distribution of MXL 4 was transformed into a negative log-normal similar to the estimated waiting time coefficient to bound the distribution in exchange for less flexibility. The dimensions of the polynomials were set to K_F = 3 and K_{WT} = 2, and u was drawn from a standard normal distribution because this specification was the most appropriate considering the resulting shape of the distributions and estimation stability.

Tables 4 and 5 show a comparison of the performance of each model specification for Nagoya data and Kyoto data, respectively. The value of the log-likelihood and adjusted $\bar{\rho}^2$ for all models demonstrated satisfactory goodness of fit. The mixed

Table 6 Estimation results: semi-nonparametric distributions

	BL		MXL 3		MXL 4	
	Estimate	Rob. s.e.	Estimate	Rob. s.e.	Estimate	Rob. s.e.
Nagoya						
ASC: request	2.347**	0.077	2.743**	0.095	2.854**	0.083
Estimated trip characteristics						
Fare (1000 JPY)						
Fixed parameter	-0.344**	0.054	—	—	—	—
Random parameter						
$\alpha_{F,0}$	—	—	-0.735**	0.077	-0.211	0.113
$\alpha_{F,1}$	—	—	0.522*	0.212	-0.472	0.645
$\alpha_{F,2}$	—	—	0.942**	0.186	-2.476**	0.469
$\alpha_{F,3}$	—	—	-0.555**	0.041	-0.530**	0.106
Observed heterogeneity						
Weekends	0.052	0.038	0.082	0.049	-0.108	0.056
Peak hours	-0.064	0.058	-0.115	0.068	0.214**	0.069
Raining	0.066	0.043	0.067	0.056	-0.045	0.058
time-distance ratio	-0.003	0.038	-0.127*	0.045	0.229**	0.021
Travel time (minute)	-0.021	0.119	0.019	0.014	0.032*	0.012
Waiting time (minute)						
Fixed parameter	-0.034**	0.009	—	—	—	—
Random parameter						
$\alpha_{WT,0}$	—	—	-2.477*	1.101	-1.41**	0.213
$\alpha_{WT,1}$	—	—	-11.389*	4.527	-6.566*	2.479
$\alpha_{WT,2}$	—	—	-24.457**	6.531	-33.267**	9.568
Other factors						
Weekends (ref = weekdays)	-0.232**	0.063	-0.216**	0.072	-0.242**	0.072
Peak hours (ref = off-peak hours)	0.446**	0.081	0.484**	0.093	0.625**	0.101
Raining (ref = not raining)	-0.183*	0.069	-0.168*	0.079	-0.136	0.079
Previous choice (ref = not requested)	0.937**	0.402	0.398**	0.047	0.362**	0.05
Kyoto						
ASC: request	2.180	0.083	2.37**	0.092	2.738**	0.082
Estimated trip characteristics						
Fare (1000 JPY)						
Fixed param.	-0.195	0.067	—	—	—	—
Random param.						
$\alpha_{F,0}$	—	—	-0.763**	0.121	-1.611**	0.250
$\alpha_{F,1}$	—	—	-0.574	0.64	-2.329**	0.476
$\alpha_{F,2}$	—	—	1.42**	0.377	-1.081**	0.355
$\alpha_{F,3}$	—	—	0.608**	0.065	-0.176*	0.076
Observed heterogeneity						
Weekends	-0.01	0.050	-0.005	0.061	-0.08	0.059
Peak hours	-0.081	0.058	-0.166*	0.073	0.101	0.073
Raining	0.016	0.051	0.04	0.064	-0.053	0.054
time-distance ratio	-0.011	0.042	-0.053	0.045	0.198**	0.026
Travel time (minute)	-0.013	0.009	0.006	0.01	0.034**	0.008
Waiting time (minute)						
Fixed param.	-0.054	0.009	—	—	—	—
Random param.						
$\alpha_{WT,0}$	—	—	-12.006**	3.077	-8.317**	2.025
$\alpha_{WT,1}$	—	—	17.761**	.572	-29.197**	8.025
$\alpha_{WT,2}$	—	—	-7.083**	1.783	-28.497**	7.865
Other factors						
Weekends (ref = weekdays)	-0.094	0.070	-0.048	0.08	-0.147	0.07
Peak hours (ref = off-peak hours)	0.346	0.079	0.415**	0.093	0.323**	0.086
Raining (ref = not raining)	-0.041	0.072	-0.070	0.083	-0.091	0.068
Previous choice (ref = not requested)	1.015	0.039	0.412**	0.045	0.391**	0.047

— Not relevant; ** Significant at 1% level; * Significant at 5% level.

logit models improved model performance as expected. Among the mixed logit models, MXL 1 had the best goodness of fit although it produced positive price coefficients, followed by MXL 4. Furthermore, MXL 4 was significantly superior to MXL 2 in terms of goodness of fit, which indicates that using a polynomial to control the heavy tail property of the log prior distribution was effective.

Table 6 shows the estimation results for the BL and two flexible MXL models: MXL 3 and MXL 4. Overall, most of the parameters were statistically significant when the significant level was at least 5%. The models demonstrated that in both cities, sessions tended to be requested during peak hours, when it

Table 7 Price elasticity distribution characteristics

	Mean	SD	Min	Q_1	Q_2	Q_3	Max
Nagoya	-0.192	0.260	-2.448	-0.202	-0.099	-0.055	-0.003
Kyoto	-0.144	0.215	-3.781	-0.137	-0.07	-0.041	-0.003

was not raining, and when the rider had previously experienced ridesourcing. The difference lay in the day of week, where the riders in Nagoya tended to request a session on weekends while the riders in Kyoto tended to request on weekdays.

a) Positive travel time coefficient

The estimation indicated that the travel time coefficients were positive in the proposed model, which normally indicates that the longer the travel time, the greater the tendency of riders to request a ride. However, this was not the case for the proposed models. The proposed models focused on the decision of whether to request a ride rather than on the choice of transportation mode. When a session was created, the origin and destination were already fixed, and the estimated travel time was approximated based on them. Choosing to make a request when the estimated travel time is long does not mean that the rider prefers a longer travel time; it simply means that the destination is far from the pick-up point. Therefore, in the case of the proposed model, where other transportation modes were not considered, positive travel time coefficients indicated that sessions with longer travel times tended to be requested.

b) Observed heterogeneity in the price preferences

The observed heterogeneity in the price preferences was captured by interactions between the price coefficient and the factors that describe when each session was created, that is, weekends or weekdays, peak hours or off-peak hours, and raining or not raining. Additionally, interaction with the time-distance ratio was introduced to capture the change in price preference caused by the speed of the service compared with the distance traveled. Although they were not significantly significant, they suggested some valuable information.

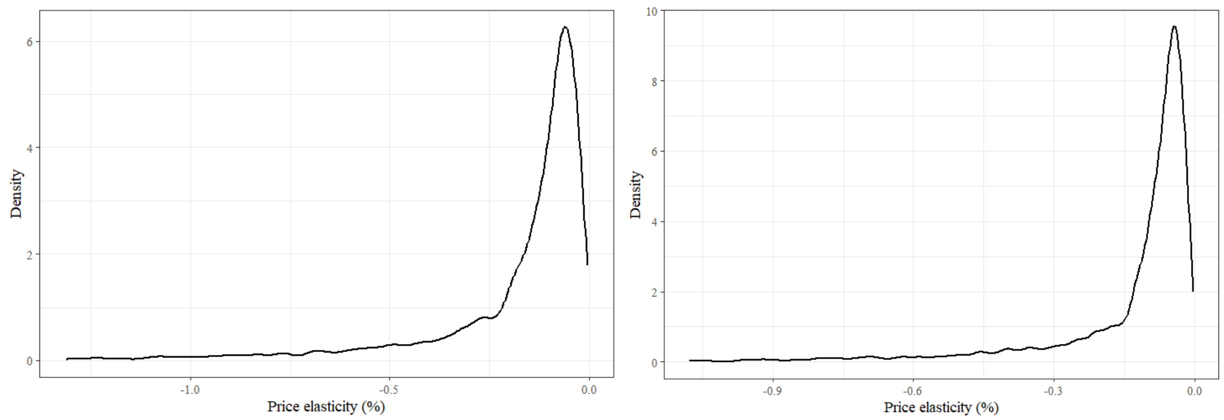
In both cities, the estimation result suggested that riders perceived price less negatively on weekends and when it was raining. This might be because, in these scenarios, riders place more significance on convenience than price. By contrast, they perceived prices more negatively during peak hours, possibly because of traffic congestion. Additionally, the interaction with the time-distance ratio shows that riders perceived price more negatively if the travel time was considered to be long compared with the distance, which suggests that, although the value of travel time was not estimated, riders were willing to pay more to reduce the travel time.

(2) Price elasticity estimation results

In this section, the price elasticity estimation results are discussed. Although MXL 4 was not the best model in terms of the goodness of fit, its results were used to estimate the price elasticities for Nagoya and Kyoto because of its flexibility and interpretability. The distribution of the price coefficient of MXL 4 was less flexible than that of MXL 3, but provided more meaningful results because it was bounded.

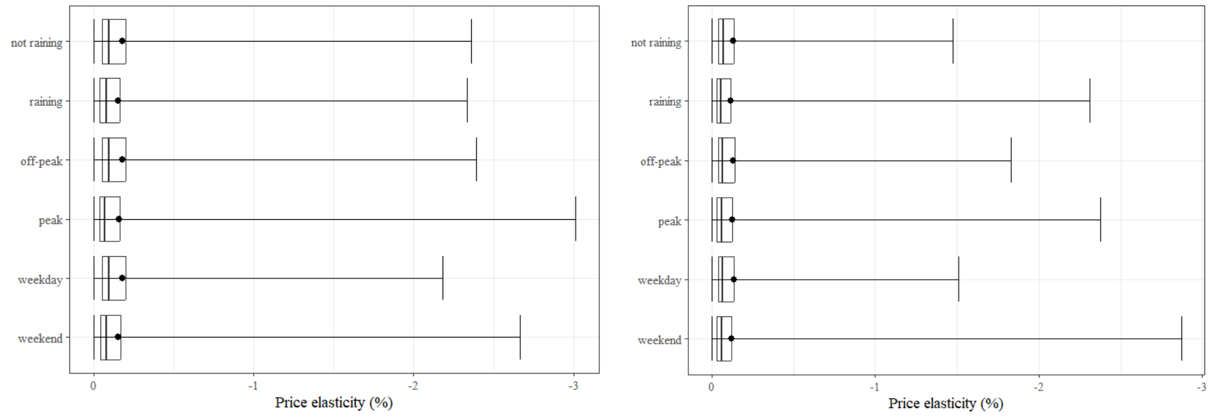
Table 7 summarizes the descriptive statistics of the price elasticities, which ranged from -2.448 to -0.003 in Nagoya and -3.781 to -0.003 in Kyoto, with an average of approximately -0.2 and -0.1, respectively. The maximum and minimum values were somewhat extreme because they were too high and too low for price elasticities. These values belonged to a small portion of the riders who created many sessions, but mostly did not request or mostly requested sessions regardless of price. However, it is not appropriate to discard them because this could be their real behavior. Fig 2 shows the distributions of price elasticities trimmed at the 99th percentile for each city to observe only the majority of riders. Around 98% to 99% of the riders in both cities had a magnitude of price elasticity smaller than 1, which indicates that the riders were inelastic to the Uber fare; that is, price may not have been the main factor in the decision of whether to request a ride for most people. This is possibly because of the Japanese context, where people tend to prioritize trip purposes over price. Additionally, Fig 3 shows the box plot corresponding to Fig 2 segmented by the session characteristics that interacted with the price coefficient. In both cities, the range between the maximum and minimum was the largest during peak hours and weekends. This suggests that if a pricing scheme was implemented, disparities would become most obvious during these situations.

Furthermore, using the sample enumeration method, the aggregate price elasticities of the entire population from Nagoya data and Kyoto data were -0.183 and -0.1399, respectively. They were relatively smaller compared with the majority of the results of past studies on ridesourcing price elasticity, but more consistent with the results of studies on taxi price elasticity. The results indicated a similarity with those of the studies conducted by Castillo (2018), who found that UberX price elasticity was 0.186, and Rose and Hensher (2013), who found that taxi price elasticity was 0.1042 and. This was expected because the focus of the present study was taxi-type



(a) (b)

Fig 2 Distribution of price elasticities among (a) Nagoya riders and (b) Kyoto riders trimmed at the 99th percentile



(a) (b)

Fig 3 Price elasticity segmented by the scenario in (a) Nagoya and (b) Kyoto trimmed at the 99th percentile

ridesourcing.

6. CONCLUSION AND FUTURE WORKS

In this study, the effects of price on ridesourcing demand were empirically investigated through the estimation of price elasticities. Using the session data of Uber taxis from Uber Japan's experiments in Nagoya and Kyoto, the behavior of the rider can be observed in disaggregated manner and allowed the taste heterogeneity to be considered. Furthermore, a semi-nonparametric approach proposed by Fosgerau and Mabit (2013) was also applied, which allowed the taste distribution to be estimated flexibly. With the revealed preference data and the methodology used in this study, reliable and less biased price elasticities were estimated.

The study provides valuable information to help TNCs and policymakers to understand the impact of pricing on ridesourcing users. The findings from Nagoya and Kyoto suggest almost the same trend. Initially, the proposed choice models suggested that riders tended to request a ride during peak hours, and when it was not raining. There was a difference in the preference of the day of week where Nagoya riders tended to request on weekends while Kyoto riders tended to request on weekdays. This provides information for TNCs when demand is high and indicates how the supply should be managed. We also found that riders who have already experienced ridesourcing once tend to request a ride. Furthermore, although not statistically significant, the interactions introduced in the price coefficient suggested that, in both cities, riders perceived price less negatively on weekends, during off-peak hours, and when it was raining. In both cities, we also found the effect of the service speed on the price elasticity.

The results demonstrated that the sample average of the price elasticities in Nagoya and Kyoto was approximately -0.2 and -0.1, respectively, where the majority of price elasticities were less than -1. The distributions of price elasticities show a largest gap between maximum and minimum value to be during peak hours and weekends in both cities. This could imply that disparity issues can be the most serious, and more caution is needed when implementing any pricing scheme during these situations. Additionally, the aggregate price elasticities were estimated to be -0.183 and -0.1399 in Nagoya and Kyoto, respectively, which suggests that ridesourcing users, as a population, are inelastic to price, and pricing schemes might not significantly affect demand. Our results are consistent with some studies (i.e., Castillo (2018) and Rose and Hensher, 2013) but somewhat different in terms of the value of price elasticities from previous studies on ridesourcing price elasticities which could be because we focused on taxi-type ridesourcing in the present study.

The limitations of this study are as follows: The existence of other transportation modes was neglected when the choice of riders was modeled, which may have resulted in the underestimation of price elasticities. Moreover, the variation of price elasticities in different scenarios was examined, but not for different groups of people. In future work, a mode choice model should be considered to increase reliability, and the inclusion of sociodemographic factors may provide greater policy implications. Furthermore, the data included sessions during the COVID-19 pandemic, which may have had unobserved effects on demand and should be investigated further.

ACKNOWLEDGEMENTS: We would like to express our sincere gratitude to Uber Japan for authorizing us to use the data and for their support during the research process.

REFERENCES

1) Alemi, F., Circella, G., Handy, S., & Mokhtarian, P. (2018). What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. *Travel Behaviour and Society*, *13*, 88–104.
2) Ata, B., Barjesteh, N., & Kumar, S. (2019). Spatial Pricing: An Empirical Analysis of Taxi Rides in New York City. *Working Paper*.
3) Booz Allen Hamilton. (2003). Appraisal of Taxi Fare Structure Issues. In *Booz Allen Hamilton* (ISBN 1 877049 61 1).
4) Castillo, J. R. (2018). Who Benefits from Surge Pricing? *Social Science Research Network*.
5) Cats, O., Kucharski, R., Danda, S. R., & Yap, M. (2022). Beyond the dichotomy: How ride-hailing competes with and complements public transport. *PLOS ONE*, *17*(1), e0262496.
6) Chalermpong, S., Kato, H., Thaithatkul, P., Ratanawaraha, A., Fillone, A., Hoang-Tung, N., & Jittrapirom, P. (2022). Ride-hailing applications in Southeast Asia: A literature review. *International Journal of Sustainable Transportation*, 1–21.
7) Clewlow, R. R., & Mishra, G. S. (2017). Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States. *Research Papers in Economics*.
8) Cohen, P., Hahn, R., Hall, J., Levitt, S., & Metcalfe, R. (2016). Using Big Data to Estimate Consumer Surplus: The Case of Uber. *Working Paper*.
9) Dias, F. F., Lavieri, P. S., Garikapati, V. M., Astroza, S., Pendyala, R. M., & Bhat, C. R. (2017). A behavioral choice model of the use of car-sharing and ride-sourcing services. *Transportation*, *44*(6), 1307–1323.
10) Filistrucchi, L., Geradin, D., Van Damme, E., & Affeldt, P. (2014). Market definition in two-sided markets: Theory and practice. *Journal of Competition Law and Economics*, *10*(2), 293–339.
11) Fosgerau, M., & Mabit, S. L. (2013). Easy and flexible mixture distributions. *Economics Letters*, *120*(2), 206–210.
12) Gómez-Lobo, A., Tirachini, A., & Gutierrez, I. (2022). Optimal prices for ridesourcing in the presence of taxi, public transport and car competition. *Transportation Research Part C: Emerging Technologies*, *137*, 103591.
13) Granados, N., Gupta, A., & Kauffman, R. J. (2012). Online and Offline Demand and Price Elasticities: Evidence from the Air Travel Industry. *Information Systems Research*, *23*(1), 164–181.
14) Guda, H., & Subramanian, U. (2019). Your Uber Is Arriving: Managing On-Demand Workers Through Surge Pricing, Forecast Communication, and Worker Incentives. *Management Science*.
15) Hall, J. D., Palsson, C., & Price, J. (2018). Is Uber a substitute or complement for public transit? *Journal of Urban Economics*, *108*, 36–50.
16) Han, C. P., & Li, J. (2009). Evaluating Estimation Techniques of Transportation Price Elasticity. *Transportation Research Record: Journal of the Transportation Research Board*, *2115*(1), 94–101.
17) Hess, S., & Palma, D. A. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. *Journal of Choice Modelling*, *32*, 100170. https://doi.org/10.1016/j.jocm.2019.100170
18) Hubert, M., & Vandervieren, E. (2008). An adjusted boxplot for skewed distributions. *Computational Statistics & Data Analysis*, *52*(12), 5186–5201. https://doi.org/10.1016/j.csda.2007.11.008
19) Kong, H., Zhang, X., & Zhao, J. (2020). How does ridesourcing substitute for public transit? A geospatial perspective in Chengdu, China. *Journal of Transport Geography*, *86*, 102769.
20) Lam, C. T., Liu, M., & Hui, X. (2021). The geography of ridesharing: A case study on New York City. *Information Economics and Policy*, *57*, 100941.
21) Lavieri, P. S., & Bhat, C. R. (2019). Investigating objective and subjective factors influencing the adoption, frequency, and characteristics of ride-hailing trips. *Transportation Research Part C: Emerging Technologies*, *105*, 100–125.
22) Lehmann, N., Sloot, D., Ardone, A., & Fichtner, W. (2022). Willingness to pay for regional electricity generation – A question of green values and regional product beliefs? *Energy Economics*, *110*, 106003.
23) Liang, Z., Kong, W., Zhan, X., & Xiao, Y. (2022). Analysis of the Impact of Ride-Hailing on Urban Road Network Traffic by Using Vehicle Trajectory Data. *Journal of Advanced Transportation*, *2022*, 1–18.
24) Litman, T. (2022). Understanding Transport Demands and Elasticities: How Prices and Other Factors Affect Travel Behavior. *Victoria Transport Policy Institute*.
25) Madanizadeh, S. A., Joshaghani, H., & Moradi, R. (2022). Ride-hailing demand elasticity: a regression discontinuity method. *Journal of Industrial and Business Economics*.
26) Rayle, L., Dai, D., Chan, N., Cervero, R., & Shaheen, S. (2016). Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. *Transport Policy*, *45*, 168–178.
27) Rose, J. M., & Hensher, D. A. (2013). Demand for taxi services: new elasticity evidence. *Transportation*, *41*(4), 717–743.
28) Tang, B. J., Li, X. Y., Yu, B., & Wei, Y. M. (2019). How app-based ride-hailing services influence travel behavior: An empirical study from China. *International Journal of Sustainable Transportation*, *14*(7), 554–568.

29) Tirachini, A., & Del Río, M. (2019). Ride-hailing in Santiago de Chile: Users' characterisation and effects on travel behaviour. *Transport Policy*, *82*, 46–57.
30) Toner, J. P. (2010). The Welfare Effects of Taxicab Regulation in English Towns. *Economic Analysis and Policy*, *40*(3), 299–312.
31) Train, K. E. (2009). *Discrete Choice Methods with Simulation* (2nd ed.). Cambridge University Press, Cambridge.
32) Wang, H., & Yang, H. (2019). Ridesourcing systems: A framework and review. *Transportation Research Part B: Methodological*, *129*, 122–155.
33) Young, M., & Farber, S. (2019). The who, why, and when of Uber and other ride-hailing trips: An examination of a large sample household travel survey. *Transportation Research Part A: Policy and Practice*, *119*, 383–392.
34) Zheng, H., Chen, X., & Chen, X. M. (2019). How Does On-Demand Ridesplitting Influence Vehicle Use and Purchase Willingness? A Case Study in Hangzhou, China. *IEEE Intelligent Transportation Systems Magazine*, *11*(3), 143–157.