Transportation Research Interdisciplinary Perspectives 16 (2022) 100722

Contents lists available at ScienceDirect

Transportation Research Interdisciplinary Perspectives

journal homepage: www.sciencedirect.com/journal/transportation-research-interdisciplinary-perspectives

Ridesourcing mode choice: A latent class choice model for UberX in Chile

Gustavo García-Melero [a,*], Rubén Sainz-González [b], Pablo Coto-Millán [b], Alejandra Valencia-Vásquez [a]

[a] *Escuela de Ingeniería de Transporte, Pontificia Universidad Católica de Valparaíso, 2340025 Valparaíso, Chile*
[b] *Departamento de Economía, Universidad de Cantabria, 39005 Santander, Spain*

ARTICLE INFO

Keywords:
ridesourcing
mode choice
latent class choice model (LCCM)
UberX
subjective value of time (SVT)

ABSTRACT

As shared mobility expands, ridesourcing has become its most popular manifestation. However, users' mode choice has not yet been sufficiently explored. Thus, this study aims to model ridesourcing mode choice across different latent classes to ascertain who chose ridesourcing and why.

We conducted a mode choice study by collecting revealed preference surveys from UberX users in Viña del Mar, Chile, in 2017. We then determined the existence of two latent classes and modeled the mode choice using a latent class choice model. Ultimately, we characterized individuals belonging to each latent class and calculated the subjective value of time (SVT).

Most UberX users were highly educated and aged 20–35 years. Further, UberX gained users principally from public transport (80%). Likewise, the two latent classes differed by socioeconomic characteristics and SVTs. A latent class grouped the highest-educated and highest-earning users, who also offered the highest SVT.

In summary, two latent classes, differentiated by educational level and income, formed the ridesourcing market. Besides, they offered distinct ridesourcing choice behavior based on the widely dissimilar SVTs. There was also a strong substitution effect between ridesourcing and transit use. The results imply that policymakers and transportation planners could have increased the competitiveness of the public transit system by improving rapidity and safety, having room to increase the fares to defray the improvements. Further, they could have used information related to the latent classes to customize relevant policies and marketing strategies (routes, frequency, fares, etc.) for every latent class.

Introduction

Shared mobility has expanded, and ridesourcing has become its most popular manifestation (Iqbal, 2020). The boost in UberX market share—the most favored ridesourcing service—has inspired an increase in research related to ridesourcing (Tirachini, 2020), such as the motivational factors behind ridesourcing use (Rayle et al., 2016) or its impact on modal substitution (Nie, 2017), traffic externalities (Ward et al., 2019), and mode choice (Asgari and Jin, 2020; Dong, 2020; Yan et al., 2019).

Regardless, the inherently disruptive shared mobility markets and the distinct contexts have not yet allowed us to thoroughly understand the competition between ridesourcing and conventional urban modes (Habib, 2019). Most importantly, there is still a need to address the lack of ridesourcing research in developing countries (Circella and Alemi, 2018). Some authors have also stated that market segmentation may help determine the factors behind ridesourcing use (Ho et al., 2018). For instance, Alemi et al. (2018b) used a latent class choice model (LCCM) to determine the factors affecting ridesourcing adoption in California. They concluded that socioeconomic characteristics and adoption variables differed across the latent classes. Above all, they showed that ridesourcing market segmentation was fundamental to understanding how the ridesourcing market worked. Thus, ridesourcing users' mode choice has not yet been sufficiently explored. Indeed, there is no research modeling ridesourcing users' mode choice in developing countries using revealed preference surveys and LCCM.

This study aims to model ridesourcing mode choice across different latent classes to ascertain who chose ridesourcing and why. We conducted a mode choice study by collecting revealed preference surveys from UberX users in Viña del Mar, Chile, in 2017. We then determined

* Corresponding author.
E-mail addresses: gustavo.garcia@pucv.cl, gustavogarciamelero@hotmail.com, gustavo.garcia@pucv.cl (G. García-Melero), sainzru@unican.es (R. Sainz-González), pedro.coto@unican.es (P. Coto-Millán), alejandra.valencia@pucv.cl (A. Valencia-Vásquez).

https://doi.org/10.1016/j.trip.2022.100722
Received 23 September 2021; Received in revised form 20 July 2022; Accepted 9 November 2022
Available online 20 November 2022

the number of latent classes and modeled the mode choice using an LCCM. Finally, we characterized individuals belonging to each latent class and calculated the subjective value of time (SVT). These results allowed us to establish ridesourcing behavior from a historical dataset.

The remainder of this paper is organized as follows. Section 2 presents the literature review. In Section 3, the LCCM estimation methodology is proposed. Section 4 describes the survey and the dataset analysis. Section 5 presents the results of the LCCM estimation, SVTs, latent class characterizations, and discussion. Finally, conclusions are presented in Section 6.

Literature review

Shared mobility has allowed users to have short-term access to low-speed modes of transportation—bikes, cars, vans, etc.—when individuals require a trip (Shaheen and Chan, 2016). Some services have permitted the shared use of a vehicle (car-sharing, scooter-sharing, or bike-sharing), whereas others have facilitated sharing a ride. The latter have been grouped into traditional ridesharing services (carpooling and vanpooling) and on-demand services (ridesourcing, ridesplitting, e-hail, and microtransit) (Shaheen and Cohen, 2019). Likewise, ridesourcing, also known as ridehailing, ridebooking, or transportation network company (TNC), has provided individualized services. In turn, ridesplitting, also known as dynamic ridesharing, has provided shared passenger rides for a reduced fare.

TNCs have proliferated in different geographic areas. Besides Uber and Lyft, which have dominated the American market, we might find other providers such as Didi in China, Grab in South Asian countries, and Ola in India. In addition, ridesourcing has been the fastest growing and most used on-demand service among shared mobility users (Circella and Alemi, 2018). For example, TNCs served over 170,000 trips on a typical weekday in San Francisco (Cooper et al., 2018); in contrast, ridesplitting achieved a low percentage (6–7%) of the total number of TNCs trips in China, based on data provided by Didi (Li et al., 2019).

Regardless, TNCs have profoundly affected the mobility sector. In the context of public transport, there was a substitution effect between ridesourcing and transit use for weekday trips in Austin (Lavieri et al., 2018). Likewise, TNCs performed best when the transit service was the least convenient in Chicago, such as on trips requiring transfers or walking long distances (Schwieterman and Smith, 2018). However, ridesourcing complemented and competed with public transit under other circumstances (Jin et al., 2018; Rayle et al., 2016). Although the empirical results have diverged, the substitution effect has been more significant than the complementarity effect (Tirachini, 2020). In the context of traditional taxis, the conclusions have also been contradictory. TNCs negatively impacted taxicab ridership in Las Vegas (Contreras and Paz, 2018) and Toronto (Young and Farber, 2019). In addition, drivers' earnings in conventional taxis decreased across metropolitan areas in the U.S. (Berger et al., 2018). Nevertheless, the competition between taxis and ridesourcing was an incentive to perform more effectively in China (Nie, 2017). Finally, we also found differing empirical results in the context of traffic externalities. TNCs occasionally reduced vehicle kilometers traveled, emissions, energy consumption, private vehicle ownership, and accidents (Martinez and Viegas, 2017; Ward et al., 2019; Yu et al., 2017). However, other studies have shown that TNCs increased traffic externalities (Nie, 2017; Tirachini, 2020; Tirachini and Gomez-Lobo, 2020).

The factors driving the use of ridesourcing have also been studied (Acheampong et al., 2020). Rayle et al. (2016) found that in San Francisco, ridesourcing users were generally younger and better educated than the average population. Likewise, Yu and Peng (2020) concluded that high population, employment and road densities, pavement completeness, land-use mix, and job accessibility by transit produced more ridesourcing demand in Austin. Lee et al. (2018) suggested that perceived risks and benefits, trust in the platform, and perceived platform qualities were significant predictors of users' intention to use Uber in Hong Kong.

In addition, adoption models have allowed us to determine the factors driving the use of ridesourcing. Alemi et al. (2018a) concluded that highly educated and older millennials (individuals born between 1981 and 1997) were more likely to use on-demand ride services than other groups in California. They also found that individuals with stronger pro-environmental, technology-embracing, and variety-seeking attitudes were more inclined to use ridehailing. Later, they delved deeper into the factors affecting the adoption of ridehailing in California by estimating a three latent-class adoption model that captured the heterogeneity in individuals' tastes and preferences (Alemi et al., 2018b). The first class comprised highly educated and independent millennials with the highest adoption rate. The adoption of ridehailing was influenced by the frequency of long-distance leisure and business-related trips made by non-car modes. The second-highest adoption rate belonged to the latent class composed of affluent individuals, either dependent millennials or older members of Generation X (individuals born between 1965 and 1980), living with their families. The frequency of use of transportation-related apps and the share of long-distance leisure air trips affected their adoption of ridehailing. Finally, the third class comprised the least affluent individuals, with the lowest education and adoption rate. The adoption of ridehailing was affected by household income, frequency of long-distance non-car business trips, transit accessibility, and use of taxis and car-sharing. In short, they demonstrated the need to segment the population because socioeconomic characteristics and variables affecting ridehailing adoption were distinct across classes.

Finally, mode choice models were used to ascertain why users chose the TNCs. Few studies have aimed to model ridesourcing users' mode choices due to the lack of data to estimate these models (Table 1). On the one hand, a few articles have modeled mode choice exclusively among public transport, ridesourcing, and ridesplitting. Chavis and Gayah (2017) surveyed public transport users in Baltimore. The choice set included traditional fixed-route transit systems, flexible-route systems with shared vehicles, and individual on-demand services like Uber or Lyft. They concluded that the mode choice was determined by the expected travel time (access, waiting, in-vehicle) and cost. Similarly, Dong (2020) surveyed Uber and Lyft users in Philadelphia, considering a choice set formed by transit and ridehailing. Higher-income individuals over 30 years of age, females, and less frequent transit users were increasingly willing to choose ridehailing. Travel time and cost were significant deterrents for travel in both modes. Azimi et al. (2020) surveyed transit and auto users in the U.S. with choice scenarios considering transit, ridesplitting, and ridesourcing. For transit users, ridesourcing choice was highly affected by travel time and cost. Travel time and cost were also significant for auto users, as were reliability, convenience, comfort, and stress relief.

On the other hand, several studies have modeled the mode choice among conventional urban modes, ridesourcing, and ridesplitting.

Table 1
Papers focused on mode choice for TNCs *

Author	Data	Choice	Explanatory variables
Asgari and Jin (2020)	SP	ECL	Travel time and cost, socioeconomic, attitudinal
Azimi et al. (2020)	SP	ML	Travel time and cost, socioeconomic, attitudinal
Dong (2020)	SP	ML	Travel time (access, waiting, travel) and cost, transfers, socioeconomic
Habib (2019)	RP	IAL, CMNL	Travel time and cost, socioeconomic
Yan, Levine, and Zhao (2019)	SP/RP	ML	Travel time (access, waiting, in-vehicle) and cost, socioeconomic, attitudinal
Chavis & Gayah (2017)	SP	MNL, HL, BL	Travel time (access, waiting, in-vehicle) and cost, socioeconomic, attitudinal

* RP: Revealed Preferences; SP: Stated Preferences; MNL: Multinomial Logit; ML: Mixed Logit; HL: Hierarchical Logit; BL: Binary Logit; IAL: Independent Availability Logit; CMNL: Constrained Multinomial Logit; ECL: Error Component Logit

Habib (2019) utilized a large-scale household travel survey in Toronto, considering a choice set composed of car drive, car passenger, transit, walking, bicycle, taxi, and Uber. Urban taxis were the leading competitors of Uber, although younger people preferred Uber independent of sex. Travel time and cost were crucial for determining the mode choice. Likewise, Asgari and Jin (2020) surveyed individuals in 10 metropolitan areas in the U.S. with choice scenarios considering conventional modes (car and public transit), exclusive on-demand services (ridesourcing), and shared on-demand services (ridesplitting). Habits acted as a barrier to behavioral changes but increasing the private mobility expenses could help overcome them. Yan et al. (2019) obtained commuting mode choices at the University of Michigan, Ann Arbor. The choice set included walking, biking, driving, transit, and an integrated ridesourcing and bus-line system. They concluded that mode choice was determined fundamentally by the expected travel time (access, waiting, in-vehicle) and cost.

In conclusion, the lack of data to estimate ridesourcing users' mode-choice models has impeded delving into the topic. Most researchers have implemented stated preference surveys, given the obstacles to obtaining data from the TNCs. Thus, mode choice modeling has considered hypothetical decisions instead of revealed choices, so the competition between TNCs and conventional modes is not yet well understood. This study aims to conduct ridesourcing mode choice modeling in developing countries using revealed preference surveys and segment the market with an LCCM.

Methodology

We used the LCCM to capture observed and unobserved heterogeneity by grouping individuals into latent classes that were not directly identifiable from the data (Walker and Ben-Akiva, 2002). LCCMs differ from discrete choice models with latent variables. Both belong to the generalized random utility model presented by Walker and Ben-Akiva (2002), but LCCMs do not need to incorporate latent variables. LCCMs have been utilized only in a few studies on mode choice, for example, under choice sets of electric bicycle, car, or public transport (Fu and Juan, 2017; Hurtubia et al., 2014; Ma et al., 2015), or in only one instance, including ridesharing as an alternative available (Saxena et al., 2019). However, no ridesourcing mode choice research has been conducted using LCCMs with data from revealed preference surveys in developing countries.

The LCCM assumed that individuals could be placed only into a latent class, but their membership was unknown beforehand. Thus, we had to set the number of latent classes in advance to test the individual inclusion using a class membership model. The membership model allowed us to predict the probability of belonging to a latent class based on individuals' observable or unobservable characteristics (Kim and Mokhtarian, 2018). In addition, the LCCM assumed that individuals belonging to different latent classes might exhibit unique choice behaviors. Therefore, the LCCM also included a mode choice model that provided a class-specific utility equation. In other words, LCCM offered distinct parameters in the utility equation for each latent class (heterogeneous preferences) but equal parameters for individuals in the same latent class (homogeneous preferences).

The LCCM theoretical formulation (Greene, 2012; Greene and Hensher, 2003) stated that individual i's choice among J alternatives in choice situation t was the one with maximum utility, given that individual i was in class c, and the utility functions were

$$U_{jit/c} = \beta_c^{'} x_{jit} + \varepsilon_{jit} \tag{1}$$

where $U_{jit/c}$ is the utility of alternative j to individual i in choice situation t, given that individual i is in class c; x_{jit} is the set of attributes considered in all utility functions; ε_{jit} is the unobserved heterogeneity for individual i and alternative j in choice situation t; and $\beta_c^{'}$ is the class-specific parameter vector. Similarly, the choice probabilities within the class were generated using a multinomial logit model:

$$Prob[y_{it} = j | class = c] = \frac{\exp(\beta_c^{'} x_{jit})}{\sum_{j=1}^{J_i} \exp(\beta_c^{'} x_{jit})} \tag{2}$$

In addition, the membership model formulation used a multinomial logit form that allowed us to determine the class probabilities as follows:

$$Prob[class = c] = Q_{ic} = \frac{\exp(\theta_c^{'} z_i)}{\sum_{c=1}^{C} \exp(\theta_c^{'} z_i)}, \theta_c = 0 \tag{3}$$

where z_i is the set of observable characteristics. If no such features are observed, the only element in z_i would be the class-specific constant. Namely, class-specific probabilities would be a set of fixed constants that would add up to 1. Consequently, the likelihood of choosing alternative j for individual i is the expected value (over classes) of class-specific probabilities:

$$\begin{aligned} Prob[y_{it} = j] &= E_c\left[\frac{\exp(\beta_c^{'} x_{jit})}{\sum_{j=1}^{J_i} \exp(\beta_c^{'} x_{jit})}\right] \\ &= \sum_{c=1}^{C} Prob(class = c)\left[\frac{\exp(\beta_c^{'} x_{jit})}{\sum_{j=1}^{J_i} \exp(\beta_c^{'} x_{jit})}\right] \end{aligned} \tag{4}$$

Finally, the LCCM estimation was carried out using NLOGIT©, which allowed us to consider the existence of between two and five latent classes (Greene, 2012). Thus, we estimated several LCCMs with distinct explanatory variables and different numbers of latent classes. We then selected the best model based on the goodness of fit and interpretability of the results, which is the standard procedure in the literature (Alemi et al., 2018b; Greene and Hensher, 2003).

Data

To collect data on ridesourcing users and trips, we conducted 2,000 online revealed preference surveys with UberX users in Viña del Mar, Chile, from August to October 2017. The survey was designed according to the literature on ridesourcing mode choice and the motivational factors behind ridesourcing use (Alemi et al., 2018a; Asgari and Jin, 2020; Azimi et al., 2020; Chavis and Gayah, 2017; Dong, 2020; Habib, 2019; Ho et al., 2018; Rayle et al., 2016; Yan et al., 2019). The survey asked 20 questions regarding a trip made by UberX, individual socio-economic characteristics, and motivational factors behind UberX use. In addition, we asked participants to consider only trips made on Tuesdays, Wednesdays, and Thursdays in those weeks when there were no public holidays. Although we initially asked for monetary values in Chilean Peso, this study provides all the monetary values in euro (€) based on the average exchange rate in August 2017 (1€ = CLP$759.15).

Of the 2,000 individuals approached to participate in the survey, 1,912 answered the questionnaire (response rate of 95.6%). However, we performed a data quality improvement analysis to identify and address inconsistent (e.g., two or more cells that should display the same information do not), erroneous (e.g., impossible numerical or character values given the data definition), suspicious (e.g., plausible data but odd in some way, including outlier' analysis), and missing data (e.g., missing values for one or more cells). Eventually, we obtained a sample of 1,536 valid observations (successful response rate of 76.8%) using simple random sampling (95% confidence level, 5% error).

Demographics of UberX users

Table 2 compares the demographics of the sample and Viña del Mar according to the data from the 2017 Chilean census (National Statistics Institute of Chile, 2017). However, the census did not include the income data, so we considered the national data from the Income Supplementary survey as a reference (National Statistics Institute of Chile, 2015).

The sex distribution was similar, even though UberX users were

Table 2
Demographics of UberX user survey respondents compared to Viña del Mar census **

Variable	Category	UberX users (%)	Viña del Mar population (%)
Gender	Female	50.8%	52.5%
	Male	49.2%	47.5%
Age	15-20	2.8%	8.9%
	20-35	70.8%	30.5%
	36-50	20.8%	21.2%
	51-65	4.6%	21.6%
	> 65	1.0%	17.8%
Education	Primary	0.2%	23.0%
	Secondary	10.8%	42.1%
	Technical	16.8%	9.3%
	University	72.2%	25.6%
Occupation	Work	71.2%	57.1%
	Study	23.6%	15.2%
	Homemaker	1.6%	10.3%
	Unemployed	2.0%	4.0%
	Retired	1.6%	13.5%
Household size	1	1.2%	33.6%
	2	10.5%	21.2%
	3 or more	88.3%	45.3%
Relationship	Head of the family	40.4%	43.0%
	Partner	22.3%	21.1%
	Child	37.3%	35.9%
Monthly income	< 265 €	25.0%	20.2%
	265-530 €	18.6%	37.3%
	530-1,060 €	34.0%	27.0%
	1,060-1,590€	20.8%	8.7%
	> 1,590€	1.6%	6.8%
Number of vehicles	0	37.4%	n/a
	1	48.8%	n/a
	2	12.8%	n/a
	3	1.0%	n/a
Driver's license	Yes	46.8%	n/a
	No	53.2%	n/a
Parking at home	Yes	78.2%	n/a
	No	21.8%	n/a
Parking at work	Yes	37.4%	n/a
	No	62.6%	n/a

** Sources: 2017 Chilean census (monthly income extracted from the 2015 Income Supplementary survey)

younger and better educated than the population of Viña del Mar. Likewise, most people in the sample worked or studied, reaching a higher level than in the census, and lived in households with three or more co-habitants, predominantly as the head of the family or child. The average UberX user was a young female worker aged 20–35 years (millennial), highly educated, and living in a household with three or more co-habitants. These results confirmed the conclusions of Alemi et al. (2018a), who determined that millennials with higher education had the highest ridehailing adoption rate in California. Similarly, Rayle et al. (2016) concluded that the primary ridesourcing users in San Francisco were highly educated young people. Thus, even though the contexts differed significantly, the ridesourcing user profiles were similar.

Higher-income levels did not show a relevant market share (1.6%). The most significant percentage were middle-income users (34%), as in previous studies conducted in Chile (Tirachini and del Río, 2019). In addition, many ridesourcing users either did not own a car (37%) or had one vehicle (49%). Rayle et al. (2016) obtained a similar percentage distribution in San Francisco, suggesting that different contexts may not significantly differ in motorization characteristics. Likewise, almost half of the survey respondents also had a driver's license (47%), and around 80% had parking facilities at home but not at work (63%). Namely, UberX users were mostly captive to public transport and soft modes (walking, cycling) or owned private vehicles but preferred not to drive.

Reasons for choosing UberX

Figure 1 shows the reasons why the survey respondents chose UberX. The main reason for this was the speed of reaching the destination (32%). Other relevant reasons included personal safety (18%) and time adequacy (15%). Quality of service (12%) and comfort (11%) were also important, but cost (6%) and reliability (6%) were not crucial. UberX users appreciated using a rapid and safe mode of transportation according to their mobility needs at any given time. Thus, the factors driving ridesourcing use were associated with perceived risks and benefits, trust in the platform, and perceived platform quality, as stated by Lee et al. (2018). Rayle et al. (2016) and Dong (2020) also concluded that the rapidity of getting to the destination was the fundamental reason behind ridesourcing use, but not personal safety. Consequently, the context was crucial in determining the reasons for ridesourcing use.

Other studies in Chile, carried out in Santiago, concluded different reasons than Viña del Mar. Tirachini and Gomez-Lobo (2020) concluded that the most relevant reasons were ease of payment and trip fare rather than rapidity and safety. However, in a previous paper, Tirachini and del Río (2019) determined that safety was the most important reason. The reasons for choosing ridesourcing were affected by different contexts, surveying times, and sample members. Hence, these studies should be conducted regularly because the results might change.

Trip purpose

Figure 2 shows the reported trip purposes from the UberX user survey. The most relevant trip purposes were returning home (38%) and leisure (22%). The remainder of the trip purposes were smaller, that is, commuting (11%), visiting someone (9%), studies (7%), shopping (5%), procedures (4%), and health (4%). UberX users mainly linked ridesourcing trips to the search for a non-stressful context. Although the results are similar to those of previous studies (Habib, 2019), leisure was the primary trip purpose found in other studies (Rayle et al., 2016; Tirachini and del Río, 2019; Tirachini and Gomez-Lobo, 2020). Thus, the context was also a determinant of trip purpose.

Modal shift and externalities

Figure 3 shows how users would have made the trip if UberX were unavailable. A large number stated that they would have used buses (48 %) or shared taxis (31%), while cars (10%), taxis (5%), walking (4%), bicycles (1%), and metro (1%) were not highly significant. Namely, UberX gained users principally from public transport (80%), while the market shares captured from cars (10%) and taxis (5%) were minor. In short, UberX threatened public transport because there was a substitution effect between ridesourcing and transit, as in previous studies (Lavieri et al., 2018; Schwieterman and Smith, 2018).

However, the results differed from those of other studies (Rayle et al., 2016), including those from Chile (Tirachini and del Río, 2019; Tirachini and Gomez-Lobo, 2020), which concluded that taxi was the most affected mode. Thus, the modal shift also depends on the context. Although there was a common feature in all of them, modal substitution for car trips was marginal. In addition, the modal shift pattern profoundly impacted traffic externalities because ridesourcing significantly increased the number of private vehicles on the road. Namely, UberX boosted traffic and congestion, as concluded in other studies (Nie, 2017; Tirachini, 2020).

Trip distance, cost, and time

Table 3 lists the features associated with UberX trips. The cost mostly came to 7.8 € (94%), while the percentage of trips over 7.8 € was negligible (6%). Likewise, UberX trips mostly lasted less than 20 minutes (60%) and seldom lasted over 30 minutes (13%). In addition, short-distance trips (less than 5 kilometers) dominated (57%), whereas trips

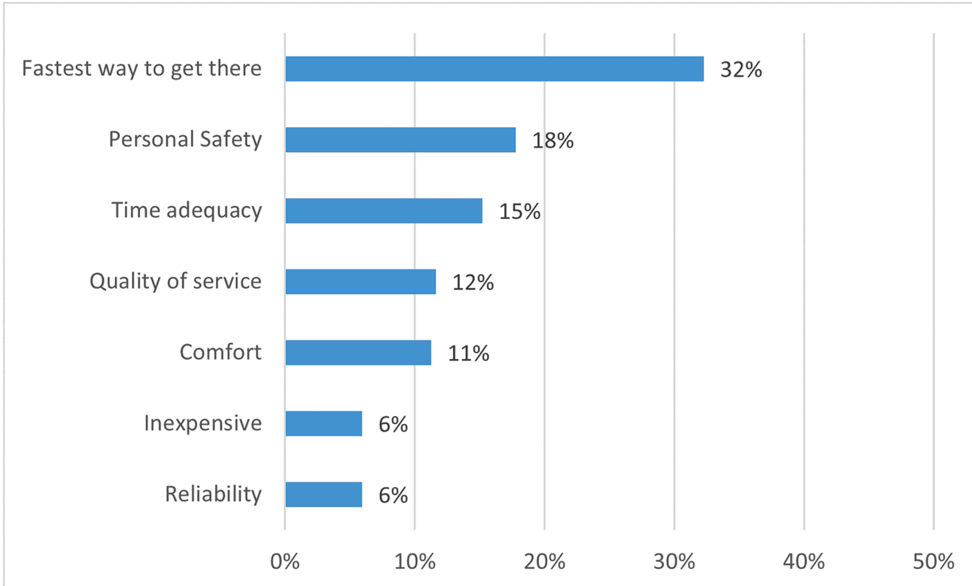

Figure 1. Percentage of responses to "What were the reasons for choosing UberX over other modes?" (Respondents could choose up to three options, n = 2,645).

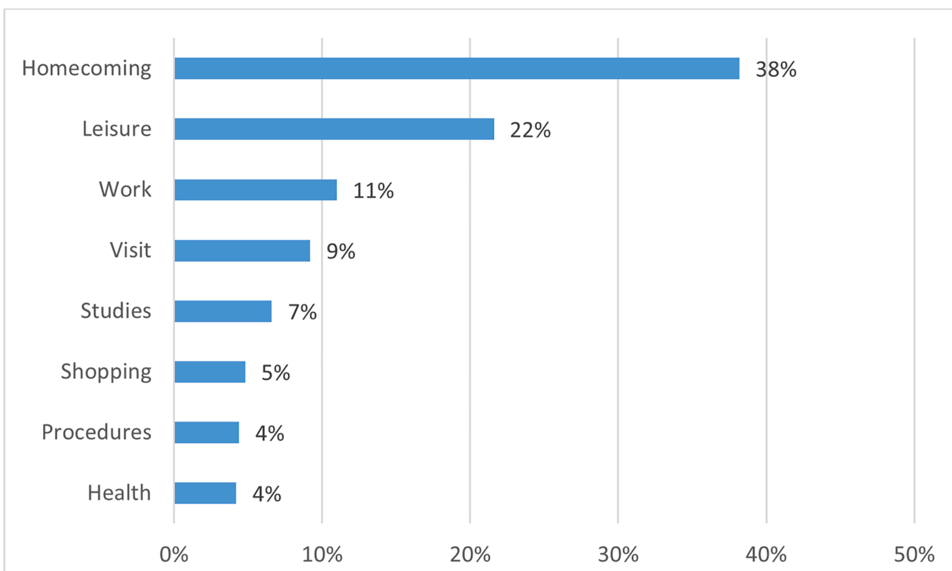

Figure 2. Percentage of responses to "What was your trip purpose?" (Respondents could choose one option, n = 1,536).

over 10 kilometers were scarce (13%). In other words, an affordable, short-distance, and quick UberX trip pattern emerged.

Results and Discussion

This study aims to model ridesourcing mode choice across different latent classes of UberX users to ascertain who chose ridesourcing and why. In this section, we provide the estimation results of the LCCM, which simultaneously allowed us to identify the number of latent classes of users and estimate a mode-choice model for each latent class. In addition, we offer the individuals' characterization belonging to each latent class, SVTs for each latent class, and a discussion of the findings.

The LCCM considered a choice set formed by four alternatives: UberX, bus, shared taxi, and car. Initially estimated many LCCMs with distinct explanatory variables and different numbers of latent classes. We then selected the best model based on the goodness of fit and interpretability of the results, which is the standard procedure in the literature (Alemi et al., 2018b; Greene and Hensher, 2003). The final LCCM consisted of two latent classes, the alternative specific constant (ASC) for three modes (UberX, bus, and shared taxi), and two explanatory variables: travel cost (TC) and travel time (TT). In addition, we considered generic parameters for each explanatory variable and latent class— equal for the same variable — for utility functions across all alternatives in the choice set. Although we considered all the available explanatory variables and between two and five latent classes, the final model allowed us to interpret the results better and provide better goodness of fit. Further, travel time and cost were the unique explanatory variables used in all ridesourcing mode choice studies (see Table 1). Thus, the utility functions of LCCM are as follows:

$$U_{UberX} = ASC_{UberX} + \theta_{TT}TT_{UberX} + \theta_{TC}TC_{UberX} + \varepsilon_{UberX} \quad (5)$$

$$U_{Bus} = ASC_{Bus} + \theta_{TT}TT_{Bus} + \theta_{TC}TC_{Bus} + \varepsilon_{Bus} \quad (6)$$

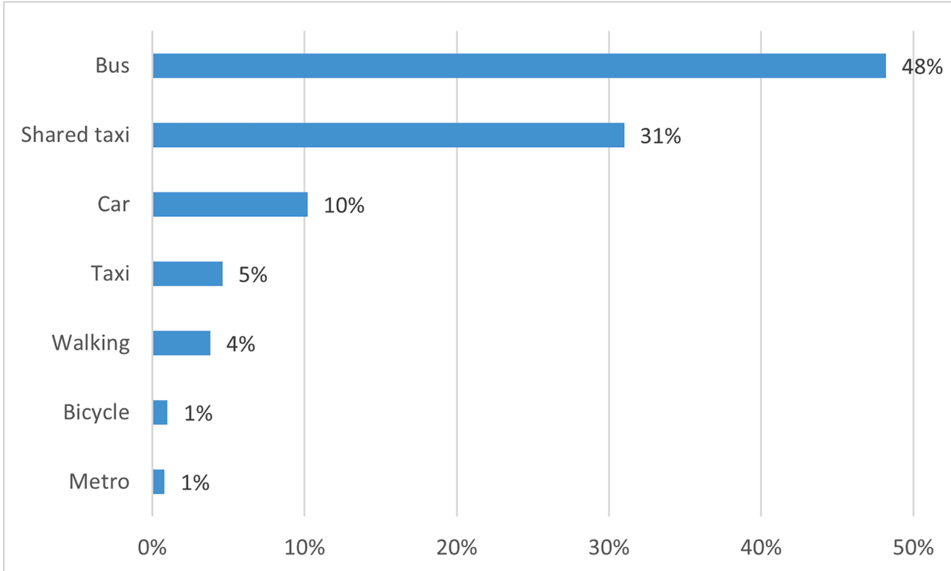

Figure 3. Percentage of responses to “How would you have made the trip if UberX was not available?” (Respondents could choose one option, n = 1,536).

Table 3
Percentage distribution of trip attributes for sampled UberX trips (n = 1,536)

Variable	Category	%
Cost	< 2.6 €	9.4%
	2.6-5.2 €	61.0%
	5.2-7.8 €	23.4%
	7.8-10.4 €	4.2%
	>10.4 €	2.0%
Time	< 10 min	7.2%
	10-15 min	27.2%
	16-20 min	25.8%
	21-25 min	14.2%
	26-30 min	13.0%
	31-35 min	7.8%
	36-40 min	2.8%
	> 40 min	2.0%
Distance	< 5 Km	57.4%
	5-10 Km	29.8%
	> 10 Km	12.8%

$$U_{Sharedtaxi} = ASC_{Sharedtaxi} + \theta_{TT} TT_{Sharedtaxi} + \theta_{TC} TC_{Sharedtaxi} + \varepsilon_{Sharedtaxi} \tag{7}$$

$$U_{Car} = \theta_{TT} TT_{Car} + \theta_{TC} TC_{Car} + \varepsilon_{Car} \tag{8}$$

Table 4 presents the final LCCM, including the results for the class membership and the mode choice models. The latent class probability determined that 57% of the sample belonged to class 1, whereas 43% belonged to class 2. All estimates provided suitable signs (time and cost were negative) and were statistically significant at 95% confidence.

Table 4
LCCM estimation results (n = 1,536)

	Latent class 1		Latent class 2	
Parameter	Estimate	t-Test	Estimate	t-Test
$\mathbf{ASC_{UberX}}$	9.61	15.80	6.28	14.46
$\mathbf{ASC_{Sharedtaxi}}$	5.37	15.47	5.51	14.71
$\mathbf{ASC_{Bus}}$	6.00	17.07	3.62	9.18
θ_{TT}	-0.00297	-5.54	-0.00255	-5.69
θ_{TC}	-0.03703	-19.56	-0.0098	-10.57
Latent class probability	0.57	16.47	0.43	12.45
SVT	4.81 €/hour		15.61 €/hour	
Log-likelihood	-1269.08			

Similarly, we obtained the SVTs for the two latent classes in €/hour. SVTs showed the willingness to pay for reducing the travel time by one hour in each latent class. The SVT of class 1 (4.81 €/hour) was much lower than that of class 2 (15.61 €/hour); therefore, ridesourcing users in latent class 2 offered a higher willingness to pay to reduce travel time. Both latent classes, especially latent class 2, presented SVTs greater than the average wage rates of 4 €/hour (latent class 1) and 8 €/hour (latent class 2). Given the reasons behind ridesourcing use (see Figure 1), these SVTs suggested that ridesourcing users were not exclusively paying for the trip but also for personal safety. Thus, further research should include a latent variable related to personal safety to test the relevance of ridesourcing mode choices. In addition, these SVTs explained why ridesourcing users did not opt for cheaper but less rapid alternatives: the fastest way to reach the destination was the priority and foremost reason for choosing UberX (see Figure 1). Regardless, most ridesourcing mode choice studies have not reported ridesourcing users' SVTs (Asgari and Jin, 2020; Chavis and Gayah, 2017; Habib, 2019; Saxena et al., 2019; Yan et al., 2019), and only one has considered a sample of ridesourcing users (Dong, 2020). However, ridesourcing users' SVT could be calculated based on the final model provided in the mode choice research carried out in Philadelphia by Dong (2020). SVT was roughly 24$/h, considering generic parameters as we did in this study, which suggested that ridesourcing users offered a high SVT under different contexts. Although further research should investigate SVTs, the results implied that ridesourcing users showed higher SVTs than average wage rates.

Based on the final LCCM, we extracted the individual probability of belonging to a latent class, identified the individuals who belonged to each class, and analyzed the socioeconomic characteristics in each latent class (Table 5). Females primarily formed latent class 1, whereas males dominated latent class 2. In addition, most individuals in both latent classes were under 35 years old, but latent class 2 offered a higher percentage of individuals between 36 and 50 years of age and a slightly higher average age. Likewise, both latent classes comprised highly educated people, but latent class 2 offered the highest percentage of university education. In addition, nearly the entire sample in latent class 2 worked. Similarly, individuals mostly lived in two-member households, but latent class 2 provided a higher percentage of living alone and a lower average number of household members. Eventually, the average income in each latent class differed significantly, with latent class 2 grouping the upper-middle income levels. In summary, the average profile of each latent class presented different socioeconomic characteristics. Latent class 1 mainly consisted of highly educated women

Table 5
Demographics, the reason for choosing, and the trip purpose for every latent class (n = 1,536)

	Latent class 1	Latent class 2
Female	54%	43%
Male	46%	57%
Average age	31	33
<35 years old	76%	71%
36-50 years old	18%	23%
>51-65 years old	6%	6%
Secondary	14%	2%
Technical	20%	18%
University	66%	81%
Work	64%	97%
Study	30%	1%
Homemaker	2%	0%
Unemployed	2%	0%
Retired	2%	2%
1 household member	21%	30%
2 household members	46%	43%
3 or more household members	32%	27%
Average income	680 €	1,300 €
0 vehicles	38%	37%
1 vehicle	48%	50%
2 or more vehicles	14%	13%
Driver's license (yes)	46%	54%
The fastest way to get there	34%	30%
Personal safety	16%	20%
Time adequacy	14%	17%
Quality of service	11%	10%
Comfort	14%	13%
Inexpensive	6%	5%
Reliability	5%	6%
Homecoming	36%	40%
Leisure	21%	23%
Work	11%	15%
Visit	10%	8%
Studies	7%	2%
Shopping	5%	4%
Procedures	4%	4%
Health	5%	4%

under 35 years with a monthly income of 680 €, predominantly working or studying, and living with another person. In turn, latent class 2 broadly comprised highly educated men under 35 years old with a monthly income of 1,300 €, almost everyone employed, and living either with another person or alone. Latent class 2 grouped the best-educated and higher-income individuals in the sample, who also offered the highest SVT. Thus, we distinguish distinct segments in the ridesourcing market based on education level and income. Consequently, highly educated and high-income individuals, previously associated with the highest adoption rate (Alemi et al., 2018b), also offered the highest SVT.

Table 5 also shows the characteristics of motorization and travel in each latent class. Most individuals owned one car, although the percentage of individuals with zero vehicles was extremely high in both the latent classes. In addition, approximately 50% of the individuals had driver's licenses in both latent classes. Likewise, for both latent classes, the most common reason for choosing UberX was the speed to get to the destination, and the second one was personal safety. Finally, the most common trip purpose in both latent classes was returning home, following a similar percentage distribution for the remainder. In short, both latent classes presented similar user profiles regarding specific characteristics associated with motorization and travel. In particular, in both latent classes, individuals predominantly owned one or no vehicle and chose UberX for speed and safety, primarily to return home. However, there was a difference in the driver's license, as latent class 1 was dominated by unlicensed individuals, whereas class 2 was the opposite. In other words, latent class 1 grouped UberX users captive of public transport and soft modes (walking and cycling), while latent class 2 grouped individuals who owned private vehicles but preferred not to drive.

In short, we found that two latent classes, differentiated by educational level and income, formed the ridesourcing market. They offered distinct ridesourcing choice behaviors based on widely dissimilar SVTs. Thus, ridesourcing market segmentation seemed fundamental for understanding regular operations. Moreover, there was a strong substitution effect between ridesourcing and transit use.

Policymakers and transportation planners need to understand how ridesourcing has changed travel patterns to design current and future transportation systems. This research drew several relevant results for transportation planning and policymaking. Ridesourcing harmed transit ridership and increased traffic and congestion in a medium-sized town; therefore, policymakers and transportation planners should have improved the competitiveness of the public transit system. Based on our results, there should have been two fundamental improvements: transit should have been faster and safer. In addition, UberX users provided high SVTs, so they were willing to pay a higher fare in exchange for faster and safer bus services. Hence, transportation planning and policymaking had room to increase fares to defray for improvements in rapidity and safety. Private transport solutions created environmental and social damage, and we should have disincentivized them because they competed with public transit systems. Similarly, our results showed that UberX users preferred not to use cars or drive, which would be another factor driving the increase in transit ridership. In addition, our results provided two different latent classes of users. Policymakers and transportation planners could have used this information to determine the targeted individuals as they deployed different policy and marketing strategies. They would have been able to customize the policies and marketing strategies (routes, frequency, fares, etc.) for each latent class.

Finally, we highlight some aspects of further ridesourcing mode choice analyses. We obtained high SVTs compared to the average wage rates for both latent classes, and personal safety was the main reason for ridesourcing use. Further research should include a latent variable related to personal safety to test its relevance. In addition, other latent or attitudinal variables and explanatory variables associated with context and lifestyle might also be relevant (Alemi et al., 2018b). Further research should also investigate the impact of autonomous vehicles and include psychological factors related to social distancing given the COVID-19 pandemic. Our research showed that ridesourcing use depended on context, data collection time, and sample members. Thus, our results might not fully represent current behavior because the ridesourcing market and society have changed since our data collection. We envision regularly conducting studies on the services included in shared mobility to test this hypothesis.

Conclusion

This study aims to model ridesourcing mode choice across different latent classes to ascertain who chose ridesourcing and why. We estimated an LCCM with two latent classes using mode choice revealed preference surveys collected in 2017 from UberX users in Viña del Mar, Chile. We then characterized individuals who belonged to each latent class and the entire sample and calculated the SVT.

Most UberX users were male workers aged 20–35 years (millennial), highly educated, and belonged to the middle-income segment. In addition, UberX users were mostly captive to public transport and soft modes (walking and cycling) or owned private vehicles but preferred not to drive. The most relevant trip purposes were returning home and leisure. Similarly, the main reasons for choosing UberX were the rapidity of getting to the destination, personal safety, and time adequacy. UberX users appreciated using a rapid and safe mode of transportation according to their mobility needs at any given time.

An affordable, short-distance and quick UberX trip pattern has emerged. 80% of the users would have made the trip by bus or shared taxi if UberX were unavailable, while the market share captured from car and taxi was minor, thereby threatening public transport. In addition, the modal shift pattern impacted traffic externalities because ridesourcing significantly increased the number of private vehicles circulating. In other words, UberX boosted traffic and congestion.

Both latent classes offered distinct socioeconomic characteristics. Latent class 2 grouped the best-educated and higher-income individuals, who provided the highest SVT and owned private vehicles but preferred not to drive. Both latent classes provided high SVTs compared to average wage rates. However, both latent classes exhibited similar motorization characteristics.

In short, we found that two latent classes, differentiated by educational level and income, formed the ridesourcing market. In addition, they offered distinct ridesourcing choice behavior based on widely dissimilar SVTs. Similarly, there was a strong substitution effect between ridesourcing and transit use. The results imply that policymakers and transportation planners could have increased the competitiveness of the public transit system by improving rapidity and safety, having room to increase the fares to defray the improvements. In addition, they could have used information related to the latent classes to customize the policies and marketing strategies (routes, frequency, fares, etc.) for each latent class.

CRediT authorship contribution statement

Gustavo García-Melero: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Rubén Sainz-González:** Conceptualization, Methodology, Validation, Writing – original draft, Writing – review & editing, Supervision. **Pablo Coto-Millán:** Conceptualization, Methodology, Writing – original draft, Supervision. **Alejandra Valencia-Vásquez:** Investigation, Resources, Data curation, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank two reviewers for suggesting substantial improvements to the original manuscript. Any persistent mistake is the responsibility of the authors. This research received no specific grant from the public, commercial, or not-for-profit funding agencies.

References

Acheampong, R.A., Siiba, A., Okyere, D.K., Tuffour, J.P., 2020. Mobility-on-demand: An empirical study of internet-based ride-hailing adoption factors, travel characteristics and mode substitution effects. Transp. Res. Part C Emerg. Technol. 115, 102638 https://doi.org/10.1016/j.trc.2020.102638.

Alemi, F., Circella, G., Handy, S., Mokhtarian, P., 2018a. What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. Travel Behav. Soc. 13, 88–104. https://doi.org/10.1016/j.tbs.2018.06.002.

Alemi, F., Circella, G., Mokhtarian, P., Handy, S., 2018b. Exploring the latent constructs behind the use of ridehailing in California. J. Choice Model. 29, 47–62. https://doi.org/10.1016/j.jocm.2018.08.003.

Asgari, H., Jin, X., 2020. Incorporating habitual behavior into Mode choice Modeling in light of emerging mobility services. Sustain. Cities Soc. 52, 101735 https://doi.org/10.1016/j.scs.2019.101735.

Azimi, G., Rahimi, A., Asgari, H., Jin, X., 2020. Role of Attitudes in Transit and Auto Users' Mode Choice of Ridesourcing. Transp. Res. Rec. J. Transp. Res. Board 2674, 1–16. https://doi.org/10.1177/0361198120941841.

Berger, T., Chen, C., Frey, C.B., 2018. Drivers of disruption? Estimating the Uber effect. Eur. Econ. Rev. 110, 197–210. https://doi.org/10.1016/j.euroecorev.2018.05.006.

Chavis, C., Gayah, V.V., 2017. Development of a Mode Choice Model for General Purpose Flexible-Route Transit Systems. Transp. Res. Rec. J. Transp. Res. Board 2650, 133–141. https://doi.org/10.3141/2650-16.

Circella, G., Alemi, F., 2018. Transport Policy in the Era of Ridehailing and Other Disruptive Transportation Technologies. In: Kamargianni, M., Shiftan, Y. (Eds.), Preparing for the New Era of Transport Policies: Learning from Experience. Elsevier Science & Technology, pp. 119–144.

Contreras, S.D., Paz, A., 2018. The effects of ride-hailing companies on the taxicab industry in Las Vegas, Nevada. Transp. Res. Part A Policy Pract. 115, 63–70. https://doi.org/10.1016/j.tra.2017.11.008.

Cooper, D., Castiglione, J., Mislove, A., Wilson, C., 2018. Profiling Transport Network Company Activity using Big Data. Transp. Res. Rec. J. Transp. Res. Board 2672, 192–202. https://doi.org/10.1177/0361198118798459.

Dong, X., 2020. Trade Uber for the Bus?: An Investigation of Individual Willingness to Use Ride-Hail Versus Transit. J. Am. Plan. Assoc. 86, 222–235. https://doi.org/10.1080/01944363.2019.1687318.

Fu, X., Juan, Z., 2017. Accommodating preference heterogeneity in commuting mode choice: an empirical investigation in Shaoxing. China. Transp. Plan. Technol. 40, 434–448. https://doi.org/10.1080/03081060.2017.1300240.

Greene, W.H., 2012. NLOGIT Version 5 Reference Guide. Econometric Software Inc, Plainview, NY.

Greene, W.H., Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transp. Res. Part B Methodol. 37, 681–698. https://doi.org/10.1016/S0191-2615(02)00046-2.

Habib, K.N., 2019. Mode choice modelling for hailable rides: An investigation of the competition of Uber with other modes by using an integrated non-compensatory choice model with probabilistic choice set formation. Transp. Res. Part A Policy Pract. 129, 205–216. https://doi.org/10.1016/j.tra.2019.08.014.

Ho, C.Q., Hensher, D.A., Mulley, C., Wong, Y.Z., 2018. Potential uptake and willingness-to-pay for Mobility as a Service (MaaS): A stated choice study. Transp. Res. Part A Policy Pract. 117, 302–318. https://doi.org/10.1016/j.tra.2018.08.025.

Hurtubia, R., Nguyen, M.H., Glerum, A., Bierlaire, M., 2014. Integrating psychometric indicators in latent class choice models. Transp. Res. Part A Policy Pract. 64, 135–146. https://doi.org/10.1016/j.tra.2014.03.010.

Iqbal, M., 2020. Uber Revenue and Usage Statistics (2020) [WWW Document]. URL https://www.businessofapps.com/data/uber-statistics/ (accessed 9.1.20).

Jin, S.T., Kong, H., Wu, R., Sui, D.Z., 2018. Ridesourcing, the sharing economy, and the future of cities. Cities 76, 96–104. https://doi.org/10.1016/j.cities.2018.01.012.

Kim, S.H., Mokhtarian, P.L., 2018. Taste heterogeneity as an alternative form of endogeneity bias: Investigating the attitude-moderated effects of built environment and socio-demographics on vehicle ownership using latent class modeling. Transp. Res. Part A Policy Pract. 116, 130–150. https://doi.org/10.1016/j.tra.2018.05.020.

Lavieri, P.S., Dias, F.F., Juri, N.R., Kuhr, J., Bhat, C.R., 2018. A Model of Ridesourcing Demand Generation and Distribution. Transp. Res. Rec. J. Transp. Res. Board 2672, 31–40. https://doi.org/10.1177/0361198118756628.

Lee, Z.W.Y., Chan, T.K.H., Balaji, M.S., Chong, A.-Y.-L., 2018. Why people participate in the sharing economy: an empirical investigation of Uber. Internet Res. 28, 829–850. https://doi.org/10.1108/IntR-01-2017-0037.

Li, W., Pu, Z., Li, Y., (Jeff) Ban, X.,, 2019. Characterization of ridesplitting based on observed data: A case study of Chengdu, China. Transp. Res. Part C Emerg. Technol. 100, 330–353. https://doi.org/10.1016/j.trc.2019.01.030.

Ma, T.-Y., Gerber, P., Carpentier, S., Klein, S., 2015. Mode choice with latent preference heterogeneity: a case study for employees of the EU institutions in Luxembourg. Transp. A Transp. Sci. 11, 441–463. https://doi.org/10.1080/23249935.2015.1007175.

Martinez, L.M., Viegas, J.M., 2017. Assessing the impacts of deploying a shared self-driving urban mobility system: An agent-based model applied to the city of Lisbon. Portugal. Int. J. Transp. Sci. Technol. 6, 13–27. https://doi.org/10.1016/j.ijtst.2017.05.005.

National Statistics Institute of Chile, 2017. Resultados de comunas – Censo 2017 [WWW Document]. accessed 7.17.21. http://www.censo2017.cl/descargue-aqui-resultados-de-comunas/.

National Statistics Institute of Chile, 2015. Encuesta suplementaria de ingresos [WWW Document]. accessed 7.17.21. https://www.ine.cl/estadisticas/sociales/ingresos-y-gastos/encuesta-suplementaria-de-ingresos.

Nie, Y. (Marco), 2017. How can the taxi industry survive the tide of ridesourcing? Evidence from Shenzhen, China. Transp. Res. Part C Emerg. Technol. 79, 242–256. https://doi.org/10.1016/j.trc.2017.03.017.

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. Transp. Policy 45, 168–178. https://doi.org/10.1016/j.tranpol.2015.10.004.

Saxena, N., Hossein Rashidi, T., Auld, J., 2019. Studying the tastes effecting mode choice behavior of travelers under transit service disruptions. Travel Behav. Soc. 17, 86–95. https://doi.org/10.1016/j.tbs.2019.07.004.

Schwieterman, J., Smith, C.S., 2018. Sharing the ride: A paired-trip analysis of UberPool and Chicago Transit Authority services in Chicago. Illinois. Res. Transp. Econ. 71, 9–16. https://doi.org/10.1016/j.retrec.2018.10.003.

Shaheen, S., Chan, N., 2016. Mobility and the Sharing Economy: Potential to Overcome First- and Last-Mile Public Transit Connections, UC Berkeley: Transportation Sustainability Research Center. https://doi.org/10.7922/G2862DN3.

Shaheen, S., Cohen, A., 2019. Shared ride services in North America: definitions, impacts, and the future of pooling. Transp. Rev. 39, 427–442. https://doi.org/10.1080/01441647.2018.1497728.

Tirachini, A., 2020. Ride-hailing, travel behaviour and sustainable mobility: an international review. Transportation (Amst). 47, 2011–2047. https://doi.org/10.1007/s11116-019-10070-2.

Tirachini, A., del Río, M., 2019. Ride-hailing in Santiago de Chile: Users' characterisation and effects on travel behaviour. Transp. Policy 82, 46–57. https://doi.org/10.1016/j.tranpol.2019.07.008.

Tirachini, A., Gomez-Lobo, A., 2020. Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile. Int. J. Sustain. Transp. 14, 187–204. https://doi.org/10.1080/15568318.2018.1539146.
Walker, J., Ben-Akiva, M., 2002. Generalized random utility model. Math. Soc. Sci. 43, 303–343. https://doi.org/10.1016/S0165-4896(02)00023-9.
Ward, J.W., Michalek, J.J., Azevedo, I.L., Samaras, C., Ferreira, P., 2019. Effects of on-demand ridesourcing on vehicle ownership, fuel consumption, vehicle miles traveled, and emissions per capita in U.S. States. Transp. Res. Part C Emerg. Technol. 108, 289–301. https://doi.org/10.1016/j.trc.2019.07.026.
Yan, X., Levine, J., Zhao, X., 2019. Integrating ridesourcing services with public transit: An evaluation of traveler responses combining revealed and stated preference data. Transp. Res. Part C Emerg. Technol. 105, 683–696. https://doi.org/10.1016/j.trc.2018.07.029.
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. Transp. Res. Part A Policy Pract. 119, 383–392. https://doi.org/10.1016/j.tra.2018.11.018.
Yu, B., Ma, Y., Xue, M., Tang, B., Wang, B., Yan, J., Wei, Y.-M., 2017. Environmental benefits from ridesharing: A case of Beijing. Appl. Energy 191, 141–152. https://doi.org/10.1016/j.apenergy.2017.01.052.
Yu, H., Peng, Z.-R., 2020. The impacts of built environment on ridesourcing demand: A neighbourhood level analysis in Austin. Texas. Urban Stud. 57, 152–175. https://doi.org/10.1177/0042098019828180.