Causal Inference on E-Scooter Adoption: A Study on Sociodemographic and Behavioral Factors

MJ Javadinasr
mjavad2@uic.edu
University of Illinois at Chicago

Pranav Veldurthi
pveldu2@uic.edu
University of Illinois at Chicago

Milind Kumra
mkurma2@uic.edu
University of Illinois at Chicago

ABSTRACT

This progress report presents the initial phase of a study on the causal determinants of e-scooter adoption, focusing on sociodemographic and behavioral factors. Utilizing data from a survey of 2,126 respondents in Chicago, we have conducted thorough data cleaning and descriptive analysis in preparation for Structural Equation Modeling (SEM). Our literature review identifies a significant gap in understanding how various factors interact to influence e-scooter adoption. With the dataset ready for SEM analysis, this study is poised to contribute substantial knowledge to urban mobility discourse and provide strategic guidance for urban planners and policymakers in fostering sustainable transportation alternatives.

1 INTRODUCTION

In recent years, the rise of shared micromobility options, including electric bikes and scooters, has significantly changed the way people move around in cities. These modes have swiftly grown in popularity, as evidenced by the sharp increase in ridership numbers in the United States, which jumped from 321,000 trips in 2010 to an impressive 136 million trips by 2019 [1]. The launch of electric scooter-sharing systems (ESSs) in 2017 played a key role in this expansion, with e-scooters quickly becoming the favorite among micromobility choices and helping to double the number of micromobility trips between 2017 and 2018 [1]. The advantages of electric scooter-sharing systems (ESSs) are varied: they operate without emissions, save energy, and provide an effective approach to reducing traffic in cities [2]. Additionally, they integrate smoothly with public transportation, addressing the challenge of covering the distance to or from stations and stops, known as the "first-mile/last-mile" issue. This integration of ESSs into city infrastructures holds promise for improving urban connectivity and supporting the environmental sustainability of city landscapes. Given the rapid uptake of e-scooters, a comprehensive understanding of what drives individuals to use e-scooters is of paramount importance. It is not just the adoption that is critical but also the continuance of their use that ensures sustainable development in urban mobility. The challenge for micromobility providers, particularly those in the competitive private sector, is not just to attract new customers but to retain existing users by encouraging their long-term continuance usage. Research directed at understanding continuance intention and loyalty towards different mobility modes is therefore critical for the growth of micromobility companies and for sustainable urban development. This project seeks to delve into the underpinnings of this mobility change, with a focus on the factors influencing the sustained adoption of shared e-scooters. To achieve this, Structural Equation Modeling (SEM) is employed, a sophisticated statistical technique capable of modeling complex relationships between observed and latent variables. SEM stands out for its ability to infer causal connections, rather than mere associations, thereby offering a comprehensive understanding of the determinants that drive individual behaviors toward e-scooter usage. At the core of our analytical approach lies the Technology Acceptance Model (TAM), a theoretical model that has been extensively validated as a predictor of technology adoption [3]. TAM suggests that the attitudinal constructs of perceived usefulness and ease of use are critical to an individual's decision to embrace new technologies. Our research extends this model by incorporating a suite of sociodemographic variables, and other factors. The goal is to construct a nuanced causal framework that elucidates the adoption of e-scooters, a topic that, despite the burgeoning number of e-scooter users, remains relatively unexplored. For this study, a comprehensive dataset of 2126 valid e-scooter user responses were gathered to test our hypothesized model. By intertwining SEM's causal inference capabilities with the TAM's solid theoretical foundation, this study not only strives to shed light on the initial adoption but also on the enduring use of e-scooters. Such insights are crucial for micromobility companies and urban policymakers who attempt to attract new users while retaining current ones. Moreover, understanding these factors can help in promoting sustainable urban development and achieving long-term goals of reducing traffic congestion and lowering environmental impact.

2 PROBLEM DESCRIPTION

This study employs the Technology Acceptance Model (TAM) as a foundational framework to investigate the factors influencing the adoption and continued use of e-scooters. The TAM suggests that user behavior can be predicted by two primary beliefs: Perceived Ease of Use (PEU) and Perceived Usefulness (PU). PEU is the degree to which a person believes that using a particular system would be free from effort. PU is the degree to which a person believes that using the system will enhance his or her job performance. In our context, the system refers to e-scooters as a mode of transportation. Let $E = e_1, e_2$ represent the core constructs of the TAM where e_1 is PEU and e_2 is PU. We expand this model by including a set of sociodemographic variables , such as age, income, education level, and other relevant demographic factors. These variables are hypothesized to influence the adoption and usage of e-scooters either directly or by moderating the effects of PEU and PU. The outcome variable, Y, indicates e-scooter adoption and usage, defined as before. The expanded TAM model can be formally described as follows:

$$Y = f(E, X, \epsilon)$$

where f represents the functional relationship that we seek to estimate, and is the error term. The objective of the study is to assess not only the direct impact of PEU and PU on e-scooter adoption but also to understand how sociodemographic factors influence

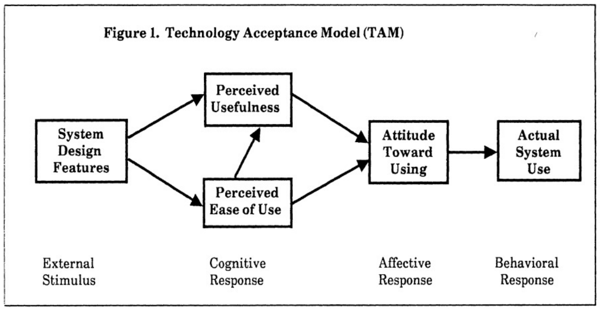

Figure 1: TAM flow diagram

these core TAM constructs and the final adoption behavior. The resulting model will look to quantify: The direct effects of PEU and PU on e-scooter usage (Y). The direct effects of sociodemographic variables (X) on Y. The interaction effects between PEU, PU, and sociodemographic variables on Y.

3 RELATED WORKS

The Technology Acceptance Model (TAM) has served as a pivotal framework in various domains, offering a proven construct for understanding the acceptance of new technologies. Our study is situated within this broader scholarly dialogue, seeking to extend the application of TAM to the domain of e-scooter adoption, a rapidly emerging mode of urban transport.

3.1 Insights from TAM and User Adoption Patterns

Expanding on the socioeconomic dimensions, Félix et al.,2023 [4] assessed shared e-scooters, weighing their benefits against potential externalities. While their work provides a comprehensive outlook on the benefits and drawbacks, it stops short of integrating the TAM framework to unpack the interrelation between user perceptions and e-scooter adoption.

Further, research by Kopplin et al.,2021[7] delved into consumer acceptance, a fundamental component of TAM, for shared e-scooters in urban and short-distance travel. Although their findings enriched our understanding of user acceptance, they did not fully leverage the predictive capabilities of TAM in explaining the nuances of e-scooter adoption.

3.2 E-Scooters as a Component of the Urban Mobility Ecosystem

The integration of e-scooters into urban mobility is a topic that has garnered attention, with studies such as Zuniga-Garcia et al. (2022)[9] examining their role as a last-mile solution in transit systems. Such research underscores the practical significance of e-scooters in enhancing urban connectivity. Our research echoes these findings while aiming to distill the causal mechanisms, as guided by TAM, that influence the sustained adoption of e-scooters for last-mile connectivity.

3.3 Advancing TAM in Urban Mobility and E-Scooter Research

This study aims to enrich the discourse in the field of urban mobility and technology acceptance by mapping the TAM framework as shown in fig.1 onto the specific context of e-scooter usage. The integration of sociodemographic variables and attitudes, alongside traditional TAM constructs, will offer a more nuanced understanding of the factors influencing e-scooter adoption. By doing so, we aspire to provide actionable insights for urban planners and policymakers to craft interventions that bolster the adoption and continual use of e-scooters, aligning with sustainability goals and evolving urban transportation landscapes.

3.4 Integrating Causal Analysis and Feature Selection with the Technology Acceptance Model

This research enhances the Technology Acceptance Model (TAM) by integrating causal discovery algorithms to clarify the directional impact among core constructs such as 'Perceived Ease of Use,' 'Perceived Usefulness,' and 'Actual Usage.' This methodological approach is key in substantiating the causal connections postulated by TAM, specifically in the adoption context of e-scooters [8].

In parallel, feature selection methods are employed to distill the analysis to essential variables, a step that is indispensable in data-rich environments. This helps in reducing redundant data, which could otherwise lead to model overfitting, thereby bolstering the generalizability and clarity of the model's outcomes. By concentrating on the most consequential predictors, these techniques bolster the model's ability to predict outcomes, providing stakeholders with actionable insights to enhance e-scooter adoption.

Furthermore, the exploration of comprehensive datasets with myriad potential explanatory variables necessitates the use of advanced feature selection methods. These methods are implemented to sharpen the focus on vital predictors, thereby enhancing the model's predictive accuracy and intelligibility. Such refined computational approaches are indispensable for delineating clear strategies to support the uptake of e-scooters.[6]

4 DATA DESCRIPTION

The dataset used in this work is a data survey of 2166 respondents who utilized Divvy mobility services in Chicago during the year 2020. The survey consists of 80 questions ranging from demographic information such as race, gender, and income to behavioral questions such as " How strongly do you agree or disagree with the following statement? I am excited by the possibilities offered by new technologies". The emphasis of this survey was placed on individuals who utilize shared e-scooters, enabling a concentrated insight into the behaviors, preferences, and attitudes of this subset of urban travelers. The CSV dataset of the survey contains 277 variables as some of the questions' responses are divided into multiple variables, and additionally, the dataset also contains some metadata from the user.

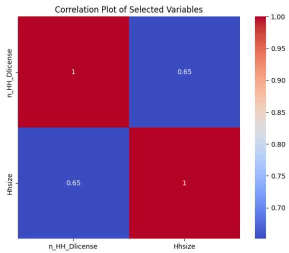

Figure 2: Correlation plot between the size of household and no people with driving licenses

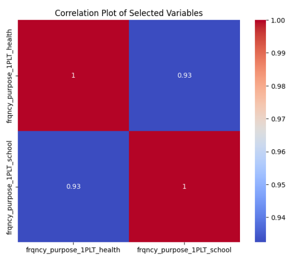

Figure 3: Correlation plot between frequency of use of E-scooter during first pilot for school and health appointments

5 PROGRESS SO FAR

5.1 Data cleaning

The Dataset contains many irrelevant variables, distinct categorical variables, and redundant variables, which need to be addressed through thorough data cleaning. The data cleaning is done through the following steps: the dataset contains empty variables such as last name, which is empty throughout the dataset, and this column is dropped. The dataset contains certain variables that contain the same value throughout the dataset, such as progress, which is 100 throughout, or user language, which is EN throughout the 2166 responses, these columns are subsequently dropped. There are certain variables that we assume are irrelevant for analysis, such as the IP address of the user, the start and end time of each question's response as the data already contains the time taken for the entire response, the coordinates of the user's location as all the responses are from the Chicago area and subsequently, these columns are dropped. The variables with categorical data, such as the frequency of using CTA buses, are converted from categorical never - often to numerical 1-7. The variables, such as race, which is stored in 7 different binary variables for each race, are converted into a single numerical variable. Through the above-mentioned data cleaning techniques, the number of variables is reduced from 277 to 145 variables.

5.2 Data Analysis

We computed correlation matrices of all variables to observe any distinct correlations in the data. It can observed in fig.2, the distinct correlations between size of the household and no people with driving licenses in the household, and in fig.3 the correlations between the frequency of use of E-scooter during the first pilot for school and frequency of use of E-scooter during first pilot for school for health appointments. However, it cannot be used to infer directly from correlations, as correlation doesn't equate to causation. Correlations also enabled identifying and removing redundant variables representing the same data as responses from certain questions were stored in different representations in multiple variables.

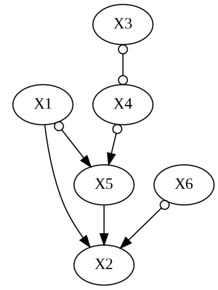

Figure 4: Causal model of specific demographic factors where X6 is the intention to continue using e-scooters

5.3 Causal Discovery

Causal discovery is used to learn structural causal model, in this work we use Fast Causal Inference(FCI) algorithm for causal discovery. Causal discovery is essential for causal estimation by identifying causal relations. FCI works by starting with a fully connected graph and iteratively eliminating edges connecting conditional variables. It is particularly useful in detecting latent confounders. In some cases it only outputs an equivalence class unlike a complete causal structure. It is highly inefficient in high dimensional data, as the dimension of the data increases the number of conditional independence tests increase exponentially. It also struggles with collinearity in data as it causes singularity in correlation matrix ,so we have to remove collinearities prior to FCI[5]. We performed FCI on specific demographic variables and the intention to continue using E-scooters in the future. In fig. 4 X1, X2, X3, X4, X5 and X6 represent age category, education category, race category, car ownership category, size of household and intention to continue using E-scooters in the future respectively.

6 STEPS TO BE DONE

- Expand TAM: Integrate demographic factors with TAM to understand their influence on e-scooter adoption.
- Apply SEM: Use Structural Equation Modeling on survey data to test the modified TAM and analyze causal relationships.
- Policy Recommendations: Translate findings into strategies for increasing e-scooter adoption, targeting interventions by demographic insights.

REFERENCES

[1] [n. d.]. Shared Micromobility in the U.S.: 2019. NationalAssociationofCityTransportationOfficials[WWWDocument],2019. URLhttps://nacto.org/shared-micromobility-2019/. Accessed: 2021-10-10.

[2] Yijie Cao and Dan Shen. 2019. Contribution of shared bikes to carbon dioxide emission reduction and the economy in Beijing. *Sustainable Cities and Society* 51 (2019), 101749.

[3] Fred Davis. 1985. A Technology Acceptance Model for Empirically Testing New End-User Information Systems. (01 1985).
[4] Rosa Félix, Mauricio Orozco-Fontalvo, and Filipe Moura. 2023. Socio-economic assessment of shared e-scooters: do the benefits overcome the externalities? *Transportation research part D: transport and environment* 118 (2023), 103714.
[5] Clark Glymour, Kun Zhang, and Peter Spirtes. 2019. Review of causal discovery methods based on graphical models. *Frontiers in genetics* 10 (2019), 524.
[6] Isabelle Guyon and André Elisseeff. 2003. An introduction to variable and feature selection. *Journal of machine learning research* 3, Mar (2003), 1157–1182.
[7] Cristopher Siegfried Kopplin, Benedikt Martin Brand, and Yannick Reichenberger. 2021. Consumer acceptance of shared e-scooters for urban and short-distance mobility. *Transportation research part D: transport and environment* 91 (2021), 102680.
[8] Jonas Peters, Dominik Janzing, and Bernhard Schölkopf. 2017. *Elements of causal inference: foundations and learning algorithms.* The MIT Press.
[9] Natalia Zuniga-Garcia, Mauricio Tec, James G Scott, and Randy B Machemehl. 2022. Evaluation of e-scooters as transit last-mile solution. *Transportation research part C: emerging technologies* 139 (2022), 103660.