Case Studies on Transport Policy 21 (2025) 101484

Contents lists available at ScienceDirect

Case Studies on Transport Policy

journal homepage: www.elsevier.com/locate/cstp

Comparing socio-demographic and experiential differences in willingness to pool rides: A hierarchical Bayesian approach

Ariane Kehlbacher *, Kerstin Stark, Laura Gebhardt, Jan Weschke

Institute of Transport Research, German Aerospace Center, Germany

ARTICLE INFO

Keywords:
Willingness to pool
Ridesourcing
Stated preferences
Contingent valuation
Hierarchical Bayes

ABSTRACT

Increasing pooling rates has benefits in terms of reducing emissions and traffic volume; however, trip data show that most rides are done solo. Users of ridesourcing services can choose whether to share a ride or, if they prefer not to, pay more for a solo ride. This study conducts a contingent valuation study to estimate latent willingness to pool (WTPO) a ride for a large number of customer segments. The results show that experience with ridesourcing is an important factor in explaining differences in WTPO. Those with experience are also less willing to pool rides, in particular younger men with higher incomes. Meanwhile, women and older age groups are predicted to be more amenable to pool rides, but so far their use of ridesourcing services is low. Recruiting them as customers may help increase pooling rates.

1. Introduction

Ridesourcing provides additional travel options to passengers. In this study the term ridesourcing refers to the process of passengers using a digital platform to source a ride from a driver who provides transportation services for a fee. Ridescourcing can take the form of solo rides or of ride pooling where rides are shared with other passengers. Focus of this study is the pooling of rides. It is beneficial for society because it reduces the number of car trips and therefore congestion and pollution, and also because it can improve access to and connectivity within the public transport system. However, to realise these benefits, passengers actually need to pool rides. Higher pooling rates imply increased mobility without disproportionally increasing traffic. If ridesourcing service providers understand who is and would be amenable to pooling they can be target their services appropriately.

Pooling rates can be predicted by factors like trip purpose and travel time (Hou et al., 2020). However, pooled rides make up only a small percentage of total trips served. For example, Henao and Marshall (2019) estimate for Denver, United States (US), an average vehicle occupancy rate of 1.4 passengers per trip, while Hou et al. (2020) estimate an average occupancy of 1.3 passengers per trip for the US city Chicago. People are likely differ in terms of what factors they consider relevant, when deciding to pool or not. Someone who generally dislikes sharing space with other passengers, may need a much lower price point, than someone who generally does not mind sharing the vehicle with others.

This study aims to inform ridesourcing customer segmentation to increase pooling rates. It uncovers differences in willingness to pool in a (sub-)urban region in Germany. A contingent valuation (CV) study is conducted. Respondents are asked how much they are willing to pay to prevent other passengers from joining a ride. Latent willingness to pool (WTPO) is estimated for different groups. The contribution of this study is that it uses a relatively simple valuation method to obtain respondents' total valuations. These are then segmented according to socio-demographic characteristics and experience level. Despite data being limited for some groups, robust WTPO estimates are obtained for all segments using hierarchical Bayes methods.

We find that accounting for experience is important when estimating WTPO. Those with experience of ridesourcing have lower WTPO estimates, while those with no experience are more willing to pool rides. The implication of this is that current reports of low pooling rates may be partly explained by the composition of the current customer base. It tends to be younger, male and with higher incomes. Pooling rates could be possibly increased by adding the customer groups who we identify as being more amenable to pooling.

2. Literature on preferences for ride pooling

Research on preferences for ride pooling is relatively recent. Hou et al. (2020) find different explanations for why people may not want to pool rides. They take longer than solo rides, but because they are cheaper customers have to weigh cost savings against increased travel time and discomfort associated with sharing. Sarriera et al. (2017) conducted a survey of Transportation Network Companies users

* Corresponding author.Correspondence to: DLR Institute of Transport Research, Rudower Chaussee 7, 12489 Berlin, Germany.
E-mail address: ariane.kehlbacher@dlr.de (A. Kehlbacher).

https://doi.org/10.1016/j.cstp.2025.101484
Received 19 October 2023; Received in revised form 30 January 2025; Accepted 14 May 2025
Available online 31 May 2025

through the online market place Amazon Mechanical Turk in the United States. They find that social interactions were relevant to mode choice, but not as much as traditional factors such as time and cost. The possibility of having a negative social interaction and safety concerns were identified as deterrents in particular for women. Much of the literature on ridesourcing pays particular attention to the resultant preference heterogeneity. Arguably this is because preference heterogeneity is more consequential in the context of ridesourcing, than in the context of public transport. Individuals who dislike sharing a vehicle with others can act on these preferences by paying more and book an solo ride. In the following we review RP and SP studies investigating differences in preferences for ride pooling.

Only few studies investigate ride pooling behaviour using RP data from the US. Kang et al. (2021) explain that RP data on ridesourcing are generally difficult to obtain, and ride pooling is available only in certain metropolitan cities (Henao and Marshall, 2019). Kang et al. (2021) collect both SP and RP data on pooled and solo rides for the city of Austin, US. Their analysis find that ethnicity and education are important predictors for propensity to pool. Henao and Marshall (2019) conduct a quasi-natural experiment in the city of Denver, US. They drove for Uber and Lyft to collect primary data on trips and real-time passenger feedback. Their analysis focuses on ride pooling only in so far as they provide an estimate for vehicle occupancy, which is 1.4 passenger per ride. Hou et al. (2020) use data from a transportation network provider in the city of Chicago, US. They asked individuals to indicate their willingness to pool and then investigate which socioeconomic, spatio-temporal, and trip characteristics are associated with it. Since indicating willingness to pool does not necessarily mean that the trip will actually be pooled, experienced users may have behaved strategically by indicating their willingness to pool to obtain solo rides with lower fares[1] Sarriera et al. (2017) conduct a survey through Amazon Mechanical Turk with Uber and Lyft service users in metropolitan areas in the US. Their analysis focuses on attitudes towards and perceptions of other passengers and the potential for discrimination. It highlights that the reasons for not wanting to pool rides are varied. A general drawback of RP data is that they include only existing customers of ridesourcing services. Customers who are currently not able or willing to use these services are missing.

Accordingly, several studies use stated preference (SP) methods, mostly choice experiments, to investigate preferences for ride pooling. This way they can include both current and future users of ridesourcing services. Al-Ayyash et al. (2016) investigate gender differences, in relation to the alternative-specific constant for shared taxi services relative to public transport. They conclude that women who use public transport are more likely to choose shared-ride taxi services than men. Lavieri and Bhat (2019) study the choice between pooled and solo rides with an autonomous vehicle. They analyse differences in WTPO, during commuting and leisure trips in the US. They identify gender, ethnicity, education, income, and vehicle availability as factors explaining differences in WTPO. Alonso-González et al. (2021) investigate preferences for different attributes related to ridesourcing in (sub)urban areas in the Netherlands. In their analysis the authors account for socio-demographic differences in relation to predicted latent class membership. Their analysis is of particular relevance to this study because they also investigate how utility changes as the number of passengers rises. Overall their results suggest the existence of some latent classes who prefer pooling rides more than others. The authors attribute this preference for pooling to the cost-saving characteristic of the pooled alternative. Al-Ayyash et al. (2016) also estimate how the marginal utility of a ride changes as the number of passengers increases. They find that Lebanese students' preferences for shared taxi service are lower when they have to share with 4–6 passengers as opposed to 1–3 passengers. Kang et al. (2021) estimate relative preferences for pooled versus solo rides. Their study is relevant to this study in that they expressly account for differences in self-reported familiarity with ridesourcing in their analysis. They estimate a multivariate model to allow for correlation in unobserved factors that explain differences in familiarity and differences in preferences for pooled versus individual rides. Most socio-demographic factors considered by them are investigated in relation to the latent constructs. Another study that is relevant to our work is Schatzmann et al. (2023). They conduct a mode choice experiment in the same area, namely the German city of Hamburg. The alternatives in their experiment include inter alia ridesourcing as a feeder for public transport and ridesourcing as a direct service. However, no differentiation is made between pooled and solo rides when choosing the ridesourcing alternative.

To summarise, only few studies investigate (future) customer preferences for ride pooling. Two studies investigate preferences for additional passengers joining a ride (Al-Ayyash et al., 2016; Alonso-González et al., 2021). The other studies compare utilities of pooled and solo rides in the context of mode choice, or they compare the utility of ridesourcing with that of other transport modes. This study adds to this literature by asking people directly how much they are willing to pay to prevent other passengers from boarding. Their responses are used to compute latent willingness to pool (WTPO). Existing studies obtain estimates of average utility either overall or for some latent classes, the membership to which is a function of socio-demographic factors. Differently, this study estimates latent WTPO of highly differentiated customer segments, including those who currently do not use ridesourcing services.

3. Data and methods

3.1. The contingent valuation method

Contingent valuation (CV) is a stated preference method for measuring the value of a non-market "good contingent on there being a market" (Mitchell and Carson, 2013). It sets a hypothetical market for the change in the quantity of a non-market good. Respondents are asked how much they are willing to pay for an increase (or willing to accept for a reduction) of the good. In theory, willingness to pay (WTP) and willingness to accept (WTA) should be equal, but the fact that WTA tends to be substantially larger than WTP has been widely acknowledged in the literature (Horowitz and McConnell, 2002). Because the WTP elicitation format tends to produce smaller estimates than the WTA elicitation format, practitioners of CV recommend using the former (Arrow et al., 1993).

Different questions formats exist for respondents to express maximum WTP, namely open-ended, single- or double-bounded dichotomous choice formats (Hanemann, 1984), and the payment card format (Mitchell and Carson, 2013). This study uses the open-ended format: respondents are asked to state their maximum WTP to prevent one additional passenger from joining the ride.

Benefit measures in CV have been developed by Hicks (1943). He suggested two measures where utility is held constant at the initial utility level (compensating variation and surplus), and two measures where utility is held constant at some specified alternative level (equivalence variation and surplus). The economic model in this study assumes that a change in the number of passengers, Q, has a direct impact on utility of the individual, whilst the price of the ride, which is the market good, stays the same. We are interested in measuring the potential benefits of a solo ride, as measured from the passenger's current level of utility. The relevant welfare measure in this case is the equivalent variation (EV) as given by

$$V(P^0, Q^0, M^0 - EV) = V(P^0, Q^1, M^0) \tag{1}$$

where V is the indirect utility function, P is price, Q is the number of additional passengers, and M is money income. The superscript 0 denotes initial levels and 1 denotes new levels. EV can be interpreted

[1] We thank the reviewer for raising this point.

Table 1
Distribution (in %) of age, gender, monthly net household income, and experience of ridesourcing in the sample, and their corresponding population statistics for the states Schleswig-Holstein (SH), Hamburg (HH), Lower-Saxon (LS) and Germany overall.

Age	Sample	SH	HH	NS
<20 years	2	12	9	12
21–30 years	12	12	16	12
31–40 years	18	12	17	12
41–50 years	19	19	18	18
51–60 years	23	15	13	16
61–70 years	17	13	11	12
70+ years	9	17	15	17
Gender	Sample	SH	HH	NS
Female or diverse	51	49	48	49
Male	49	51	52	51
Income	Sample	Germany		
Income EUR <1500	18	18		
Income EUR 1500–2999	37	43		
Income EUR 3000–4999	32	17		
Income EUR 5000+	13	44		
Use of ridesourcing				
1–7 times per week	3.6			
1–3 times per month	7.4			
Monthly or less	11.9			
(almost) never	77.1			

Table 2
Mean stated WTP amounts for riding alone, and preventing the first and second passenger from boarding by age, income, gender, experience. The bottom row shows the percentage of respondents who indicated WTP = 0 to the given question (N = 2348).

		WTP (in EUR) for sharing ride with...		
		0 passengers	1 passenger	2 passengers
Gender	Female/div	2.1	1.7	1.4
	Male	2.9	2.3	2.2
Age	$ < $20 yr	6.3	5.4	5.7
	21–30 yr	4.5	4.0	3.2
	31–40 yr	3.1	2.8	2.7
	41–50 yr	1.9	1.4	1.4
	51–60 yr	2.0	1.5	1.0
	61–70 yr	1.6	1.1	0.9
	70+ yrs	1.0	0.6	0.5
Income	EUR $ < $1500	2.2	1.7	1.6
	EUR 1500–2999	2.6	2.2	2.0
	EUR 3000–4999	2.0	1.6	1.3
	EUR 5000+	3.9	3.3	2.7
Experience	1–7 per week	5.3	4.9	4.2
	1–3 per month	5.2	4.6	4.8
	Monthly or less	3.4	3.0	2.7
	Never used	2.0	1.5	1.3
WTP = 0		46%	54%	62%

as the monetary equivalent of a change in utility that would arise as an additional passenger boards the vehicle. It is elicited by asking for the maximum WTP for avoiding an increase in the number of an additional passenger from Q^0 to Q^1.

A number of biases can affect CV based value estimates. Bias arise when values that are elicited in a hypothetical context differ from those elicited in a real context. Hypothetical bias affects SP studies in general (e.g. Brownstone and Small, 2005; Beck et al., 2016), although it is more pronounced in open-ended CV studies as respondents are free to choose the amount. As a consequence, value estimates are likely to be inflated. Another potential bias is scope insensitivity. It arises when respondents fail to offer a significantly higher payment amount for larger amounts of a good. Strategic bias arises when respondents provide a biased answer in order to influence a particular outcome (Carson and Mitchell, 1993).

3.2. Data collection and sample

Data collection took place from February to March 2021. Study area is the German city of Hamburg (64% of sample) and its adjacent districts, which are part of the states Lower Saxony (20% of sample) and Schleswig-Holstein (16% of sample). An online survey was administered to 1000 respondents out of which 987 completed the survey. After removing missing observations, the final sample consists of 783 respondents and $N = 2348$ observations. Table 1 compares sample distributions of age and gender to relevant state level statistics (German Census Data Bank, 2011), and net household income is compared to national statistics (German Federal Agency for Civic Education, 2020). The sample is representative in terms of gender and for middle age ranges. Very young, very old and those in high income households are under-represented, and those in households with moderately high income levels (EUR 3000-4999) are over-represented. Table 1 also shows the distribution of responses to the question "How often did you usually use shared on-demand services prior to the Covid-19 pandemic?". The majority of respondents has never used ridesourcing (77%). In our analysis, they are classified as having no experience of ridesourcing, and their responses are interpreted accordingly.

Table 2 reports mean stated WTP according to gender, age, income and experience, in response to the following questions: "Assuming you are making a trip with an on-demand shuttle that takes between 20 and 30 min and costs EUR 5. How much would you be willing to pay extra to ensure that you are alone/there are is only one other/there are only two other passengers in the vehicle?". The questions elicit WTP to prevent additional passengers from joining a ride. Mean stated WTP is highest for a solo ride and it decreases for preventing the first and second passenger to join the ride. Men tend to state higher WTP amounts than women or diverse genders. Younger respondents tend to state higher WTP amounts than older respondents. For example, mean stated WTP of respondents under 20 years is EUR 6.3. This constitutes a more than 100% increase in the costs of travel that was presented in the scenario. Looking at differences in WTP between different household incomes, respondents in households with incomes above 5000 Euro per month have the highest mean stated WTP, followed by respondents with household income ranging from 1500 to 2999 Euros. In the sample, some stated WTP amounts had values of up to 30 Euros. This is high considering that the cost of the ride was 5 Euros. Section 3.4 discusses how regularisation is used in the estimation to lessen the impact of this type of extreme observations on group WTP estimates.

3.3. Econometric model specification

Because WTP bids are censored at zero, we specify a Tobit model (Tobin, 1958) with the following likelihood

$$y_i = \begin{cases} y_i^* = A_j + \beta' x_i + \varepsilon_i, & \text{if } y_i^* > 0 \\ 0, & \text{if } y_i^* \leq 0 \end{cases} \qquad \varepsilon_i \sim \mathrm{N}(0, \sigma^2) \tag{2}$$

where y_i is stated WTP, and y_i^* is latent WTP to avoid pooling. The negative of latent WTP can be interpreted as willingness to pool (WTPO), which is the focus of this analysis. The $N \times K$ matrix x_i contains covariates indicating public transport use (1 = daily or up to 3 times per week, 0 = otherwise), and access to a car (1 = never, 0 = otherwise). Parameter β is a 2×1 vector, which is estimated. Parameter σ is the standard deviation of the normally distributed error term, which is also estimated. The term A_j relates to the hierarchical model specification, which is explained in the next section, where j is an index for groups $j = 1, \ldots, J$.

3.4. Hierarchical Bayes model specification and estimation

We aim to understand differences in latent WTPO of customer segments, which differ in terms of age, income, gender and experience of ridesourcing. One possible approach is to use interaction terms and regress them on stated WTP. However, the resultant regression coefficient captures the combined effect of age, gender, income and experience. Treating the covariates as continuous, would assume that the modification of the effect of one covariate on the dependent variable by another covariate occurs in a linear fashion. For example, the effect of age on WTPO would be assumed to change linearly as income changes. Using dummy coding for different factor levels would lead to a large number of different combinations of indicator variables, some of which are difficult to estimate due to lack of available data. Concretely, in our data there are seven age groups, two gender categories, four household income levels and two experience levels (see Table 1). This results in a total of $7 \times 2 \times 4 \times 2 = 112$ possible combinations. Obtaining estimates can be challenging as the number of observations per group can be small. Estimates will be noisy due to small sample size and/or because of the influence of extreme stated WTP amounts on group means. Moreover, data are available for only 94 of the 112 groups, but we still want to make predictions for all groups.

To deal with this, we use hierarchical Bayes (HB) (see e.g. Gelman et al., 2013) to estimate the Tobit model as described by the likelihood in Eq. (2). HB models have a lower level representing the parameters that govern groups, and a higher level with prior parameters governing the distribution of the group parameters. Variation in the data is used to estimate the variance parameter of the prior distribution for the group-specific intercepts. Specifying and estimating the prior distribution in this way has the advantage that information across groups is "pooled" such that an estimate for one group is partially informed by estimates in other groups. This process is called regularisation. A group with fewer observation has its estimate "pulled" towards mean latent WTPO based on other groups. This gives less weights to groups for which less information is available, whilst still allowing to estimate parameters for all groups. Extreme values become less probable, unless there is strong evidence in the data. In this study regularisation helps reduce the influence of very high stated WTP amounts on group parameters, and it enables estimation of groups with few or no observations. An advantage of using a Bayesian approach is that parameter uncertainty is transmitted between hierarchies in the HB model (see e.g. Gelman et al., 2013).

The hierarchical prior specification is introduced via A_j in (2), which is specified as follows:

$$A_j = \alpha + \alpha_{e[i]} + \alpha_{e[i]j[i]} \tag{3}$$

where α is the overall intercept, $\alpha_{e[i]}$ are varying intercepts for experience levels, where $e[i]$ denotes experience condition e for person i; $\alpha_{e[i]j[i]}$ are deviations from $\alpha_{e[i]}$ in terms of socio-demographic groups, where $j[i]$ denotes group j for person i. There are $J = 56$ socio-demographic groups and $e = 1, \dots, E$ and $E = 2$ experience levels. Experience levels are grouped into none and otherwise. Overall, there are 112 possible combinations of experience levels and socio-demographic characteristics.

Bayesian inference requires the specification of the full probability model for the data in terms of likelihood function and prior distributions of its parameters. For each parameter in the likelihood specified in Eqs. (2) and (3) we specify a prior distribution

$$\begin{aligned} \beta &\sim \mathrm{N}(0,1) \\ \sigma &\sim \text{Student-t}^{+}(3,0,2.5) \\ \alpha &\sim \mathrm{N}(0,1) \\ \alpha_e &\sim \mathrm{N}(0,\tau_e) \\ \alpha_{ej} &\sim \mathrm{N}(0,\tau_{ej}) \\ \tau_e &\sim \mathrm{N}^{+}(0,s_0) \\ \tau_{ej} &\sim \mathrm{N}^{+}(0,s_0) \end{aligned} \tag{4}$$

where σ is the overall error standard deviation, which follows a positive truncated Student-t distribution. The parameter τ_e is the standard deviation of the varying intercepts for experience levels and τ_{ej} is the standard deviation for the varying intercepts of all combinations of socio-demographic groups and experience levels. The standard deviations of the varying intercepts are positive truncated Normal distributions. They provide some regularisation to improve convergence and sampling efficiency, whilst at the same allowing differing degrees of variation of the group specific intercepts. The value of the hyperparameter, s_0, is determined using prior predictive checks. This involves estimating the model and making predictions based on the prior distributions of the parameters only, whilst ignoring the likelihood. The idea is to identify prior specifications that yield distributions for the data with some mass around extreme but plausible data sets and less mass on completely implausible data sets (Gabry et al., 2019). This approach, where an observational model for the data is developed first, and then constrained via a soft containment prior to prevent potentially problematic behaviours, is called defencive prior modelling by Betancourt (2022).

The specification in this study is based on the authors' experience of what constitutes reasonable costs for ridesourcing in Hamburg. We expect some people to be willing to pay extra to prevent others from boarding a ride, but for most people WTP to be close to zero. This means average WTP should be centred around zero. Given that the cost of the ride in the scenario given to respondents was EUR 5.00, we decide that a 97.5% probability of predicted average observed WTP ranging from EUR 0.00 to 5.00 is realistic. A soft containment prior model is implemented in model $\mathcal{M}_2$, where we set $s_0 = 2$.

Next, prior distributions are combined with the likelihood to obtain the joint posterior distribution, which is estimated using Markov Chain Monte Carlo (MCMC) methods in RStan (Stan Development Team, 2021) via the R package brms version 2.16 (Buerkner, 2017). The sampler was run with four independent Markov chains, with 2000 warm-up iterations and 2000 sampling iterations each, resulting in a total of 8000 posterior draws. Convergence of the estimation procedure is checked by monitoring the stationarity of the Markov chains via the split potential scale reduction factor (Rhat). It measures the ratio of the average variance of samples within each chain to the variance of the pooled samples across chains. A value of 1 indicates that chains have converged to the same level as they are comparable to each other. Reliability of the estimates is monitored by checking bulk and tail effective sample sizes, which, if large, indicate that estimates are based on large numbers of independent samples from the posterior distribution and therefore reliable. Finally, out-of-sample predictive accuracy is evaluated using Pareto smoothed importance sampling cross-validation (Vehtari et al., 2016). An estimate of the shape parameter of the Pareto distribution, k, of larger than 0.7 is reason for concern (see Vehtari et al., 2016) because it indicates that a data point is highly influential to the posterior distribution, and therefore has the potential to negatively affect future predictions.

4. Results

Before presenting the results from the model estimation, we check the out-of-sample prediction accuracy of the model. Fig. 1 shows Pareto k estimates for all data points. The x-axis indicates the number of the observation, the y-axis indicates the magnitude of an observation's Pareto k estimate. Its magnitude is influenced by the extent to which the joint posterior density is expected to change, if the model had been estimated without a given observation. Larger values indicate that the joint posterior distribution is expected to change a lot, and vice versa. If the posterior changes a lot, when an observation is omitted, this means that generalising from sample to population may not be appropriate. A generally accepted threshold for Pareto k estimates is 0.7. All observations are below this threshold. There is not any one observation with particularly strong influence on the posterior distribution. While this is important for generalising our findings to the population, it also shows that the model can accommodate the

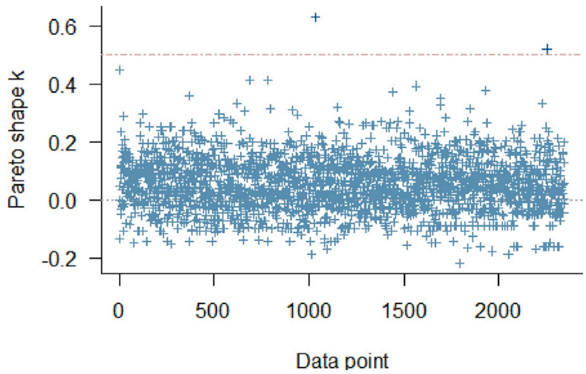

Fig. 1. Scatter plot of Pareto shape parameter estimates k. They show for each data point how much the joint posterior would be affected, if it were omitted from the estimation.

Table 3
Posterior means of the population regression parameters and their 95% credible intervals, potential scale reduction factors (Rhat) and effective bulk sample sizes (ESS).

Covariate	Mean	l-95% CI	u-95% CI	Rhat	ESS
Intercept	1.057	−0.71	2.81	1.00	6461
Prevent 1st passenger	−0.923	−1.59	−0.25	1.00	11 708
Prevent 2nd passenger	−1.782	−2.46	−1.13	1.00	11 212
No access to car	0.961	0.16	1.76	1.00	9788
Public transport often	−0.619	−1.25	0.01	1.00	11 375
Standard deviation τ_e	2.270	0.95	4.47	1.00	5057
Standard deviation τ_{ej}	3.468	2.77	4.27	1.00	2046
Standard deviation σ	6.288	6.00	6.59	1.00	11 499

high stated WTP amounts mentioned in Section 3.2, without them being unduly influential. Table 3 reports the parameter estimates of the overall regression. All potential scale reduction factors, Rhat, are below the critical threshold 1.05 indicating that the MCMC sampler has converged. Effective bulk and tail sample sizes are all large indicating that estimates are based on large numbers of independent draws from the posterior distribution and therefore reliable. First, looking at the standard deviation of the varying intercepts for experience levels, its magnitude indicates that WTP varies because of groups' different levels of experience of ridesourcing. Parameter τ_e is centred at 2.3, but values of up to 4.5 are probable. It is only slightly smaller than the standard deviation related to socio-demographic differences within experience conditions, τ_{ej}, which has a posterior mean of 3.5, with values up to 4.3 being probable. For comparison, variation in stated WTP due to unobserved factors is estimated at 6.3 with values up to 6.6 being probable. This highlights that variations in both experience and socio-demographic characteristics contribute significantly to the heterogeneity in WTPO.

As for the regression parameters, the posterior mean of the overall intercept can be interpreted as average stated WTP for riding alone. The small value of its posterior mean is due to the fact that most respondents stated zero WTP amounts. Next, we look at the effect of increasing the number of passengers prevented from boarding on stated WTP. Both effects for the first and second passenger are negative, implying WTP = 0. Preventing the second passenger from boarding is worth less (−1.8) than preventing the first one (−0.9) from boarding. This makes sense, the marginal utility from preventing one more passenger from boarding is expected to decrease with increasing passenger numbers. It also suggest that there is evidence against scope sensitivity bias (see Section 3.1). As for the effect public transport use, the parameter estimate's 95% uncertainty interval just about includes zero. So there could be no difference between those who use public often and those who do not. The fact that the mass of the posterior is in the negative region, suggests some that those who use public transport often could be more inclined to pool rides, as their stated WTP amounts are lower. By contrast, having access to a car is estimated to lead to higher stated WTP implying less willingness to pool.

To obtain posterior predictive distributions of each group's average latent WTPO, estimated posteriors of the overall intercept and relevant varying group-intercepts are combined. For groups without data (see also Section 3.4), predictions are based on the joint prior distribution, which is estimated from data for other groups with the same characteristics. For groups with data, predictions are based on both data and information coming from the hierarchical prior distribution. Latent WTPO is defined as the inverse of stated WTP for a solo ride. A positive value implies willingness to pool, a negative value implies that a group is unwilling to pool. If uncertainty intervals of predicted latent WTPO are positive and exclude zero, the group is predicted to be willing to pay extra to prevent having to pool a ride (WTP > 0). Posterior distributions of predicted latent WTPO are reported in Figs. 2 to 5.

There are four figures in total. Two figures show results for women and diverse genders with and without experience of ridesourcing. Two figures show results for men with and without experience. In a given figure, there are four panes, one for each income level. They range from the lowest income level in (a) to the highest income level in (d). A given pane shows the posterior predictive density of latent WTPO of a given group. The y-axes indicate the magnitude and the x-axes indicate age groups.

First, looking at the results for women with no experience in Fig. 2, we find that average latent WTPO of older groups tends to be higher than that of younger groups. This is consistent across income levels. Younger groups across all income levels are predicted to have negative latent WTPO, but uncertainty of the marginal posterior predictive WTPO distribution is high. Thus, although younger women with no experience of ridesourcing dislike pooling more than older women, they are not predicted to be willing to pay extra to avoid pooling.

Second, looking at results for women with experience of ridesourcing in Fig. 3 we notice that estimates are particularly uncertain for older age groups as indicated by the width of their posterior predictive distributions of latent WTPO. This happens because these groups' estimates are based on fewer responses and relatively more informed by the joint hierarchical prior distribution for all groups. The pattern of older groups being more willing to pool than younger groups across income levels is less pronounced, than it is for women without experience in Fig. 2. Older groups with experience are predicted to be less willing to pool than their counterparts with no experience. For two income groups in panes (b) and (d) the dislike of pooling of age groups 21–30 and 31–40 is such the posterior predictive distribution of latent WTPO excludes zero. This means they are predicted to be willing to pay extra to avoid pooling.

Third, Fig. 4 depicts predicted average latent WTPO of men with no experience of ridesourcing. Similar to women, there is a pattern where younger groups are less willing to pool than older groups. This pattern is more pronounced among lower income groups as can be seen in panes (a) and (b). In particular men below 20 years with household incomes below EUR 1500 are unwilling to pool and predicted to be willing to pay extra to avoid it.

Fourth, looking at results for men with experience of ridesourcing in Fig. 5, we find again that older groups are predicted to be more willing to pool than younger groups. The posterior predictive distribution of average latent WTPO of the younger groups excludes zero. Therefore, groups up to the age of 40 are predicted to be willing to pay extra to avoid pooling. As is the case for their female counterparts, uncertainty of latent WTPO estimates is higher for older groups because there are fewer observations for these groups.

To summarise, the magnitudes of the standard deviations of the hierarchical prior distributions indicate that differences in socio-demographic characteristics and in experience are important sources of heterogeneity in latent WTPO. Average latent WTPO is generally low and although groups experience disutility from pooling, it is

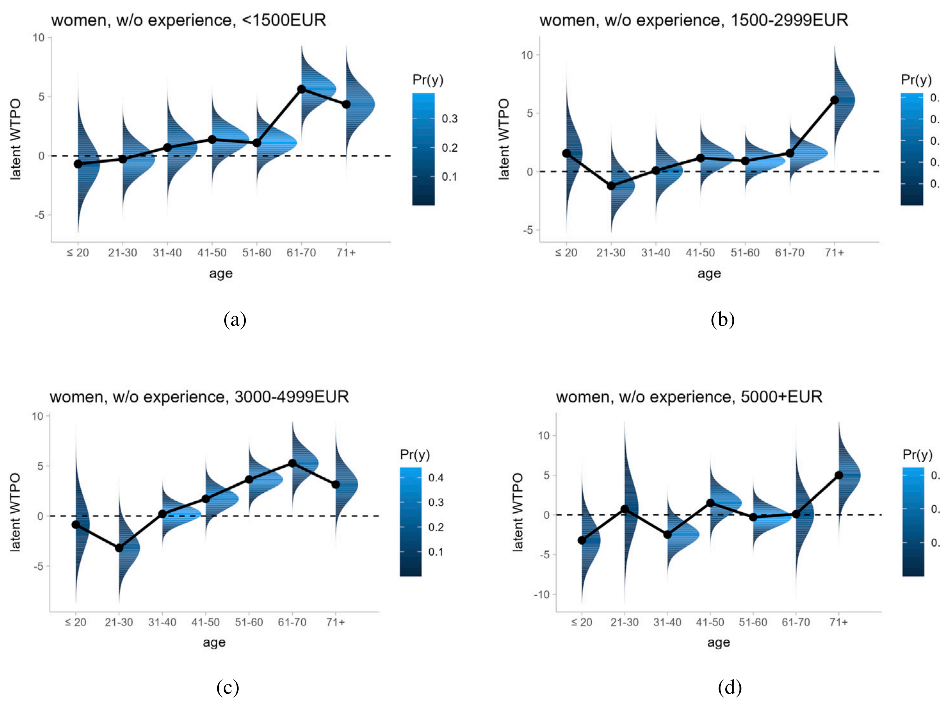

Fig. 2. Posterior predictive distributions of average latent WTPO of women with no experience of ridesourcing who have no access to a car and use public transport less often than daily at different age and income levels.

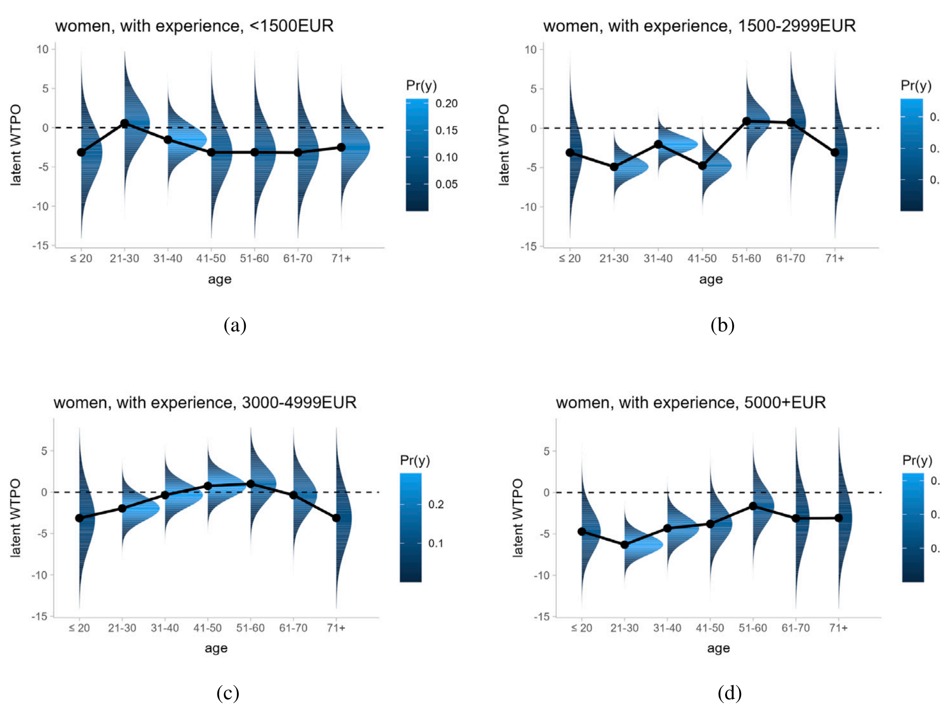

Fig. 3. Posterior predictive distribution of average latent WTPO of women with experience of ridesourcing and who have no access to a car and use public transport less often than daily at different age and income levels.

mostly not large enough to be willing to pay extra to avoid it in the given scenario. The exception are younger men with experience of ridesourcing. Regardless of income level, their latent WTPO is negative and they are predicted to be willing to pay extra to avoid pooling. Another pattern we find is that older groups with no experience, are consistently more willing to pool than their younger counterparts. But this difference between older and younger groups is less pronounced if the groups have experience of ridesourcing. Overall, our results highlight the importance of allowing for nonlinear associations when modelling differences in latent WTPO across socio-demographic groups and experience levels. Visually, non-linearity in relation to age can be easily detected. Finally, the widths of the posterior predictive distributions of average latent WTPO show that parameter uncertainty is higher for groups with experience of ridesourcing, in particular for groups including older women and older men, as well as low income groups. This reflects the fact that data on these groups are limited.

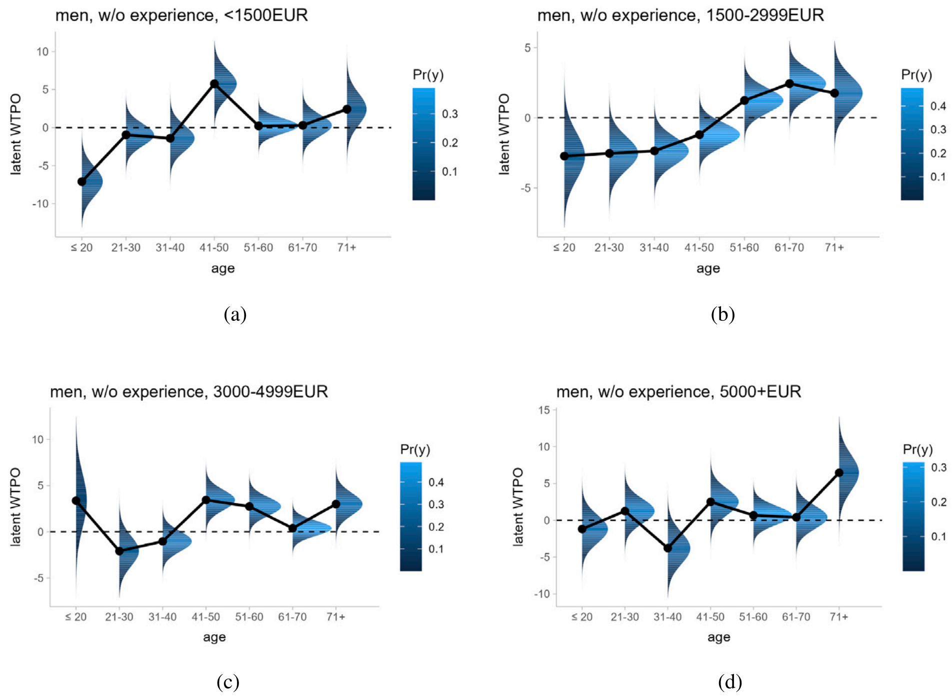

Fig. 4. Posterior predictive distribution of average latent WTPO of men with no experience of ridesourcing who have no access to a car and use public transport less often than daily at different age and income levels.

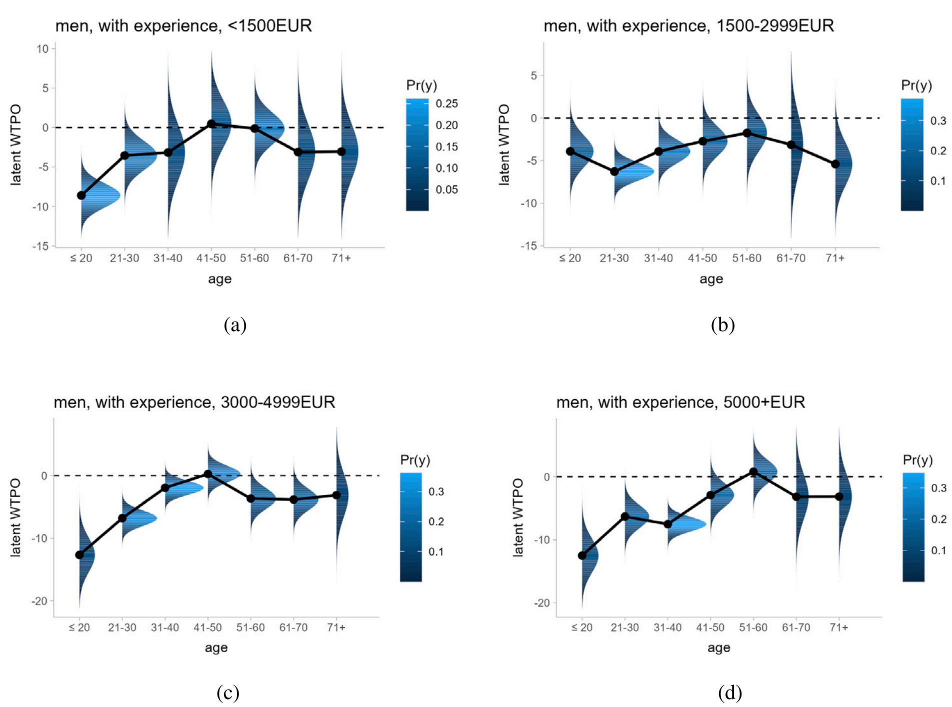

Fig. 5. Posterior predictive distribution of average latent WTPO for a 25 min ride of men with experience of ridesourcing who have no access to a car and use public transport less often than daily at different age and income levels.

5. Discussion

This study aims to investigate how preferences for ride pooling differ between a large number of customer segments, including those who have not yet used ridesourcing. It employs the CV method whereby WTP amounts are elicited directly. Unlike SP experiments, where effects of changes in other attributes can be taken into account, the CV method elicits valuations only for a given travel time and cost scenario. This is restrictive in terms of being able to understand interdependencies with other attributes. SP experiments offer this complexity in terms of trade-offs with other attributes. However, in their SP experiment (Lavieri and Bhat, 2019) find that WTP to avoid travelling with strangers presents a fixed cost that appears to be independent of travel time. This aligns with the approach chosen in this study where travel time and cost are fixed, only the number of passengers varies. The direct value elicitation format means a model with stated WTP as dependent variable can be specified. This in turn allows modelling complex interactions of socio-demographic characteristics and experience conditions. Each combination of socio-demographic factors and experience levels is assigned its own intercept. This way nonlinear relationships between combinations of levels of socio-demographic characteristics and experience conditions can be accommodated by the model. When estimating the varying intercept parameters, information is shared between groups through a hierarchical prior. As a consequence, differences in latent WTPO between a larger number of (potential) customer segments can be identified. It is also possible to obtain estimates for groups for which data are limited, for example, older groups who use ridesourcing or low income groups. This is helpful given that these groups are currently underrepresented among current ridesourcing users in the study area (e.g. Schatzmann et al., 2023)

The following discussion focuses on differences in latent WTPO. Absolute value estimates are less reliable due to hypothetical bias. In the absence of a general theory of respondent behaviour that can explain hypothetical bias (Loomis, 2011), we assume that hypothetical bias affects average value estimates of all groups equally. Its effects on the predictions are attenuated. We use defencive prior modelling whereby implausible values are assigned lower probability mass, and we estimate a HB model whereby extreme group estimates are regularised towards the overall mean (see Section 3.4).

Results from the overall regression parameters are reported in Table 3. On average, individuals with no access to a car are less willing to pool rides compared to those with access to a car, all else equal. This suggests that those without a car may view ridesourcing as similar to a taxi service, where they hire a driver to bring them to their destination without having to deal with other passengers. Furthermore, there is some evidence to suggest that regular public transport users may be more willing to pool rides. However, the true effect size in this instance could be zero, so this is not conclusively supported by our analysis. Still, as the mass of the posterior density is in the negative region, a tentative explanation is provided. Because regular public transport users are accustomed to sharing rides with others, they may be less hesitant to do so during ridesourcing as well.

Overall, we find that experience of ridesourcing is an important driver of differences in WTPO based on the magnitude of the standard deviation of the hierarchical prior distribution for the experience conditions. This is in line with Kang et al. (2021). Their model jointly estimates familiarity with and choice of ridesourcing whilst allowing unobserved factors to affect both equations. They also conclude that explicitly considering familiarity when investigating the choice process for pooled ridesourcing is important. Different to Kang et al. (2021), we focus on how the effect of experience on WTPO varies between socio-demographic groups. It makes a difference in two ways to our results. First are the uncertainty intervals of the estimates. They are narrower for groups of men with experience of ridesourcing at mid and high income levels. For these groups more observations are available, which is why we can be more certain about whether and how much they would be willing to pay extra for being able to ride solo. WTPO estimates are more uncertain for groups that are older and/or have lower household incomes and who also have experience of ridesourcing. Fewer respondents fall into these categories. Lavieri and Bhat (2019) also find that a higher household income is associated with having experience with ridesourcing. And they also find that older individuals are less likely to have used ridesourcing than younger individuals. Although observations of older and low income groups with experience of ridesourcing are limited, it is still possible to make valid inferences about their preferences. The HB model obtains group-specific estimates for these groups by partially drawing on information from all other groups in addition to the (albeit limited) observations for these groups. This way our study can make predictions for groups who are currently missing from RP studies and whose WTPO is therefore unknown. The second way in which experience affects results is in terms of the extent to which groups are willing to pool. Groups with experience of ridesourcing are predicted to have lower WTPO. For comparison, Kang et al. (2021) find that higher familiarity with pooled ridesourcing to be associated with a high propensity for sharing. However, results may not be comparable. Kang et al. (2021) study the choice between pooled and individual rides. There may be unobserved factors other the number of additional passengers that are related to this choice. Another SP study also finds that experience affects valuations. Hensher et al. (1991) find that experience changes stated preferences for traffic management devices. They suggest this may happen because of changes in awareness of the attributes. Jensen et al. (2013) use repeated SP experiments and find that preferences for attributes of an electric vehicle change after a real life experience. A potential explanation for this is that respondents change their reference points in the choice or valuation task, something which is not foreseen in neo-classical economics, but addressed by Prospect theory (Kahneman and Tversky, 1979). In the context of this study, this would imply that individuals with experience of ridesourcing have a different reference point – namely, their real-life experience – while individuals without experience may use taxi services or public transport as a reference point. Based on Kang et al. (2021) finding that if someone has pooled rides in the past, they are likely to do also so in the future, another explanation for our results is possible. Those without experience also intend to not use ridesourcing in the future. They are therefore happy to state that they are willing to share rides with other passengers, possibly reflecting social desirability bias their answers.

Next we discuss differences in WTPO between socio-demographic groups. Starting with age, a noticeable pattern is that older groups are more willing to pool than younger groups. This holds true across income, gender, and experience levels. A possible explanation is provided by Lavieri and Bhat (2019) who find that younger adults (18–34 years of age) display greater levels of privacy-sensitivity compared to older adults when using ridesourcing services.

As for gender differences, womens' WTPO tends to be higher than that of men. This is different to Lavieri and Bhat (2019). They find that women are less likely to pool rides for a commute trip, but find no difference to men when it comes to leisure trips. Lavieri and Bhat (2019) look at preferences for pooling rides using autonomous vehicles whereas this study investigates preferences for pooling when using a vehicle with a driver. For this context, Sarriera et al. (2017) find that women report more often than men that they feel safer having another person in the car other than the driver. Thus, a possible explanation for our finding is that women being relatively more willing to pool rides than men so as to not be alone with the driver. In the interplay between gender and age, we find that the above mentioned gap between WTPO of younger and older groups is more pronounced for men than for women. Partly this is explained by the higher uncertainty of the latent WTPO estimates for women with experience, but it may warrant further investigation.

Looking at income, we find no systematic differences in WTPO associated with income, once age, gender and experience are taken into account. In the literature, evidence on differences in WTPO in relation to income is mixed. SP studies like Alonso-González et al. (2021) find that individuals with higher income tend to have lower WTPO, and Lavieri and Bhat (2019) find that high-income individuals are less likely to adopt pooled ride-hailing, whereas Brown (2018), who analyse trip data in the US, find a higher share of pooled trips in high income areas than in low income areas. By contrast, also looking at trip data, Hou et al. (2020) find that trips starting or ending in higher income areas to have lower ratio of pooled to solo rides.

To conclude, while most groups experience disutility from pooling, it is not enough, to be willing to pay extra to avoid it. This is in line with Lavieri and Bhat (2019) who find that cost-time trade-offs are more important than sharing disutility. Similarly, Kang et al. (2021) find rather small average WTP estimates to prevent pooling ranging from USD 0.6 to USD 1.7. Given this, it is interesting to note that RP studies like Henao and Marshall (2019) find most rides to be done without pooling. In their study only 13% of requests were for pooling services. This contradiction may be partly explained by the fact that current users of ridesourcing services are part of those groups, which we predict to be unwilling to pool — namely men up to 40 and across all income levels. Indeed, the sample of recorded rides by Henao and Marshall (2019) includes mostly younger adults and it has an even split between men and women. Their sample is non-representative in terms of income distribution, although it is not clear in what sense. Similarly, Rayle et al. (2016) in their survey on ridesourcing find that respondents are generally younger than the average population, that the lowest income group is underrepresented, and men are slightly over-represented. Kang et al. (2021) investigate stated familiarity with pooled rides. They also find that younger and higher income individuals are more likely to be familiar with pooled rides, but they do not differentiate by gender in their analysis.

In addition to the already discussed hypothetical bias, there are a number of limitations that need to be taken into account when interpreting our findings. First, while our sample is representative in terms of gender and of individuals in the middle age ranges, very young and very old individuals are underrepresented. Individuals in high income households are also under-represented, whereas persons in households with moderately high income levels (EUR 3000-4999) are over-represented. However, since HB models can help mitigate some of the effects of a non-representative sample by allowing estimates for groups with underrepresented data to "borrow strength" from groups with more data. Second, the model assumes exchangeability, which means unobserved factors explaining deviations from the population level intercept for a given group are assumed to be systematically unrelated to those of other groups. Arguably, this may not hold, if groups are adjacent in terms of age and income, but given the relatively granular decomposition of our groups, assuming exchangeability still seems reasonable, and making this assumption makes it possible to investigate deep interactions in the first place. Third, our findings pertain to suburban and urban areas. WTPO in rural areas may well be different, for example, in that the benefits of access outweigh any disutility from having to share the vehicle with other passengers. Finally, our model allows us to learn how WTPO varies across socio-demographic groups and experience conditions, but it does not tell us why preferences vary. We see this as a potential avenue for future research.

6. Conclusion

Ridepooling is still relatively uncommon in Germany, and consequently, data on ridepooling are also limited. This SP study investigates differences in WTPO for a large number of socio-demographic groups with and without experience of ridesourcing. The estimation approach allows for nonlinear relationships between socio-demographic and experience conditions to uncover patterns between combinations of factors. We find that most groups mildly dislike pooling, but their dislike is not strong enough to be willing to pay extra to avoid it. Not having access to a car is also associated with a stronger dislike of pooling. In line with previous work, we find that accounting for experience is important. It is an important source of heterogeneity, and groups with experience of ridesourcing tend to be less willing to pool. A notable finding is that older groups state to be more willing to pool than younger groups. This applies across income, gender and experience levels. At the same time, younger groups and men are more likely to have experience of ridesourcing. These groups also tend to particularly dislike pooling to the extent that they are predicted to be willing to pay extra to avoid it. An implication of this is that evidence from RP studies, which find that most rides are not pooled, is partially explained by the fact that current users belong to those groups, which this study identifies as particularly disliking pooling, namely younger men with higher incomes.

7. Policy implications and recommendations

It is concerning that so few ridesourcing trips currently involve pooling. This study finds that most customer segments, while not liking to pool rides, appear to willing to accept them. This means a pricing strategy where solo rides are substantially more expensive than pooled rides may increase pooling rates, without losing too many customers. However, this prediction is for the scenario where the customer base is not so heavily skewed as it is currently the case towards younger and/or male users, as these groups are relatively more unwilling to pool and may switch to another transport mode if the price of solo rides is too high. A recommendation for service providers is to target new customer segments such as older people and women, as they are more inclined to pool rides. For policymakers trying to reduce the number of cars on the road by increasing pooling rates, a recommendation is to create conditions that facilitate access to ridesourcing for these groups. Given that there is some (albeit inconclusive) evidence that regular public transport users are more willing to pool rides, another recommendation is to support the provision of ridesourcing services in the vicinity of public transport hubs.

For researchers we see two interesting lines of investigation. First, from a practical point of view, future research on groups who are currently not using ridesourcing but are predicted to be willing to pool, namely women and older groups, appears useful. Understanding how these groups' access to and use of ridesourcing can be improved helps service providers tailor their marketing efforts in terms of messaging, tone, and offers, and helps policy makers understand how to support ridesourcing services provisions for these groups. Second, given that ignoring experience may have resulted in too optimistic predictions for certain groups in terms of predicting future market demand for ride pooling, we see methods of accounting for experience when eliciting preferences for ridesourcing as an interesting issue to be further explored.

CRediT authorship contribution statement

Ariane Kehlbacher: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Kerstin Stark:** Writing – original draft, Investigation, Conceptualization. **Laura Gebhardt:** Writing – original draft, Investigation. **Jan Weschke:** Writing – original draft, Investigation, Data curation.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used OpenAI in order to check grammatical correctness and clarity of some of the sentences. After using this service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

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

Acknowledgements

The authors gratefully acknowledge funding by the German Federal Ministry for Digital and Transport as part of the project RealLabHH.

References

Al-Ayyash, Z., Abou-Zeid, M., Kaysi, I., 2016. Modeling the demand for a shared-ride taxi service: An application to an organization-based context. Transp. Policy 48, 169–182. http://dx.doi.org/10.1016/j.tranpol.2016.02.013.

Alonso-González, M.J., Cats, O., van Oort, N., Hoogendoorn-Lanser, S., Hoogendoorn, S., 2021. What are the determinants of the willingness to share rides in pooled on-demand services? Transp. 48 (4), 1733–1765.

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., 1993. Report of the NOOA panel on contingent valuation. natural resource damage assessment under the oil pollution act of 1990. Fed. Regist..

Beck, M.J., Fifer, S., Rose, J.M., 2016. Can you ever be certain? Reducing hypothetical bias in stated choice experiments via respondent reported choice certainty. Transp. Res. Part B: Methodol. 89, 149–167.

Betancourt, M., 2022. Prior modeling. https://betanalpha.github.io/assets/case_studies/prior_modeling.html#License. (Accessed 02 August 2022).

Brown, A.E., 2018. Ridehail Revolution: Ridehail Travel and Equity in Los Angeles. University of California, Los Angeles.

Brownstone, D., Small, K.A., 2005. Valuing time and reliability: assessing the evidence from road pricing demonstrations. Transp. Res. Part A: Policy Pr. 39 (4), 279–293, Connection Choice: Papers from the 10th IATBR Conference.

Buerkner, P., 2017. Brms: An R package for Bayesian multilevel models using Stan. J. Stat. Softw. 80 (1), 1–28.

Carson, R.T., Mitchell, R.C., 1993. The issue of scope in contingent valuation studies. Am. J. Agric. Econ. 75 (5), 1263–1267.

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., Gelman, A., 2019. Visualization in Bayesian workflow. J. R. Stat. Soc.: Ser. A (Stat. Soc.) 182 (2), 389–402.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., Rubin, D., 2013. Bayesian Data Analysis, third ed. Chapman & Hall/CRC.

German Census Data Bank, 2011. Census data bank. ergebnisse2011.zensus2022.de/datenbank/online.

German Federal Agency for Civic Education, 2020. Einkommen privater Haushalte. bpb.de/nachschlagen/zahlen-und-fakten/soziale-situation-in-deutschland/61754/einkommen-privater-haushalte.

Hanemann, W.M., 1984. Welfare evaluations in contingent valuation experiments with discrete responses. Am. J. Agric. Econ. 66 (3), 332–341.

Henao, A., Marshall, W., 2019. The impact of ride-hailing on vehicle miles traveled. Transp. 46, 2173–2194.

Hensher, D.A., Battellino, H.C., Gee, J.L., 1991. The role of stated preferences and discrete choice models in identifying community preferences for traffic management devices.

Hicks, J.R., 1943. The four consumer's surpluses. Rev. Econ. Stud. 11 (1), 31–41.

Horowitz, J.K., McConnell, K.E., 2002. A review of WTA/WTP studies. J. Environ. Econ. Manag. 44 (3), 426–447.

Hou, Y., Garikapati, V., Weigl, D., Henao, A., Moniot, M., Sperling, J., 2020. Factors influencing willingness to pool in ride-hailing trips. Transp. Res. Rec. 2674 (5), 419–429.

Jensen, A.F., Cherchi, E., Mabit, S.L., 2013. On the stability of preferences and attitudes before and after experiencing an electric vehicle. Transp. Res. Part D: Transp. Environ. 25, 24–32. http://dx.doi.org/10.1016/j.trd.2013.07.006.

Kahneman, D., Tversky, A., 1979. Prospect theory: An analysis of decision under risk. Econom. 47, 263–291.

Kang, S., Mondal, A., Bhat, A.C., Bhat, C.R., 2021. Pooled versus private ride-hailing: A joint revealed and stated preference analysis recognizing psycho-social factors. Transp. Res. Part C: Emerg. Technol. 124, 102906.

Lavieri, P.S., Bhat, C.R., 2019. Modeling individuals' willingness to share trips with strangers in an autonomous vehicle future. Transp. Res. Part A: Policy Pr. 124, 242–261.

Loomis, J., 2011. What's to know about hypothetical bias in stated preference valuation studies? J. Econ. Surv. 25 (2), 363–370.

Mitchell, R.C., Carson, R.T., 2013. Using Surveys to Value Public Goods: The Contingent Valuation Method. Rff Press.

Rayle, L., Dai, D., Chan, N., Cervero, R., Shaheen, S., 2016. Just a better taxi? A survey-based comparison of taxis, transit, and ridesourcing services in San Francisco. Transp. Policy 45, 168–178. http://dx.doi.org/10.1016/j.tranpol.2015.10.004.

Sarriera, J.M., Álvarez, G.E., Blynn, K., Alesbury, A., Scully, T., Zhao, J., 2017. To share or not to share: Investigating the social aspects of dynamic ridesharing. Transp. Res. Rec. 2605 (1), 109–117.

Schatzmann, T., Zwick, F., Axhausen, K.W., 2023. Investigating the preferences for the use of urban ridepooling. In: 11th Symposium of the European Association for Research in Transportation. HEART 2023, IVT, ETH Zurich.

Stan Development Team, 2021. RStan: the r interface to Stan. URL: http://mc-stan.org/, R package version 2.21.1.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econom. 26 (1), 24–36.

Vehtari, A., Gelman, A., Gabry, J., 2016. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. 27 (5), 1413–1432. http://dx.doi.org/10.1007/s11222-016-9696-4.