Unveiling Transport Dynamics in Developing Countries, a Data-Driven Agent-Based Modeling Approach

Kathleen Salazar-Serna[1,2][0000−0003−3824−7044], Lorena Cadavid[2][0000−0002−6025−5940], Carlos J. Franco[2][0000−0002−7750−857X], and Kathleen M. Carley[3][0000−0002−6356−0238]

[1] Pontificia Universidad Javeriana, Calle 18 N°118-250, Colombia
kathleen.salazar@javerianacali.edu.co
[2] Universidad Nacional de Colombia, Medellin Av.80 N°65-223, Colombia
{kgsalaza,dlcadavi,cjfranco}@unal.edu.co
[3] Carnegie Mellon University, Pittsburgh, 4665 Forbes Ave, 15213, USA
kathleen.carley@cs.cmu.edu

Abstract. This paper proposes a comprehensive framework for modeling transport mode choices with a focus on developing countries where motorcycles are highly prevalent. Using an agent-based modeling approach, the framework captures the decision-making processes of urban commuters who choose between cars, motorcycles, and public transit. A comprehensive survey conducted in a Colombian city provides the sociocultural insights necessary to parameterize the model. The framework is applied to develop a model that simulates transport dynamics and evaluates policies designed to improve mobility. The results demonstrate the model's ability to represent real-world patterns and forecast impacts of policies, like fare-free public transit and increased frequency, which reduced emissions and accidents by promoting modal shift. This data-driven approach provides a robust tool for analyzing transport issues and policy interventions, offering insights into how to promote sustainable mobility transitions.

Keywords: Survey data · Urban mobility · Policy modeling.

1 Introduction

Agent-based modeling (ABM) is a widely used simulation method for policy analysis in fields like public health [1], resource management [2], and transportation [3]. It allows for the representation of diverse agents and their interactions in a defined space to study complex behaviors that result from these interactions. ABM helps evaluate policies by simulating individual and collective decisions that impact and are affected by the environment [4]. It has been applied to study transport planning and travel behavior, particularly in urban areas [5], but most research has focused on developed countries where transport dynamics

differ significantly from those of developing regions.

Transportation systems in developing countries, commonly lacking in quality public transit services, heavily rely on motorcycles. This affordable and flexible vehicle is primarily used by the young male working class [9]. The increasing popularity of motorcycles has raised concerns about accident rates, traffic congestion, and pollution [18]. Two-wheelers account for more than 60% of the national fleets in developing countries such as Colombia and Vietnam. Around 83% of the world's population resides in developing countries with low disposable incomes per capita [10], which explains the significant proportion of motorcycles concentrated in these regions. Despite these figures, little research has been done on policies to address these issues in developing countries [6].

This paper proposes a framework for modeling transport mode choices using ABM to analyze policy impacts. Mode choices are influenced by context and sociocultural factors, requiring a realistic representation of cultural behavior in the models [13]. We first analyze the mode choice drivers to identify the factors that need to be included in the simulation to represent urban commuters. In parallel, a conceptual model was developed to define the decision rules of the agents. The model was implemented in NetLogo 6.4.0 and parameterized using data from a city in Colombia that is representative of the transport dynamics in developing countries. Statistical and census data were used for model initialization. In addition, a survey was designed and conducted to gather sociocultural information from transport users, informing those variables that are not usually measured in household or origin-destination surveys. The model was then calibrated and validated using historical city data to ensure accuracy in simulating the transport system.

In the subsequent sections, the framework for modeling transport mode choices is described, along with the results of applying this framework to a case study city in Colombia.

2 Background

Traditional modeling of transport mode choice has relied on discrete choice models and more recently, classification models [19]. While these methods can predict choices, they struggle to capture the nonlinear relationships and constant interactions in complex transportation systems, making it challenging to understand travel behavior and the impacts of changes on factors influencing mode choice [22]. Previous attempts using agent-based models have fallen short in adequately representing human behavior, with a bigger focus on traffic simulations. The MOSH framework [17] stands out as an exception, proposing to integrate human behavior theories with ABM to represent individual travelers in a university setting in a developed country. This framework draws inspiration from the CONSUMAT approach [20], a cognitive-inspired model that combines

socio-psychological theories to portray consumer behavior. It provides different elements to represent decision makers using social-oriented heuristics, social network structures for agents' interactions and cognitive processes in human decision making. Building on these foundations, we developed an extended framework for modeling transport systems that allows for feedback between agents' decisions and system reactions, enabling a more comprehensive evaluation of transport policies' impacts. Our approach focuses on laying down the groundwork for modeling transport policies in developing countries.

3 Proposed framework to model transport mode choices

Urban transport planning is a complex process involving various actors with different institutional arrangements, belief systems, and implications for the environment. Policymakers face challenges in developing effective strategies to address transport issues in these coupled infrastructure systems [11]. ABM serves as a valuable tool to model transport dynamics and travel behavior of urban commuters considering spatial characteristics, population heterogeneity, and sociocultural aspects [12]. Fig. 1 illustrates a framework for analyzing individual travel behavior, modal shift phenomena, and promoting sustainable transport options through public policies.

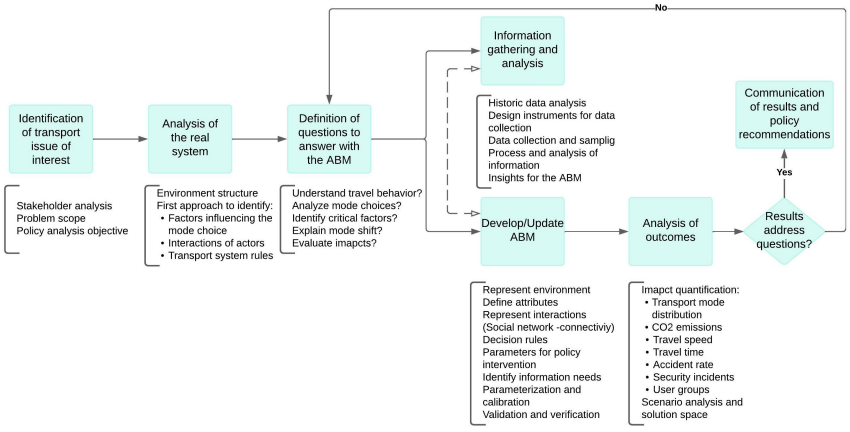

Fig. 1. Framework to model transport mode choices.

The initial step involves recognizing a problematic situation or a policy need by stakeholders. It is then essential to comprehend the territory's characteristics, such as its economic, cultural, political, technological, demographic, infrastructure, and normative aspects. This understanding helps identify actors, their relationships, and the factors influencing transport mode choices. The questions posed in the initial step should then be translated into simulation objectives.

These objectives will guide the model development process based on the system's level of comprehension. If stakeholders have identified the factors influencing specific behaviors and are solely interested in a policy's impact, the model's focus can be on impact evaluation, with details on the sensitivity of decision-making factors possibly being excluded. Developing models involves iterative progress through several stages: conceptualization, design and implementation, verification, validation, and application [7]. Information gathering and analysis are commonly considered to be part of these stages; however, given the empirical nature of this simulation, a significant amount of information needs to be gathered, and thus it is presented as a separate component in the framework. However, there is an intrinsic relationship and continuous feedback between model development and information gathering components.

Based on the literature review and the conceptual model, we identified variables to be considered in the simulation. These variables are categorized into attributes related to agents, transport modes, and the city. Fig. 2 outlines the key variables associated with each element. The attributes of transport modes relate to the technical aspects of cars, motorcycles, and public transit. A market analysis with various vehicle brands was carried out to calculate initial relative scores for variables like acquisition and operating cost; efficiency and emissions, are normally distributed using the average values. At the city level, understanding the initial distribution of the fleet and spatial layout of neighborhoods based on socioeconomic status and congested areas is crucial. Factors like population density and traffic regulations including speed limits are also considered. As for agent attributes, they primarily consist of sociodemographic data sourced from census records. Nonetheless, certain parameters, like perceptions of transport modes or criteria for their selection, are not typically included in these surveys. This required the creation and implementation of a survey to collect such sociocultural aspects essential for representing travel mode decisions in the model. The subsequent section discusses this survey.

Once the model has been parameterized and validated for the city of interest, the evaluated policies are compared to the base-case scenario using different indicators to measure changes in the system. Finally, results are communicated with recommendations for policy implementation.

3.1 Survey Design and Implementation

In creating the survey, our main objectives were to study the factors influencing transport mode decision-making and to collect data for the parameterization of the simulation model. The questionnaire, derived from previous research on sociodemographics [15], travel habits [16], and user attitudes [17], was customized for our study by adapting and supplementing existing questions. Additional details, including the survey dictionary link, can be found in the supplementary material.

City	Transport modes	Users
Fleet distribution, road accident rates, neighborhoods' socioeconomic situation and spatial distribution, speed limit, average distances between neighborhoods, congestion spots.	Average possible/real speed, CO2 emissions, fuel efficiency, acquisition cost, operating cost. Public system capacity, wait time, average travel time.	Age, gender, socio-economic level, neighborhood, average time and distance of travel, weights for transport attributes, risk perception, comfort perception, satisfaction threshold, uncertainty threshold.

Fig. 2. Variables considered in the simulation identified by city, transport modes and agents.

The survey was structured into various sections to gather information that enables the characterization of transport users, identifying their travel patterns, and stated preferences. The initial section gathers sociodemographic details anonymously. The subsequent part delves into the current commuting mode, encompassing usage duration, ownership, satisfaction level, and reasons for selection. In terms of travel behavior, inquiries cover average time and distance, common origins and destinations, and sharing practices. A following section explores the significance users attribute to transport features such as cost and comfort. The final section inquires about perceptions of various modes, including perceived crime risk, opinions on costs, road safety, comfort, travel time, and pollution.

The focus of this research is on the analysis of public policies to improve mobility in developing countries. For this purpose, we selected Cali, a Colombian city that like many other cities in the Global South, exhibits high accident rates, a large number of motorcycles per capita, and traffic congestion issues. Participants were chosen through simple random sampling, with the sample size determined using Cochran's formula at a 95% confidence level and a $\pm 5\%$ margin of error, assuming the highest estimated variance. A sum of 970 volunteers participated in the study in 2023. The survey data was analyzed and the factors that affect mode selection were validated using multinomial logit models (MNL). The results of the survey are described in Section 4.

3.2 Transport Users Decision Module

The conceptual model that guides agents in choosing their commuting transport mode is based on the CONSUMAT approach [20]. This consumer behavior framework integrates various behavioral theories to elucidate product or service selection. As per this approach, decisions are influenced by the internal mental state (current transport mode satisfaction) and behavioral control factors, including uncertainty tolerance and desired satisfaction levels (see Fig. 3). Agents initiate with a mode of transport determined by their demographic char-

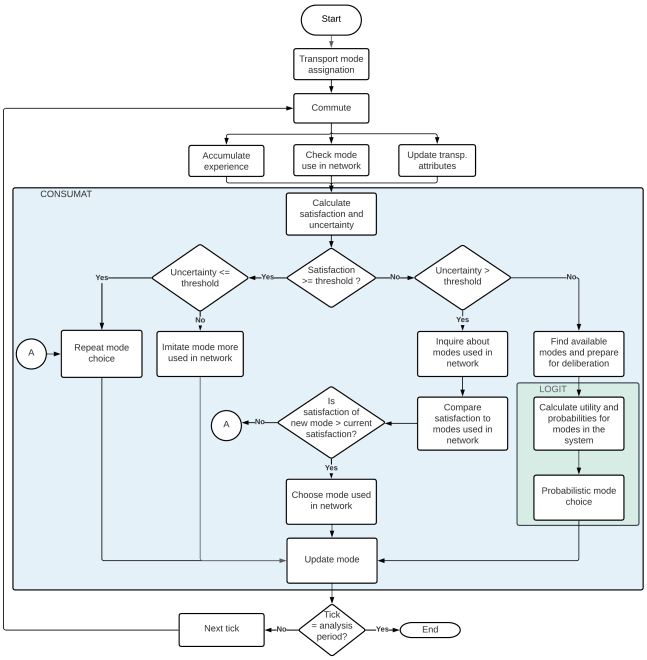

Fig. 3. Agents' decision-making flow.

acteristics. They accumulate information from their journey and the status of the system, evaluating the satisfaction obtained with each transport attribute considered in the mode decision-making process. Depending on the importance agents assign to these attributes, an overall weighted satisfaction is calculated. In addition, agents accumulate experience using the transport mode over time steps. The more knowledge they gain, the less uncertainty they face about future satisfaction. The uncertainty is calculated as a combination of accumulated personal and collective experience (from peers in the social network), weighted by a factor that determines the balance between uncertainty avoidance and collectivism. For the Colombian case, this was parameterized based on the cultural dimensions studied by Hofstede et al. [21].

Based on the comparison between satisfaction and uncertainty levels with their personal thresholds, users apply one of the following strategies to make decisions: repetition, deliberation, imitation, or inquiry. The first two strategies involve individual reasoning processes, whereas the latter two involve social engagement, where agents interact with their social network.

4 Survey Results Informing the Model

Fig. 4 showcases results from the survey. The findings reveal that 40% of users favor cars, with 23.9% opting for public transportation, 16.6% for motorcycles, and 19.4% for other modes of transport. Most participants are from the middle class (51%), while 25% represents the high class, and 24% claim to be in the low-income bracket. These trends reflect the demographic distribution of the 2018 census.

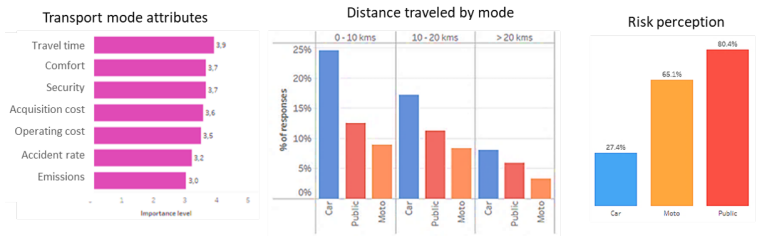

Fig. 4. Summary of survey results.

Using MNL models on the survey results, we found that the transport attributes identified in the literature review and focus groups significantly influence the choice of mode. Age, sex, and socioeconomic status were included as individual agent attributes. Acquisition cost, operating cost, comfort, road safety, personal security, travel time, and pollution were significant factors in mode choice and were used to calculate agents' satisfaction function.

Statistical analysis revealed that the importance placed on different factors varies among users of different socioeconomic classes (see supplementary material); while travel time stands out as the most crucial aspect for all users, costs rank as the second most important factor for individuals in the low-income bracket, with comfort taking precedence for those in the high and middle classes. This observation prompted us to differentiate the weights used for calculating agents' satisfaction in the simulation based on the socioeconomic groups. Additional parameters based on survey questions, such as the perceived risk of encountering security issues while using different modes of transport and the comfort levels associated with each alternative, were integrated into the model. The first was introduced into the model as the likelihood of experiencing crime incidents, while the latter represents a comfort rating influenced by probabilistic weather conditions and traffic congestion levels over time. Questions about travel behavior such as travel time and distance were used to calibrate the model, confirming that agents exhibit similar averages per simulated trip according to their chosen transport mode and residential area. Finally, satisfaction levels per mode as gauged by the survey, are being utilized as normally distributed thresholds for agents. Overall, the comprehensive utilization of survey findings in developing

the agent-based model reflects a data-driven approach, bolstering the credibility and relevance of our simulation in comprehending and forecasting transportation dynamics within urban settings.

5 Running the Simulation for Cali City

After implementing the conceptual model, collecting historical data, and calculating parameters from survey results, the model was initialized and calibrated to reflect the system's status in 2017. By comparing simulated data with historical transport user patterns, we confirmed that the simulation accurately represents real system behavior. Furthermore, we compared the model's forecast for future years with the projected Bass diffusion curve of motorcycles in the city, yielding very similar results. Refer to [8] for further information on the validation process.

Experiments simulating a 10-year period were conducted 80 times - the point at which the coefficient of variation stabilizes - to establish a baseline scenario for 2022. Without intervention, private commuting was predicted to rise by 86% by 2032, while public transit users would decrease 50%. This shift would worsen pollution, congestion, and accidents. Subsequently, two policies were introduced to promote public transit use: fare-free and improved capacity/frequency for BRT. Implementing policies led to a 5% reduction in CO2 emissions and a 11% decrease in accidents annually due to a 5% rise in public transit users. Both policies yielded comparable outcomes, effectively deterring the shift of commuters from public transit to private vehicles or motorcycles (See Fig. 5). This can be elucidated by increased satisfaction, particularly with the cost and time factors important to low- and middle-income groups.

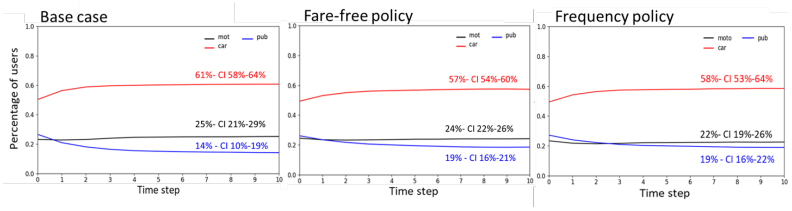

Fig. 5. Summary of survey results.

6 Conclusion and future work

In this research, we presented a comprehensive framework for modeling urban commuters' transport-mode choices, focusing on developing countries where motorcycles are prevalent and cause high road accidents, traffic congestion, and air

pollution. The framework was successfully applied to develop a simulation model for a Colombian city, which accurately represents the typical transport dynamics in such contexts.

The model was parameterized using data from a tailored survey, which revealed key mode choice factors. These findings informed the model's satisfaction functions, reflecting the city's sociocultural behavior. Differentiated perception parameters enhanced the simulation's realism. The base case scenario demonstrated the model's accuracy in reflecting real system behaviors, making it useful for predicting policy impacts. Implemented policies positively influenced decisions, reducing accidents, emissions, and congestion.

The framework can be applied to analyze transport dynamics in contexts beyond developing countries, while the simulation model serves as a testbed for territories with urban settings similar to Cali. Future work will include additional survey questions to analyze the dynamics between genders, facilitating the development of gender-based policies. Moreover, incorporating additional characteristics of the environment – such as road capacity or speed limits – can further expand the model's ability to explore a wider range of policies.

Acknowledgments. This study was funded by MinCiencias Colombia through the Technology and Innovation Program (CTeI). GRANT: BPIN2022000100068.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Supplementary Materials Complementary information can be found at: `https://lc.cx/4ZjSCl`

References

1. Blakely, T., et al. Association of simulated COVID-19 policy responses for social restrictions and lockdowns with health-adjusted life-years and costs in Victoria, Australia. In JAMA Health Forum. American Medical Association **2**(7), e211749–e211749 (2021)
2. Sun, D., Zhang, M., Jung, D. Policy evaluation of economic–environmental tradeoffs in regulating industrial water use: An agent-based model. Journal of Environmental Management, **346**, 118988 (2023)
3. Delcea, C., Chirita, N. Exploring the applications of agent-based modeling in transportation. Applied Sciences, **13**(17), 9815 (2023)
4. Ahanchian, M., Gregg, J. S., Tattini, J., Karlsson, K. B. Analyzing effects of transport policies on travelers' rational behaviour for modal shift in Denmark. Case studies on transport policy, **7**(4), 849-861 (2019) doi: 10.1109/ITSC.2018.8569232
5. B. H. Park, H. M. Abdul Aziz, A. Morton, R. Stewart, High performance Data Driven Agent-based Modeling Framework for Simulation of Commute Mode Choices in Metropolitan Area, 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, USA, 3779-3784 (2018)

6. Cadavid, L., Salazar-Serna, K. Mapping the Research Landscape for the Motorcycle Market Policies: Sustainability as a Trend—A Systematic Literature Review. Sustainability, **1**(19), 10813 (2021)
7. Ligmann-Zielinska, A., Siebers, P. O., Magliocca, et al. 'One size does not fit all': A roadmap of purpose-driven mixed-method pathways for sensitivity analysis of agent-based models. Journal of Artificial Societies and Social Simulation, **23**(1) (2020)
8. Salazar-Serna, K., Cadavid, L., Franco, C. Analyzing Transport Policies in Developing Countries with ABM. In Proceedings of Annual Modeling and Simulation Conference (ANNSIM) 2024. Preprint ArXiv:2404.19745 (2024)
9. PAHO 2017. "The Pan American Health Organization. La motocicleta en el tránsito en las Américas". `https://www.paho.org/es/documentos/hoja-informativa-motocicleta-transito-americas`, last accessed 2024/05/15
10. UN 2022. "Now 8 billion and counting". `https://unctad.org/data-visualization/now-8-billion-and-counting-where-worlds-population-has-grown-most-and-why`, last accessed 2024/05/15
11. Anderies, J. M., Janssen, M. A., Schlager, E. Institutions and the performance of coupled infrastructure systems. International Journal of the Commons, **10**(2), 495-516 (2016)
12. Kagho, G. O., Balac, M., Axhausen, K. W. Agent-based models in transport planning: Current state, issues, and expectations. Procedia Computer Science, 170, 726-732 (2020)
13. Salazar-Serna, K., Ng, L. H. X., Carley, K., Cadavid, L., Franco, C. J. Simulating the social influence in transport mode choices. In 2023 Winter Simulation Conference (WSC) 3154–3165 (2023) IEEE.
14. Chiu, B. Y., Guerra, E. What predicts motorcycle ownership, mode choice, and use for utilitarian travel? A literature review. Transport reviews, **43**(6), 1165-1189 (2023)
15. Ashalatha, R., Manju, V. S., Zacharia, A. B. Mode choice behavior of commuters in Thiruvananthapuram city. Journal of Transportation Engineering, **139**(5), 494-502 (2013)
16. Madhuwanthi, R. A. M., Marasinghe, A., RPC, J., Dharmawansa, A. D., Nomura, S. Factors influencing to travel behavior on transport mode choice-A case of Colombo metropolitan area in Sri Lanka. International Journal of Affective Engineering, **15**(2), 63-72 (2016)
17. Faboya, O. T., Ryan, B., Figueredo, G. P., and Siebers, P.-O.: Using agent-based modelling for investigating modal shift: The case of university travel. Computers and Industrial Engineering **139**, (2020). `https://doi.org/10.1016/j.cie.2019.106077`
18. Suatmadi, A. Y., Creutzig, F., Otto, I. M. On-demand motorcycle taxis improve mobility, not sustainability. Case Studies on Transport Policy, **7**(2), 218-229 (2019)
19. Liu, J., Xiao, L. Socioeconomic differences in effect size: predicting commuting mode choice of migrants and locals using a light gradient boosting approach. Transportation, **51**(1), 1-24 (2024)
20. Jager, W., and Janssen, M.:An updated conceptual framework for integrated modeling of human decision making: The Consumat II. In: Workshop Complexity in the Real World @ECCS, pp. 1–18, Brussels,(2012). https://doi.org/10.1016/j.cie.2019.106077
21. Hofstede, G. J., 2023. Hofstede Insights: Country Comparison Tool. `http://word.tips.net/Pages/T000273_Numbering_Equations.html`, last accessed 2023/05/15
22. Bazzan, A. and Klügl, F.: A review on agent-based technology for traffic and transportation The Knowledge Engineering Review. **29**(3), 375-403 (2014)