Agent-Based Modeling of Heterogeneous Travel Preferences with MATSim

Christian Rakow[1*], Kai Nagel[1]

[1]Technische Universität Berlin, Chair of Transport Systems Planning and Transport Telematics, Straße des 17. Juni 135, 10623 Berlin, Germany

Abstract

This work presents a framework for integrating taste heterogeneity into agent-based transport simulations, specifically within MATSim. By incorporating principles from Random Utility Maximization theory, we address a critical limitation in current approaches that often fail to account for preference heterogeneity among travelers. Our methodology combines a two-stage model estimation process with a novel implementation that fixes individual preferences and handles random utility terms consistently. Validation against Berlin travel survey data demonstrates significant improvements in predictive accuracy, with dramatically improved log-likelihood scores and more realistic distance distributions for different travel modes. Most notably, our enhanced model produces more realistic policy elasticities than the baseline model. These results highlight the importance of accounting for taste variations to improve both the behavioral realism and policy sensitivity of agent-based transport simulations.

Key words: MATSim; mode choice; taste variations; taste heterogeneity.

Introduction

Agent-based transport simulations aim to model the behavior of individual travelers making decisions about their daily mobility. A fundamental challenge in developing these models is ensuring that simulated agents make realistic choices that reflect actual human decision-making patterns. This requires addressing several aspects: implementing theoretically sound behavioral models, ensuring consistency with empirical observations, and accurately capturing traveler responses to system changes.

While Random Utility Maximization (RUM) provides a robust framework for modeling discrete choices (McFadden, 1981; Train, 2009), implementing this within agent-based contexts presents unique challenges. Current approaches often fail to fully account for heterogeneity in preferences, which leads to unrealistic behavioral patterns and exaggerated policy responses, as we demonstrate in this paper.

This research explores the integration of established choice modeling theory with agent-based simulations, specifically focusing on fixing individual preferences and random utility terms according to RUM assumptions. Despite being a promising direction outlined in previous research (Nagel et al., 2016), there has been limited systematic exploration of this approach in the context of MATSim or other agent-based simulation frameworks (Hörl, 2021; Hörl et al., 2018).

Methodology

Our approach consists of three main components: model validation, choice model estimation, and simulation implementation.

[*] Corresponding author. Address: Technische Universität Berlin, Chair of Transport Systems Planning and Transport Telematics, Straße des 17. Juni 135, 10623 Berlin, Germany. Email: rakow@vsp.tu-berlin.de

For validation, we established a framework that evaluates how well an agent-based model predicts observed decision-making behavior. By creating a reference population that mirrors characteristics of actual people, we enable direct comparison between simulated choices and observed behavior. The log-likelihood (LL) metric serves as our primary evaluation measure, as it considers the probabilistic nature of choices and penalizes confident but incorrect predictions more heavily than uncertain ones.

Our choice model estimation follows a two-stage process. First, a trip-based Multinomial Logit (MNL) model estimates parameters that do not vary across decision-makers. Then, a plan-based Mixed Logit (MXL) model accounts for taste heterogeneity, calculating utility for complete daily plans rather than individual trips. The utility function incorporates mode-specific constants, travel time, costs (with income interactions), and additional mode-specific parameters. For the MXL model, parameters can be person-specific and distributed according to a random distribution, capturing unobserved factors that influence an individual's inherent preferences for different modes.

To integrate the estimated models into the simulation, person-specific parameters are sampled once for each agent at simulation start. Pseudo-random errors are generated on-the-fly when needed, while ensuring that identical choice situations always yield the same values (Horni et al., 2016). A seed calculation algorithm ensures consistency across iterations while maintaining the required variance properties. This implementation ensures that agent choices align with RUM theory while maintaining computational efficiency.

Results

We evaluated our approach by comparing a baseline model (without taste heterogeneity) to our enhanced model using travel diary survey data (Hubrich et al., 2019) and the OpenBerlin[2] model. The log-likelihood improved dramatically, indicating that our model captures underlying choice behavior substantially better than the baseline.

The enhanced model significantly improved the match of distance distributions compared to the baseline. In the baseline model, the attractiveness of walking decreased too steeply with distance, resulting in too few walking trips between 1 and 4 km. Similarly, the change in bike mode share with increasing distance was too steep. Both distributions improved by incorporating taste heterogeneity. The distribution of mode users per day also improved, with better alignment between the simulation and reference data regarding the percentage of the population using each mode at least once daily.

Perhaps most importantly, the enhanced model demonstrates more plausible elasticities compared to the baseline. For example, a 2 km/h speed increase for bicycles leads to a 31.4% increase in bike mode share in the enhanced model, which aligns closely to empirical findings of previous studies (Li et al., 2018).

Our research demonstrates how mode choice models with taste heterogeneity can be effectively estimated and integrated into agent-based transport simulations in a manner consistent with RUM theory. The significant improvements across all evaluated metrics highlight the importance of incorporating random utility and taste variations as an essential component of agent-based choice models.

[2] https://github.com/matsim-scenarios/matsim-berlin

References

Horni, A., Nagel, K., & Axhausen, K. W. (2016). Destination Innovation. In *The Multi-Agent Transport Simulation MATSim: The Multi-Agent Transport Simulation MATSim*. https://depositonce.tu-berlin.de/handle/11303/8615

Hubrich, S., Ließke, F., Wittwer, R., Wittig, S., & Gerike, R. (2019). *Methodenbericht zum Forschungsprojekt. „Mobilität in Städten – SrV 2018``*. https://orlis.difu.de/handle/difu/259350

Hörl, S. (2021). Integrating Discrete Choice Models with MATSim Scoring. *Procedia Computer Science*, *184*, 704–711. https://doi.org/10.1016/j.procs.2021.03.088

Hörl, S., Balac, M., & Axhausen, K. W. (2018). A First Look at Bridging Discrete Choice Modeling and Agent-Based Microsimulation in MATSim. *Procedia Computer Science*, *130*, 900–907. https://doi.org/10.1016/j.procs.2018.04.087

Li, H., Ding, H., Ren, G., & Xu, C. (2018). Effects of the London Cycle Superhighways on the Usage of the London Cycle Hire. *Transportation Research Part A: Policy and Practice*, *111*, 304–315. https://doi.org/10.1016/j.tra.2018.03.020

McFadden, D. L. (1981). Econometric Models of Probabilistic Choice. In *Structural Analysis of Discrete Data with Econometric Applications: Structural Analysis of Discrete Data with Econometric Applications*. MIT Press.

Nagel, K., Axhausen, K. W., Kickhöfer, B., & Horni, A. (2016). Research Avenues. In *The Multi-Agent Transport Simulation MATSim: The Multi-Agent Transport Simulation MATSim* (pp. 533–542). Ubiquity Press. https://www.research-collection.ethz.ch/handle/20.500.11850/119104

Train, K. E. (2009). *Discrete Choice Methods With Simulation* (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511805271