Capturing behavioral differences between regular and seldom users of ride-pooling in a cross-nested logit model

Nadine Kostorz-Weiss*[1], Michael Heilig[1], Martin Kagerbauer[1], Peter Vortisch[1]

[1]Institute for Transport Studies, Karlsruhe Institute of Technology (KIT), Kaiserstr.12, 76131 Karlsruhe, Germany

SHORT SUMMARY

Within this study, we estimate a Cross-Nested Logit (CNL) model incorporating ride-pooling and other shared modes using combined revealed and stated preference data. The CNL allows to captures complex substitution patterns that cannot be represented by simpler MNL or NL models. Results show that ride-pooling correlates with multiple nests and that behavior differs significantly between frequent and seldom users. Frequent users (using ride-pooling at least once a month) would rather replace public transport and cycling, while seldom users would shift to other shared modes. This could provide evidence that frequent users already incorporated ride-pooling in daily travel routines while seldom users are still in the exploring phase. Overall, the study highlights the importance of accounting for behavioral heterogeneity and target-group specific analysis, when investigating new modes or setting up models to support decision-makers with scenario analyses.

Keywords: *discrete choice modelling, ride-pooling, ride-splitting, shared mobility,*

1. INTRODUCTION

Travel demand models are a powerful tool for researchers and policy makers, as they enable informed decision-making by providing insights into potential future developments under different circumstances. In the context of emerging mobility services, numerous simulation studies have demonstrated that ride-hailing services (e.g., Uber) increase vehicle kilometers traveled (e.g., Henao and Marshall, 2019, Schaller, 2021). This effect is driven not only by the substitution of public transport trips, but also by the empty vehicle movements between passengers. Ride-pooling (also called ride-splitting), in which multiple passengers share a vehicle (Shaheen and Cohen, 2019), has the potential to mitigate these negative impacts and contribute to a more efficient transport system (e.g., Tirachini and Gómez-Lobo, 2020, Zwick et al., 2021). However, the current usage of such shared services remains comparatively low (e.g., Tang et al., 2020 or Schaller 2021).

A key challenge in analyzing and forecasting the impacts of ride-pooling lies in the strong interdependency between supply and demand: increasing demand leads to higher fleet utilization, which in turn results in longer waiting and travel times for additional users. Consequently, both demand behavior and service characteristics must be adequately represented in modeling approaches.

On the one hand, many studies investigate optimal fleet management algorithms using static demand for future planning (Bischoff et al., 2017, Engelhardt et al. 2019, Zwick et al. 2021). On the other hand, researchers have investigated user preferences to better understand the use of shared on-demand services. However, most studies rely on stated preference surveys targeting potential users (e.g., Liu et al., 2019, Alonso-Gonzáles et al., 2020, Morsche et al., 2019, Krauss et al., 2022), while actual users are less often included. Yet, empirical evidence suggests that users and non-users evaluate service attributes differently (König et al., 2018). Further, current user analysis provides evidence that different user groups substitute different travel modes (e.g., Wang and

Shen, 2024). Only a limited number of studies explicitly targeted the users of ride-pooling to develop mode choice models: Schatzmann et al. (2023) surveyed users of MOIA in Hamburg and investigated the intermodal use of MOIA as a feeder for traditional public transport. Schlenther et al. (2025) implemented the model to investigate the effects of different pricing strategies. Goerguelue et al. (2025) surveyed users of the ride-pooling service fips in Mannheim. However, their research focus lied on the effect of day- and night time on mode choice. Both studies estimated MNL models without user specification.

To close this research gap, this article presents a detailed mode choice model explicitly capturing ride-pooling usage. Combining data from the German national household travel survey MiD (Nobis and Kuhnimhof, 2018) and survey data, a Cross-Nested Logit model (CNL) is estimated to account for complex substitution patterns between transport modes. Special attention is given to differences between frequent and occasional users, revealing significant heterogeneity in substitution effects.

2. METHODOLOGY

Data collection and processing

The research presented is part of the MOIA accompanying research (Kagerbauer et al. 2021). This project studied the use of ride-pooling at the example of MOIA in Hamburg, Germany. The mode choice model for the microscopic travel demand model in this study is estimated from joint revealed preference (RP) and stated preference (SP) data. The RP data is a MID subset (over 41,000 trips) from Hamburg, Germany, and its surroundings from Lower Saxony and Schleswig-Holstein, that corresponded to the project's planning area (Nobis and Kuhnimhof, 2018). The SP data (over 111,000 decisions) comes from the stated choice experiment (SCE) that was part of a comprehensive online survey conducted in 2019. A detailed overview of the empirical research activities within this project is provided by Kostorz-Weiss et al. (in press), including information on the SCE set-up.

The SCE focused primarily on the use of shared mobility services to better understand underlying trade-offs when using these modes, as the use of traditional modes is adequately captured by the household survey. The choice situations were developed using D-efficient designs with NGENE (Choice Metrics 2024).

Table 1 Encountered attributes within the SCE

Options	Cost	Travel time	Waiting time	Access time	Egress time
Walking		X			
Bicycle		X			
Car as Driver	X	X		X	X
Car as Passenger	X	X	X		
Public Transport	X	X	X	X	X
Taxi	X	X	X		
Bike-sharing	X	X		X	
Car-sharing	X	X		X	
E-Scooter-sharing	X	X		X	
Ride-pooling (MOIA)	X	X	X	X	X



Figure 1 Exemplary visualization of one SCE decision in the online survey (Kostorz-Weiss, in press)

In total, every respondent faced ten decisions (five decisions per trip purpose) including six mode alternatives (if license was available). Every decision included all shared mobility options (bike-sharing, e-scooter-sharing, ride-pooling, car-sharing free-floating (ff)) and two traditional modes (walking, car as driver/passenger, bicycle, public transport, taxi). To ensure a somewhat realistic scenario, the respondents were asked, which mode was their predominantly mode used for the trip purpose presented. This alternative was presented in all choice sets for the trip purpose as were the shared modes. The sixth alternative iterated among the other traditional modes to ensure enough trade-off situations captured with all other modes. The trip purposes were randomly drawn accounting for the respondent's employment status (e.g. work only if employed). The choice alternatives were visually displayed in trip chains to facilitate the processing of all information given (see Figure 1). Table 1 presents the attributes encountered for every mode.

The survey was distributed via the MOIA newsletter and an additional sample of 1,000 people from Hamburg was recruited via Dynata's market research panel to capture also the non-user perspective. In total, over 12,000 respondents took part in the survey, over 11,000 remained after data cleaning for further analysis. Differences between users and non-users from Hamburg, typical usage, commuting and the use of people with disabilities were described by Kostorz et al. (2021).

To replicate the decision-making process from the RP data base, travel times for alternative modes were fed into the trips reported (Nobis and Kuhnimhof, 2018). On the one hand, geocoded trips were imported into the PTV Visum model for the planning area. This allowed us to extract travel time information on cell level, comparable to the travel-demand model. The geocoded data-set was lacking short trips below 2km. Therefore, we took the non-geocoded data, extracted trips below 2km and calculated travel times for mode alternatives based on reported trip distances assuming a constant average speed. The same procedure was also applied to trips within a cell to account for different trip lengths reported. Moreover, sociodemographic information from both datasets was harmonized before matching the two datasets.

Modeling

Multinomial Logit (MNL) models are still state-of-the-art when it comes to practice-oriented travel demand models due to their computational efficiency. The Independence of Irrelevant Alternatives (IIA) assumption is a major shortcoming of MNL models, which might lead to overestimates of modal shifts. More complex nested model structures, such as Nested Logit (NL) models, account for observed and unobserved correlations between transportation alternatives. Current research, however, does not suggest a clear correlation between ride-pooling and other modes as substitutional behavior varies widely among the users (e.g., Wang and Shen, 2024, Gödde et al., 2023, Tang et al., 2020). As in NL models, one alternative can only be assigned to one nest, it is questionable if this set-up would capture all correlation effects adequately. Cross-Nested Logit (CNL) models allow to assign alternatives to more than one nest. The choice set C_n is divided into *M* nests C_{mn} and each mode alternative *i* is allocated to each nest by a_{im}. See e.g. Ben-Akiva and Bierlaire (1999) for further details on the modeling theory and underlying assumptions. The utility functions of our model follow the structure given in Equation 1. They incorporate alternative-specific constants and attributes of time and cost, with components flexibly interacting with socio-demographic, mobility-related, and trip-specific variables via shift parameters (see Table 3).

$$U_i = asc_i + \sum b_{shift_asc_i} * x_{shift_asc_i} + \sum (b_{time_i} + (b_{shift_{time_i}} * x_{shift_{tt_i}})) * x_{time_i} + \sum (b_{cost_i} + (b_{shift_{cost_i}} * x_{shift_{tt_i}})) * x_{cost_i} \quad (1)$$

3. RESULTS AND DISCUSSION

The goal of our research was to build a mode choice model, that reflects mode substitution in response to supply adjustments of ride-pooling and other potential policies adequately. We developed a CNL model as illustrated in Figure 2, in which ride-pooling is part of several nests. Thereby, the interactions between ride-pooling and other modes can be modeled specifically well. For model estimation, we used the R apollo package (Hess and Palma, 2019). As we expected the likelihood function of the CNL model to be non-monotonic due to non-linearity, we used the start search option of apollo with 20 random start value sets in order to prevent local maxima.

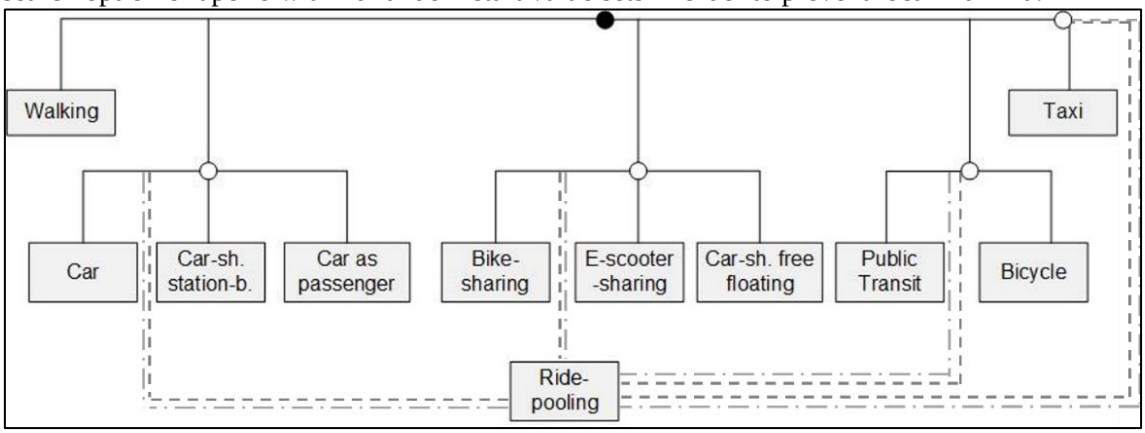


Figure 2 Final nest structure of the CNL model with different ride-pooling allocations for different user-types

Table 2 Goodness of fit measures for MNL, NL and CNL models

Measure	MNL	NL	CNL	CNL sign. only
LL	-139,357.4	-139,098.5	-139,068.5	-138,871.7
BIC	280,926.2	280,444.2	280,455.9	279,619.9

We developed MNL and NL models to benchmark the final CNL. Table 2 reports the goodness-of-fit measures for all models. Since the MNL and NL models were estimated to compare the parameters with the final CNL model and the final significant parameters of the CNL were not known beforehand, all possible parameters (including insignificant ones) were retained in the MNL and NL models. For the CNL the measures for both models (with and without insignificant parameters) are shown. As expected, the NL outperforms the MNL due to its ability to capture correlations. The CNL showed a better log likelihood than the NL, indicating that the interrelations between the nests need to be considered. In contrast, the BIC of the CNL is higher than the BIC of the NL showing that the improvement in the log likelihood comes with the cost of many additional parameters. Restricting the CNL to significant parameters reduces the BIC. The log likelihood of the CNL with significant factors only cannot be compared to the others though.
Table 3 shows the CNL parameters. The different estimated lambda values <1 confirm the correlation among alternatives within each nest - smaller values correspond to higher within-nest correlation. (Ben-Akiva and Bierlaire, 1999) The final model consists of five nests in total (see Figure 2). Ride-pooling could partly be allocated to all nests except walking. This allocation is represented by the alpha-parameters. In our case, we could even estimate different alphas for different ride-pooling user groups, representing different behavior (though not at a 5% level). Still, estimates show that frequent users of ride-pooling (at least one to three times per month) are more likely to shift toward PT or bicycle whenever conditions for ride-pooling change. The seldom users (using ride-pooling less often than once per month) shift more likely to other new modes. This might indicate that frequent users have integrated ride-pooling in their daily routines as they switch to traditional modes and seldom users see ride-pooling more as an exciting part of new mobility services rather than an actual alternative in their daily routines.

Table 3: CNL specific parameters

Parameter	Estimate	t-ratio(0)
lambda_new_mob	0.668	34.346
lambda_cars	0.695	46.302
lambda_pt_bike	0.859	55.868
lambda_taxi	0.518	5.797
alpha_freq_newmob	-0.322	-1.296
alpha_freq_cars	-0.523	-1.614
alpha_freq_pt_bike	0.000	NA
alpha_freq_taxi	-1.679	-5.056
alpha_seldom_newmob	1.597	2.717
alpha_seldom_cars	-0.616	-0.895
alpha_seldom_pt_bike	0.000	NA
alpha_seldom _taxi	0.785	1.085

Table 4 shows the estimated parameters for the utility functions, all parameters shown are significant at 5% level. The cost and time parameters are all significant with the proper sign. Users opt for ride-pooling for more often than non-users. The effect is stronger for frequent users. Further, frequent users are less cost-sensitive. The multitude of significant “shifting parameters” highlights the heterogeneity of preferences for different purposes or among different groups of people.

Table 4 Estimated parameters for CNL utility functions

		walk	bicycle	car as driver	car as passenger	public transport	taxi	bike-sharing	car-sharing (ff)	car-sharing (sb)	e-scooter-sharing	ride-pooling
	asc	0.00	-3.34	-4.69	-5.43	-5.12	-6.27	-7.87	-8.70	-7.79	-6.33	-4.99
socio-demoraphic shifts	female			-0.05	0.89	0.14		-0.74	-0.28	-0.93		
	age 0-17	2.34	2.79		3.11	2.85						
	age 18-29	-0.29	0.39		0.23	0.36	-0.46	0.88		0.23	1.31	
	age 30-49	-0.33	0.11			0.08		0.38			0.32	-0.13
	age 50-69							-0.20		-0.17	-0.39	0.24
	age 70(+)											0.84
	economical status			-0.09		-0.08	0.10					-0.07
shifts mobility-tools	license				-0.66	-0.72	-1.25				-0.74	-1.00
	public transit pass		-0.14	-0.42	0.10	1.40	0.27					
	car in household	-0.60	-0.85		-0.09	-1.08	-0.82	-0.52	-0.93		-0.29	-0.52
	CS membership								2.13			
	seldom RP user											0.92
	frequent RP user											1.87
shifts trip purpose	leisure	1.17	0.86	-0.14		0.78		1.02			0.50	1.59
	work	-0.50	1.75	-1.39	1.20	-1.54		0.76				0.30
	business	-0.78	1.28		-0.43	1.75	1.82	0.50				
	shopping	-1.06		0.48		-0.68	-1.98				-0.88	
	pick-up/drop-off	-1.76	-1.72		-1.08	-0.73		-0.44			-1.02	0.28
	b_travel time	-0.15	-0.13	-0.04	-0.04	-0.03	-0.03	-0.11	-0.04	-0.04	-0.10	-0.06
	b_waiting time				-0.04	-0.03	-0.03					-0.14
	b_access/egress time			-0.09				-0.06	-0.11	-0.25		-0.19
shifts on travel time	b_tt_pt_pass					-0.01						
	b_tt_leisure	-0.03		0.02	0.01				0.03			
	b_tt_work	-0.05	-0.04	0.02		-0.02						1.20
	b_tt_business	-0.03	-0.04		-0.02	-0.05						-0.02
	b_cost	-	-	-0.07	-	-0.17	-0.17	-0.17	-0.17	-0.17	-0.17	-0.17
shifts on cost	b_cost_econ_status_high			0.03		0.03	0.03	0.03	0.03	0.03	0.03	0.03
	b_cost_business					0.08	0.08	0.08	0.08	0.08	0.08	0.08
	b_cost_freq_user											0.04

4. CONCLUSIONS

This study investigates the mode choice behavior of current ride-pooling users. Based on RP data from the national German household travel survey and SP data from a stated choice experiment that was part of a comprehensive online survey with MOIA ride-pooling users a Cross-Nested Logit (CNL) model is estimated. The CNL allows to allocate ride-pooling in different nests to capture the potential substitutional effects on travel behavior in the future the best way possible. Furthermore, we found that the nest allocation varies for regular users (using ride-pooling at least once per month) and seldom users. Considering the distinction between frequent and seldom users within the group of ride-pooling members leads to a more realistic prediction of modal shifts when the number of users might increase in the future.

ACKNOWLEDGEMENTS

This research was funded by the MOIA GmbH.

REFERENCES

Alonso-González, M. J., Van Oort, N., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2020). Value of time and reliability for urban pooled on-demand services. Transportation Research Part C: Emerging Technologies, 115, 102621. https://doi.org/10.1016/j.trc.2020.102621

Ben-Akiva, M., & Bierlaire, M. (1999) Discrete Choice Methods and their Applications to Short Term Travel Decisions.In: Hall, R. (Publisher), Handbook of Transportation Science. Springer.

Bischoff, J., M. Maciejewski and K. Nagel (2017) City-wide shared taxis: A simulation study in Berlin, IEEE Conference on Intelligent Transportation Systems, Proceedings, ITSC

ChoiceMetrics (2024) Ngene 1.4 USER MANUAL & REFERENCE GUIDE

Engelhardt, R., F. Dandl, A. Bilali and K. Bogenberger (2019) Quantifying the Benefits of Autonomous On-Demand Ride-Pooling: A Simulation Study for Munich, Germany, paper presented at the *2019* IEEE Intelligent Transportation Systems Conference (ITSC), 2992–2997.

Goerguelue, M. E., Kostorz-Weiss, N., Voss, A.-S., Kagerbauer, M., & Vortisch, P. (2025). Effects of Daytime vs. Nighttime on Travel Mode Choice and Use Patterns: Insights from a Ride-Pooling Survey in Germany. *Applied Sciences*, *15*(14), 7774. https://doi.org/10.3390/app15147774

Gödde, J., Ruhrort, L., Allert, V., & Scheiner, J. (2023). User characteristics and spatial correlates of ride-pooling demand – Evidence from Berlin and Munich. *Journal of Transport Geography*, *109*, 103596. https://doi.org/10.1016/j.jtrangeo.2023.103596

Henao, A., Marshall, W.E., (2019) The impact of ride-hailing on vehicle miles traveled. Transportation 46, 2173–2194. doi:10. 1007/s11116-018-9923-2.

Hess, S., Palma, D., (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. Journal of choice modelling 32, 100170.

Kagerbauer, M., Kostorz, N., Wilkes, G., Dandl, F., Engelhardt, R., Glöckl, U., Fraedrich, E., & Zwick, F. (2021). *Ridepooling in der Modellierung des Gesamtverkehrs—Methodenbericht zur MOIA Begleitforschung*. Karlsruher Institut für Technologie (KIT). https://doi.org/10.5445/IR/1000141282

König, A., Bonus, T., & Grippenkoven, J. (2018). Analyzing Urban Residents' Appraisal of Ride-pooling Service Attributes with Conjoint Analysis. *Sustainability*, *10*, 3711. https://doi.org/10.3390/su10103711

Kostorz, N., Fraedrich, E., Kagerbauer, M., (2021). Usage and user characteristics—insights from moia, europe's largest ridepooling service. Sustainability 13, 958. doi:10.3390/su13020958.

Kostorz-Weiss, N., Kagerbauer, M., Vortisch, P. (in press) Investigating ride-pooling use through a mixed methods approach, presented at the 13th International Conference on Transport Survey Methods, Da Nang, Vietnam, 30.04.2025-04.04.2025

Krauss, K., Krail, M., Axhausen, K., (2022). What drives the utility of shared transport services for urban travellers? a stated preference survey in german cities. Travel Behaviour and Society 26, 206–220. doi:https://doi.org/10.1016/j.tbs.2021.09.010.

Liu, Y., Bansal, P., Daziano, R., Samaranayake, S., (2019). A framework to integrate mode choice in the design of mobility-on demand systems. Transportation Research Part C: Emerging Technologies , 648–665 URL: https://arxiv.org/abs/1805.06094, doi:10.1016/j.trc.2018.09.022.

Morsche, W., La Paix Puello, L., Geurs, K., (2019). Potential uptake of adaptive transport services: An exploration of service attributes and attitudes. Transport policy 84, 1–11. doi:10.1016/j.tranpol.2019.09.001.

Nobis, C., Kuhnimhof, T., (2018). Mobilität in Deutschland (MiD 2017) (English: Mobility in Germany). infas, DLR, IVT and infas 360.

Schatzmann, T., Zwick, F., Axhausen, K.W., (2023). Investigating the preferences for the use of urban ridepooling. hEART 2023:11th Symposium of the European Association for Research in Transportation

Schlenther, T., Volotskiy, T., Leich, G., Zwick, F., Kuehnel, N., Smirnov, J., & Nagel, K. (2025). *Ridepooling and Public Transit: How Pricing Schemes Reveal the Tradeoff between Intermodality and On-Demand Efficiency*. https://depositonce.tu-berlin.de/handle/11303/26160

Shaheen, S., & Cohen, A. (2019). Shared ride services in North America: Definitions, impacts, and the future of pooling. *Transport Reviews*, *39*(4), 427–442. https://doi.org/10.1080/01441647.2018.1497728

Schaller, B. (2021). Can sharing a ride make for less traffic? Evidence from Uber and Lyft and implications for cities. *Transport Policy*, *102*, 1–10. https://doi.org/10.1016/j.tranpol.2020.12.015

Tang, B.-J., Li, X.-Y., Yu, B., & Wei, Y.-M. (2020). How app-based ride-hailing services influence travel behavior: An empirical study from China. *International Journal of Sustainable Transportation*, *14*(7), 554–568. https://doi.org/10.1080/15568318.2019.1584932

Tirachini, A. and A. Gómez-Lobo (2020) Does ride-hailing increase or decrease vehicle kilometers traveled (VKT)? A simulation approach for Santiago de Chile, International Journal of Sustainable Transportation, **14** (3) 187–204.

Wang, Y., & Shen, Q. (2024). A latent class analysis to understand riders' adoption of on-demand mobility services as a complement to transit. *Transportation*, *51*(3), 1043–1061. https://doi.org/10.1007/s11116-022-10351-3

Zwick, F., N. Kuehnel, R. Moeckel and K. W. Axhausen (2021) Ride-pooling efficiency in large, medium-sized and small towns - Simulation assessment in the Munich metropolitan region, Procedia Computer Science, 184, 662–667.