Uncovering the link between intra-individual heterogeneity and variety seeking: the case of new shared mobility

Fangqing Song · Stephane Hess · Thijs Dekker

Abstract Preferences can vary both across respondents (i.e. inter-respondent preference heterogeneity) and across choice tasks within respondents (i.e. intra-respondent preference heterogeneity). Ignoring the existence of intra-respondent preference heterogeneity could bias preference elicitation and demand forecast. Thus far, most studies covering inter- and intra-respondent preference heterogeneity have applied the mixed multinomial logit (MMNL) model. Meanwhile, the behavioural explanations for such preference variations remain under-explored. This paper accommodates inter- and intra-respondent preference heterogeneity through a two-layer latent class modelling structure, where the continuous random distributions are replaced with discrete mixtures in both layers. A latent variable representing variety-seeking is included to explain class membership probabilities, offering additional behavioural insights concerning the source of preference heterogeneity both across and within respondents. Two aspects associated with variety-seeking are examined: novelty-seeking (i.e. the inclination to adopt new modes) and alternation (i.e. the tendency to vary one's behaviour regularly by selecting different modes continuously). In the context of new shared mobility, this paper finds the role of both aspects in preference heterogeneity. Specifically, novelty seekers are found to be more likely to fall into the class with higher probabilities of switching from existing modes to the new air taxi service than novelty avoiders, and alternation seekers are more likely to belong to the class with higher proba-

Fangqing Song
Institute for Transport Studies & Choice Modelling Centre, University of Leeds, UK
E-mail: F.Song1@leeds.ac.uk (F.Song)

Stephane Hess
Institute for Transport Studies & Choice Modelling Centre, University of Leeds, UK
E-mail: S.Hess@leeds.ac.uk (S. Hess)

Thijs Dekker
Institute for Transport Studies & Choice Modelling Centre, University of Leeds, UK
E-mail: T.Dekker@leeds.ac.uk (T. Dekker)

bilities to exhibit intra-respondent preference heterogeneity than alternation avoiders. This paper, therefore, provides empirical evidence to identify the target customers of the new air taxi service.

Keywords inter- and intra-respondent preference heterogeneity · latent variable · latent class model · variety-seeking · vertical take-off and landing · urban air mobility

1 Introduction

Whilst a great deal of attention has been paid to preference variation over choices in revealed-preference (RP) data, for example, day-to-day variability (Cherchi and Cirillo, 2014), preference homogeneity is usually assumed across choice tasks in repeated stated choice (SC) data. This is supported by the fact that, unlike RP surveys which can collect data over a longer time span where preference variation might arise, SC surveys are usually conducted in a single sitting so that respondents' preferences are normally considered stable throughout the SC survey. Nevertheless, an increasing number of studies have demonstrated the presence of preference variations within a respondent (i.e. intra-respondent preference heterogeneity) in SC surveys (Hess and Rose, 2009; Hess and Train, 2011; Hess and Giergiczny, 2015; Becker et al., 2018).

Despite the growing interest in accommodating intra-respondent preference heterogeneity on top of inter-respondent preference heterogeneity, there remain research gaps to be bridged. Firstly, the common practice to account for inter- and intra-respondent preference heterogeneity is establishing the model within a Mixed Multinomial Logit (MMNL) framework by incorporating two layers of preference heterogeneity, i.e. one across respondents and another one across choice tasks. However, this is achieved at a high computational cost because calculating the resulting log-likelihood involves integration at the two layers (Hess and Train, 2011). Secondly, existing studies on inter- and intra-respondent preference heterogeneity still lack an explicit behavioural explanation of the sources of the intra-respondent preference heterogeneity. Therefore the main objective of the present paper is to accommodate inter- and intra-respondent preference heterogeneity at a lower computational cost whilst providing a behavioural explanation for intra-respondent preference heterogeneity.

In this paper, we hypothesise that preference heterogeneity can be associated with a latent construct of variety-seeking. Regardless of different modelling methods, variety-seeking can reflect the tendency to experience new things (i.e. novelty-seeking) or to vary choices over a period of time (i.e. alternation) (McAlister and Pessemier, 1982; Ha and Jang, 2013). While some people intrinsically prefer exploring novel experiences, others would be more inclined to avoid changes and stick to their habitual travel experiences; moreover, some people have stronger tendencies to vary their choices over time, whereas others' choices remain relatively more stable. Our adopted modelling approach treats variety-seeking as an underlying personality trait. As such, the novelty-seeking aspect of variety-seeking relates to preference heterogeneity across respondents, while the alternation aspect of variety-seeking is connected with the preference heterogeneity across choices.

Variety seeking might arise, especially when new alternatives are introduced to the market. We test our hypotheses on novelty seeking and alternation in the context of a mode choice experiment where new shared mobility is introduced. In each choice task, existing ground-based modes are presented together with an upcoming novel travel mode, i.e. air taxi (also known as "flying taxi"). This is an on-demand vertical take-off-and-landing (VTOL) service and a vital element of the broader concept of "Urban Air Mobility" (UAM). Although UAM has been gaining substantial investment interest in recent years, commercial air taxi products are still in development[1] and travel behaviour analysis remains limited compared to other modes.

This research thereby has a triple contribution. Methodologically, this research provides empirical evidence of the presence of inter- and intra-respondent preference heterogeneity through a modified latent class modelling structure. From a behavioural perspective, this paper offers behavioural explanations of inter- and intra-respondent preference heterogeneity and contributes to the application of variety-seeking theory in the transport realm. In addition, this paper provides empirical evidence about consumer preferences towards the upcoming air taxi service, which can be helpful to policymakers in designing market strategies and improving the level of services.

The remainder of this paper is organised as follows. Section 2 reviews existing literature about intra-respondent preference heterogeneity, variety-seeking and urban air mobility. Section 3 describes how the survey was carried out and presents a descriptive analysis of the data. Our approach to account for inter- and intra-respondent preference heterogeneity is explained in Section 4, followed by a discussion of the estimation results in Section 5. Conclusions are presented in the last section.

2 Literature review

2.1 Intra-respondent preference heterogeneity

With regard to recovering preference heterogeneity using repeated SC data, most studies assume that preferences of a respondent remain stable across choices (i.e. intra-respondent preference homogeneity) whilst allowing for variations in preferences across respondents (i.e. inter-respondent preference heterogeneity). Ignoring the existence of intra-respondent variations could mislead preference elicitation and demand forecasts (Ben-Akiva et al., 2019).

Typically, studies accounting for inter- and intra-respondent preference heterogeneity incorporate two layers of preference heterogeneity within the mixed multinomial logit (MMNL) model. That is, for a given preference parameter, a continuous mixing density across respondents and an additional continuous mixing density across observations are specified. This specification essentially assumes random variations around the sample-level average preference both across respondents (i.e. the panel) and across choice scenarios (i.e. the cross-sectional). Examples can be found in Hess and Rose (2009); Hess and Train (2011); Hess and Giergiczny (2015).

[1] For example, Airbus is leading the European commission's Urban Air Mobility Initiative; and NASA aims to establish and expand the UAM network encompassing air shuttle, air taxi and air ambulance, each fitting a specific area of the wider UAM spectrum (Goyal, 2018)

The accommodation of inter- and intra-respondent preference heterogeneity is achieved at a high computational cost because evaluating the log-likelihood involves integration over random distributions at both inter- and intra-respondent layers (Hess and Train, 2011). Recently, efforts have been made to accommodate inter- and intra-respondent preference heterogeneity through other modelling frameworks or estimation methods. For example, given that both MMNL and LC models can accommodate preference heterogeneity whilst the latter is relatively easier to estimate, Hess (2014) raised the question "whether replacing one layer with weighted summation through a latent class structure would be beneficial". It is suggested that the preference heterogeneity across respondents can be replaced by a latent class structure, leaving only one layer of integration over observations in estimation. However, this idea has not been implemented in an empirical analysis yet, nor has it been extended to replacing both layers of continuous mixtures with discrete mixtures to reduce the computational cost to a greater extent.

Apart from this strategy, Bayesian analysis has been used to quicken the estimation when integration is needed at both layers. For example, Dekker et al. (2016) investigated the impact of decision uncertainty through an integrated choice and latent variable (ICLV) model, where the latent uncertainty was introduced at the choice task level while inter-respondent preference variation was accounted for in the alternative specific constants (ASC). Becker et al. (2018) also introduced a Hierarchical Bayes estimator for MMNL models with inter- and intra-respondent preference heterogeneity through Markov Chain Monte Carlo (MCMC) estimation rather than the commonly used maximum simulated likelihood estimation, leading to a substantial reduction in computational time. Krueger et al. (2019) further derived a Variational Bayes method for posterior inference in MMNL models that account for inter- and intra-respondent preference heterogeneity. Zhu et al. (2020) uncovered the inter-respondent preference heterogeneity with a collaborative learning structure and the intra-respondent preference heterogeneity with a time-dependent model based on data collected from an online stated choice experiment.

Meanwhile, a growing effort can be seen in the existing studies on uncovering the behavioural explanations of this intra-respondent preference heterogeneity in SC experiments. Hess and Rose (2009) suggested that the preferences of a given individual may change over stated choice tasks because of learning effect, cognitive burden, etc. In the presence of a new alternative, its unique attributes may also lead to ambiguity in interpreting their meanings. A recent study on environmental services by Hess and Giergiczny (2015) showed that the preference instability across SC tasks could be higher for attributes which respondents are unfamiliar with. Moreover, Dekker et al. (2016) inferred from their analysis that greater uncertainty would not only decrease the scale of utility but also increase the likelihood of choosing the status-quo or opt-out option.

2.2 Variety-seeking

McAlister and Pessemier (1982) and Pessemier (1985) suggest that respondents' varied behaviour can be attributed to external triggers and intrinsic direct motives. Variety seeking behaviour can be classified as an intrinsic direct motive, because individuals may have a desire for exploring something unfamiliar, or alternate among familiar options (Trijp et al., 1996; Ha and Jang, 2013). Henceforth, we refer to *'novelty-seeking'* as an individual's tendency to explore something new and unfamiliar and define *'alternation'* as the phenomenon of a respondent choosing a different alternative from their choice set over time due to the utility derived from the change itself. The latter utility is irrespective of the alternative that the decision-maker switches to or from (Borgers et al., 1989; Givon, 1984). Both aspects of variety-seeking have been widely addressed in consumer and psychology research (e.g. (Givon, 1984; Borgers et al., 1989; Chintagunta, 1998)). However, they are rarely accommodated in discrete choice analyses using stated choice data in the transport realm.

Regarding methods of analysis, some variety-seeking studies explicitly specify the mathematical structure of switching. For example, Givon (1984) proposed an alternation-based model assuming that the probability of switching choices depend on the preference for the currently chosen alternative and the preference for switching. Borgers et al. (1989) focused on transition probabilities in recreational choices, assuming that the probability of choosing differently in two consecutive occasions was a function of the (dis)similarity between the currently and previously chosen alternatives. Chintagunta (1998) developed a brand switching model based on the hazard function, which allowed the brand choice probabilities to vary over time and found that variety seekers are more likely to purchase a brand positioned farthest away from the previously purchased brand.

In another stream of work, psychometric scales have been created as tools to measure variety-seeking tendencies. Most psychometric scales are context-specific (e.g. Pearson (1970); Pessemier and Handelsman (1984); Lee and Crompton (1992); Wills et al. (1994); Baumgartner and Steenkamp (1996); Trijp et al. (1996)). Variety-seeking is commonly treated as a personality trait that varies across respondents. On the one hand, this means that the preference to stick to old habits, resistance to changes, and uncertainty might be stronger for some respondents, whereas others favour unfamiliarity and novelty. On the other hand, this means some people might have a stronger desire for alteration and hence would choose a broader range of different alternatives compared to others (i.e. alternation aspect). Nevertheless, the statements in the scales of variety-seeking usually do not clearly distinguish between the novelty-seeking and alternation aspects as these two aspects are essentially correlated and intertwined.

Responses to psychometric scales can be used to segment markets (e.g. Van Trijp and Steenkamp (1992); Assaker and Hallak (2013)). Such responses can also be used in Structural Equation Models to analyse the correlation between variety-seeking tendencies and other constructs. For example, Jang and Feng (2007) examined the relationship between novelty-seeking and tourists' intentions to revisit destinations. Responses to psychometric scales have also been included in discrete choice models. Rieser-Schüssler and Axhausen (2012) and Song et al. (2018) both treated variety-seeking as a latent variable explaining choices and the responses to the statements from a psychometric scale on variety-seeking. Neither paper accounted for the alternation aspect of variety seeking.

2.3 Urban air mobility

Urban Air Mobility is a new form of shared mobility.[2] It describes an air transportation system that enables on-demand, point-to-point and highly automated passenger or package-delivery air travel services at a low altitude within and around populated urban areas (Goyal, 2018). Ultimately, the UAM system could enable travellers to find an "air taxi" nearby through mobile apps and possibly to share the space and travel cost with other air-poolers on the same aerial vehicle, just like ride-sourcing service on land.[3]

Electric or hybrid Vertical Take-off and Landing (VTOL) is recognised as the primary type of aerial vehicle for UAM in the near future[4]. The deployment of VTOL would not take up much valuable urban space for constructing "airports", "runways" etc., as high buildings' rooftops can be transformed into take-off and landing pads. Additionally, autonomous VTOL is beneficial to solve a shortage of pilots. In general, VTOLs are expected to minimise travel time, mitigate traffic congestion on the ground, reduce operation errors and contribute to zero emissions (Holden and Goel, 2016).

Various methods have been adopted to evaluate the impacts of on-demand ride services on urban development, to assess or optimise the system performance of on-demand ride service networks, and to improve the understanding of individual behaviour in the new context accordingly, etc. However, the research predominantly focuses on ground-based services. In constrast, little effort has been devoted to UAM, and there is a lack of such empirical evidence in the context of air taxi. Mode choice studies between air and other

[2] According to Shaheen et al. (2016), shared mobility refers to *"an innovative transportation strategy that enables users to gain short-term access to transportation modes on an as-needed basis."*

[3] Air-taxi is different from "flight-sharing". The latter (e.g. Wingly, Coavmi) allows certified private pilots to carry passengers such that the travel cost could be split among passengers including the pilots. In the European Union, flight-sharing is permitted on a non-commercial basis (EASA, 2018), whereas flight-sharing has been completely banned in the U.S., which has caused much criticism (Koopman and Dourado, 2017).

[4] On-demand helicopter platforms already exist (e.g. Voom by Airbus in São Paulo and Mexico City). However, it is recognised that distributed electric propulsion and autonomous operation technologies, which are features of VTOL, are the key to addressing the major barriers to the large-scale commercialised operation of UAM, such as safety, noise, emission and vehicle performance (Holden and Goel, 2016). Ultimately, drones will be adopted to transport passengers, which are expected to create zero emissions.

modes (e.g. high-speed rail) for medium-to-long distance intercity travel have been conducted widely (e.g. Park and Ha, 2006; Román et al., 2007; Hess et al., 2018). Regarding urban travel, flying has rarely been treated as an option as scheduled airline services are usually considered not competitive for short-distance travel.

In light of the introduction of the new air taxi service, fit-for-purpose empirical analyses need to be conducted with the help of specifically-designed stated choice data to explain individual preferences and the impact on travel demand. Some studies calibrated (rather than estimated) a multinomial logit model based on existing travel surveys excluding the new on-demand air service and then applied the obtained coefficients to compute aggregate mode shares for the new market with the hypothetical on-demand air service (e.g. Pu et al. 2014; Joshi et al. 2014; Baik et al. 2008). Thus, empirical analysis is needed to verify the assumptions about sensitivities towards various level-of-service attributes and explain the behavioural mechanisms behind individual choices.

Peeta et al. (2008) estimated a binary choice model based on stated choice data to analyse the probability of switching to the new on-demand "very light jet" service rather than the novel UAM services. More recently, Fu et al. (2018) used stated choice data to examine mode choice behaviour amongst private car, public transit, autonomous vehicle and autonomous VTOL air taxi via MNL models. However, the model specification could have been improved to better account for preference heterogeneity across respondents. For example, although the author had collected information related to respondents' attitudes towards adopting new autonomous transportation modes, this information was not accommodated in the model. Binder et al. (2018) and Garrow et al. (2019) are also empirical studies on mode choices between electric VTOL air taxi and other modes. However, the experimental design on mode choices lacks sufficient variations in the attribute levels, and the study was only focused on survey design without qualitative and modelling analysis. This work was later extended in Garrow et al. (2020) where factor analysis was performed followed by cluster analysis to explore market segmentation. Al Haddad et al. (2020) lately developed multinomial logit (MNL) models and ordered logit models with stated preference data to explore the factors influencing respondents' adoption and use of VTOL, where the adoption time horizon was treated as the dependent variable rather than the conventional mode alternatives. To the best of our knowledge, no other empirical analyses explored the preferences for on-demand aerial services, particularly in the new context of Urban Air Mobility, where air taxi is expected to be powered by (autonomous) VTOL vehicles.

3 Survey and data

3.1 UberAIR service context

This paper uses data provided by Uber on mode choice amongst different alternatives, including its upcoming on-demand electric VTOL air taxi service, i.e. UberAIR.[5][6]

It is expected to cut existing door-to-door travel times by an estimated 30% to 60% and create zero emissions and low levels of noise (Holden and Goel, 2016). Flights may be shared with other riders, leading to a reduced cost per individual. Passengers will be able to book UberAIR services with the same mobile app as existing ground-based services. Moreover, Uber's air and ground services may be integrated and coordinated in operation, such that passengers can book door-to-door trips through a single request and payment and be driven by ground service like UberX to/from the UberAIR take-off/landing pads. Fig. 1 illustrates the UberAIR service.

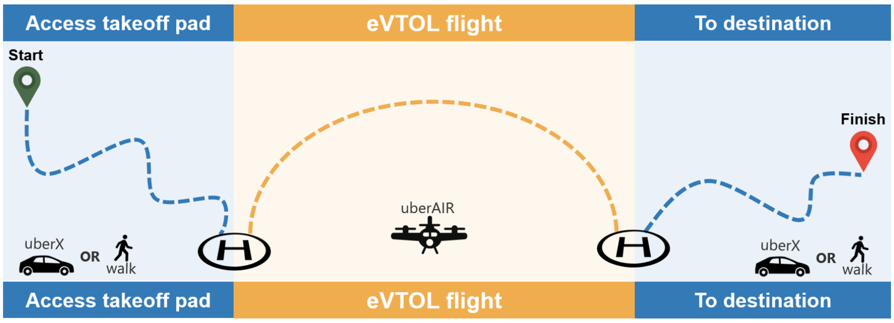

Fig. 1 Illustration of UberAIR service.

3.2 Questionnaire and respondent sampling

Since the commercialised operation of UberAIR has not yet been realised, we cannot use revealed preference (RP) data to analyse people's preferences and

[5] The University of Leeds, UK, was provided with anonymised data by Uber Technologies, Inc. ("Uber"). Neither the University of Leeds nor the authors received funding or financial support from Uber. The views, opinions, and conclusions expressed in this article are those of the authors and do not constitute any representation of Uber.

[6] Uber Elevate planned to launch its "UberAIR" service with commercial flight operations in Dallas-Fort Worth and Los Angeles in 2023. However, in December 2020, it was announced that Uber Elevate would be acquired by the start-up Joby Aviation and the respective services of both companies would be integrated. As our data was collected in 2018 and the paper was initially submitted in 2019, the new air taxi service is still referred to as "UberAIR" in the present paper.

trade-offs between different level-of-service attributes. Instead, a stated choice (SC) survey was conducted.

The survey took around 15min to complete and was mainly comprised of five components: 1) screening questions; 2) trip experience; 3) SC survey; 4) attitudinal statements; and 5) socio-demographic characteristics.

The survey was aimed at people living in the greater Dallas-Fort Worth or Los Angeles areas. Respondents were invited from four groups: LA online panel, DFW online panel, LA Uber customer list, and DFW Uber customer list. The online panel was general population and was representative of resident Census demographics, screening only for a qualifying trip within the region. The screening questions were related to respondents' recent trip experiences. If the respondent could not meet all of the criteria below, they would be disqualified. As to respondents from Uber customer lists, apart from the requirements mentioned below, they would also be disqualified if they had not used a ride-sourcing service in the past month. The sampling criteria are:

- Home zip code match qualifying zip code for the targeted location (Dallas-Fort Worth or Los Angeles MSAs);
- Having used at least one of the following transportation modes and services within the last month - Personal or household vehicle; Rent vehicle; Car-share service; Bus; Light rail, metro, or subway; Commuter rail; Taxicab; Ride-sourcing;
- Having completed at least one ground trip that took place in, around, or through the Dallas-Fort Worth/Los Angeles area;
- The trip was between 7-75 miles (one-way);
- The trip took at least 30 minutes in total (one-way);
- The trip purpose was one of the following purposes - Work commute; Other work-related business; Go to/from school; Go to/from airport; Shopping; Social or recreational; Entertainment event; Other personal business.

Disqualified respondents did not need to take the SC survey but were branched directly to the attitudes and socio-demographics so that they could finish the survey. Regarding qualified participants, their qualified trips would be regarded as the "reference trips" which would feed into the following SC survey. In the SC survey, individual-specific reference mode was always shown as the first alternative; meanwhile, UberX, UberPOOL and the new UberAIR were always presented in the SC survey. The modelling work only makes use of the responses from qualified participants who completed the whole questionnaire. The responses obtained from disqualified respondents were not used for model estimation in the current study, even though they were presented with the attitudinal statements.

A total of 2,607 qualified respondents finished the entire survey. It needs to be noted that only a limited number of people used rental vehicle/car-share services, taxicab, other ride-sourcing services or UberBLACK/UberSELECT for their reference trips, accounting for much smaller shares (7.2% altogether) compared to the other modes. This leads to a situation where these four alternatives were rarely available in the SC survey compared to the other modes. Therefore, in order to improve model efficiency, the discrete choice models included in this paper are all estimated on a subset of the qualified sample, where only respondents using personal/household vehicle, transit, UberX or UberPOOL for their reference trips are involved. Those who travelled by rental vehicle/car-share service, taxicab, other ride-sourcing service or UberBLACK/UberSELECT in their reference trips were excluded. Consequently, 2,419 respondents are used for model estimation. The analysis and discussion in the remainder of this paper are all established on these 2,419 respondents.

Table 1 illustrates the sampling results among these 2,419 respondents. It can be found that different trip purposes were almost evenly distributed among the sample. Over 60% of respondents used personal/household vehicles in the reference trip, whereas TNC services (i.e. UberX and UberPOOL) dominated the remaining 40% of the sample and the rest used public transport for their reference trips. This sample is, of course, not necessarily representative of the real-world travelling population and is potentially biased towards existing users of Uber services. However, the purpose of the present study is exploratory and focused on specific behavioural traits rather than seeking representative findings for policy work.

Table 1 Reference trips of sampled respondents

		Count	Percentage(out of 2,419 respondents)
Trip purpose	Work commute	310	12.8%
	Other work-related business	307	12.7%
	Go to/from school	274	11.3%
	Go to/from airport	315	13.0%
	Shopping	308	12.7%
	Social or recreational	306	12.6%
	Entertainment event	294	12.2%
	Other personal business	305	12.6%
Trip mode	Personal/Household vehicle	1,540	63.7%
	Transit	142	5.9%
	UberX	542	22.4%
	UberPOOL	195	8.1%

3.3 Trip experience and socio-demographic characteristics

Each qualified respondent was required to provide further information about the reference trip, including departure time, total duration, delay experience, etc. These questions were tailored for respondents based on what the reference mode was. For example, if the reference mode was personal/household vehicle or ride-sourcing, the respondent needed to suggest whether they experienced a delay due to traffic congestion on the trip, how many people were in the vehicle on the trip, etc.

Table 2 summarises the reference trip among the 2,419 selected respondents. Although the average trip distance varies across different reference modes, the average trip time calculated by Google for each reference mode group is around 30min. However, due to delay time, waiting time, access/egress time, etc., the actual door-to-door trip time is much more diverse across reference modes, with transit taking the longest time (86min) and UberX costing just over half of the transit time (45min). Comparing the personal/household vehicle group and UberX group, it can be found that with similar Google-calculated trip distance and trip time, UberX leads to a quarter less total travel time on average than personal/household vehicle, which might be due to the time saving from parking. Moreover, we can also discover that in comparison to UberPOOL, UberX can allow respondents to reach 8.1km farther with 6min less on average, which can be largely attributed to the time spent matching other ride sharers and detouring to their destinations for UberPOOL trips.

Table 2 Descriptive summary of reference trip experience for the focus sample used in modelling (total amount: 2419)

Reference mode	**Personal/ Household vehicle**	**Transit**	**UberX**	**UberPOOL**
Total respondents #	1,540	142	542	195
Respondents # who experienced delay	1,006 (65%)	NA	304 (56%)	134 (69%)
Average total delay time (min)	15	NA	11	17
Average Google-calculated trip distance (mile)	25.5	18	22.7	14.6
Average Google-calculated trip time (min)	33	27	32	26
Average total trip duration (min)	60	86	45	51

Table 3 describes the distribution of various socio-demographic characteristics. Respondents from the Dallas area and Los Angeles area are relatively similar. Females account for two-thirds of the population. A sufficient number of respondents in each age band were approached, with a slight and steady decrease in proportion as age increases, except for the youngest band. Over 93% of the respondents have at least one vehicle in the household. Additionally, while the official statistics show that the median household income (in 2017 inflation-adjusted Dollars) in 2017 is \$54,501 in Los Angeles city and \$47,285 in Dallas city (U.S. Census Bureau, 2018), our sample has a mean household income of \$100,615 and a median household income of \$62,500. This means that our sample contains a higher proportion of rich people than the census. Nevertheless, given that on-demand VTOL air taxi services would inevitably be more expensive, at least initially, than its ground competitors, we think approaching more high-income people is appropriate.

It needs to be noted that this paper mainly aims to accommodate inter- and intra-respondent preference heterogeneity and apply the theory of variety-seeking to investigate the behavioural explanation of this heterogeneity. Uber's mode choice data incorporating air taxi presented a suitable opportunity to delve into this research objective. This paper, however, does not aim to accurately forecast the travel demand of air taxi or calculate the modal split among different modes when air taxi enters the market. Therefore, not having a representative sample does not affect the objective of this paper.

Table 3 Descriptive summary of the focus sample

Socio-demo characteristics	Level	Amount	Percentage (out of 2,419 respondents)
Residence	Dallas	1,101	45.5%
	LA	1,318	54.5%
Gender	Female	1,616	66.8%
	Male	777	32.1%
	Prefer not to say	26	1.1%
Age	18-24	308	12.7%
	25-29	351	14.5%
	30-34	338	14.0%
	35-39	287	11.9%
	40-44	243	10.0%
	45-49	195	8.1%
	50-54	184	7.6%
	55-59	168	6.9%
	60-64	140	5.8%
	65-69	108	4.5%
	70 or older	97	4.0%
Household vehicle	None	151	6.2%
	1 vehicle	809	33.4%
	2 vehicles	962	39.8%
	3 vehicles	331	13.7%
	4 vehicles	114	4.7%
	5 or more vehicles	52	2.1%
Household annual income	<$35,000	479	19.8%
	$35,000-$49,999	335	13.8%
	$50,000-$74,999	416	17.2%
	$75,000-$99,999	368	15.2%
	$100,000-$149,999	341	14.1%
	$150,000-$199,999	153	6.3%
	$200,000-$249,999	75	3.1%
	$250,000-$499,999	62	2.6%
	>$500,000	38	1.6%
	Prefer not to say	152	6.3%

3.4 Stated choice survey

After a brief introduction to UberAIR, each respondent was presented with 10 hypothetical scenarios and was required to choose the most preferred alternative in each scenario. D-efficient experimental design was adopted to generate the stated choice experiment. The experimental design used priors only for the explanatory variables (time, cost, etc.), which were obtained from past non-academic studies, and not for the constants for different modes. As a result, the fact that UberAir does not yet exist is not a problem. Besides, in order to make the choice scenarios more realistic, the hypothetical choice scenarios were framed around the reference trip reported by each respondent about the travel information of a most recent qualified trip.

In each choice task, the first alternative was always related to the reference trip alternative, and the last alternative was always UberAIR. While this potentially introduces ordering effects, this approach was outside the control of the analysis team. Besides, UberX and UberPOOL were always included in each choice task. Hence, if a respondent used a private vehicle or transit as the reference mode, then UberX and UberPOOL would serve as the second and the third alternatives, respectively. In cases where UberX or UberPOOL was the reference mode, UberX or UberPOOL would only appear as the reference mode, i.e. only three alternatives would be available to be selected from. Fig. 2 gives an example of a stated choice task where UberPOOL was identified as the reference mode.

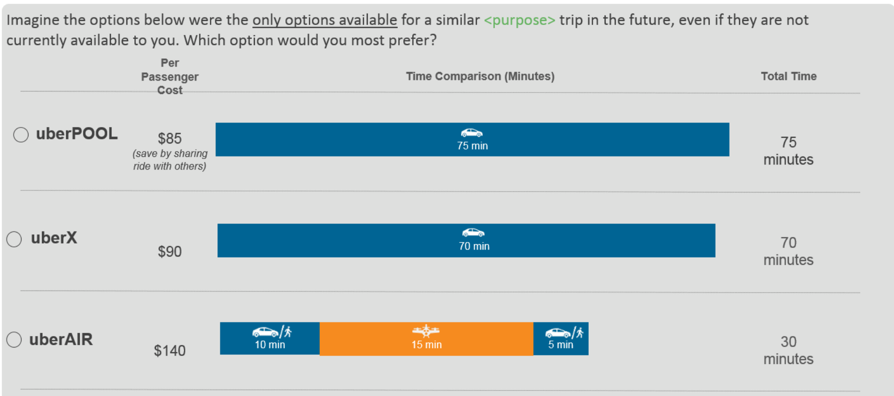

Fig. 2 Example of SC tasks.

A total of 5 attributes, including “travel cost”, “travel time”, “flight time”, “access time”, and “egress time”, were involved in the SC survey, not all of which apply to every alternative. Travel cost was used to describe the other alternatives except for personal/household vehicle. Travel time served as an attribute for all the existing ground-based modes, capturing the total travel time. UberAIR’s total travel time was split into flight time, access time and egress time. The cost levels were chosen to be realistic given the market plans for the new mode. Table 4 gives each attribute’s median and mean values for each alternative across observations. We notice that the distributions of travel time in the SC survey are comparable to the actual travel time in the reference trip shown in Table 2. The travel cost for the car option was set to 0 in the experimental design conducted by Uber. This assumption was made because the cost for the other non-car alternatives is usually paid on a per-trip basis, while the cost associated with a car trip is more complex and less easy to perceive on a per-trip basis as it involves fuel cost, maintenance cost, insurance cost etc.

Table 4 Summary of stated choice tasks

	Alternatives				
	private vehicle	transit	UberX	UberPOOL	UberAIR
Attributes (median, mean)					
travel cost ($)	-	(3, 8)	(35, 40)	(28, 32)	(70, 88)
travel time (min)	(58, 70)	(87, 99)	(51, 62)	(55, 68)	-
flight time (min)	-	-	-	-	(12, 15)
access time (min)	-	-	-	-	(7, 9)
egress time (min)	-	-	-	-	(7, 9)

3.5 Attitudinal statements

In order to capture the influence of underlying psychometric constructs on choice behaviour, attitudinal statements were used to measure these unobserved factors. We excluded statements #4, #9 and #12 on Table 5 from factor analysis as they were considered closely related to brand loyalty and lexicographic decision and environmental-friendliness in respective, and thus irrelevant to the other statements. The remaining statements were used in exploratory factor analysis. The scree plot obtained via parallel analysis (see Fig. 3) shows 5 observed eigenvalues lie above or very close to the corresponding simulated/resampled eigenvalues, suggesting that 2-5 factors could be suitable. We tested different factor solutions and found that loading the remaining 9 statements on 3 factors with a cut-off point of 0.5 gives the most interpretable results. Seven statements were identified, explaining 53% of the variance of the sample. That is, #8 and #10 for "variety-seeking", #1 and #6 for "comfort of flying", and #2, #7 and #11 for "dissatisfaction for status-quo". Although statement #5 was thought to be related to variety-seeking, its loading was below the cut-off point and therefore was excluded.

Table 5 Attitudinal statements used for factor analysis.

#	Attitudinal statements	Underlying constructs
1	I am comfortable with flying in a small aircraft	Comfort of flying
2	Traffic congestion is a major problem in my area	Dissatisfaction for status-quo
3	I wouldn't mind pooling with other people on eVTOL flights	-
4	Uber is my preferred rideshare service	✗
5	I would use an autonomous vehicle if it is available	-
6	I am comfortable with flying in a battery-powered aircraft	Comfort of flying
7	My current travel options for long-distance trips (50-100 miles) take too long	Dissatisfaction for status-quo
8	**I am one of the first to adopt new technology**	**Variety-seeking**
9	I usually take the cheapest mode of transportation available to me	✗
10	**I'm excited for eVTOL travel to become available in my area**	**Variety-seeking**
11	I wish travel times were more consistent and predictable in my area	Dissatisfaction for status-quo
12	I am concerned about my impact on the environment	✗

One objective of this paper is to examine the role of variety-seeking in mode choices when a novel service enters the market; thereby, we only discuss the statements loaded onto the construct of variety-seeking, which are statements #8 and #10 in Table 5. Their Chronbach's alpha estimate is 0.7, and Guttman's Lambda 6 estimate is 0.54, suggesting relatively good internal consistency between these two statements. Table 6 selectively presents 4 indices

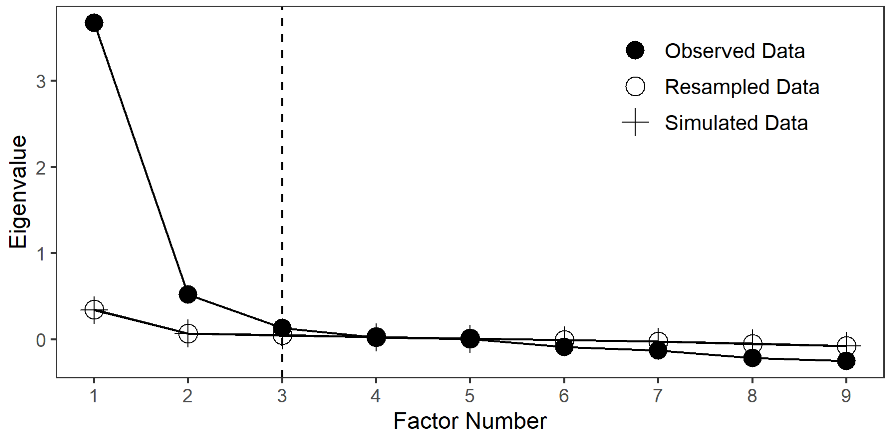

Fig. 3 Parallel analysis scree plots for the factor analysis.

that reflect variety-seeking in the mode choice experiences and stated choice tasks and shows the average value for each index by the score of statements #8 and #10. It can be observed that stronger agreement with these two statements is related to a broader choice of ride-sourcing companies in the past and alternatives in the SC survey, as well as a higher frequency of choosing the new UberAIR option and a lower frequency of selecting the reference mode in the SC survey.

Table 6 Relation between the responses to attitudinal statements and mode choice experience/ stated choices

	Alternation		Novelty-seeking	
Score	Ride-sourcing companies used in real life (mean)	Different alternatives chosen across SC tasks (mean)	Times UberAIR chosen in SC tasks (mean)	Times reference mode chosen in SC tasks (mean)
statement #8				
1	0.6	1.6	0.9	7.5
2	0.8	1.8	1.3	6.1
3	1.0	2.0	1.7	5.0
4	1.3	2.2	2.8	3.8
5	1.5	2.3	3.7	1.9
statement #10				
1	0.6	1.4	0.7	7.3
2	0.7	1.6	0.6	7.2
3	0.9	1.9	1.2	5.6
4	1.1	2.2	2.6	4.3
5	1.5	2.3	3.8	2.2

4 Methodology

4.1 Hypothesis

This section discusses the approach we proposed to accommodate intra-respondent preference heterogeneity on top of inter-respondent preference heterogeneity and explores the role of variety-seeking in mode choice behaviour in the new context of air taxi. All models discussed in this section are established on the random utility maximisation (RUM) assumption.

In the present paper, variety-seeking is regarded as an unobservable personality trait. As mentioned in section 2, variety-seeking can be reflected or driven by novelty-seeking and (or) alternation. Hence, we aim to distinguish and discern both aspects. Two hypotheses are put forward with respect to the novelty-seeking aspect and the alternation aspect of variety-seeking:

Hypothesis 1: Stronger novelty-seeking is linked to a higher propensity to adopt the upcoming air taxi mode, i.e. UberAIR in our case.

Hypothesis 2: Stronger alternation would relate to a higher tendency to exhibit unstable preferences over choice tasks of a SC survey.

As such, part of unobserved preference heterogeneity across respondents (i.e. inter-respondent preference heterogeneity) is explained by the novelty-seeking aspect of variety-seeking tendencies. Meanwhile, the alternation aspect is associated with preference heterogeneity over choices within a given individual (i.e. intra-respondent preference heterogeneity).

We hence explore the role of variety-seeking in a stated choice setting by addressing three key questions:

1) Can variety-seeking reflect itself through the novelty-seeking aspect and whether variety seekers have a higher probability of showing a higher inclination to adopt the new air taxi service?
2) Can variety-seeking reflect itself through the alternation aspect and whether variety seekers have higher tendencies to switch their choices more often over time?
3) If the impact of variety-seeking is detected, what type of respondents are more likely to be variety-seekers?

Enlightened by the discussion by Hess (2014), we propose two new models in this paper. The first new model involves an additional layer to account for intra-respondent preference heterogeneity on top of inter-respondent preference heterogeneity. The other new model further introduces a latent variable of variety-seeking to explain what causes the preference heterogeneity across respondents and within respondents, leading to behavioural benefits. Briefly speaking, we resemble the conventional way of accommodating inter-and-intra heterogeneity within a latent class model framework and further incorporate variety-seeking as a latent variable to explain class allocation probabilities.

In these two new models, respondents can be probabilistically classified into "novelty-seeker" class and "novelty-avoider" class, and each can continue to be segmented into "alternation-seeker" class and "alternation-avoider" class.

This two-step segmentation allows us to capture preference variations across respondents. Meanwhile, the alternation effect is controlled only within the "alternation-seeker" class by implementing probabilistic allocation on discrete distributions over choice tasks, i.e. allowing for intra-respondent preference heterogeneity. In the second new model, variable-seeking is introduced into the model as a latent variable to explain the class segmentation functions. The details about these two models can be found in section 4.3 and section 4.4.

4.2 Basic Latent Class (LC) model

The Multinomial Logit (MNL) model (McFadden, 1973) has been widely used in understanding choice behaviour. It assumes all the preference heterogeneity is captured deterministically, e.g. through interactions between sensitivity parameters with socio-demographic characteristics. However, there exists preference heterogeneity that cannot be explained deterministically. Two typical methods to capture unobserved preference heterogeneity are the Mixed Multinomial Logit (MMNL) model (Boyd and Mellman, 1980; Cardell and Dunbar, 1980) and Latent Class (LC) model (Kamakura and Russell, 1989; Gupta and Chintagunta, 1994). While the former incorporates unobserved preference heterogeneity by using continuous distributions in parameters, the latter uses discrete distributions. Thus, the LC model does not need to make specific assumptions about the distribution of parameters. In a latent class model, preference heterogeneity can be captured by probabilistically assigning membership to each respondent (Walker and Ben-Akiva, 2002). [7]

A basic LC model is developed with an underlying MNL model. Essentially, this basic LC model resembles the MMNL model with the assumption of inter-respondent preference heterogeneity. It assumes that there are a finite number of classes S with different values for the parameters (including ASC vector δ_s and sensitivities vector β_s) in each class. Given class membership s, decision maker n derives an unobserved utility $U_{int,s}$ from alternative i in choice task t. This utility $U_{int,s}$ consists of a deterministic portion $V_{int,s}$ and unobserved and random disturbance $\varepsilon_{int,s}$. Thus, the utility function is written as:

$$U_{int,s} = V_{int,s} + \varepsilon_{int,s} = \delta_{i,s} + \beta_s' x_{int} + \varepsilon_{int,s}, \tag{1}$$

where $V_{int,s}$ typically follows a linear-in-parameter specification with an alternative-specific constant (ASC) $\delta_{i,s}$. x_{int} is a vector of explanation variables for alternative i which is presented to respondent n in task t. A vector of to-be-estimated parameters β_s explains the sensitivities, and is treated as homogeneous across choice tasks. The random error term $\varepsilon_{int,s}$ is independently and identically distributed (IID) type I extreme value distribution.

[7] Comparisons between the latent class model and mixed logit model can be found in some literature (e.g. Greene and Hensher (2003); Shen (2009)). Moreover, latent and mixed logit can be combined to allow for continuous randomness in preference heterogeneity within a class by specifying a random parameter latent class model (Greene and Hensher, 2013).

In our case, we allow for two classes of respondents, i.e. $s \in (1, 2)$ in Eq.1. This was found to give adequate gains in fit without undue increase in complexity. Following common practice, the class allocation model for two classes of respondents is specified in a binary logit form. We start from the basic specification, which assumes the class allocation functions to be constant across respondents. The probability π_s of a given respondent n falling into class s can be computed by:

$$\begin{aligned} \pi_1 &= \frac{e^{\gamma_1}}{e^{\gamma_1} + 1}, \\ \pi_2 &= 1 - \pi_1 \end{aligned} \tag{2}$$

such that $\sum_{s=1}^{S} \pi_s = 1$ and $0 \leq \pi_s \leq 1$, where γ_1 is the class-specific constant in the class allocation functions. The unconditional likelihood of making a sequence of choices by respondent n can be obtained by taking a weighted summation of the conditional likelihood given the class membership across classes, such that:

$$P(y_n) = \sum_{s=1}^{S} \pi_s \left(\prod_{t=1}^{T} P\left(y_{nt} \mid \delta_s, \beta_s\right) \right). \tag{3}$$

The log-likelihood function is given by: $LL(y) = \sum_{n=1}^{N} \ln P(y_n)$.

4.3 New model 1: Two-layer Latent Class (2L-LC) model

Now we elaborate on how the new latent class model with two layers of heterogeneity is constructed to resemble the structure of the two-layer MMNL model. This is achieved by replacing the continuous mixture with a discrete mixture at both inter-respondent and intra-respondent layers, which can substantially reduce the computational burden. The alternation effect is controlled at the intra-respondent layer to manifest preference variation across choice tasks. Fig. 4 illustrates how the sample is probabilistically classified at the inter-respondent layer and how the alternation effect is controlled at the intra-respondent layer. The model with latent variety-seeking is discussed in the section 4.4 but still follows this structure.

4.3.1 inter-respondent layer

At the inter-respondent layer, respondents are first of all probabilistically segmented into S classes, each class carrying different preference parameters. This segmentation is the same as the basic LC model in section 4.2. That is, a given respondent has a probability of π_s to belong to class s with ASC δ_s and sensitivities β_s which are specific to class s. In our case, $S = 2$ as we expect to discern one class of "novelty-avoiders" and one class of "novelty-seekers".

We continue to segment class s into $Q = 2$ subclasses based on the assumption that while some respondents have consistent preference across choice tasks

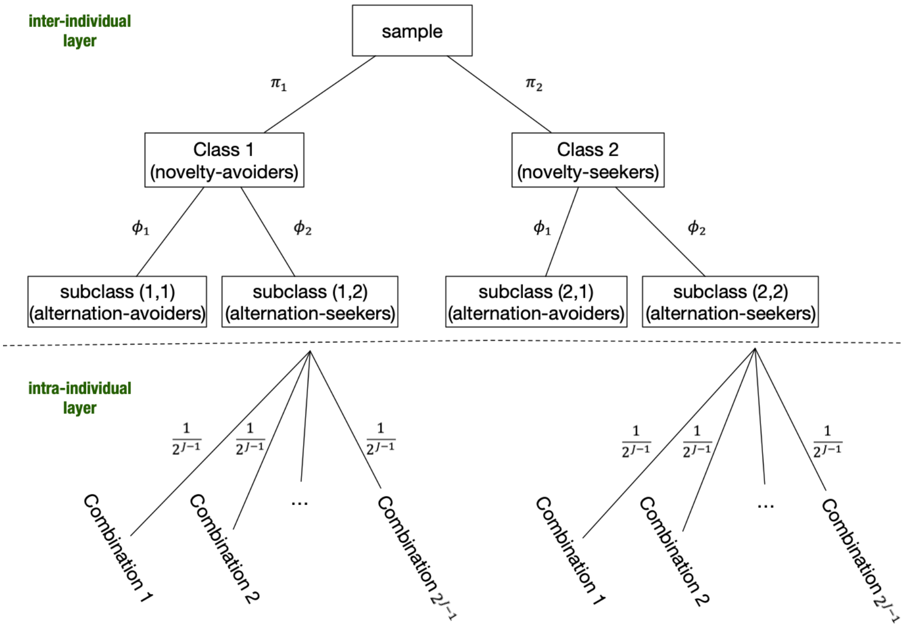

Fig. 4 Structure of the 2L-LC model.

(i.e. alternation-avoiders), others experience preference variation in the course of completing choice tasks (i.e. alternation-seekers). That is, for each class s, it is further segmented into a “alternation-avoiders” subclass with a probability of ϕ_1, and a “alternation-seekers” subclass with a probability of ϕ_2. Herein, we use (s, q) to denote the class membership, with $q = 1$ standing for a “alternation-avoiders” subclass, and $q = 2$ for a “alternation-seekers” subclass. As shown in the upper part of Fig. 4, we eventually obtain four subclasses of respondents, among which $(1, 1)$ and $(2, 1)$ are “alternation-avoiders” subclasses with stable preference to alternatives across tasks, whereas $(1, 2)$ and $(2, 2)$ are “alternation-seekers” subclasses exhibiting heterogeneous preference over tasks.

Therefore, while keeping the class allocation model at the upper part the same as in Eq. 2, we further adopt another binary logit model to determine the class allocation probability at the lower part of the inter-respondent layer such that:

$$\begin{aligned} \phi_1 &= \frac{e^{\lambda_1}}{e^{\lambda_1}+1}, \\ \phi_2 &= 1-\phi_1 \end{aligned} \tag{4}$$

where λ_1 is the constant specific to “alternation-avoiders” subclasses in the class allocation function and is generic in any class s. Herein, λ_1 (and so is ϕ_1) is kept generic in any class s to facilitate the identification of the 2L-LC model (and also the more complex 2L-LV-LC model to be discussed in section 4.4). We acknowledge that this restriction may overlook the differences regarding the alternation probabilities between the novelty-seekers class and novelty-avoiders class. We will leave this for future research to improve the examination of the role of the novelty-seeking aspect and alternation aspect.

As to the “alternation-avoiders” subclasses (i.e. $q = 1$), they are characterised with the baseline preference parameters δ_s and β_s at each choice. Thus, the utility function for alternative i given the class membership (s,1) is written as:

$$U_{int,(s,1)} = \delta_{i,(s,1)} + \beta'_{(s,1)} x_{int} + \varepsilon_{int,(s,1)} = \delta_{i,s} + \beta'_s x_{int} + \varepsilon_{int,(s,1)}, \quad s \in (1,2). \tag{5}$$

Moreover, the conditional likelihood of observing a choice made by individual n at task t is:

$$P\left(y_{nt} \mid \delta_{(s,1)}, \beta_{(s,1)}\right) = P\left(y_{nt} \mid \delta_s, \beta_s\right). \tag{6}$$

As to the “alternation-seekers” subclassess (i.e. $q = 2$), $\delta_{i,(s,2)}$ is not a constant value at the task level. We discuss how intra-respondent preference heterogeneity is accommodated for these subclasses in section 4.3.2.

4.3.2 intra-respondent layer

As stated earlier, we associate the alternation effect with the tendency to exhibit intra-respondent preference heterogeneity. Intra-respondent preference heterogeneity is only accommodated for the ‘alternation-seekers” subclasses (i.e. $q = 2$). Contrary to this, preferences are kept stable across choice tasks if allocated to a “alternation-avoiders” subclass.

Specifically, intra-respondent preference heterogeneity in “alternation-seekers” subclasses (i.e. $q = 2$) is implemented by letting the ASC parameters $\delta_{(s,2)}$ shift around the baseline values by $\mathit{\Delta}$ at the observation level, such that the intrinsic preferences towards each alternative vary across choice tasks. However, the marginal utilities $\beta_{(s,2)}$ are fixed to the baseline values of β_s over tasks, i.e. no intra-respondent heterogeneity in the marginal utility parameters.[8]

We replace the continuous distributions across choices used in the MMNL model with discrete mixtures at the intra-respondent layer. More precisely, we assume that each $\delta_{i,s}$ has an equal probability to either have an alternative-specific shift term $\mathit{\Delta}_i$ added or deducted, where $\mathit{\Delta}_i$ is kept generic in any class s. Thus, we specify:

$$\delta_{i,(s,2)} = \delta_{i,(s,2),m_i} = \delta_{i,s} + \mathit{\Delta}_i (m_i == 1) - \mathit{\Delta}_i (m_i == 2), \tag{7}$$

where m_i is an alternative-specific indicator showing whether the shift term is added or deducted.

This specification allows us to achieve an analogue of the MMNL model with inter- and intra-respondent preference heterogeneity. For a given random parameter in the MMNL model, an additional continuous distribution is

[8] This specification is more in line with the definition of alternation, as alternation is more closely related to the instability of choices rather than the instability of sensitivities towards specific attributes. Hence, we allow variations in ASCs instead of the marginal utilities.

specified over choice tasks on top of the continuous distribution over decision-makers. The mean is captured by the distribution at the inter-respondent layer, while the variance is estimated for the distribution at the intra-respondent layer. In our case, given subclass membership $(s, 2)$, Eq. (7) enables preference variation at the choice level while keeping the mean of ASC for alternative i the same as in the corresponding "alternation-avoiders" subclass $(s, 1)$, which equates to $\delta_{i,s}$.

Given J alternatives in a choice set, alternative J is used as the base for normalisation with the corresponding ASC $\delta_{J,s}$ fixed to 0. Thus, we only account for intra-respondent variation for the remaining $J-1$ non-zero ASCs. In particular, we take into account all the possible combinations for the vector $\left(\delta_{1,(s,2),m_1}, \delta_{2,(s,2),m_2}, \cdots, \delta_{J-1,(s,2),m_{J-1}}\right)$, such that all the combinations amount to 2^{J-1} in total for a given individual at a given choice task. The lower part of Fig. 4 presents the treatment at the intra-respondent layer, where the discrete mixture is taken over 2^{J-1} combinations.

Then we average the probability over the 2^{J-1} possible situations and use it as the conditional choice probability for respondent n at task t given the membership of a "alternation-seekers" subclass i.e. $q = 2$, such that:

$$
\begin{aligned}
&P(y_{nt} \mid (\delta_{(s,2)}, \beta_{(s,2)})) \\
=&\frac{1}{2^{J-1}} \sum_{m_1=1}^{2} \sum_{m_2=1}^{2} \cdots \sum_{m_{J-1}=1}^{2} P\left(y_{nt} \mid \left(\delta_{1,(s,2),m_1}, \delta_{2,(s,2),m_2}, \cdots, \delta_{J-1,(s,2),m_{J-1}}\right), \beta_s\right),
\end{aligned} \tag{8}
$$

Combined with Eqs. 6 - 8, we can get the unconditional likelihood of observing a sequence of choices for a given respondent n by replacing Eq. 3 with:

$$
P(y_n) = \sum_{s=1}^{S} \pi_s \sum_{q=1}^{Q} \phi_q \left(\prod_{t=1}^{T} \left(P\left(y_{nt} \mid \delta_{(s,q)}, \beta_{(s,q)}\right)\right)\right). \tag{9}
$$

4.4 New model 2: Two-layer Latent Variable Latent Class (2L-LV-LC) model

Now we delve deeper into the drivers of inter- and intra-respondent preference heterogeneity, i.e. variety-seeking. We treat variety-seeking as a latent variable to reduce the risk of endogeneity and measurement errors. It is incorporated in both class allocation functions at the inter-respondent layer, with two different parameters τ_{NS} and τ_{AT} capturing the novelty-seeking effect and alternation effect, respectively. By doing so, people can be probabilistically segmented into different classes as functions of the latent construct (Hess et al., 2013; Motoaki and Daziano, 2015). Due to the concern that the two aspects of variety-seeking are related and intertwined, we do not explicitly specify two separate latent variables. Fig. 5 illustrates the modelling framework of the 2L-LV-LC model, showing how the latent variable of variety-seeking is introduced into the 2L-LC model. Apart from having the latent variety-seeking in explaining class membership probabilities and the responses to selected indicators, the two-layer structure is maintained to be the same as in the 2L-LC model (see Fig. 4). This section hence only explains the differences against the 2L-LC model.

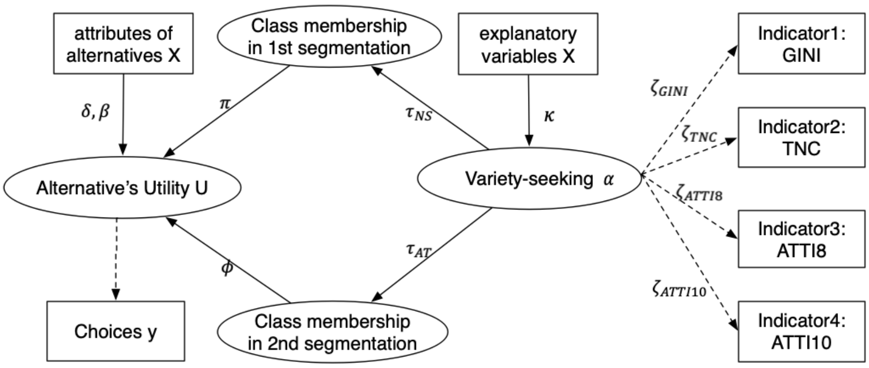

Fig. 5 Modelling framework of the 2L-LV-LC model.

4.4.1 Structural equations for latent variable

We define a latent variable α_n to describe the underlying construct of variety-seeking in the structural equation. It is explained by selected socio-demographic characteristics in the structural equations as:

$$\alpha_n = \kappa' Z_n + \eta_n, \tag{10}$$

where η_n follows a standard Normal distribution across respondents. Z_n denotes the vector of selected covariates, with the vector κ measuring its impact on the latent variable for respondent n.

4.4.2 Latent variables in class allocation functions

To account for the impact of latent variety-seeking in the two-layer latent class model, we rewrite the class allocation probabilities specified in Eq. 2 and in Eq. 4 as:

$$\begin{aligned} \pi_{n,1} &= \frac{e^{\gamma_1 + \tau_{\mathrm{NS}} \alpha_n}}{e^{\gamma_1 + \tau_{\mathrm{NS}} \alpha_n} + 1}, \\ \pi_{n,2} &= 1 - \pi_{n,1} \end{aligned} \tag{11}$$

and

$$\begin{aligned} \phi_{n,1} &= \frac{e^{\lambda_1 + \tau_{\mathrm{AT}} \alpha_n}}{e^{\lambda_1 + \tau_{\mathrm{AT}} \alpha_n} + 1}, \\ \phi_{n,2} &= 1 - \phi_{n,1} \end{aligned} \tag{12}$$

such that the class allocation probabilities $\pi_{n,s}$ and $\phi_{n,q}$ vary across respondents. Parameters τ_{NS} and τ_{AT} measure whether and to what extent the novelty-seeking and alternation aspects influence class membership probabilities, respectively. Providing that a higher value of the latent variable α_n is associated with a stronger variety-seeking tendency, we would expect to see significant negative τ_{NS} and τ_{AT}. This implies that variety-seekers have higher probabilities of falling into the class with a stronger inclination to seek novelty (i.e. $s = 2$), and variety-seekers are more likely to belong to the class with preference heterogeneity over tasks (i.e. $q = 2$). Of course, the same result also applies if both taus are positive, given that a higher latent variable is associated with a lower variety-seeking tendency.

Consequently, the conditional likelihood for the choice model component given the value of latent variety-seeking for respondent n can be written as:

$$P(y_n \mid \alpha_n) = \sum_{s=1}^{S} (\pi_{n,s} \mid \alpha_n) \sum_{q=1}^{Q} (\phi_{n,q} \mid \alpha_n) \left(\prod_{t=1}^{T} \left(P\left(y_{nt} \mid \delta_{(s,q)}, \beta_{(s,q)}\right)\right) \right), \tag{13}$$

where $P\left(y_{nt} \mid \delta_{(s,1)}, \beta_{(s,1)}\right)$ and $P\left(y_{nt} \mid \delta_{(s,2)}, \beta_{(s,2)}\right)$ follow the specifications in Eq. 6 and Eq. 8, respectively.

4.4.3 Latent variables in measurement equations

In the meantime, the latent variable of variety-seeking is used in the measurement model components to explain four selected observable indicators.

Drawing on the concept of the Gini coefficient, we first calculate an inequality index $I_{n,\mathrm{GINI}}$ as a measure of variety in mode choice in real-world travel experience by:

$$I_{n,\mathrm{GINI}} = \left(\sum_{k=1}^{K} \sum_{r=1}^{K} |g_{nk} - g_{nr}| \right) \Big/ \left(2 \sum_{k=1}^{K} \sum_{r=1}^{K} g_{nr} \right) \tag{14}$$

where g_{nk} stands for a "score of exposure" towards mode k for respondent n which takes a value of 2, 1, and 0 for the response of *"used mode k within the last month"*, *"used mode k over one month ago"* and *"never used before"* respectively. $K = 8$ as this exposure information is available for 8 modes, encompassing personal/household vehicle, rental vehicle, bus, light rail/metro/subway, commuter rail, taxicab, ride-sourcing service, and car-sharing service. Similar to the interpretation of the classical Gini coefficient, a higher value of the indicator $I_{n,\mathrm{GINI}}$ is linked with greater inequality in exposure among different modes, meaning that the respondent has less diversity in mode choices and presumably only relies on a small set of modes.

$I_{n,\mathrm{GINI}}$ is treated as a continuous dependent variable in a simple linear regression function (Ben-Akiva et al., 2002). Specifically, we centre it on 0 and then use a Normal density so that the mean of the Normal distribution does not need to be estimated (Hess and Stathopoulos, 2013), such that:

$$I_{n,\mathrm{GINI}} - \overline{I_{\mathrm{GINI}}} = \zeta_{\mathrm{GINI}}\alpha_n + \sigma_{I_{\mathrm{GINI}}}\xi_{I_{\mathrm{GINI}}}, \tag{15}$$

with $\overline{I_{\mathrm{GINI}}}$ being the mean of $I_{n,\mathrm{GINI}}$ across respondents. Parameter ζ_{GINI} measures the role of latent variety-seeking in explaining the responses towards the "Gini" indicator. The variance is estimated by $\sigma_{I_{\mathrm{GINI}}}$, with $\xi_{I_{\mathrm{GINI}}}$ distributed a standard Normal. Thus, the likelihood of observing $I_{n,\mathrm{GINI}}$ is given by:

$$P(I_{n,\mathrm{GINI}} \mid \alpha_n) = \frac{1}{\sigma_{I_{\mathrm{GINI}}}\sqrt{2\pi}}\left(e^{-\frac{\left(I_{n,\mathrm{GINI}}-\overline{I_{\mathrm{GINI}}}-\zeta_{\mathrm{GINI}}\alpha_n\right)^2}{2\sigma^2_{I_{\mathrm{GINI}}}}}\right). \tag{16}$$

We also count the number of ride-sourcing companies (i.e. TNC, including Uber/Lyft/Others) used in the past as another indicator, which is denoted as $I_{n,\mathrm{TNC}}$ and can take any integer from 0 to 3. It suggests "no experience with ride-sourcing services", "one company", "two companies" and "more than two companies" if $I_{n,\mathrm{TNC}}$ takes a value of 0, 1, 2 and 3, respectively.[9] The remaining two indicators are the responses to the two attitudinal statements described in section 3.5. As shown in Table 6, higher agreement toward these two statements is associated with a wider choice of alternatives in the SC survey and a higher frequency of choosing the new UberAIR alternative. We denote these two indicators as $I_{n,\mathrm{ATTI8}}$ and $I_{n,\mathrm{ATTI10}}$, accordingly.

We deal with $I_{n,\mathrm{TNC}}$, $I_{n,\mathrm{ATTI8}}$ and $I_{n,\mathrm{ATTI10}}$ in a different way by accounting for the ordered characteristics of them, as omitting this nature would result in less behavioural explanation power (Daly et al., 2012; Dekker et al., 2016). Following Daly et al. (2012), we specify an ordered logit model for each ordinal indicator. We denote L_c as the number of levels that indicator c can take, and use ζ_c to measure the impact of latent variety-seeking α_n on the value of $I_{n,c}$. Thus, the probability of observing indicator $I_{n,c}$ taking the value of level l ($l \in (1, \cdots, L_c)$) for respondent n is written as:

$$P(I_{n,c} = l \mid \alpha_n) = \frac{e^{\mu_{c,l}-\zeta_c\alpha_n}}{1+e^{\mu_{c,l}-\zeta_c\alpha_n}} - \frac{e^{\mu_{c,l-1}-\zeta_c\alpha_n}}{1+e^{\mu_{c,l-1}-\zeta_c\alpha_n}}, \tag{17}$$

where $\mu_{c,l}$ is the threshold parameter for indicator c and level l. For normalisation purpose, we set $\mu_{c,0} = -\infty$ and $\mu_{c,L_c} = +\infty$, and each indicator only needs $L_c - 1$ thresholds to be estimated. As such, the likelihood of observing the responses towards the four indicators by respondent n given the value of α_n is written as:

$$P(I_n \mid \alpha_n) = P(I_{n,\mathrm{GINI}} \mid \alpha_n)P(I_{n,\mathrm{TNC}} \mid \alpha_n)P(I_{n,\mathrm{ATTI8}} \mid \alpha_n)P(I_{n,\mathrm{ATTI10}} \mid \alpha_n) \tag{18}$$

[9] This indicator is created according to the 15 binary responses towards 15 different types of ride-sourcing services provided by Uber, Lyft and other companies, including both basic economic services and expensive premium services. If a respondent has not used any of the 15 types or claimed to "I don't know" about these ride-sourcing services, we assume they have no experience with ride-sourcing services.

4.4.4 Log-likelihood function

Combining Eq. 13 and Eq. 18, the log-likelihood function of observing all the stated choices and the indicators across all the respondents can be obtained by taking the integral over all possible values of the random latent variable of α_n, such that:

$$
\begin{aligned}
&LL(y, I) \\
&= \sum_{n=1}^{N} \ln \int_{\alpha_n} \left(\sum_{s=1}^{S} (\pi_{n,s} \mid \alpha_n) \sum_{q=1}^{Q} (\phi_{n,q} \mid \alpha_n) \prod_{t=1}^{T} \left(P\left(y_{nt} \mid \delta_{(s,q)}, \beta_{(s,q)}\right)\right) \right) P\left(I_n \mid \alpha_n\right) \\
&\quad f(\pi_n, \phi_n \mid \alpha_n) \mathrm{d}\alpha_n.
\end{aligned}
\tag{19}
$$

Since no closed-form expression can be obtained for the resulting LL function due to the integral over the random latent variable, we use simulated log-likelihood to approximate the true LL.

5 Estimation and results

Maximum simulated likelihood estimation (MLE) was adopted for each model. All the models in this paper were estimated in R using the package Apollo (Hess and Palma, 2019). The estimation results are summarised in Table 7. Moving from left to right, the specification complexity increases and each new model uses the estimates of the previous model as starting values in estimation.

In each model, UberX was chosen as the base alternative with the corresponding ASC parameters (including $\delta_{\text{uberx},1}$, $\delta_{\text{uberx},2}$, and $\varDelta_{\text{uberx}}$) fixed to 0 and not shown in Table 7. This is due to that UberX was shown to each respondent in each choice task, and that UberX has the lowest variance in the unidentified MMNL model that estimates the variance of all the alternatives (Walker et al., 2007). Before discussing the estimation results in detail, it needs to be noted that as part of the confidentiality agreement, the estimates from which the market shares could be inferred are not shown in Table 7 (i.e. ASCs). Consequently, this section does not discuss the differences in individual preferences across alternatives. Instead, $\delta_{i,1}$ for the first class in each model are hidden and marked with "$\star$". Meanwhile, we show how much the ASCs shift in the second class against the first class for the same alternative. The t-ratio statistics indicating the significance of the difference in ASCs between classes are also presented. Nevertheless, a positive/negative difference in ASC for the same alternative does not necessarily imply a higher/lower market share for that alternative in Class 2 than Class 1.

We further conducted post-estimation analysis for each model to better illustrate the differences across models and (sub)classes within each latent class model. The results are presented in Table 8. To state more precisely:

- Firstly, we calculated the value of travel time (VTT, \$/min) for each time component. The VTT estimates were computed both over the sample and within each class. As to model 2 and model 3, only ASCs vary at the task level, whereas all the sensitivity parameters are kept constant across choice tasks given class membership. Thus, VTT results are the same for an "alternation-seekers" subclass and an "alternation-avoider" subclass if they are grouped under the same class s at the inter-respondent layer. It needs to be noted that as a non-linear specification of travel cost is adopted in each model, VTT depends on the travel cost. Herein, we used the price of the chosen alternative in calculating VTT estimates.

- Secondly, we computed the market share for each alternative by averaging the choice probabilities for each alternative across all the tasks using the model estimates. These market shares were calculated within each class for the basic latent class model (i.e. model 1). Regarding model 2 and model 3, we can obtain four different sets of within-class choice probabilities, each for one subclass. Additionally, for the "alternation-seekers" subclass, the choice probability for each alternative at a given choice task is obtained by averaging across all the $2^{J-1} = 16$ combinations.
  Again, we cannot present detailed market shares across alternatives due to confidentiality restrictions. Instead, we illustrate the order of market shares for the same alternative across (sub)classes. Specifically, we hide the market shares for the first (sub)class in each latent class model (i.e. Class 1 in model 1, and subclass (1,1) in model 2 and model 3), marked with "$\star$". Moreover, we indicate how the market share in each of the remaining (sub)classes changes relative to the first (sub)class for a given alternative. The minus symbol "$-$" and the plus symbol "$+$" suggest that the market share in the corresponding (sub)class is lower and higher than that in the starred first (sub)class, respectively. When there are more than two classes, and using the example where the value is highest in the first class, a single dash "$-$" indicates the second highest value for that ASC, a double-dash "$--$" the third highest, etc.

5.1 Model 1: Basic LC model

Model 1 is a basic latent class model, where preference heterogeneity is accommodated solely across respondents.

5.1.1 Sample-level results

Egress time has the highest VTT over the sample in model 1 (and is relatively consistent in all models), indicating that the convenience of moving from landing pads to final destinations plays a crucial role in determining the attractiveness of UberAIR. This implies the significance of integrating and coordinating the existing ground-based services with UberAIR.

Table 7 Estimation results of choice model and class allocation models

	model 1: basic LC			model 2: 2L-LC			model 3: 2L-LV-LC		
parameter#	19			24			47		
LL	-16929.74			-15625.74			whole: -24443.96, SC component: -15613.48		
ρ^2	0.4611			0.5026			SC component: 0.5030		
BIC	34051.26			31493.73			whole model: 49362.33		
		est.	**rob. t-rat.**		**est.**	**rob. t-rat.**		**est.**	**rob. t-rat.**
	$\beta_{\text{access},1}$	-0.099	-7.10	$\beta_{\text{access},1}$	-0.140	-4.92	$\beta_{\text{access},1}$	-0.137	-4.88
	$\beta_{\text{egress},1}$	-0.122	-7.93	$\beta_{\text{egress},1}$	-0.170	-6.10	$\beta_{\text{egress},1}$	-0.169	-6.12
	$\beta_{\text{flight},1}$	-0.078	-8.90	$\beta_{\text{flight},1}$	-0.117	-6.80	$\beta_{\text{flight},1}$	-0.115	-6.81
	$\beta_{\text{travel},1}$	-0.040	-11.38	$\beta_{\text{travel},1}$	-0.058	-7.22	$\beta_{\text{travel},1}$	-0.057	-7.27
	$\beta_{\text{cost},1}$	-3.530	-11.71	$\beta_{\text{cost},1}$	-6.670	-15.05	$\beta_{\text{cost},1}$	-6.654	-14.59
	$\delta_{\text{car},1}$	⋆	⋆	$\delta_{\text{car},1}$	⋆	⋆	$\delta_{\text{car},1}$	⋆	⋆
	$\delta_{\text{transit},1}$	⋆	⋆	$\delta_{\text{transit},1}$	⋆	⋆	$\delta_{\text{transit},1}$	⋆	⋆
	$\delta_{\text{uberpool},1}$	⋆	⋆	$\delta_{\text{uberpool},1}$	⋆	⋆	$\delta_{\text{uberpool},1}$	⋆	⋆
	$\delta_{\text{uberair},1}$	⋆	⋆	$\delta_{\text{uberair},1}$	⋆	⋆	$\delta_{\text{uberair},1}$	⋆	⋆
	$\beta_{\text{access},2}$	-0.018	-2.56	$\beta_{\text{access},2}$	-0.061	-5.08	$\beta_{\text{access},2}$	-0.062	-5.04
	$\beta_{\text{egress},2}$	-0.044	-5.57	$\beta_{\text{egress},2}$	-0.088	-5.24	$\beta_{\text{egress},2}$	-0.091	-5.10
	$\beta_{\text{flight},2}$	-0.021	-4.87	$\beta_{\text{flight},2}$	-0.044	-5.29	$\beta_{\text{flight},2}$	-0.046	-5.19
	$\beta_{\text{travel},2}$	-0.021	-8.37	$\beta_{\text{travel},2}$	-0.044	-10.62	$\beta_{\text{travel},2}$	-0.045	-10.34
	$\beta_{\text{cost},2}$	-1.740	-15.93	$\beta_{\text{cost},2}$	-3.156	-16.75	$\beta_{\text{cost},2}$	-3.185	-16.55
	$\delta_{\text{car},2} - \delta_{\text{car},1}$	2.499	2.40	$\delta_{\text{car},2} - \delta_{\text{car},1}$	4.072	3.11	$\delta_{\text{car},2} - \delta_{\text{car},1}$	4.030	3.04
	$\delta_{\text{transit},2} - \delta_{\text{transit},1}$	-7.148	-5.96	$\delta_{\text{transit},2} - \delta_{\text{transit},1}$	-14.826	-6.22	$\delta_{\text{transit},2} - \delta_{\text{transit},1}$	-14.597	-6.16
	$\delta_{\text{uberpool},2} - \delta_{\text{uberpool},1}$	3.348	19.07	$\delta_{\text{uberpool},2} - \delta_{\text{uberpool},1}$	4.768	16.20	$\delta_{\text{uberpool},2} - \delta_{\text{uberpool},1}$	4.868	16.40
	$\delta_{\text{uberair},2} - \delta_{\text{uberair},1}$	-1.017	-3.31	$\delta_{\text{uberair},2} - \delta_{\text{uberair},1}$	-3.600	-6.86	$\delta_{\text{uberair},2} - \delta_{\text{uberair},1}$	-3.545	-6.66
	γ_1	0.280	3.78	γ_1	0.452	6.54	γ_1	0.444	5.95
	-			-	-	-	τ_{NS}	-0.523	-9.24
				Δ_{car}	3.315	10.00	Δ_{car}	3.332	9.81
				Δ_{transit}	11.205	9.45	Δ_{transit}	11.244	9.32
				Δ_{uberpool}	-5.008	-10.25	Δ_{uberpool}	-5.037	-10.07
				Δ_{uberair}	8.761	23.35	Δ_{uberair}	8.851	22.97
				λ_1	0.738	11.49	λ_1	0.798	11.61
				-	-	-	τ_{AT}	-0.325	-5.27
							ζ_{ATTI8}	1.616	12.78
							ζ_{ATTI10}	1.555	12.69
							ζ_{GINI}	-0.068	-13.17
							ζ_{TNC}	1.111	12.64
							σ_{GINI}	0.206	75.22
							$\mu_{\text{ATTI8\_1}}$	-3.250	-22.72
							$\mu_{\text{ATTI8\_2}}$	-1.145	-14.42
							$\mu_{\text{ATTI8\_3}}$	0.794	12.10
							$\mu_{\text{ATTI8\_4}}$	3.004	22.58
							$\mu_{\text{ATTI10\_1}}$	-3.500	-23.46
							$\mu_{\text{ATTI10\_2}}$	-2.246	-21.05
							$\mu_{\text{ATTI10\_3}}$	0.121	2.01
							$\mu_{\text{ATTI10\_4}}$	1.991	20.71
							$\mu_{\text{TNC\_noexperience}}$	-0.850	-16.18
							$\mu_{\text{TNC\_one}}$	0.671	11.82
							$\mu_{\text{TNC\_two}}$	5.226	22.10
							κ_{age}	-1.185	-12.87
							κ_{income}	0.213	10.16
							κ_{female}	-0.660	-11.23
							κ_{delay}	0.200	3.79
							κ_{vehicles}	-0.094	-3.36

Table 8 Value-of-time estimates and choice probabilities

		model 1: basic LC			**model 2: 2L-LC**					**model 3: 2L-LV-LC**				
parameter #		19			24					47				
$LL(whole)$		-16929.74			-15625.74					-24443.96				
$LL(SC)$										-15613.48				
ρ^2		0.4611			0.5026					SC component: 0.5030				
BIC		34051.26			31493.73					whole model: 49362.33				
		Class 1 (novelty-avoiders)	**Class 2 (novelty-seekers)**	**All**	**Class 1 (novelty-avoiders)**		**Class 2 (novelty-seekers)**		**All**	**Class 1 (novelty-avoiders)**		**Class 2 (novelty-seekers)**		**All**
Value of time (VTT, $/h)	access time	36.20	13.43	26.40	27.05		24.74		26.15	26.66		25.07		26.03
	egress time	44.77	32.28	39.40	32.93		36.14		34.18	32.89		36.78		34.43
	flight time	28.36	15.32	22.75	22.57		18.11		20.83	22.22		18.53		20.75
	travel time	14.51	15.87	15.09	11.30		18.14		13.96	11.15		18.16		13.94
					alternation-avoiders (1,1)	**alternation-seekers (1,2)**	**alternation-avoiders (2,1)**	**alternation-seekers (2,2)**		**alternation-avoiders (1,1)**	**alternation-seekers (1,2)**	**alternation-avoiders (2,1)**	**alternation-seekers (2,2)**	
Market share changes	car	$\star$	-		$\star$	-	- -	- - -		$\star$	-	- -	- - -	
	transit	$\star$	-		$\star$	-	- - -	- -		$\star$	-	- - -	- -	
	UberX	$\star$	-		$\star$	- - -	-	- -		$\star$	- - -	-	- -	
	UberPool	$\star$	+		$\star$	+	+++	++		$\star$	+	+++	++	
	UberAir	$\star$	+		$\star$	++	+	+++		$\star$	++	+	+++	

5.1.2 *Class-specific results*

As shown in Table 7, the constant γ_1 (est.=0.280, rob.t=3.78) in the class allocation function implies a probability of 56.95% for respondents to fall into Class 1 and a probability of 43.05% to be in Class 2. Comparing the model estimates of the two classes, we can find that Class 2 is associated with significantly lower sensitivities towards all the attributes, including travel cost.

If further looking at the VTT results in Table 8, we can see that Class 2 shows much lower VTT for all the time components, except for travel time which is almost similar between classes. Generally, Class 1 exhibits higher VTT than Class 2 in model 1.

The distinction in preferences towards different alternatives across classes can be manifested by the within-class choice probability of each alternative. For example, as shown in Table 8, Class 2 shows a higher probability of selecting the UberPOOL and UberAIR options than Class 1. In contrast, car, transit and UberX all have lower proportions in Class 2 than Class 1. Since UberPOOL was unavailable in reality in the Dallas area during the data collection period, the UberPOOL alternative can also be seen as a new mode for respondents recruited there. In this sense, we can infer from model 1 that Class 2 respondents are more likely to try new service(s) than Class 1 respondents.

5.2 Model 2: 2L-LC model

Model 2 accounts for intra-respondent preference heterogeneity in addition to inter-respondent preference heterogeneity, resulting in four subclasses in total. The findings concerning the VTT and choice probabilities over the sample in model 2 do not present many differences against model 1. However, model 2 can give more insight into preference patterns and market segmentation (see section 5.2.4).

5.2.1 *Model estimates*

We first look at the sensitivity parameters at the inter layer in Table 7. Similarly to model 1, marginal utilities for most of the attributes in Class 2 are significantly lower than the corresponding parameters in Class 1. The only exception is travel time, of which the difference is insignificant between classes (diff.=-0.014, rob.t=-1.51, by delta method calculation).

Turning to the model estimates at the intra layer, the significant estimates of the shift terms Δ for all the ASCs suggest that the 2L-LC models can successfully detect the variation and instability of preference over choice tasks for a given respondent. For example, compared to the base alternative UberX, people's preferences towards transit and UberAIR are much more unstable across choice tasks, whereas the preference disturbance for car and UberPOOL is relatively milder.

The two class allocation models are both solely explained by a constant. Parameter γ_1 (est.=0.452, rob.t=6.54) results in a generic probability of 61.11% to fall into Class 1 (i.e. novelty-avoiders) and a generic probability of 38.89% to fall into Class 2 (i.e. novelty-seekers). Parameter λ_1 (est.=0.738, rob.t=11.49) leads to a generic probability of 67.66% in belonging to a "alternation-avoiders" subclass and 32.34% in being assigned to a "alternation-seekers" subclass.

5.2.2 Value-of-time results

Regarding the VTT patterns shown in Table 8, Class 1 presents a higher value of access time and flight time but a lower value for egress time from landing pads and time spent in vehicles on land, compared to Class 2. It appears that we cannot, like in model 1, detect distinctive VTT patterns between classes in model 2 (and also in model 3), which accounts for the instability of preferences towards alternatives across choice tasks.

5.2.3 Within-class choice probabilities

Nevertheless, the within-class choice probabilities for different alternatives can provide sufficient indications with respect to the characteristics of each class. Similar to the results of model 1, we can see that Class 2 respondents (including both subclass $(2,1)$ and subclass $(2,2)$) present higher probabilities of adopting the new UberAIR alternative as well as the UberPOOL alternative. Meanwhile, Class 1 respondents (including both subclass $(1,1)$ and subclass $(1,2)$) are much more prone to stick to the other existing ground-based modes, particularly personal/household vehicle and transit. These results imply that Class 2 respondents are more likely to try the new service(s) than Class 1 respondents.

Furthermore, to illustrate the differences between "alternation-avoiders" and "alternation-seekers" subclasses under a same set of sensitivities, we calculate the mean of chosen probability for each subclass which is averaged over all the observations. It is found that the "alternation-avoiders" subclasses $(1,1)$ and $(2,1)$ have higher average chosen probabilities (i.e. 66.04% and 55.88%) than "alternation-seekers" subclasses $(1,2)$ and $(2,2)$ (i.e. 45.85% and 30.30%), respectively. This suggests that respondents who fall into the "alternation-seekers" class are associated with less deterministic choices, which is in accordance with our expectation.

5.2.4 Classes' profiles

Combining the discussions above, we can obtain the profiles as well as the allocation probabilities for all the four different subclasses of respondents as:

- Subclass $(1,1)$: 41.35%
    - Low tendency to try new modes including UberAIR (i.e. avoid novelty)
    - Stable preference across choice tasks (i.e. avoid alternation)

- Subclass $(1, 2)$: 19.77%
  - Low tendency to try new modes including UberAIR (i.e. avoid novelty)
  - Unstable preference across choice tasks (i.e. seek alternation)
- Subclass $(2, 1)$: 26.31%
  - High tendency to try new modes including UberAIR (i.e. seek novelty)
  - Stable preference across choice tasks (i.e. avoid alternation)
- Subclass $(2, 2)$: 12.58%
  - High tendency to try new modes including UberAIR (i.e. seek novelty)
  - Unstable preference across choice tasks (i.e. seek alternation)

5.3 Model 3: 2L-LV-LC model

As a final step, we report the results of model 3, which uses latent variety-seeking as an additional explanatory variable in explaining class allocation probabilities across the respondents. Overall, model 3 presents similar patterns to model 2, in terms of model estimates, VTT results and within-class choice probabilities. Herein, we only discuss the unique characteristics of model 3, i.e. the impact of latent variety-seeking.

5.3.1 Variety-seeking in class allocation models

As shown in Table 7, the constants γ_1 and λ_1 at the inter-respondent layer are very close to those in model 2. The negative and significant τ_{NS} (est.=-0.523, rob.t=-9.24) means that a higher value of the latent variable α would result in greater propensity to fall into Class 2, which features stronger willingness to choose the new UberAIR service. Similarly, the negative and significant τ_{AT} (est.=-0.325, rob.t=-5.27) implies a decrease in probability of belonging to "alternation-seekers" subclasses $(1, 1)$ and $(2, 1)$ with an increase in the latent variable α. Thus, the probabilities of falling in a given subclass vary across respondents in model 3, depending on the value of α.

5.3.2 Variety-seeking in measurement model component

Now we jointly examine the role of the latent variable α in the class allocation functions and the measurement equations. The threshold parameter $\mu_{c,l}$ presents a monotonically increasing trend as the level l goes up for each ordinal indicator c. From the positive and significant parameters ζ_{ATTI8}, ζ_{ATTI10} and ζ_{TNC}, we can see that an increase in the latent variable α would lead to a stronger agreement towards the attitudinal statements ATTI8 and ATTI10, as well as a larger number of ride-sourcing companies experienced in the past. In terms of the "Gini" coefficient, the negative and significant ζ_{GINI} implies that a stronger α is associated with a lower Gini coefficient, suggesting less inequality and less uniqueness in mode choice experience. These results infer that the latent variable α can indeed be interpreted as "variety-seeking", such that a larger value in α corresponds to stronger variety-seeking.

Combining the interpretation of the latent variable α and the class allocation models, we can confirm our hypothesis. The results suggest that variety-seeking plays a role in both inter-respondent and intra-respondent preference heterogeneity. Specifically, compared to people with lower variety-seeking tendencies, people perceiving higher variety-seeking tendencies are more likely to fall into the class with higher probabilities of switching to the novel UberAIR and UberPOOL options and lower probabilities of choosing the long-existing car and transit alternatives (i.e. falling into novelty-seekers class). This is in line with an earlier study of variety-seeking in the context of intermodality between air and high-speed rail, where variety seekers were found more likely to select the new intermodal service (Song et al., 2018). It also aligns with another study in the context of ride-sourcing services, where variety-seekers were found more inclined to use ride-sourcing services (Alemi et al., 2018). Additionally, we discovered that people with higher variety-seeking tendencies also have higher propensities to belong to the "alternation-seekers" subclasses, where preferences for alternatives are unstable and less deterministic across choice tasks. This implies that in the course of completing a SC survey, people with stronger variety-seeking are more likely to switch their mode choices among different alternatives continuously.

Consequently, the classification of respondents and profiles of different subclasses discussed in section 5.2.4 can be retrieved by model 3. Notably, due to the significant role of latent variety-seeking, the probability of falling into each of the four subclasses varies across respondents rather than being generic.

5.3.3 Structural equation for variety-seeking

After regressing the responses towards attitudinal statements related to variety-seeking on different socio-demographic and trip characteristics, we adopt *age*, *income*, *the number of owned vehicles*, *gender* and *whether experienced delay* as explanatory variables in the final specification for Eq. 10. All these covariates are centred on 0, so the latent variable has a mean of 0. Age, income and the number of owned vehicles are treated as continuous variables, while the remaining two variables are treated as binary ones. To avoid incomparable scales between different covariates, we divide the age and income variables by the original mean values.

Parameters κ in Table 7 show how these explanatory variables affect the value of latent variety-seeking. As expected, the negative κ_{age}, κ_{female} and κ_{vehicles} show that older people, female respondents and people with more vehicles are characterised by weaker variety-seeking tendency. Meanwhile, the positive κ_{income} and κ_{delay} suggest that people with more income and who have experienced delays on the same trip in the past have a stronger variety-seeking tendency.

5.4 Comparisons of model fit

Moving from model 1 to model 2, we can see that model fit improves as the model specification becomes more complex, in terms of the log-likelihood, ρ^2 values and the Bayesian Information Criterion (BIC). This improvement over models can also be confirmed by the likelihood ratio test, of which the p-value is 0 when comparing model 2 against model 1. All these reflect the significant benefits obtained from better accommodation of preference heterogeneity, both across respondents and within respondents.

It is reasonable to see that both log-likelihood and BIC for the whole model in model 3 are much worse than in other simpler models, as model 3 simultaneously explains the observations of indicators of latent variety-seeking in the measurement model component. We acknowledge that Vij and Walker (2016) have demonstrated that incorporating latent variables in the choice model cannot result in a better fit than a corresponding reduced form model without latent variables. In the present paper, neither explanatory variables nor random terms are incorporated in the allocation functions in model 2, meaning that model 2 does not have the same flexibility as model 3 does and should not be regarded as the reduced form of model 3. Thus, it is reasonable to achieve a slight improvement in fit for the choice component in model 3.

6 Conclusions

It is crucial to improve the accommodation of unobserved preference heterogeneity in discrete choice modelling analysis. Growing effort in recent years has been devoted to uncovering intra-respondent preference heterogeneity on top of inter-respondent preference heterogeneity in stated choice data. These models usually are based on mixed multinomial logit (MMNL) with an additional layer of randomness that varies across choice tasks to account for intra-respondent preference heterogeneity. This practice is computationally demanding because of the additional layer of randomness, and the behavioural explanations of this inter- and intra-respondent preference heterogeneity still require further exploration. Therefore our paper accommodates intra-respondent preference heterogeneity in a less computationally demanding way and provides additional behavioural insights. The SP data we got from Uber on their upcoming new mobility "UberAir" provides us with a proper context to look into this issue. In the meantime, we take this chance to explore the impact of both aspects of variety-seeking, i.e. novelty-seeking and alternation-seeking, as neither has been sufficiently discussed in existing transport studies.

This paper proposed a two-layer latent class (latent variable) modelling approach to accommodate the unobserved preference heterogeneity both across respondents and across choice tasks. At the inter-respondent layer, respondents were first probabilistically segmented into two classes, one exhibiting a higher propensity to adopt the new UberAIR service than the other. Then, given class membership, respondents were further probabilistically segmented into two subclasses - one with stable preferences towards alternatives and another with preference variations across choice tasks. Intra-respondent preference heterogeneity was only accommodated for the "alternation-seekers" subclasses through an additional layer of discrete mixture, with variations in ASCs across choice tasks. This model essentially replaced continuous distributions used in the MMNL models (Hess and Rose, 2009) with discrete distributions at both layers, which can reduce the computational burden.

We also contributed to the behavioural explanation of unobserved preference heterogeneity across respondents as well as the application of variety-seeking theory. We treated variety-seeking as an underlying personality construct and introduced it into the model as a latent variable. Specifically, each step of segmentation was a function of the latent variable of variety-seeking. On the one hand, we associated the novelty-seeking aspect of variety-seeking with inter-respondent preference heterogeneity, assuming that stronger variety-seeking would lead to a stronger inclination to try the new alternative. On the other hand, we related the alternation aspect of variety-seeking with intra-respondent preference heterogeneity, presuming that stronger variety-seeking would contribute to a higher propensity to exhibit unstable preference towards different alternatives across choice tasks.

This paper additionally contributed to the urban air mobility literature with empirical evidence on mode choice behaviour when the new air taxi service enters the market. We believe this work is relevant to the context of air taxi and can be applied in situations where we need to understand the adoption and preferences towards other new mobility services when they enter the market. Moreover, the proposed new approaches can be extended to a non-transport setting to account for consumers' uptake of new products at the initial stage of the diffusion process.

The results confirmed the two hypotheses and answered the three research questions identified in the Introduction in a mode choice experiment involving the upcoming air taxi service. A significant impact of variety-seeking was discerned in each class allocation function, which supports our presumption about the roles that the novelty-seeking and alternation aspects of variety-seeking would play on mode choices. We found that compared to people with lower variety-seeking tendencies, people with stronger variety-seeking tendencies are not only more likely to adopt the new UberAIR service, but also more likely to exhibit unstable preferences towards alternatives across choice tasks than. It is also discovered from the structural equation component that people with higher income and those who had experienced delays on the same trip have stronger variety-seeking tendencies than those with lower income and without delays experience. In the meantime, the estimates in the measurement question component showed that those variety-seekers scored stronger agreement in attitudinal statements describing their interest in adopting new technologies. They were found to be associated with broader exposure to ride-sourcing services and other types of ground-based transport modes in the past.

Policy insights can be derived from these results. Firstly, this work quantified the impact of various factors influencing people's mode choices between the novel air taxi service and other conventional modes of transport. The value-of-time estimates suggested that people would be relatively more sensitive to the time spent accessing or egressing from the take-off-landing pads than to the time spent on the flight or other ground-based vehicles. Hence, enhancing the accessibility to air taxi services is paramount to forging an attractive air taxi product. Secondly, the latent class framework could help policymakers identify which group(s) of people are most likely to become early adopters of a newly introduced or to-be-introduced mode. For example, our results indicated that younger and high-income people are prone to exhibit stronger variety-seeking tendencies and hence show a stronger willingness to adopt the new air taxi mode. Thirdly, the coexistence of inter-respondent and intra-respondent preference heterogeneity unveiled the complex impact of unobserved preference heterogeneity on choice decisions. Recognising that preference homogeneity across choices might not hold within individual respondents would stimulate transport practitioners to maintain a consistently high standard of travel services.

We acknowledge the shortcomings of the proposed two-layer latent class framework. This mainly relates to our estimation method, i.e. maximum simulated likelihood estimation. Thus a model built within this framework might struggle with local optimum issue and the estimation results could be sensitive to the starting values. We have tried to minimise the impact of these issues by using the estimates of a more constrained model as the starting values of a more general model with a more complex specification. Nevertheless, it would be worth testing the model with other alternative estimation methods, e.g. EM algorithms (Train, 2008). We also acknowledge that the implications related to variety-seeking in our paper are obtained from repeated stated choice data rather than longitudinal revealed preference data. Hence novelty-seeking and alternation aspects' impacts might be not significant in real-life situations. However, we cannot test this assumption with our data. We will leave the work of validating the role of variety-seeking in real life to future research, provided suitable longitudinal RP data is available.

Future research potentials include replicating this work in other choice contexts and testing the performance of this new two-layer latent class model with (or without) latent variables in explaining inter- and intra-respondent preference heterogeneity. In addition, a two-layer latent class model can have more than two classes at each level, so it could be tailored to meet the requirement of a specific study. Finally, it is worth exploring whether novelty-seeking is a purely short-term effect or also works in the longer run as a counterpart to habits, e.g. examine the adoption and diffusion of new technology (El Zarwi et al., 2017).

Acknowledgements Fangqing Song acknowledges the support of the China Scholarship Council while Stephane Hess was supported by the European Research Council through the consolidator grant 615596-DECISIONS. The authors are thankful to RSG and Uber for providing the stated choice data.

Conflict of interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Al Haddad C, Chaniotakis E, Straubinger A, Plötner K, Antoniou C (2020) Factors affecting the adoption and use of urban air mobility. Transportation research part A: policy and practice 132:696–712

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

Assaker G, Hallak R (2013) Moderating effects of tourists' novelty-seeking tendencies on destination image, visitor satisfaction, and short-and long-term revisit intentions. Journal of Travel Research 52(5):600–613

Baik H, Trani AA, Hinze N, Swingle H, Ashiabor S, Seshadri A (2008) Forecasting model for air taxi, commercial airline, and automobile demand in the united states. Transportation Research Record 2052(1):9–20

Baumgartner H, Steenkamp JBE (1996) Exploratory consumer buying behavior: Conceptualization and measurement. International journal of Research in marketing 13(2):121–137

Becker F, Danaf M, Song X, Atasoy B, Ben-Akiva M (2018) Bayesian estimator for logit mixtures with inter-and intra-consumer heterogeneity. Transportation Research Part B: Methodological 117:1–17

Ben-Akiva M, Walker J, Bernardino AT, Gopinath DA, Morikawa T, Polydoropoulou A (2002) Integration of choice and latent variable models. Perpetual motion: Travel behaviour research opportunities and application challenges pp 431–470

Ben-Akiva M, McFadden D, Train K, et al. (2019) Foundations of stated preference elicitation: Consumer behavior and choice-based conjoint analysis. Foundations and Trends® in Econometrics 10(1-2):1–144

Binder R, Garrow LA, German B, Mokhtarian P, Daskilewicz M, Douthat TH (2018) If you fly it, will commuters come? a survey to model demand for evtol urban air trips. In: 2018 Aviation Technology, Integration, and Operations Conference, p 2882

Borgers A, Van Der Heijden R, Timmermans H (1989) A variety seeking model of spatial choice-behaviour. Environment and Planning A 21(8):1037–1048

Boyd JH, Mellman RE (1980) The effect of fuel economy standards on the us automotive market: an hedonic demand analysis. Transportation Research Part A: General 14(5-6):367–378

Cardell NS, Dunbar FC (1980) Measuring the societal impacts of automobile downsizing. Transportation Research Part A: General 14(5-6):423–434

Cherchi E, Cirillo C (2014) Understanding variability, habit and the effect of long period activity plan in modal choices: a day to day, week to week analysis on panel data. Transportation 41(6):1245–1262

Chintagunta PK (1998) Inertia and variety seeking in a model of brand-purchase timing. Marketing Science 17(3):253–270

Daly A, Hess S, Patruni B, Potoglou D, Rohr C (2012) Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. Transportation 39(2):267–297

Dekker T, Hess S, Brouwer R, Hofkes M (2016) Decision uncertainty in multi-attribute stated preference studies. Resource and Energy Economics 43:57–73

EASA (2018) Flying in the EU: Ops is in the air. https://www.easa.europa.eu/sites/default/files/dfu/EASA_GA_LEAFLET_AIR_OPS_2018_EN.pdf, accessed: 2019-04-15

El Zarwi F, Vij A, Walker JL (2017) A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services. Transportation Research Part C: Emerging Technologies 79:207–223

Fu M, Rothfeld R, Antoniou C (2018) Exploring preferences for transportation modes in an urban air mobility environment: Munich case study. Transportation Research Record p 0361198119843858

Garrow LA, German B, Mokhtarian P, Glodek J (2019) A survey to model demand for evtol urban air trips and competition with autonomous ground vehicles. In: AIAA Aviation 2019 Forum, p 2871

Garrow LA, Mokhtarian P, German B, Boddupalli SS (2020) Commuting in the age of the jetsons: A market segmentation analysis of autonomous ground vehicles and air taxis in five large us cities. In: AIAA AVIATION 2020 FORUM, p 3258

Givon M (1984) Variety seeking through brand switching. Marketing Science 3(1):1–22

Goyal R (2018) Urban air mobility (uam) market study

Greene WH, Hensher DA (2003) A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological 37(8):681–698

Greene WH, Hensher DA (2013) Revealing additional dimensions of preference heterogeneity in a latent class mixed multinomial logit model. Applied Economics 45(14):1897–1902

Gupta S, Chintagunta PK (1994) On using demographic variables to determine segment membership in logit mixture models. Journal of Marketing Research 31(1):128–136

Ha J, Jang SS (2013) Variety seeking in restaurant choice and its drivers. International Journal of Hospitality Management 32:155–168

Hess S (2014) 14 latent class structures: taste heterogeneity and beyond. In: Handbook of choice modelling, Edward Elgar Publishing Cheltenham, pp 311–329

Hess S, Giergiczny M (2015) Intra-respondent heterogeneity in a stated choice survey on wetland conservation in belarus: first steps towards creating a link with uncertainty in contingent valuation. Environmental and Resource Economics 60(3):327–347

Hess S, Palma D (2019) Apollo: a flexible, powerful and customisable freeware package for choice model estimation and application

Hess S, Rose JM (2009) Allowing for intra-respondent variations in coefficients estimated on repeated choice data. Transportation Research Part B: Methodological 43(6):708–719

Hess S, Stathopoulos A (2013) Linking response quality to survey engagement: a combined random scale and latent variable approach. Journal of choice modelling 7:1–12

Hess S, Train KE (2011) Recovery of inter-and intra-personal heterogeneity using mixed logit models. Transportation Research Part B: Methodological 45(7):973–990

Hess S, Shires J, Jopson A (2013) Accommodating underlying pro-environmental attitudes in a rail travel context: application of a latent variable latent class specification. Transportation Research Part D: Transport and Environment 25:42–48

Hess S, Spitz G, Bradley M, Coogan M (2018) Analysis of mode choice for intercity travel: Application of a hybrid choice model to two distinct us corridors. Transportation Research Part A: Policy and Practice 116:547–567

Holden J, Goel N (2016) Fast-forwarding to a future of on-demand urban air transportation. San Francisco, CA

Jang SS, Feng R (2007) Temporal destination revisit intention: The effects of novelty seeking and satisfaction. Tourism management 28(2):580–590

Joshi A, DeLaurentis D, Peeta S, Agusdinata DB (2014) Integrated framework and assessment of on-demand air service in multimodal context. Journal of Aircraft 51(2):402–411

Kamakura WA, Russell GJ (1989) A probabilistic choice model for market segmentation and elasticity structure. Journal of marketing research 26(4):379–390

Koopman C, Dourado E (2017) Defining common carriers: Flight sharing, the faa, and the future of aviation

Krueger R, Bansal P, Bierlaire M, Daziano RA, Rashidi TH (2019) Variational bayesian inference for mixed logit models with unobserved inter-and intra-individual heterogeneity. arXiv preprint arXiv:190500419

Lee TH, Crompton J (1992) Measuring novelty seeking in tourism. Annals of tourism research 19(4):732–751

McAlister L, Pessemier E (1982) Variety seeking behavior: An interdisciplinary review. Journal of Consumer research 9(3):311–322

McFadden D (1973) Conditional logit analysis of qualitative choice behavior

Motoaki Y, Daziano RA (2015) A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. Transportation Research Part A: Policy and Practice 75:217–230

Park Y, Ha HK (2006) Analysis of the impact of high-speed railroad service on air transport demand. Transportation Research Part E: Logistics and Transportation Review 42(2):95–104

Pearson PH (1970) Relationships between global and specified measures of novelty seeking. Journal of Consulting and Clinical Psychology 34(2):199

Peeta S, Paz A, DeLaurentis D (2008) Stated preference analysis of a new very light jet based on-demand air service. Transportation Research Part A: Policy and Practice 42(4):629–645

Pessemier E, Handelsman M (1984) Temporal variety in consumer behavior. Journal of Marketing Research 21(4):435–444

Pessemier EA (1985) Varied individual behavior: some theories, measurement methods and models. Multivariate Behavioral Research 20(1):69–94

Pu D, Trani AA, Hinze N (2014) Zip vehicle commuter aircraft demand estimate: A multinomial logit mode choice model. In: 14th AIAA Aviation Technology, Integration, and Operations Conference, p 2411

Rieser-Schüssler N, Axhausen KW (2012) Investigating the influence of environmentalism and variety seeking on mode choice. Transportation Research Record 2322(1):31–41

Román C, Espino R, Martín JC (2007) Competition of high-speed train with air transport: The case of madrid–barcelona. Journal of Air Transport Management 13(5):277–284

Shaheen S, Cohen A, Zohdy I (2016) Shared mobility: current practices and guiding principles. Tech. rep.

Shen J (2009) Latent class model or mixed logit model? a comparison by transport mode choice data. Applied Economics 41(22):2915–2924

Song F, Hess S, Dekker T (2018) Accounting for the impact of variety-seeking: Theory and application to hsr-air intermodality in china. Journal of Air Transport Management 69:99–111

Train KE (2008) Em algorithms for nonparametric estimation of mixing distributions. Journal of Choice Modelling 1(1):40–69

Trijp HCV, Hoyer WD, Inman JJ (1996) Why switch? product category–level explanations for true variety-seeking behavior. Journal of marketing research 33(3):281–292

US Census Bureau (2018) 2013-2017 American Community Survey 5-Year Estimates. `https://factfinder.census.gov/faces/nav/jsf/pages/community_facts.xhtml`, accessed: 2019-04-21

Van Trijp HC, Steenkamp JBE (1992) Consumers' variety seeking tendency with respect to foods: measurement and managerial implications. European Review of Agricultural Economics 19(2):181–195

Vij A, Walker JL (2016) How, when and why integrated choice and latent variable models are latently useful. Transportation Research Part B: Methodological 90:192–217

Walker J, Ben-Akiva M (2002) Generalized random utility model. Mathematical social sciences 43(3):303–343

Walker JL, Ben-Akiva M, Bolduc D (2007) Identification of parameters in normal error component logit-mixture (neclm) models. Journal of Applied Econometrics 22(6):1095–1125

Wills TA, Vaccaro D, McNamara G (1994) Novelty seeking, risk taking, and related constructs as predictors of adolescent substance use: an application of cloninger's theory. Journal of substance abuse 6(1):1–20

Zhu X, Wang F, Chen C, Reed DD (2020) Personalized incentives for promoting sustainable travel behaviors. Transportation Research Part C: Emerging Technologies 113:314–331

Fangqing Song was a PhD student at the Institute for Transport Studies and the Choice Modelling Centre, University of Leeds, UK, when the paper was first submitted. Her research interest is in the application of discrete choice modelling for the analysis of travel behaviour. She is now working as a research fellow in air travel behaviour modelling at University of Leeds.

Stephane Hess is Professor of Choice Modelling in the Institute for Transport Studies and Director of the Choice Modelling Centre at the University of Leeds. He is also Honorary Professor in Choice Modelling in the Institute for Transport and Logistics Studies at the University of Sydney, Honorary Professor of Modelling Behaviour in Africa at the University of Cape Town, and has a director position at RSG, a leading North American consultancy company. His area of work is the analysis of human decision using advanced discrete choice models, and he is active in the fields of transport, health and environmental economics. Hess has made contributions to the state of the art in the specification, estimation and interpretation of such models. He is also the founding editor in chief of the Journal of Choice Modelling and the founder and steering committee chair of the International Choice Modelling Conference.

Thijs Dekker is an Associate Professor in Transport Economics at the Institute for Transport Studies at the University of Leeds. His research in the field of discrete choice modelling is on the cutting edge between econometrics, micro-economics and psychology. One of the strong features of his work is the use of both classical and Bayesian estimation techniques. In both the fields of environmental and transport economics Thijs Dekker has a proven track record of publishing in the field's leading journals such as the Journal of Environmental Economics and Management and the Journal of Transport Geography.