How Transferable Are Mode Choice Models With Autonomous Vehicles? Evidence from Seven Regions

Notelaers, L.*[1], Chen, T.D.[2], Tampère, C. [3]

[1] PhD Candidate, Department of Mechanical Engineering (Mobility and Industrial Management), KU Leuven, Belgium

[2] Professory, Department of Civil and Environment Engineering (Link Lab), University of Virginia, United States of America

[3] Professor, Department of Mechanical Engineering (Mobility and Industrial Management), KU Leuven, Belgium

SHORT SUMMARY

Simulation-based transport planning often relies on transferring mode choice models between regions, particularly when studying new or emerging transport modes such as autonomous vehicles (AVs), for which observed behavioral data are scarce. The validity of such transfers, however, remains uncertain. This study empirically assesses the transferability of discrete mode choice models including AV alternatives across Flanders, Cantabria, the Netherlands, Athens, Kansas City, Seattle, and New York City. Comparable multinomial logit models were estimated from stated preference datasets, and cross-regional equivalence was evaluated using likelihood ratio tests. Results show that behavioral parameter equivalence across regions is generally rejected, with substantial heterogeneity observed for most transcontinental comparisons and among U.S. metropolitan areas. Evidence consistent with behavioral equivalence is found primarily within Europe, particularly between Belgium and the Netherlands. These findings underscore the need for caution when applying transferred mode choice models in simulation and policy analysis.

Keywords: Discrete Mode Choice, Model Transferability, Scale Heterogeneity, (Shared) Autonomous Vehicles

1. INTRODUCTION

Simulation-based transport planning relies on behavioral models for e.g. mode choice. When local data are unavailable, model parameters are often transferred across regions. Simulation studies, often applied for example mode choice coefficients originally estimated for another region, while recalibrating the alternative-specific constants (ASCs) to reproduce local mode shares (Hörl et al. 2019; Auld et al. 2016). Although pragmatic, this approach assumes stable sensitivities across contexts, an assumption rarely tested, particularly for emerging modes like autonomous vehicles (AVs).

This challenge has become particularly relevant in the context of AVs. Their impacts depend on travelers' preferences for alternative AV service forms—such as privately-owned AVs (PAVs), shared AVs (SAVs) and pooled SAVs (pSAVs)—which are typically studied using stated preference (SP) methods (Pernestål and Kristoffersson 2019). Because SP data collection is both resource- and time-intensive, many AV studies adopt the practice of transferring behavioral models from other regions.

While many studies examine AV preferences, cross-regional evidence remains limited. Most empirical work focuses exclusively on a single region. Only a few collected SP data for multiple regions. Haboucha et al. (2017) analyzed differences between Israel and North America using a nested logit kernel model that was only partially segmented between the two regions. Their models revealed distinct socio-demographic effects: gender effects were significant only in Israel. Two further studies—Etzioni et al. (2020) and Polydoropoulou et al. (2021)—collected harmonized SP data across several European countries and developed comparable AV mode choice models. Both concluded that cross-country variation in preferences could be explained by certain socio-demographic characteristics, while Polydoropoulou et al. (2021) additionally identified more consistent behavioral patterns within southern and northern European clusters. Yet, neither study statistically tested whether behavioral parameters differed or were equal across contexts, and potential scale heterogeneity was neglected.

So, existing studies suggest socio-demographic and regional differences, but rarely test parameter equivalence or account for scale heterogeneity. Clarifying the extent to behavioral parameters can be transferred across regions is essential for credible transport planning and robust decision-making. This study contributes to this discussion by empirically assessing the transferability of mode choice models estimated based on different regional SP datasets including AVs.

2. METHODOLOGY

Data

The analysis in this paper draws on SP datasets previously collected in several regions (table 1) (Notelaers et al. 2025; Cordera et al. 2022; Winter et al. 2017; Triantafillidi et al. 2023; Jia et al., n.d.; Liu et al. 2019).

Table 1: Overview datasets

Region/city	Year	Respondents (responses)	Trip characteristics	Existing modes	AV modes	Trip purposes	Reference
Flanders (Belgium)	2023	645 (4948)	Based on reported reference trip	PT, car, bike, walk	pSAV	Commute, Leisure,	(Notelaers et al. 2025)
Cantabria (Spain)	2020	296	Based on reported reference trip	Car, bike, walk	PAV, pSAV	Commute, Leisure	(Cordera et al. 2022)
Amsterdam, Utrecht, Rotterdam, The Hague (the Netherlands)	2016	732 (7704)	Exogenous trip parameters and levels for a fictious 8 km trip	PT, car, taxi, free-floating carsharing (FFCS)	SAV	Commute	(Winter et al. 2017)
Athens (Greece)	2023	164 (1476)	Exogenous trip parameters and levels	PT, car	SAV	unspecified	(Triantafillidi et al. 2023)
Seattle region (USA)		516 (3038)	Based on reported reference trip	PT, car, taxi, bike, walk	PAV, SAV, pSAV	Commute, Leisure, Shopping, Medical, Meal, Transport	(Jia et al., n.d.)
Kansas City region (USA)	2020-2021	558 (3336)	Based on reported reference trip	PT, car, taxi, bike, walk	PAV, SAV, pSAV	Commute, Leisure, Shopping, Medical, Meal, Transport	(Jia et al., n.d.)
NewYork City (USA)	2017	1507	Based on reported reference trip	PT, car, taxi, UberPool (ptaxi)	SAV, pSAV	Commute	(Liu et al. 2019)

Following Random Utility Theory (Ben-Akiva and Lerman 1985; Train 2009), the utility U_{in} that individual n associates with alternative i consists of a systematic component V_{in} and a random error term ε_i:

$$U_{in} = V_{in} + \varepsilon_i = \beta X_{in} + \varepsilon_i \quad (1)$$

For each dataset, a multinomial logit (MNL) model is estimated, where the systematic utility of alternative i for individual n is specified as:

$$V_{in} = ASC_i + \beta_{IVTT}\, IVTT_{in} + \beta_{OVTT}\, OVTT_{in} + \beta_C\, C_{in} \quad (2)$$

where $IVTT_{in}$, $OVTT_{in}$, and C_{in} represent the in-vehicle travel time (minutes), out-of-vehicle travel time (access and waiting time in minutes) and monetary cost (USD dollar), respectively. The coefficients β_{IVTT}, β_{OVTT} and β_C are generic across all (motorized[1]) modes to ensure comparability between the datasets that differ in their alternative set.

All models were estimated separately for each region using maximum-likelihood estimation implemented in Biogeme (Bierlaire 2023).

[1] Because some SP datasets also included active travel modes (e.g., walking, cycling) with distinct behavioural attributes, a separate utility specification with a specific travel time parameter (β_{TT_active}) was estimated for these alternatives to preserve comparability across datasets.

Regularity testing framework

Model transferability concerns the extent to which a travel behavior model—or components of it—can be validly applied outside the context where it was estimated. According to Sikder et al. (2013), model transferability can be assessed at different levels, from (1) behavioral theory, (2) the mathematical structure of the model, (3) the empirical specification of variables, to (4) the estimated parameters, with transferability generally decreasing at lower levels.

The most common evaluation approaches fall into two categories:

1. Statistical tests of model equivalence, which compare parameters between estimation and application contexts;
2. Predictive-accuracy measures (e.g. goodness-of-fit) to assess the predictive accuracy of transferred models in the application context.

In Random Utility Theory, discrete choice probabilities depend only on utility differences, the overall scale of utility is not identified (Train 2009). As a result, taste coefficients estimated from separate datasets may differ not only because of genuine preference differences but also due to differences in the variance of unobserved utility.

Accounting for these effects, two models are said to be regular if their underlying preference parameters are proportional (Swait and Louviere 1993), that is:

$$\beta^{(A)} = k * \beta^{(B)}, k > 0 \tag{3}$$

where k represents the relative scale parameter between the datasets *A* and *B*.

We follow a three-step log-likelihood ratio (LLR) testing procedure to evaluate parameter comparability between datasets:

1. Separate estimation:
   Estimate region-specific models independently and record their log-likelihoods LL_A and LL_B.

2. Test for parameter equality allowing for scale heterogeneity:
   Estimate a pooled model imposing common parameters across datasets while estimating a relative scale parameter k. Compare the log-likelihood of this pooled model ($LL_P^{(k)}$) with the sum of the separate model log-likelihoods using:

   $$LLR_1 = -2 * \left[LL_P^{(k)} - (LL_A + LL_B)\right] \tag{4}$$

   LLR_1 follows a chi-squared (χ^2) distribution with degrees of freedom equal to the number of parameters constrained to be equal across datasets. If $LLR_1 > \chi^2_{df,\alpha}$ at significance level α, the null hypothesis is rejected, providing statistical evidence suggesting that the two models differ up to a scaling factor.

3. Test for equal scale parameters:
   If the first test indicates that the datasets can be pooled (regularity holds), estimate a restricted pooled model that additionally constrains $k = 1$ (equal scale). The second LLR test compares this restricted model ($LL_P^{(k=1)}$) with the previous pooled model ($LL_P^{(k)}$):

$$LLR_2 = -2 * \left[LL_P^{(k=1)} - LL_P^{(k)}\right] \quad (5)$$

which also follows a chi-squared (χ^2) distribution with 1 degree of freedom (since only the scale parameter is constrained). A non-significant LLR_2 indicates that the data do not provide sufficient evidence to reject the hypothesis that both the preference parameters and the error variances are equivalent across datasets.

If both tests reject the null hypothesis, the parameters are considered statistically different across regions[2]. If only the first test fails to reject the null hypothesis, the results suggest that preferences are comparable across regions up to a scale factor, with differences arising from differences in error variance.

To enable the testing, we restrict the cross-dataset comparisons to the subset of parameters common to all datasets and not the ASCs (Swait and Louviere 1993).

While the Swait and Louviere (1993) procedure evaluates whether parameter equality across datasets can be rejected, failure to reject equality does not guarantee predictive transferability. Therefore, predictive performance is additionally assessed using quantitative indices in another region or application context. Two complementary indicators are employed for this purpose: the transfer index (TI) and the transferred goodness-of-fit measure ($\rho^2_{transfer}$).

The transfer index (Koppelman and Wilmot 1982) quantifies the improvement in log-likelihood achieved by a transferred model relative to a base model, compared to the improvement provided by a model estimated directly for the local context. It is defined as:

$$TI = \frac{LL_T - LL_B}{LL_L - LL_B} \quad (2\text{-}6)$$

Where LL_T is the log-likelihood of the transferred model applied to the target dataset, LL_L is the log-likelihood of a model estimated directly on that target dataset, and LL_B is the log-likelihood of a selected base model. A TI value of 1 indicates that the transferred model is as accurate as the locally estimated model. The index has no lower bound; negative values indicate that the transferred model performs worse than the local base model.

In this study, the transferred model applies behavioral parameters estimated in another region—specifically the coefficients for in-vehicle time, out-of-vehicle time, and cost—while keeping the ASCs fixed at their locally estimated values. Because ASCs in SP datasets are shaped by the exogenously defined experimental design rather than by observed market conditions, transferring them across regions would not be meaningful. The base model used for comparison includes only these locally calibrated ASCs and is referred to as the market share model.

The **transferred goodness-of-fit** is expressed as:

$$\rho^2_{transfer} = 1 - \frac{LL_T}{LL_B} \quad (2\text{-}7)$$

[2] This interpretation holds only when the sample size is adequate to estimate parameters with sufficient precision for the null hypothesis to be rejectable. When this condition is not met, non-rejection should not be interpreted as evidence that the parameters are transferable.

This measure expresses the proportional improvement in model fit obtained by the transferred model over the base model (Koppelman and Wilmot 1982). The both models are similarly defined as for the TI.

3. RESULTS AND DISCUSSION

Table 2 reports the estimated coefficients for each model. The model fits are comparable to those of the original studies, confirming that the general behavioral structure was successfully reproduced.

Table 2: Mode choice model estimated for each region

Parameter	Flanders	Spain	Netherlands	Athens	Kansas	Seattle	NY
β_{OVTT}	-0.057***	-0.027	-0.056***	-0.070***	-0.033***	-0.046***	-0.025
β_{C}	-0.489***	-0.233***	-0.434***	-0.230**	-0.007	0.008	-0.102***
β_{IVTT}	-0.102***	-0.021	-0.049***	-0.056***	-0.038***	-0.032***	-0.037***
β_{TT_active}	-0.109***	-0.051***			0.025	-0.054***	
ASC_{bike}	2.010***	0.281**					
ASC_{car}	1.590***		-0.474***	0.424*	3.670***	3.380***	1.240***
ASC_{FFCS}			-0.864***				
ASC_{PAV}		0.759***			2.080***	1.870***	
ASC_{PT}	0.329	0.144*		-0.039	3.670***	3.120***	0.844***
ASC_{SAV}			-0.450***		0.589***	0.404***	0.452***
ASC_{taxi}			-1.240***		2.880***	1.740***	0.479***
ASC_{walk}	2.290***	1.640***			2.220***	3.870***	
ASC_{pTAXI}							0.116*
Observations	3903	2391	5894	1476	4759	4962	8635
LL_null	-2705.353	-3848.166	-9486.027	-1621.552	-6597.375	-6878.793	-9332.440
LL_final	-1439.048	-3514.724	-8291.143	-1409.660	-3143.330	-3567.792	-7907.749
ρ^2(adj.)	0.468	0.087	0.126	0.131	0.524	0.481	0.153

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 3 reports the pairwise likelihood-ratio (LLR_1) test statistics comparing each pooled model estimated across two regions with the corresponding local models. LLR_1 values exceeding the critical χ^2 thresholds at significance level 1% ($df = 3$; $\chi^2_{3,0.01} = 11.345$) lead to rejection of H_0, indicating behavioral differences between the two models.

Table 3: Pairwise likelihood-ratio (LLR_1) test statistics

Region	Flanders	Spain	Netherlands	Athens	Kansas	Seattle	NY
Flanders		4.10	8.82	12.78	54.20	125.66	7.94
Spain			0.18	5.40	41.96	30.36	9.40
Netherlands				37.62	54.08	90.56	50.00
Athens					44.48	64.38	4.52
Kansas						16.78	29.08
Seattle							59.82
NY							

For the Flanders–Spain (4.10), Flanders-Netherlands (8.82), Spain–Netherlands (0.18) and Spain-Athens (5.40) pairs, the test statistics fall below the critical value at the 1% significance level, and the null hypothesis of equal behavioral parameters cannot be rejected. This indicates that, for these region pairs, there is no statistical evidence against behavioral equivalence in mode-choice preferences. While the results suggest broadly comparable behavior across parts of Europe, Athens is behaviorally distinct from the northern European regions, as evidenced by rejection of parameter equality in the Athens–Flanders (12.78) and Athens–Netherlands (37.62) comparisons.

Interestingly, the New York–Flanders (7.94), New York–Spain (9.40), and New York–Athens (4.52) comparisons fail to reject the null hypothesis of parameter equality, suggesting no statistical evidence against behavioral equivalence for these pairs. The absence of statistical evidence against behavioral equivalence may be related to shared characteristics such as dense urban form and relatively high reliance on multimodal transport systems.

In contrast, most other transcontinental comparisons—particularly those involving European regions and the U.S. cities of Seattle and Kansas—yield likelihood ratio statistics well above the critical value, indicating substantial behavioral heterogeneity and limited parameter transferability across these contexts. This divergence is likely driven by pronounced differences in cost sensitivity across regions, as reflected in the estimated cost coefficients reported in table 2.

Further, significant behavioral heterogeneity is also observed within the United States. The Kansas–Seattle (16.78), Seattle–NY (59.82) and Kansas-NY (29.08) comparisons also surpass the threshold, highlighting pronounced inter-city differences despite a shared national context.

Table 4 presents the second set of likelihood-ratio tests (LLR_2), conducted for which behavioral equivalence up to a scale factor could not be rejected based on the first likelihood-ratio test. These tests compare the scaled and non-scaled pooled specifications (df = 1; $\chi^2_{1,0.01} = 6.635$) to assess whether accounting for scale heterogeneity significantly improves model fit at significance level 1%.

Table 4: Pairwise likelihood-ratio (LLR_2) test statistics (for only the combination of regions for which LLR1 is accepted)

Region	Flanders	Spain	Netherlands	Athens	Kansas	Seattle	NY
Flanders		31.82	4.82				79.32
Spain			23.36	4.80			1.18
Netherlands							
Athens							25.68
Kansas							
Seattle							
NY							

Most retained region pairs provide statistical evidence of scale heterogeneity. In particular, the Flanders–Spain (31.82), Spain–Netherlands (23.36), Flanders–New York (79.32), and Athens–New York (25.68) test statistics exceed the critical value at the 1% significance level, indicating statistically significant differences in scale across these datasets. In contrast, the equality in scale between Flanders–Netherlands and Spain–Athens cannot be rejected at the 1 % level.

Taken together, these results point to a potential north–south distinction within Europe, with stronger scale differences observed for north–south comparisons such as Flanders–Spain, and weaker differences for within-cluster pairs—for example, within northern Europe (Flanders–Netherlands) or within southern/eastern Europe (Spain–Athens). This interpretation should be viewed as indicative rather than definitive.

Interestingly, the Spain–NY combination cannot be rejected at the 1% level, suggesting no statistical evidence of differences in unobserved variance once scale is accounted for. Nonetheless, this result might be questioned as two of the three coefficients in the Spain model are not significant.

Error! Reference source not found. 5 presents the transfer indices for each combination of regions. Each row represents the application context (target dataset), while each column corresponds to the context of the behavioral parameters used (transferred model). Values close to 1 indicate that the transferred specification performs nearly as well as a locally re-estimated model and signal high predictive transferability. Region pairs with relatively high predictive transferability (transfer index ≥ 0.7) are shown in bold.

Table 5: Transfer index values (relative to the base model representing observed market shares).

	Flanders	Spain	Netherlands	Athens	Kansas	Seattle	NY
Flanders	**1.000**	0.575	**0.729**	0.268	-0.183	-0.043	-1.280
Spain	0.284	**1.000**	0.687	**0.846**	-10.100	**0.840**	**0.966**
Netherlands	**0.759**	0.663	**1.000**	0.684	-0.134	-0.235	0.314
Athens	-1.179	0.435	-0.056	**1.000**	0.537	0.122	**0.783**
Kansas	-46.052	-15.353	-35.927	-15.947	**1.000**	-0.458	-3.085
Seattle	-43.777	-15.487	-35.966	-16.790	-0.066	**1.000**	-4.344
NY	-5.018	-0.084	-3.749	0.010	0.089	-0.251	**1.000**

Table 6 reports the transferred goodness-of-fit ($\rho^2_{transfer}$) values, which provide an additional indicator of predictive performance for the transferred models. As in the transfer index table, rows denote the application context, and columns denote the source of the behavioral parameters. The diagonal values represent the fit of the locally estimated models relative to the base (market share)

model and thus serve as reference points for evaluating transfer performance. The diagonal value and the highest off-diagonal value $\rho^2_{transfer}$ are highlighted to identify region pairs for which transferred models achieve comparatively stronger goodness-of-fit.

Table 6: Transferred goodness-of-fit values (relative to the base model representing observed market shares).

	Flanders	Spain	Netherlands	Athens	Kansas	Seattle	NY
Flanders	**0.060**	0.035	**0.044**	0.016	-0.011	-0.003	-0.077
Spain	0.018	**0.063**	0.043	0.053	-0.638	0.053	**0.061**
Netherlands	**0.050**	0.043	**0.066**	0.045	-0.009	-0.015	0.021
Athens	-0.149	0.055	-0.007	**0.126**	0.068	0.015	**0.099**
Kansas	-0.442	-0.147	-0.345	-0.153	**0.010**	-0.004	-0.030
Seattle	-0.609	-0.215	-0.500	-0.233	-0.001	**0.014**	-0.060
NY	-0.263	-0.004	-0.196	0.001	0.005	-0.013	**0.052**

Consistent with the LLR results, both the transfer index (table 5) and transferred goodness-of-fit (table 6) confirm that predictive transferability is highest among the European regions. In particular, the Flanders–Netherlands transfers (in both directions) show transfer index values exceeding 0.7 in both directions and transferred goodness-of-fit values approaching the diagonal reference—i.e., the locally estimated models compared with the base market-share model. These findings indicate that within these northern European regions, both models retain most of their explanatory power when applied to each other, reflecting broadly similar behavioral sensitivities to travel time, cost, and out-of-vehicle travel time variables.

In contrast, Spain and Athens exhibit more distinct behavioral patterns. For Spain, multiple transferred models (from Athens, Seattle, and New York) achieve adequate predictive performance, likely due to the insignificance of two out of three behavioral coefficients in the Spanish model. For Athens, only the transfer from New York results in a transfer index above 0.75, suggesting partial alignment in behavioral responses but otherwise clear divergence from the northern European contexts.

Across cross-continental transfers, including those between New York and Flanders or the Netherlands, the transfer index and transferred goodness-of-fit values are markedly lower—almost always negative or below 0.3—indicating a pronounced decline in predictive performance when behavioral parameters are applied across fundamentally different socio-economic and transport environments. Even within the United States, transfers among Kansas, Seattle, and New York yield negative or very low transfer index values, reinforcing the LLR finding of substantial heterogeneity among U.S. metropolitan areas.

4. CONCLUSIONS

Overall, the preference regularity tests indicate that behavioral parameter equality is rejected for most region pairs, including both transcontinental comparisons and comparisons among U.S. cities. In contrast, the Flanders–Netherlands comparison stands out as a robust case where behavioral parameter equality cannot be rejected, providing the strongest support for model transfer between these regions.

The predictive transfer assessment metrics reinforce this conclusion: while cross-continental and U.S. intercity transfers show a significant loss in predictive performance, northern European models retain most of their explanatory power when transferred within the region.

Together, the results highlight that model transfer is most successful in contexts sharing comparable travel cultures, urban form, and multimodal accessibility, whereas transfers across fundamentally different environments require local recalibration to maintain behavioral validity.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the researchers responsible for collecting and sharing the stated preference datasets analyzed in this paper.

REFERENCES

Auld, Joshua, Michael Hope, Hubert Ley, Vadim Sokolov, Bo Xu, and Kuilin Zhang. 2016. "POLARIS: Agent-Based Modeling Framework Development and Implementation for Integrated Travel Demand and Network and Operations Simulations." *Transportation Research Part C: Emerging Technologies* 64 (March): 101–16. https://doi.org/10.1016/j.trc.2015.07.017.

Ben-Akiva, Moshe, and Steven R. Lerman. 1985. *DISCRETE CHOICE ANALYSIS : Theory and Application to Travel Demand*. MIT Press.

Bierlaire, Michel. 2023. *A Short Introduction to Biogeme (Report TRANSP-OR 230620)*.

Cordera, Rubén, Esther González-González, Soledad Nogués, Julián Arellana, and José L. Moura. 2022. "Modal Choice for the Driverless City: Scenario Simulation Based on a Stated Preference Survey." *Journal of Advanced Transportation* 2022. https://doi.org/10.1155/2022/1108272.

Etzioni, Shelly, Jamil Hamadneh, Arnór B. Elvarsson, et al. 2020. "Modeling Cross-National Differences in Automated Vehicle Acceptance." *Sustainability* 12 (22): 9765–9765. https://doi.org/10.3390/su12229765.

Haboucha, Chana J., Robert Ishaq, and Yoram Shiftan. 2017. "User Preferences Regarding Autonomous Vehicles." *Transportation Research Part C: Emerging Technologies* 78 (May): 37–49. https://doi.org/10.1016/j.trc.2017.01.010.

Hörl, Sebastian, Milos Balac, and Kay W. Axhausen. 2019. "Dynamic Demand Estimation for an AMoD System in Paris." June, 260–66. https://doi.org/10.1109/IVS.2019.8814051.

Jia, Wenjian, Donna T. Chen, Wenwen Zhang, and Kaidi Wang. n.d. *Local Context Matters: Examining Mode Choice Preferences in the Autonomous Vehicle Era in Two U.S. Metropolitan Regions*.

Koppelman, Frank S., and Chester G. Wilmot. 1982. "TRANSFERABILITY ANALYSIS OF DISAGGREGATE CHOICE MODELS." *Transportation Research Record*, no. 895: HS-035 371. https://trid.trb.org/View/189355.

Liu, Yang, Prateek Bansal, Ricardo Daziano, and Samitha Samaranayake. 2019. “A Framework to Integrate Mode Choice in the Design of Mobility-on-Demand Systems.” *Transportation Research Part C: Emerging Technologies* 105 (August): 648–65. https://doi.org/10.1016/j.trc.2018.09.022.

Notelaers, Lotte, Mohammad Ali Arman, Pieter Vansteenwegen, and Chris M. J. Tampère. 2025. “Preferences for Stop-Based Pooled Automated Vehicles in Mode Choice Models across Trip Purposes in Flanders, Belgium.” *European Transport Research Review* 17 (1): 64–64. https://doi.org/10.1186/s12544-025-00756-3.

Pernestål, Anna, and Ida Kristoffersson. 2019. “Effects of Driverless Vehicles: Comparing Sim Ulations to Get a Broader Picture.” *European Journal of Transport and Infrastructure Research* 19 (1). https://doi.org/10.18757/EJTIR.2019.19.1.4079.

Polydoropoulou, Amalia, Ioannis Tsouros, Nikolas Thomopoulos, et al. 2021. “Who Is Willing to Share Their AV? Insights about Gender Differences among Seven Countries.” *Sustainability* 13 (9): 4769–4769. https://doi.org/10.3390/su13094769.

Sikder, Sujan, Abdul Rawoof Pinjari, Sivaramakrishnan Srinivasan, and Roosbeh Nowrouzian. 2013. “Spatial Transferability of Travel Forecasting Models: A Review and Synthesis.” *International Journal of Advances in Engineering Sciences and Applied Mathematics* 5 (2–3): 104–28. https://doi.org/10.1007/s12572-013-0090-6.

Swait, Joffre, and Jordan Louviere. 1993. “The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models.” *Journal of Marketing Research* 30 (3): 305–305. https://doi.org/10.2307/3172883.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press. https://doi.org/10.1017/CBO9780511805271.

Triantafillidi, Eirini, Panagiotis G. Tzouras, Ioanna Spyropoulou, and Konstantinos Kepaptsoglou. 2023. “Identification of Contributory Factors That Affect the Willingness to Use Shared Autonomous Vehicles.” *Future Transportation* 3 (3): 970–85. https://doi.org/10.3390/futuretransp3030053.

Winter, Konstanze, Oded Cats, Karel Martens, and Bart van Arem. 2017. “A Stated-Choice Experiment on Mode Choice in an Era of Free-Floating Carsharing and Shared Autonomous Vehicles.” *TRB 96th Annual Meeting Compendium of Papers*.