Title

Causal Inference, Transportation, and Travel Demand: A Conceptual Review with Applications using Observational and Experimental Data

Author
Obeid, Hassan

Publication Date
2022

Peer reviewed|Thesis/dissertation

Causal Inference, Transportation, and Travel Demand: A Conceptual Review with Applications using Observational and Experimental Data

By

Hassan Obeid

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Civil and Environmental Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Joan Walker, Chair

Professor Scott Moura

Professor Michael Anderson

Professor Daniel Rodriguez

Fall 2022

Abstract

Causal Inference, Transportation, and Travel Demand: A Conceptual Review with Applications using Observational and Experimental Data

By

Hassan Obeid

Doctor of Philosophy in Engineering – Civil and Environmental Engineering

University of California, Berkeley

Professor Joan Walker, Chair

The field of transportation and travel behavior research has long been interested in answering causal questions. Take the recent COVID-19 pandemic as an example: the transportation sector in particular was one of the most heavily impacted sectors, and transportation researchers found themselves with a plethora of questions to answer regarding the current and future impacts of the pandemic on the transportation system. Yet, out of roughly 250 pandemic related transportation research papers we reviewed, only about 10 explicitly reference the causal inference literature and are explicit about their causal designs, despite the fact that a significantly larger portion of those papers are trying to answer causal queries. This disconnect is the motivation behind this dissertation.

It is important to acknowledge that transportation researchers have indeed used and contributed to several areas of the causal inference literature. Notable contributions include addressing self-selection bias between residential choice and travel behavior, addressing omitted variable bias through integrated choice and latent variable models (ICLVs), and addressing endogeneity between multiple travel outcomes using joint discrete choice models.

Despite those contributions, many advances in the causal inference literature have yet to enter the field of travel demand modeling. There are many reasons for this disconnect, some of which stem from long-rooted beliefs and practices within the travel behavior and demand modeling literature on model development, selection, and validation. For starters, there is a tendency for discrete choice modelers in transportation to assume their models allow for causal interpretations because they rely on a behavioral theory of human decision making, like variations of random utility theory. While this criterion typically entails endogeneity checks between the outcome and the regressors, it overlooks important nuances about the data generating process and sources of variation in the exogenous variables. This is further exacerbated by the heavy reliance in the field of transportation demand modeling on goodness-of-fit statistics and statistical tests of significance when finalizing modeling specifications, making them prone to the issue of “bad controls”. For instance, adding post-treatment or mediator variables that are exogenous to the outcome but endogenous to the treatment variable of interest, undermines the causal interpretation of the model coefficients, even if adding those variables results in improved predictive model performance and goodness-of-fit statistics. Finally, causal identification strategies rarely appear in the transportation literature. In the causal inference literature, the analyst states the assumptions before drawing any causal conclusions from the model by explicitly specifying the source of variation in the treatment of interest. Such assumptions are referred to as identification strategies: they are assumptions about the data generating process that, only if true, allow the modeler to interpret the model parameters as causal. Those strategies are rarely explicitly stated in transportation demand models, even though those models are often used to evaluate the impact of policy interventions.

To address this gap, this dissertation comprises two parts: 1) a conceptual part, and 2) an applied part.

The conceptual part consists of Chapter 1, where I elaborate in detail on the disconnect described above and point to specific examples from the transportation literature where such misconceptions about causality are most evident. Next, I provide a review of the key concepts in causal inference, and give an overview of the main causal identification strategies used in the empirical sections of this dissertation. I also give an overview of causal graphical models, an alternative causal inference framework to the more well-known potential outcomes (PO) framework which has been gaining popularity in recent years. I focus the overview on parts where I believe this framework, and Directed Acyclic Graphs (DAGs) more specifically, are most useful to transportation researchers.

The applied part consists of Chapters 2, 3 and 4, where I apply some of the causal identification strategies presented in Chapter 1 to answer three different empirical causal research questions in transportation, two of which rely on observational data, while one involves randomized experiments. Each chapter results in domain-specific empirical contributions in its respective area. The common theme across all three chapters is the explicit focus on estimating causal parameters, the clear statement of the causal identification strategies used to estimate those parameters, and the transparency about the source of variation in the treatment. This is in contrast to the common practice in travel behavior modeling where models are specified and estimated without explicitly stating the assumptions under which the estimated parameters can have causal interpretations, and can often lead to erroneous conclusions and misapplications of those models.

In Chapter 2, I quantify the causal effect of telecommuting on travel frequency and distance traveled. This question is motivated by the unprecedented rise of telecommuting in the past two years and its likely persistence in a post-pandemic world. To answer this question, I collected and used five waves of original U.S.-based survey data combined with passive smartphone Point-of-Interest (POI) data collected over the course of the pandemic. I quantify the effects of changes in the frequency of telecommuting on the total number of daily and weekly trips that a telecommuter makes, as well as their total daily and weekly distance traveled. Crucially, I overcome important limitations of related work in the literature by controlling for unobserved individual confounders, a limitation of existing research that predominantly relies on cross-sectional observational data. I do this by leveraging the longitudinal aspect of the data and implementing causal quasi-experimental designs like first-differences and two-way fixed effects regressions with individual, time, and individual by time controls. I find strong evidence that telecommuting causes the generation of new non-commute trips. Specifically, I find that individuals make an average of 1 additional non-commute trip on telecommuting days relative to commute days. This trip is on average shorter than the two-way commute trips, meaning that the net effect of telecommuting on total distance traveled is negative. Importantly, I find that this relationship persists at the weekly level, where I estimate that 1 additional day of telecommuting per week causes an increase of about 1 additional non-commute trip. This means that the additional travel on telecommuting days is the result of a newly generated trip, not a trip that has been shifted from other days of the week. The results are more robust than those in the existing literature and suggest that the trip reduction effects of telecommuting could be overestimated if telecommuting-induced new trip generation is not properly accounted for.

In Chapter 3, I quantify the causal effect of vaccines on reversing pandemic induced mobility trends. The question is motivated by the extensive literature analyzing and forecasting the stickiness of pandemic-induced behavioral and mobility trends in a post-pandemic world. Using the same data as the one used in Chapter 2, I show how the pandemic behavioral response of people in the U.S. was heterogeneous: individuals with low levels of concern about being infected with COVID-19 engaged in riskier behaviors than those with higher levels of concern, including traveling more, attending large gatherings, and using public transportation. Then, using difference-in-differences designs, I show how getting vaccinated affected those behavioral differences. Specifically, I find that getting vaccinated caused an increase in mobility, with vaccinated individuals increasing their number of weekly trips by 4.8 trips per week after getting vaccinated, compared to 1.8 trips for the unvaccinated during the same time period. The difference-in-differences estimate is 3 trips per week, or 170% of the increase in trips for the unvaccinated. The collective results provide important insights on human and travel behavior during the pandemic impact and recovery periods, and how vaccines affect those behaviors.

The third and final empirical question is answered in Chapter 4, and is on learning and optimizing user behavior at plug-in electric vehicle (PEV) charging stations. Research on consumer behavior in the PEV charging context is limited, data is lacking, and consumer preferences, especially at workplace charging stations, remain poorly understood. I address this gap by designing and implementing randomized pricing experiments, the gold standard of causal inference, and quantify key behavioral quantities like the willingness and required incentives to delay charging, and the relationship between plug-in duration and hourly prices. I then propose a novel optimization framework that incorporates those learned behaviors into the optimization objective and significantly improves the operational efficiency of PEV charging stations. My analysis shows that incorporating behavioral theory in the optimization framework results in significantly lower operational costs (up to 17%) and higher net revenues (up to 50%) for the charging station operator compared to the uncontrolled baseline, without sacrificing the user experience.

Aside from the conceptual contributions (Chapter 1) and domain-specific empirical contributions (Chapters 2, 3, and 4), the dissertation also makes data contributions through the collection of two original datasets that will soon be made publicly available to the transportation research community. I have participated in those data collection efforts playing a primary role in designing the surveys, defining the sample, and implementing randomized pricing experiments. The two datasets are:

- An extensive dataset on a panel of U.S. participants with a comprehensive set of behavioral questions over the course of the pandemic. The data comprises passively collected POI data, as well as five waves of surveys which included questions regarding respondents' employment status, travel and telecommuting behavior, vaccination status, demographic characteristics, and ideological beliefs. An anonymized and de-identified version of this dataset, along with aggregated mobility metrics from the POI data will be made publicly available for transportation researchers.

- An experimental dataset of user behavior at PEV charging stations that includes exogenous variation in the prices and the resulting real life behavioral responses of study participants. This dataset fills a critical gap in the human behavior literature in PEV charging research, where most pricing data are observational and do not allow for price perturbations. The data will be made publicly available for academic researchers along with an accompanying data manuscript that describes it.

Collectively, this dissertation emphasizes the importance of being rigorous about the design and assumptions under which researchers and modelers can draw causal conclusions from models, an important task in transportation research. After all, causal inferences are only valid if their underlying assumptions are correct, yet those assumptions are rarely discussed and formally stated in transportation models. This dissertation fills this gap.

To Mom and Dad

Mohamad and Samer

Sienna

Table of Contents

Acknowledgement

Like anything of value in life, this dissertation has been a team effort. Many people have had an essential role in not only shaping the technical work in this dissertation, but also shaping me as a human and getting me to this point.

First and foremost, I want to thank the members of my dissertation committee, whose support throughout the last three years made this work possible. Finding technical support in a place like UC Berkeley is perhaps the easy part, but I feel incredibly lucky to have had the chance to work with such amazing and wholesome human beings. To my advisor, Joan Walker: your impact on me as a scholar and a human is more than I can put in words. Thank you for your knowledge, diligence, guidance, and support. But more than all of that, thank you for being a great mentor over the last three years, and for creating a safe space that I could come to with any problem. To Scott Moura: thank you for being an example in compassion, for your wise advice, for perfectly balancing giving me growth and leadership opportunities without being overwhelming, and for always cheering me on through the ups and downs of this journey: I always left your office feeling better than how I entered it, and I truly appreciated that. To Michael Anderson: I learned something new in every conversation I had with you, and your course planted the seeds for the dissertation topic and astronomically grew my interest in causal inference. And to Daniel Rodriguez, who set an example in class and professionalism that I aspire to emulate. You all have inspired me more than you will ever know, and for that I am eternally grateful.

I am also grateful to the many friends I made throughout this journey. To Mohamed Amine Bouzaghrane, whose friendship is one of the most valuable gifts my journey at Cal has given me. Thank you for pushing me to grow every day, and sharing the ups and downs of this journey with me. I am grateful to have had you through it all and sharing your journey with you as well. To Catu, Jessica, Jason and Andrew: thank you for making the first masters years as fun as they were. And to my current and old friends in McL 116 and beyond, whose camaraderie made the weight of the PhD lighter: Juan for his light spirit, Alex for the formula 1 conversations, Anu, Qianhua, Aqshems, Max. Thank you for the memories, discussions, and fun times. And to Timothy Brathwaite, my friend and colleague with whom early conversations sparked an interest in me to explore the topic of causality. I appreciate you all so much.

To my friends and family in Lebanon, whose friendship and support shaped the person I am today. To George, thank you for being the best friend one can ask for, and for always being by my side. To Bassel, Roy, Samira, Karim, and Hashem: thank you for the memories and the continuous support. My cousins, aunts and uncles: you are the best family one can ask for, and I am so grateful to have your unconditional support.

And most of all, to my parents, Zeina and Ahmad. I will never be able to thank you enough, not in a million years. Everything I am, and everything I accomplish in life, is thanks to you. My dad, you have sacrificed so much to give us the opportunities you never had, and this success is yours more than it is mine. My mom, your strength, unconditional love, and support throughout all my years made me the person I am today. To my brothers, Mohammad and Samer, you guys are my favorite people in the world, and you will never truly know your impact on me. I love you all more than anything. And to my grandma, Layla, who has been my biggest fan since the day I was born, your impact on me is more than you will ever know.

Finally, to the love of my life, Sienna. Thank you for going through this journey with me, for accepting me, for lifting me up when I needed it the most, and for forcing me to celebrate when I did not feel like celebrating. You were the light I looked to in the darkest moments, and the first person I called to celebrate the highs. None of this would have been possible without you by my side.

Chapter 1: Introduction

1. Motivation

A common application of travel demand models is to support policy evaluations. Specifically, we often use behavioral models to evaluate the impact of external interventions on travel outcomes: How will adding a bike lane affect the mode choice of commuters in a given neighborhood? How does telecommuting affect a person's travel frequency? These evaluations are causal by definition: we are interested in how the system reacts to external interventions. Yet, when transportation models are developed or used for the purpose of policy evaluation, a formal discussion of causality and the assumptions under which the model coefficients can have causal interpretations are often missing. Furthermore, even when causal concepts are accounted for, the process is done implicitly without a clear and formal framework. This disconnect is the main topic and recurring theme of this dissertation.

Before I talk about the disconnect, it is important to acknowledge that transportation researchers have indeed used and contributed to several aspects of the causal inference literature. Specifically, the issue of self-selection bias has received considerable attention in the literature on residential choice and its impacts on travel behavior (1, 2). Integrated choice and latent variable models (ICLVs), which were first proposed in the mid-1980s in the econometrics literature (3, 4) but later popularized and extended by transportation researchers in the early 2000s (5) can also be thought of as attempting to address omitted variable bias. In addition, joint discrete choice models, like ones that estimate the joint probability of residential location as well as travel behavior outcomes (like mode choice, car ownership, etc.) were also proposed to address the endogeneity and self-selection between travel behavior and residential choice (6, 7). In fact, endogeneity specifically has explicitly received significant attention in the literature on travel demand modeling, and continues to be a hot topic of discussion (8–11), with researchers using methods such as control functions using instrumental variables, use of proxies, integration of latent variables, and the use of multiple indicators to address this issue.

My point, however, goes beyond those discussions. I am also not claiming that an expert, rigorous, and careful travel demand modeler or researcher cannot produce valid causal inferences without explicitly following the terminology and frameworks from the causal inference literature. Instead, I am arguing that the field of travel demand modeling can gain a lot by being explicit about incorporating advances and techniques from causal inference, and approaching the topic of causality more holistically, rather than focusing on specific issues that fall under the wide umbrella of causality individually. This is especially true when the parameters of the model being developed are interpreted as causal effects and used to forecast the impact of policies (12). In addition, I believe that there are misconceptions and misunderstandings about causality that are commonly observed in the field of travel demand modeling and that lead at times to erroneous model applications and interpretation of results. I elaborate more below (see (13) for another discussion of reasons behind this disconnect).

Let me start by talking about some misconceptions I see when the topic of causality is discussed in the travel demand modeling literature. I believe that the word "causal" in this literature is used loosely at best, and misused at worst. Take the following statement from a recently published discussion paper in the Journal of Choice Modeling as an example:

- "The theory-driven discrete choice models [...] are also understood as causal models. As a result of their causal structure, they are generally deemed suitable to make predictions – e.g. for policy interventions – beyond the support of the observed data, or for counterfactual analysis." (14)

By theory driven, the authors mean some variation of utility theory, like utility maximization or regret minimization. While being theory driven in this way is extremely important, especially as it allows the modeler to estimate useful parameters like willingness-to-pay and quantify trade-offs that individuals make for certain features, theory-driven models in this sense do not guarantee causality like the statement asserts. For starters, what does it mean that the models are "understood as causal?" Generally speaking, and in most cases, causal models can estimate the causal effect of one variable at a time (a notable exception is when the data is coming from a randomized experiment designed specifically for multiple treatments and treatment effect estimation, which is rarely the case in transportation research). This is because the modeling decisions needed to estimate the causal effect of one variable often conflict with those needed to estimate the effect of another. However, when discrete choice models are used for policy evaluation, multiple of the model parameters are interpreted causally at the same time, in-line with the assumption that discrete choice models are "generally understood as causal." Statements like these emphasize a general misunderstanding among transportation demand modelers of what causality means. This misconception also manifests in another common theme: transportation demand modelers very often make the case that discrete choice models are advantageous over machine learning models when evaluating the impact of policies because of their reliance on sound behavioral theory. I believe this is a strawman argument. Indeed, just like discrete choice models, there is nothing prohibiting machine learning models from estimating causal quantities if they are applied within a valid causal identification framework. This is evident in fact that, over the last decade, there has been a resurgence in machine learning research to leverage the power and flexibility of machine learning models in causal applications, which led to many important contributions to the causal inference literature (15–21). What machine learning researchers have leveraged is the high predictive accuracy of their models and ability to learn complex relationships[1]. This higher accuracy can be (and has been) used

---

[1] A common critique to machine learning models' high predictive accuracy is that it comes at the extent of interpretability. But when the goal is to remove bad variation, interpretability is not important as long as the models are estimated properly. Another critique is usually the tendency of machine learning models to overfit the data, but this is easily mitigated in a multitude of ways, from cross-validation, to regularization. A perfect example is the double/de-biased machine learning algorithm.

to isolate bad variation in the treatment and the outcome by building predictive models of each as a function of the controls, partialling out the portion of the variation that can be predicted by those controls, and using the remaining "good variation" to estimate the causal effects.

Additionally, in discussions of endogeneity, transportation researchers have not distinguished between confounders and mediators, something that the causal inference literature explicitly addresses. These two types of variables have different implications on endogeneity and their inclusion or exclusion from a model affects the causal interpretation of the model parameters in different ways. Take the following statements from a review paper assessing methods to correct for endogeneity in discrete choice models, published in Transportation Research Part A: Policy and Practice.

- "Endogeneity occurs when some explanatory variables are correlated with the error term of an econometric model due to, among other things, omitted attributes, measurement or specification errors, simultaneous determination or self-selection. [...] For example, in a mode choice model between public and private urban transportation, it is likely that the perceived level of discomfort (due to e.g. crowding) will grow with travel time. Since the perceived level of discomfort is relevant for the decision maker, but very difficult to measure by the researcher, its omission will cause endogeneity. This omission will make the model useless for assessing policies that enhance comfort, such as providing air conditioning, or redesigning the vehicle's layout. Besides, *this omission will result in poor forecasting capabilities and in an overestimation of the value of travel time savings, which will be confounded with the improvements in comfort* [emphasis added]." (10)

While it is true that omitting the perceived level of discomfort will make the model incapable of assessing policies that enhance comfort, such omission does *not* result in "poor forecasting capabilities and overestimation of the value of travel time savings." This is because the level of discomfort in this hypothetical example is a mediator through which travel time affects the utility of public transportation, not a confounder of travel time. In fact, controlling for the perceived level of discomfort by including it in the model will *underestimate* the *total* causal effect of travel time savings. This can be proven mathematically. First, under this hypothetical example, travel time, which I denote by T, is assumed to affect the perceived level of discomfort C positively, and both T and C affect the utility of public urban transportation, U. These relationships can be summarized with the following data generating process:

$$T = \epsilon_1 \qquad \text{(eq-I.1)}$$
$$C = \alpha_0 + \beta_0 * T + \epsilon_2 \qquad \text{(eq-I.2)}$$
$$U = \alpha_1 + \beta_1 * T + \gamma_1 * C + \epsilon_3 \qquad \text{(eq-I.3)}$$

Substituting (eq-I.2) in (eq-I.3), we get:

$$U = \alpha_1 + \beta_1 * T + \gamma_1 * (\alpha_0 + \beta_0 * T + \epsilon_2) + \epsilon_3$$

Rearranging the terms, we get:
$U = \alpha_1 + \gamma_1 * \alpha_0 + (\beta_1 + \gamma_1 * \beta_0) * T + \gamma_1 * \epsilon_2 + \epsilon_3$

Now, if we define:
$\alpha_1 + \gamma_1 * \alpha_0 = \alpha$,
$(\beta_1 + \gamma_1 * \beta_0) = \beta$, and
$\gamma_1 * \epsilon_2 + \epsilon_3 = \epsilon$, we get:
$$U = \alpha + \beta * T + \epsilon \quad \text{(eq-I.4)}$$

In eq-I.4 and under this data generating process, β, which is a function of β_1, γ_1, and β_0, can be interpreted as the total causal effect of travel time on the utility of public transportation, which includes the direct effect of travel time, as well as its effect through the perceived level of discomfort. Note also that the error term ϵ in eq-I.4, which omits C, is in fact independent of T, since it does not contain any terms that are dependent on T. Consequently, no endogeneity issues exist by omitting the perceived level of discomfort, C, when estimating the total *causal* effect of travel time, T, on mode choice. In fact, if we were to control for C by specifying the model as shown in eq-I.3 (assuming we can actually measure C), we will actually be estimating β_1, which is not the total causal effect of travel time. β_1 can be interpreted as the partial effect of travel time while holding the level of discomfort constant, which, in my opinion, is less meaningful: if a policy intervention were to reduce travel time, the true effect of such intervention will be β (multiplied by the change in travel time due to the intervention), not β_1. The easiest way to estimate β would be to simply exclude level of discomfort from the model and estimate the reduced model specification in eq-I.4. All of this to say, just because C is correlated with T does not mean that excluding it will cause endogeneity or omitted variable bias *if* the goal is to estimate the total causal effect of travel time. For that to happen, C needs to cause T (and U), making it a confounder of T, and not vice versa. In eq-I.1 and eq-I.2, if T was a function of C instead of the opposite, then C would be a confounder of T and omitting it from the specification will indeed cause endogeneity, and the coefficient of T estimated in a reduced form model like in eq-I.4 will not estimate the true causal effect (this can be shown using a similar proof to the one above). The example emphasizes the importance of the data generating process when thinking about causality, and that misinterpretations can happen when assumptions about the data generating process are not clearly stated or accounted for. Again, it is important to emphasize that the misinterpretation *only* arises when eq-I.3 is estimated and β_1 is interpreted as the true causal effect of travel time. Wooldridge provides an excellent in-depth discussion of those issues in (22). There, a classical example that Wooldridge uses is one where a modeler is interested in the causal effect of a beer tax on alcohol related traffic fatalities. Should the modeler control for beer consumption in a regression of fatalities on beer tax? Like the transportation example above, beer tax is correlated with beer consumption, but is not a confounder of it. Instead, beer consumption is the mediator through which the beer tax affects traffic fatalities. So, if the modeler controls for beer consumption, the coefficient on beer tax will greatly underestimate its causal effect on fatalities. Of course, there is nothing wrong with running that regression, i.e. a regression that includes beer consumption and beer tax in its regressors, as long as its parameters are interpreted correctly. In fact, that regression will likely have more predictive power than one which includes beer tax only, so if the purpose is better predictive accuracy, then it may be more appropriate. The problem only arises when the parameters of that equation are interpreted causally.

For another example, consider again the case of determining the effect of adding a bike lane on mode choice in a given neighborhood. A common practice in travel demand modeling is to use the coefficient on a variable in a model to quantify its marginal effect on the outcome. These coefficients are interpreted as marginal utilities, and are generally assumed to have causal interpretation (14). So, in this example, it is very common to interpret the coefficient on the bike lane dummy variable in a mode choice model as the causal effect of adding a bike lane on the utility of biking (23). Even in the ideal case where all the variables that affect an individual's mode choice are observed and present in the model (i.e., no omitted variable or self-selection bias), this approach can still be problematic if certain variables that are affected by the bike lane are included in the model specification. Similar to the previous example, these post-treatment or mediator variables (24) should not be controlled for if the purpose is to estimate the treatment effect of biking. For example, it is common for such models to contain variables like bike travel time, as well as the bicycle infrastructure attributes (like type of bike facilities). If the objective is to estimate the true effect of bike travel time, then one should control for the type of bicycle infrastructure, which could affect bike travel time as well as mode choice (in other words, bike infrastructure is a confounder of bike travel time and mode choice). However, if the objective is to estimate the effect of the bike infrastructure itself by interpreting its coefficient as causal, then any mediating variables (like bike travel time) should be excluded from the model in order to correctly interpret the model parameters causally, similar to the previous example on level of discomfort from (10). I delve more into this topic in the Selection on Observables section, where I explain this phenomenon more clearly using Directed Acyclic Graphs (DAGs). DAGs can provide an intuitive representation of the data generation process that is often clearer and more intuitive than the more common econometric representations using structural equations like the ones shown in equations (1-4).

To be clear, though, I want to emphasize again that there is nothing inherently wrong with the way those models are developed, and indeed, they can be appropriate for the right applications. As the famous George Box’s saying goes, “All models are wrong, but some are useful,” so claiming for example that certain models are particularly wrong would be, ironically, wrong. Instead, the focus should be on whether a given modeling framework is the right one for the objective at hand. As I argued in the previous paragraphs, if the objective of a demand model is indeed to estimate causal parameters or to be used for policy evaluation, then the modeler should be aware of the issues discussed here, and should be clear about presenting their assumptions about the causal mechanisms through which those effects manifest by explicitly stating the assumed data generating process before drawing any causal conclusions. This will greatly help a practitioner, or whoever uses the model for policy evaluations, to be informed about how to interpret the model parameters. The problem in my view is that demand models are in fact very often used to evaluate the impact of policies, yet if one looks at the Federal Highway Administration (FHWA)'s travel model improvement programs (TMIP) over the last two decades, the focus is mostly on predictive validity and accuracy, with little mention of causality issues during model development.

The rest of this chapter is organized as follows. In Section 2, I discuss issues of identification as treated in the transportation choice modeling literature and the causal inference literatures, and argue that causal identification is rarely discussed in the transportation literature. In Section 3, I give an overview of some causal identification strategies used in the empirical sections of the dissertation. In the overview, I focus on the identifying assumptions of those strategies, as opposed to specific estimation techniques. Indeed, identification and estimation of causal effects are two separate tasks in causal inference research, and there are usually a plethora of estimation techniques for any given identification strategy. In the Selection on Observables identification strategy section, I also introduce and illustrate Directed Acyclic Graphs, a powerful tool that transportation researchers can leverage to encode their assumptions and reason about the correct model specification that allows for estimation of causal parameters. In Section 4, I conclude this chapter with a summary and an outline of the rest of this dissertation, which consists of three empirical transportation questions that I answer by designing and implementing causal identification strategies.

2. Identification: Causal vs. Statistical

Identification is a broad and general concept in statistics, with pioneers like Manski being among the first people to formalize this concept. Put simply, a model is said to be identifiable if its parameters can be uniquely learned with an infinite number of observations (25, 26). A key notion in this definition is the uniqueness of the model parameters, which effectively means that different values of the model parameters must result in a different distribution of the observable variables and outcomes. Given this definition, it is easy to see why identification is an important issue of statistical learning: no statistical inferences are possible on the model parameters if those parameters cannot be uniquely identified with infinite data, since statistical inference revolves around estimating standard errors that require the asymptotic convergence of the parameter of interest. Causal identification, on the other hand, refers to the problem of whether the estimated model parameters actually estimate a causal quantity. In this section, I briefly describe the identification problem as predominantly treated in the discrete choice modeling literature and used in transportation research, and contrast it with causal identification, which receives little to no attention by transportation demand modelers.

2.1. Identification as treated in the Discrete Choice Modeling literature

Train succinctly summarizes the identification issue in discrete choice modeling with the following two statements (27):

- "Only differences in utility matter"
- "The scale of utility is arbitrary"

The identifying assumptions are rooted in the behavioral theory underlying the decision making process. This is reflected in the two statements above. The first statement is mathematically equivalent to the fact that increasing or decreasing the utility of all the alternatives by a constant does not affect the final choice of the decision maker, as the ranking and relative appeal of those alternatives remain the same. The second statement similarly implies that multiplying or dividing the utility of all alternatives by the same constant does not affect the choice of alternatives either.

Those two statements have important mathematical implications, and they result in assumptions on model specifications that are necessary for the statistical identifiability of those models. Consider a discrete mode choice model with *J* choice alternatives: the fact that one of the model's alternative specific constants (ASC) is fixed to a certain arbitrary number (often set to zero) is because there are infinite combinations of *J* ASCs that result in the same differences in alternative utility, since there are only *J-1* such differences. Setting one of the ASCs to zero (or another number) is thus a necessary condition under which the remaining ASCs can be uniquely identified and estimated. Similarly, the fact that the absolute scale of the utility is arbitrary requires other modeling decisions on the error terms in order to uniquely estimate the parameters of the model, which is often accomplished by normalizing the variances of those error terms. Refer to Section 2.5 of (27) for a more detailed discussion.

Beyond the concepts above, the issue of statistical identification and its necessary conditions in discrete choice modeling has received a lot of attention in the transportation literature (see for example (28, 29) for dealing with identification in extended discrete choice models). For the purposes of this chapter, however, it is sufficient to understand that the type of identification issues that are typically dealt with in the travel demand modeling literature refers strictly to the set of assumptions that provably result in an identified model, i.e. a model in which all parameters can be uniquely estimated given enough data. In the next subsection, I discuss the issue of causal identification and how it is different from the identification issues discussed in this section.

2.2. Identification in Causal Inference

The concept of identification in causal inference goes beyond the statistical identifiability of parameters and the specification issues that are dealt with within the travel demand modeling literature. Indeed, a model can be correctly specified to ensure that its parameters are identifiable in the statistical sense (i.e., uniquely estimable given the data), yet those parameters can still be void of any causal meaning. For a parameter to be causal, a certain set of assumptions need to be made about the data generating process under which the value estimated can have a causal interpretation. Those assumptions are often untestable due to the unobservability of counterfactuals, a problem often referred to as the fundamental problem of causal inference. In the transportation literature, the closest concepts to causal identification are the concepts of unbiasedness or consistency of the model parameters, which are typically used to refer to the ability of the model to recover the "true" population parameters asymptotically, where true is generally understood as causal. Those concepts are not discussed with a formal causal framework, however, which can lead to mistakes, as shown in the Motivation subsection earlier in this chapter.

To make this concept clearer, let's start by introducing some notation. Denote by D_i a treatment variable representing the treatment status of individual *i*. For simplicity, let's assume that D is a binary treatment with values 0 (which means individual *i* did not receive a treatment and is part of the control group) or 1 (which means individual *i* is in the treatment group). Also denote by Y_i the outcome of interest. An important notion in the Rubin Causal Model (RCM) is the notion of potential outcomes (for a thorough introduction to this framework, refer to (30)). More specifically, the potential outcome of individual i under treatment D, $Y_i(D)$, is the outcome that unit $i$will have if they are exposed to treatment status D. Consequently, in the binary treatment case, the causal effect of treatment D for individual i is simply the difference in the potential outcomes corresponding to the two levels of treatment status (treatment and control):

$$\tau_i = Y_i(1) - Y_i(0) \qquad \text{(eq-I.5)}$$

Embedded in eq-I.5 above is the fundamental problem of causal inference. The equation shows that the causal effect of the treatment for individual i depends on the potential outcomes under both the treatment and control, but an individual can only be assigned to one group at a time, and never to both. This means that we only observe one of the potential outcomes, and it is impossible to observe both potential outcomes for a given individual or unit at the same time. The fact that we never fully observe all the quantities necessary to estimate the causal effect of the treatment is the source of identification problems in causality. As a result, causal identification assumptions **must** be made in order to make inferences about the causal effects. The plausibility of those assumptions determines whether or not the estimated effects can have a causal interpretation. Loosely speaking, those assumptions allow the modeler to estimate what the counterfactuals would have been if the units were subjected to the other treatment level (i.e., if the control units were subjected to the treatment, and the treatment units were subjected to the control). In the next paragraph, I make this notion mathematically clear.

First, notice that τ_i in eq-I.5 is the individual treatment effect for unit i. In most applications, we are interested in the expected value of the treatment effect across a given population, or the average treatment effect (ATE), instead of the individual effects (we are sometimes interested in the conditional average treatment effect (CATE) when heterogeneity is important, but this can still be thought of as the ATE for a specific sub-population or group, so I do not address it separately). The ATE can be expressed as follows:

$$ATE = E(\tau_i) = E[Y_i(1) - Y_i(0)] = E[Y_i(1)] - E[Y_i(0)]$$
$$ATE = \pi * \{ E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1]\}$$
$$+ (1 - \pi) * \{ E[Y_i(1)|D_i = 0] - E[Y_i(0)|D_i = 0]\} \quad \text{(eq-I.6)}$$

where π is the fraction of individuals who received the treatment. Note that the ATE defined in eq-I.6 includes terms that are unobserved. Specifically, the expected value of the potential outcome under treatment, $E[Y_i(1)]$, is only observed for those who received the treatment, $E[Y_i(1)|D_i = 1]$. Similarly, the expected value of the potential outcome under control, $E[Y_i(0)]$, is only observed for those who are in the control group, $E[Y_i(0)|D_i = 0]$. We do not observe the potential outcome under treatment for control units, $E[Y_i(1)|D_i = 0]$, nor the potential outcome under control for the treated units, $E[Y_i(0)|D_i = 1]$. The fact that we can never observe those quantities irrespective of how much data we collect is the source of the causal identification problem. The identification strategies explained in the next section are in essence a set of assumptions about those counterfactuals that allow us to estimate the ATE when some of its components are not observable[2]. Keele provides an excellent review of causal identification strategies for readers interested in a more comprehensive discussion (31).

3. Causal Identification Strategies: A Review

In this section, I give an overview of some of the identification strategies used in the empirical sections of this dissertation. The focus here is on the necessary conditions needed to estimate causal effects using the data, and I do not cover the specific estimation techniques that can be used under each strategy. Indeed, there exist many estimation techniques and models that can be used under each identification strategy, but the underlying assumptions of each of those models remain unchanged. If the assumptions do not hold, however, no degree of modeling sophistication or complexity can overcome those issues and produce valid causal estimates.

3.1. Randomized Control Trial

The first and perhaps most well known identification strategy is the randomized experiment or randomized control trial (RCT). I use this identification strategy in Chapter 4, where I design and implement randomized pricing experiments to estimate the

[2] Note that I do not talk about additional identifying assumptions in this chapter, since the purpose is not to be comprehensive. One major assumption that is often necessary on top of the identification strategies explained in this chapter is the Stable Unit Treatment Value Assumption (SUTVA), which assumes that one's potential outcome is unaffected by other people's treatment status. The unfortunate name aside, the assumption simply means no interaction effects between study units.

relationship between prices and the choice of Electric Vehicle charging options, as well as the relationship between hourly prices and charging duration.

Referred to as the gold standard of causal inference, in randomized experiments, subjects are randomly assigned to treatment and control groups. This offers many advantages when estimating causal effects: in short, randomization helps achieve the ignorability assumption[3], which means that the treatment assignment, D_i, is independent of the potential outcomes, $Y_i(D_i)$. In other words, it means that there are no differences in the *potential* outcomes of treated and control units, and selection into the treatment (or the control) does not happen based on those potential outcomes. To see why this is important, note the following equalities that follow when the potential outcomes, $Y_i(D_i)$, are independent of treatment assignment, D_i (i.e. when the ignorability assumption is satisfied):

$E[Y_i(1)|D_i = 1] = E[Y_i(1)|D_i = 0] = E[Y_i(1)]$

$E[Y_i(0)|D_i = 0] = E[Y_i(0)|D_i = 1] = E[Y_i(0)]$

Thus, to identify the ATE, eq-I.6 becomes:

$ATE = E[Y_i(1)] - E[Y_i(0)] = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 0]$

Note how all the elements in eq-I.6 are observed and can be estimated from the sample. Again, the key identification strategy here is that treatment assignment is independent of the potential outcomes, and the randomization of D_i, if done correctly, ensures the validity and credibility of this strategy. Thus, to identify the ATE in a randomized experiment, it suffices to compute the difference in means between the treatment and control groups.

A few additional things to note regarding randomized experiments:

- Randomized experiments ensure internal validity of the treatment effect, meaning that the estimate of the treatment effect is valid for the study sample. Additional assumptions and steps need to be taken to ensure external validity of those estimates, i.e. that the estimates generalize to a broader population. For example, random sampling is required to ensure external validity. External validity is typically checked by comparing the marginal and joint distributions of pre-treatment variables (in most cases, these are demographic variables) between the study participants and the larger population of interests.
- Issues of attrition and non-compliance are also important for randomized experiments. It is important that individuals do not drop out from the study, or if they do drop out, it happens randomly and is not correlated with the potential

[3] It is called the ignorability assumption because the treatment assignment is ignorable, i.e. we do not need to worry about it in the analysis, when it is not correlated with the potential outcomes of the study units

outcomes. Checking for non-compliance is done by ensuring that the people in the control group are not exposed to the treatment, and that people in the treatment group are complying with the treatment rules. See (32) for a thorough discussion of those issues and a detailed introduction to the topic of randomized experiments.

- Finally, note that there are many variations of the randomized experiment described here. Often, it is not possible to perform randomization of the treatment at the individual level. In those cases, cluster randomization, or time randomization, can be more feasible. For a detailed review of those variations, please refer to (32).

3.2. Difference-in-Differences and Longitudinal Models

Another common and widely used strategy to identify treatment effects is leveraging longitudinal data. Arguably, using longitudinal data is the second best thing to randomized experiments[4] when trying to identify causal effects.

As mentioned earlier, a challenge in estimating causal effects stems from the difficulty of identifying good exogenous variation in the treatment. In cross-sectional observational data, this is hard due to selection bias (often referred to as endogeneity in transportation demand modeling): treatment assignment is not random and is often correlated with the potential outcome of the study units, and individuals with different observed and unobserved characteristics can self-select into the treatment and control groups. A famous example in transportation is that individuals who live near a bike lane may be inherently more likely to bike to work than others: in other words, their treatment status (having a bike lane), is correlated with their potential outcomes (whether or not they bike to work). When panel data are available, however, we can use within-individual variation in the treatment status over time to identify treatment effects. When using the within-individual variation in the treatment, we are effectively controlling for any individual confounders that are fixed over time. In the same bike lane example, if we observe the same individuals before and after installing the bike lane, then we can control for their inherent preferences for biking using, for example, fixed effects models.

Many empirical designs are used when working with longitudinal data. The simplest design is the Differences design, also known as a before/after study. Under this design, the treatment effect estimate is the difference in means of the study units before and after the treatment is implemented. The identification assumption in such design is that the before/after differences in the outcome would have been zero if it weren't for the treatment. This is a strong assumption, as there are often underlying time trends that could cause such differences, or there may be other factors that have changed over the same time as the treatment that could explain away the differences in the outcome. As a result, the Differences design is a weak design due to its reliance on strict and unlikely identification assumptions, and is rarely considered a reliable causal design.

[4] An argument can be made for valid natural experiments where treatment is close to randomly assigned.

A very popular extension of the Differences design is the Difference-in-differences (diff-in-diffs) design (33, 34). This design is the most common research design used to estimate the impact of policies with longitudinal data. The design extends the Differences design by including a control group that has not been exposed to the treatment, and as the name suggests, the treatment effect is estimated as the difference in the before/after differences between the control and the treatment group. The key identification strategy in the diff-in-diffs design is the parallel trend assumption, which asserts that the treatment and control groups would have trended similarly if it weren't for the treatment. In other words, the strategy assumes that the difference in the outcome between the control and treatment groups in the pre-treatment period would have remained the same in the post-treatment period had the treatment not been administered. As a result, if the difference between the control and treatment groups increases in the post-treatment period, the treatment is assumed to have a positive effect, and vice versa. In the bike lane example mentioned earlier, if we observe two group of individuals over multiple points in time, where one group gets a bike lane installed on their path to work and one does not, then we can look at the difference between the rates of biking before and after the installment of the bike lane for each of the two groups, and then compare how those differences changed before and after installing the bike lane (i.e. compute the difference-in-differences). This would be the estimate of the average causal effect of the bike lane.

To provide evidence for the plausibility of this assumption, the analyst will typically try to find multiple pre-treatment data points for the treatment and control groups and show that the differences in the outcomes between those two groups remained constant over time before the treatment was assigned. I do this in Chapter 3, where I estimate the effect of getting vaccinated on the number of weekly trips that individuals make. There, I show that the number of weekly trips for vaccinated and unvaccinated individuals trended similarly before vaccines became available, so the change in the difference in the number of weekly trips before and after the treatment group gets vaccinated can be causally attributed to the vaccine.

There are many variations and extensions of the diff-in-diffs design, and many estimation techniques that can be used to implement it. In Chapter 3, I implement a canonical diff-in-diffs when estimating the effect of getting vaccinated on the number of weekly trips that individuals make, and the effect of getting vaccinated on the frequency of using transit. In Chapter 2, I implement a variation of the diff-in-diffs design called First Differences, where I use multiple observations over the same individual to quantify the effect of telecommuting on the number of non-commute trips that individuals make. Refer to those respective chapters for additional details about those designs.

3.3. Selection on Observables

As the name suggests, selection on observables is an identification strategy which assumes that selection into the treatment group happens based on observable characteristics. In other words, even though people self-select into the treatment group, we observe all the factors that determine (i.e. predict) this selection process. Under this assumption, once we condition on those factors, treatment assignment becomes as good as random. This is a strong assumption in observational studies, but is often implied in cross-sectional regression models that interpret the coefficients as causal effects.

Mathematically, the identification assumption under the selection on observables strategy is that treatment assignment is independent of the potential outcomes once we condition on a set of covariates, X:

$\{Y_i(0), Y_i(1)\} \perp D_i | X_i$

The above assumption is also known as the conditional ignorability or unconfoundedness assumption. Again, it is very important to emphasize that the key concept in this strategy is that we observe all the factors X that affect both the treatment and the potential outcomes. Any violation of this assumption would lead to arbitrarily biased treatment effects.

When it comes to estimating treatment effects under this assumption, plenty of statistical techniques are available. In fact, this is an area that has received a lot of attention in the last few years from the machine learning literature as well, despite the common critique that machine learning models receive as being unable to handle non-predictive objectives. Going over the specific estimation techniques is beyond the scope of this chapter, but interested readers can research the terms regression adjustment, propensity score methods (weighting, matching, and blocking), Meta learners (T-learner, X-learner), Double/De-biased machine learning, Causal forests, etc. What these methods have in common is that they use the covariates X to partial out any covariation in the treatment, D, and the outcome, Y, that are induced by X. In other words, all those techniques seek to partial out bad variation in D and Y – i.e., variation that can be explained by X – and use only the good variation – i.e., exogenous variation – in D to estimate its effect on Y.

The selection on observables strategy is what is implicitly assumed in transportation discrete choice models estimated on cross-sectional data and where the model parameters are given a causal interpretation, or in any cross-sectional regression model where the parameters are interpreted causally. Even more advanced discrete choice models, like latent class choice models and mixed logit models, still fall under selection on observables – those extensions address unobserved heterogeneity in taste parameters, but still assume that all confounders are observed and controlled for in the model. The exception is integrated choice and latent variable models (ICLVs), which adjust for latent variables (typically attitudinal) in the utility equations of the choice alternatives. However, one may argue that even ICLVs fall under selection on observables, since the latent variables are themselves typically assumed to be a function of observable characteristics in the vast majority of applications, and one needs to collect (i.e. observe) additional data in the form of survey questions in order to estimate the latent variables.

While more often than not it is hard to justify the selection on observables assumption in observational data, the fact remains that it is used very frequently in practice. The main task in those designs is to find a set of regressors X that satisfies the ignorability assumption, but are there any restrictions on which variables can go in the model if the purpose is to estimate treatment effects? In other words, is there any harm in including the most comprehensive set of controls in the model? The answer to both those questions is yes: including the wrong controls, or bad controls, in the model can significantly undermine the causal interpretation of its parameters. This is an issue that is commonly observed in transportation models, since if the modeler is not careful, they risk adding post-treatment or mediator variables into the specification equations. In the next subsection, I give an overview of DAGs, a powerful tool for visualizing and thinking about how to specify models (i.e., how to select a set of controls X that satisfies the conditional ignorability assumption) when the objective is to identify causal parameters, and present a common example where demand modelers fall into the trap of conditioning on bad controls. For a crash course in good and bad controls, refer to these excellent papers that offer intuitive explanations and a more comprehensive overview (35, 36).

3.3.1. Directed Acyclic Graphs

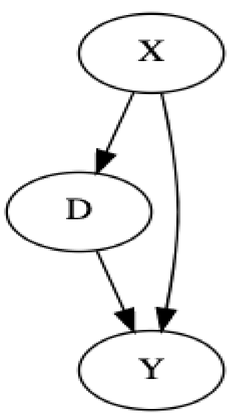
Causal diagrams and causal graphical models were introduced by Pearl in 1995 as a powerful tool for causal inference, especially in observational studies (37). In my view, the most important and useful feature of causal graphs and its added value to transportation demand modeling is the clear representation of the causal relationships between the regressors. In this section, I focus specifically on illustrating the power of DAGs to represent and encode causal relationships between variables in an intuitive and clear manner, and how that can help the modeler decide what they should and should not control for. Specifically, I illustrate the main causal structures that a modeler should pay attention to, and give an example from transportation where those structures are often present. Interested readers can refer to (36, 37) for a deeper dive into causal graphical models and the topic of good and bad controls in more complicated structures.

Before I delve more into DAGs, I believe it is important to give a word of caution. Despite the fact that Pearl describes DAGs and causal graphical models as revolutionizing the field of causal inference (and in some lectures, he goes as far as claiming DAGs “solve” causal inference), in my view (and that of notable economists (38)), their main contribution is a better, more intuitive, and explicit representation of assumptions about the data generating process, which result in (theoretically) testable implications. They allow a modeler, given a causal model in the form of a DAG, to algorithmically determine a set of good controls that allow for causal estimation. The key notion here though is that the correctness of the results and the causal estimates depend heavily on the correctness of the DAG. In the ideal case, a transportation modeler should use an identification strategy that does not make so many assumptions about the relationship between variables, like instrumental variables (IV), diff-in-diffs, regression discontinuity, etc. When that is not possible, which is often the case in transportation research where good instruments, randomization, or panel data are not available, this is where DAGs become useful. They force the modeler to be clear and transparent about their assumptions when specifying models, and they allow the reader to think about the validity of those assumptions when they are judging the validity of the results.

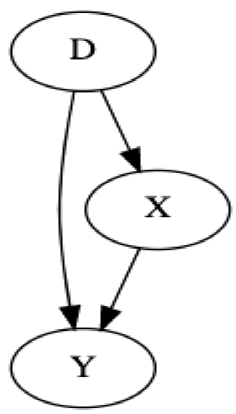
Another limitation of DAGs worth noting is their inability to handle cyclical structures. In those situations, which arise in cases where two variables cause each other. As their name suggests, the kind of relationships where DAGs are most useful are acyclic, meaning that the direction of causality flows strictly from one variable to another. This makes them unfit to handle problems where the treatment and the outcome may have a simultaneous relationship: for example, telecommuting may cause people to move farther from work, but living farther from work causes people to telecommute more as well. These kinds of problems are best handled using longitudinal data, and DAGs are not necessarily more intuitive if a time dimension is added. I face a similar situation in Chapter 2, where I quantify the causal effect of telecommuting on distance traveled. There, I do not use DAGs, and instead use variations of diff-in-diffs designs and longitudinal fixed effects models to quantify the causal impact of telecommuting on travel frequency.

Main Causal Structures in DAGs

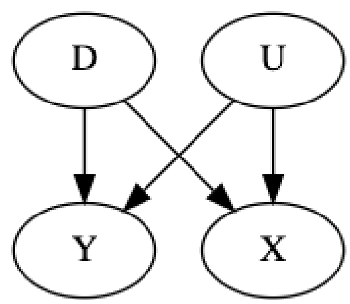
There are three main structures that form the fundamental building blocks of any DAG, all shown in Figure I-1.



a) X is a confounder with respect to D



b) X is a mediator variable with respect to D and Y



c) X is a collider variable with respect to D

Figure I-1: The building blocks of Directed Acyclic Graphs

In all three structures, suppose that we are interested in estimating the causal effect of D on Y. Figure I-1-a shows the common confounding structure, also known as the fork structure. In this example, in order to adjust for confounding from X, we should condition on (aka control for) X. Under this data generating process, X induces “bad variation” in both Y and D and should be adjusted for. This is the graphical equivalent of the conditional ignorability assumption in selection on observables. Figure I-1-b shows the mediating structure, in which part of the causal effect of D on Y is channeled through X.

If our goal is to estimate the causal effect of D on Y, then controlling for X will bias our estimate, since the mediated effect of D on Y will be consumed by the coefficient on X. Thus, one should not control for X in this case in a regression of D on Y. Figure I-1-c is the less common case of colliders. Colliders are also referred to as common effects since they are caused by the treatment of interest as well as other variables. In this case, conditioning on the collider X will induce non-causal association between D and U, an unobserved confounder that also affects Y, causing bias that would not otherwise occur if we do not condition on X. Thus, we should not control for the collider X in a regression to estimate the causal effect of D on Y. This case rarely appears in practice but is theoretically possible, so modelers need to keep an eye out for such situations when specifying their models.

Finally, an important thing to note about the assumptions encoded in a causal graph are conditional independence assumptions. For example, consider the graph shown in Figure I-1-c), and focus on the variables Y, X, and U. The path formed by these variables is of the form $X \leftarrow U \rightarrow Y$, with no direct path from X to Y. What such a structure implies is that X and Y are conditionally independent given U. The construction of DAGs can thus be viewed as encoding the modeler's conditional independence assumptions. Those assumptions can theoretically be tested using data-driven statistical tests, but those tests are typically low powered. Personally, I believe those assumptions are best checked using subject matter expertise and discussions between researchers. In that sense, DAGs would facilitate such discussions since the assumptions will be explicitly and intuitively stated, rather than being implicit or buried in (sometimes intimidating) mathematical notations. Even if agreement on those assumptions may never be universal, at the very least, being explicit about them in the form of DAGs will allow the readers and applied practitioners to judge for themselves the validity of those assumptions and whether they are plausible for the specific application at hand.

<u>Examples from Transportation</u>

Next, I give a transportation example to better illustrate how these causal structures appear in typical travel demand modeling applications, and how they may undermine the causal inferences that people associate with the model parameters if not accounted for properly. Recall the example at the beginning of this chapter about quantifying the effect of adding a bike lane on mode shares in a given neighborhood. As mentioned earlier, a very common practice within the discrete choice modeling literature in transportation is to use the coefficient on a bike lane dummy variable as the causal effect of the bike lane on the utility of biking. As I alluded to in the previous subsections, this approach is problematic. Let's start by thinking about a plausible causal graph that captures the true data generating process. If the data are observational, a possible causal graph is shown in Figure I-2. Note that I only focus on the parts that involve the bike lane, and do not present a full causal graph of all the variables that could be involved in such a problem. The graph also assumes no selection bias into the bike lane treatment, meaning that people who live near a bike lane are similar in their potential outcomes to those who do not. This is obviously a simplifying assumption unlikely to hold in practice but is a useful one to make the point of good and bad controls. I give a more realistic graph in Figure I-3.

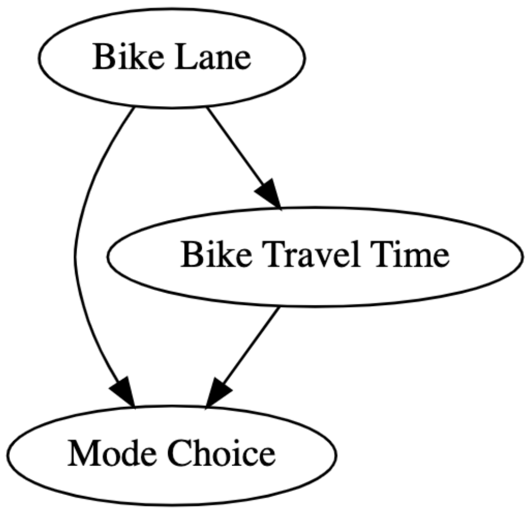


Figure I-2: DAG for the relationship between travel time, bike lane, and mode choice

What does the DAG in Figure I-2 tell us? For starters, it makes it clear that it is not possible to correctly estimate the causal effects of bike travel time and bike lane in the same model. Stated differently, if they are all included in the discrete choice model, the coefficients on the bike lane and the bike travel time cannot all be interpreted causally under this data generating process. Below, I elaborate on why that is the case.

The bike lane variable is a confounder to bike travel time, but bike travel time is a mediator for the bike lane variable. Thus, to estimate the effect of bike travel time on the mode choice, one must control for the confounding bike lane variable. However, if the goal is to estimate the causal effect of a bike lane on mode choice, then the travel time variables (for both bikes and cars) cannot be included in the model. If included, those mediators will absorb a significant portion of the effect of the bike lane, leading to significantly underestimating it. This again highlights the invalidity of general statements like "[...] choice models are generally understood as causal." A choice model that includes all three variables can be interpreted causally for the travel time variables, but not for the bike lane variable. Yet, in the transportation literature, it is typical for all three parameters to be interpreted causally (see (39) for an example).

So what's the solution? As always, it depends on the modeling objective. Below are a few scenarios:

- If the modeler is only interested in the causal effect of travel time on the mode share, then modeling specifications that include all variables are correct.
- If the modeler is only interested in the total causal effect of the bike lane, then the modeler must exclude the travel time variables from the model.
- If the modeler is interested in the causal effect of the bike lane, but is also interested in quantifying the channels of causation through the travel time variables (i.e., the modeler is interested in how much of the causal effect of a

bike is due to the effect on travel time), then the modeler will need to resort to structural equation modeling, where a model quantifying the effect of a bike lane on bike and car travel times will need to be developed. These techniques have been applied in transportation before, and have been used in other fields like psychology. The downside of such models, however, is that the final estimate of the effect of the bike lane will ultimately depend on the correctness of the functional form assumptions of all the structural models.

The example in Figure I-2 may have been simplistic to clearly distinguish between confounders and mediators. Now consider a less straightforward example of a more plausible data generating process as shown in Figure I-3.

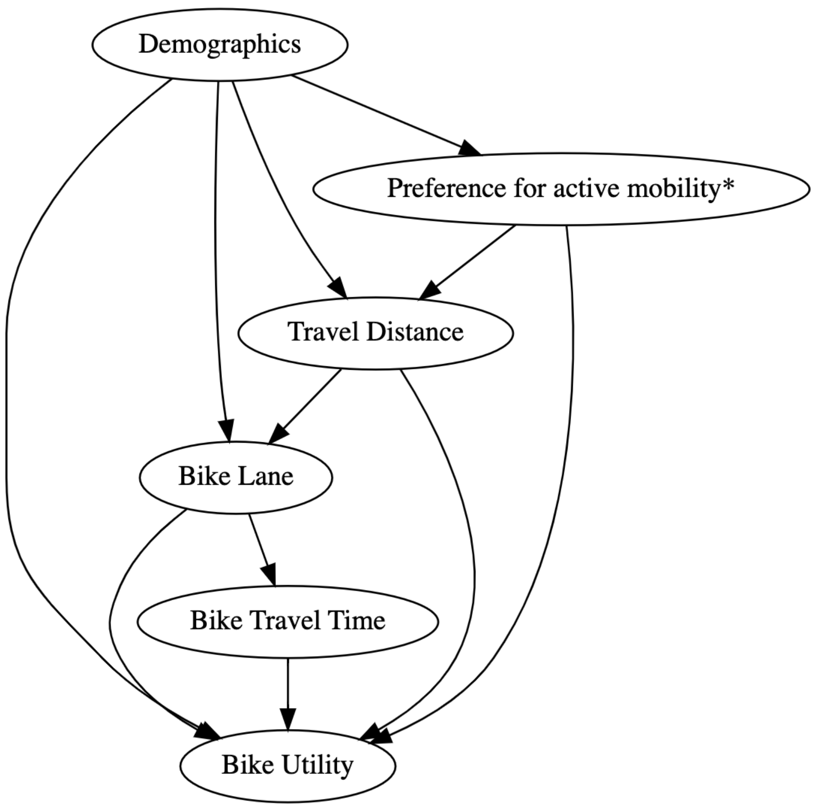


Figure I-3: DAG for the factors affecting biking utility. * means unobserved variable

In Figure I-3, suppose again we are interested in the effect of a bike lane on the utility of biking. Here, we do not assume that there is no self-selection bias of the bike lane variable. Instead, the presence of a bike lane is confounded with demographic characteristics, as well as travel distance. Like the previous example, however, bike travel time is not a confounder of the bike lane. So, to estimate the causal effect of the bike lane on the utility of biking, the modeler should adjust for the confounders (demographics, travel distance), but exclude the mediators (bike travel time in this case).

There is another important point I am making in Figure I-3. Namely, the DAG as shown assumes that there is another unobserved factor, preference for active mobility, that also indirectly confounds the bike lane variable. However, in my DAG, I assume that the effect of the unobserved preference for active mobility on the bike lane variable happens entirely through the travel distance variable, and there is no direct path of the preference for active transportation to the bike lane variable. For example, those who have this preference live in areas denser areas closer to their work location (so shorter travel distance to work), and those areas, because they are more dense, are more likely to have a bike lane. Under this assumption (that there is no direct path from the unobserved preferences to the bike lane variable), controlling for the confounding effect of travel distance is enough to adjust for those unobserved preferences. Of course, another modeler or practitioner may disagree with this assumption. For example, they may assume that people with a preference for active transportation may choose to live specifically in areas with bicycle infrastructure. In other words, they may assume that there is a direct arrow from the preference for active mobility variable to the bike lane variables. Under this assumption, not adjusting for the preference for active mobility will result in confounding (aka omitted variable) bias. What is good about presenting those assumptions in the form of DAG, however, is they encourage more productive discussions about the validity of the data generating process under which the model is estimated, and they allow the modeler and practitioner to correctly interpret the model parameters, reason about their validity, and hypothesize about the direction of bias. This will, hopefully, greatly reduce misapplications of the model and misinterpretations of its parameters. The other good thing about DAGs is that algorithms exist that determine the set of control variables that result in unbiased estimation of causal effects. So, as the complexity of the problem increases, the modeler can rely on those algorithms to figure out which controls should be included, and which should be excluded, from the final model specifications. Those algorithms are based on the theory developed in (37).

4. Summary and Dissertation Outline

In this section, I discussed the disconnect between the causal inference and transportation demand modeling literature, misconceptions leading up to it, and some erroneous modeling practices that undermine the causal interpretation of modeling parameters. I also explained the difference between statistical identification as treated in the transportation choice modeling literature, and causal identification which received little to no attention in travel demand modeling. Then, I went over some common causal identification strategies that I use in the subsequent empirical sections of the dissertation and explained the assumptions underlying them. Finally, I gave an overview of DAGs as a powerful tool to illustrate a modeler's assumptions and allow them to determine the correct specification of their models if the objective is to identify causal parameters.

In the following chapters, I essentially practice what I preach. I answer three empirical causal transportation questions using the causal inference techniques and identification strategies I reviewed in this chapter. Those chapters are independent manuscripts that have either been submitted or in the process of being submitted to refereed transportation journals. In Chapter 2, I quantify the causal effect of telecommuting on travel frequency and distance traveled. In Chapter 3, I quantify the causal effect of vaccines on reversing pandemic induced mobility trends. In Chapter 4, I design and implement randomized pricing experiments to study the charging behavior of electric vehicles’ owners at plug-in electric vehicle charging stations. I conclude in Chapter 5, where I summarize the contributions of this dissertation and propose future research directions.

Chapter 2: Does telecommuting reduce trip-making? Evidence from a U.S. panel during the COVID-19 impact and recovery periods

Executive Summary

Telecommuting has risen to unprecedented levels in the past two years and remains one of the most disrupted aspects of transportation behavior during the COVID-19 pandemic. In this paper, we investigate the transport impacts of telecommuting. We use a combination of passively collected Point of Interest (POI) data between January 2020 and December 2021 and five waves of actively collected surveys on a panel of participants to quantify the effects of changes in the frequency of telecommuting on the total number of daily and weekly trips that a telecommuter makes, as well as their total daily and weekly distance traveled. We overcome important limitations of related work in the literature by controlling for unobserved individual confounders as well as various demographic and temporal variables. Doing so, we find evidence that telecommuting results in the generation of new non-commute trips that offset a significant portion of the reduction in commute trips. We show that telecommuters make an average of roughly one additional non-commute trip on telecommute days relative to commute days. The additional trip is on average shorter than the commute trip, as we find that the total distance traveled on telecommute days is significantly shorter than on commute days for employees in our panel. At the weekly level, we also find that one additional day of telecommuting results in one additional non-commute trip. This suggests that the additional non-commute trip on telecommuting days is a newly generated trip, not a trip that has been shifted from other days of the week. Our results suggest that the trip reduction effects of telecommuting could be overestimated if telecommuting-induced new trip generation is not properly accounted for. We discuss the implications of our findings on policy-making and the literature on the transport impacts of telecommuting.

1. Introduction

The early phases of the COVID-19 pandemic has forced several sectors of the economy to shift to either a hybrid or full telecommuting model. Following this shift, researchers found that workers wanted to maintain their ability to telecommute, at least partly, after the pandemic is no longer a threat; research conducted by the Pew Research Center shows that, for individuals whose work responsibilities can be completed from home, the fraction of telecommuters increased from 20% pre-pandemic to 71% during the pandemic. Furthermore, half of those individuals expressed their desire to continue telecommuting post-pandemic, suggesting a continuation of the pandemic-induced telecommuting trend (53). Other research showed that the fraction of workers who expected to telecommute at least a few times a week post-pandemic is 26%, double that of pre-pandemic levels (54), and that vaccination does not seem to play a significant role in encouraging people to return to offices (cite our vaccine work). These historically high telecommuting rates provide a unique opportunity to study their transport impacts, and the likely persistence of high telecommuting rates in the future makes this topic timely and critical to understand, especially given its potential to affect transportation policymaking.

Generally speaking, telecommuting is widely believed to be an effective travel demand management strategy for reducing vehicle miles traveled (VMT) (71). The most tangible and direct impact of telecommuting is the elimination of commute trips, which often occur during peak travel periods. The reduction in commute trips, however, leads to savings on travel costs, both temporal and monetary. This freed-up time and money can potentially be used for other trips. The degree to which those two competing transport impacts offset one another has long been a subject of researchers' interests, with studies often providing conflicting findings (see section 2 for a review of the literature). Earlier studies generally argue that telecommuting leads to a significant reduction in both commute and non-commute trips. Methodologically, these studies relied on stated responses and self reported travel diaries of a small number of participants, and over a limited number of days. They consequently suffer from the generalizability concerns and self reporting biases that are typical to these types of analyses. More recent work tries to overcome previous research limitations by using significantly larger samples like household travel surveys, and find that a significant portion of the reduction in commute trips is offset by new trip generation. However, those studies mostly use cross-sectional observational data, and consequently suffer from other limitations like self-selection bias and unobserved confounding. For example, telecommuters in a sample may travel significantly more (or less) than non-telecommuters regardless of whether or not they telecommute, possibly due to the nature of their jobs, home locations, or other unobserved factors. These differences are hard to control for and could result in biased estimates of the impact of telecommuting on transport outcomes.

In this work, we have two main research objectives. First, we want to determine whether workers engage in more non-commute travel on telecommuting days, relative to commute days. Second, we want to determine whether this difference in non-commute travel activity between telecommuting and commute days is the result of newly generated trips, or of people shifting their existing travel from other days of the week to telecommuting days. Methodologically, we implement empirical causal designs and use individual-level variation in telecommuting days to robustly estimate the travel effects of telecommuting, leveraging a unique longitudinal dataset of passively collected POI data combined with five waves of survey data on a panel of U.S. smartphone user. In doing so, we overcome many of the methodological and sample size limitations of the existing literature that predominantly either relies on large cross-sectional data, or small self-reported travel diaries data. The longitudinal aspect of our data allows us to control for many observed and unobserved confounders, and the POI data allow us to overcome the self-reporting biases that are likely to be present when individuals self-report their travel activities. We implement quasi-experimental designs to quantify the causal effect of telecommuting on: 1) the number of trips, and 2) distance traveled by employed participants, both at the daily and weekly levels. At the daily level, we run fixed-effects regressions that allow us to control for unobserved individual and time fixed effects, as well as unobserved individual effects that vary monthly. At the weekly level, we implement a first-difference (FD) design to estimate the effect of a change in the weekly telecommuting frequency on the weekly number of trips and distance traveled, which allows us to assert whether the daily effects of telecommuting are additive or substitutional. Owing to our causal designs and longitudinal data, our findings are more robust and provide stronger evidence on the transport impacts of telecommuting than previous studies in the literature.

2. Literature

Researchers have proposed multiple theories on the mechanisms through which telecommuting affects travel (see, e.g., (72, 73) for a discussion). A review of the early literature by Nilles (1988) finds that commute trip savings due to home-based telecommuting are not completely offset by the generation of new trips (74). Mokhtarian (1998) also finds that when 1.5% of the population commutes on a given day, the resulting reduction in VMT is 1% (75). More recent studies, however, report even smaller reductions in travel as a result of telecommuting. One study found that while telecommuting partially reduces commuting trips, it increases non-commute trips for households with one vehicle per employed member (76), effectively offsetting the reduction in work travel due to telecommuting. Other work using data from the 2006 Household Travel Survey in South Korea found that the commute trip reduction effect of telecommuting is substantially offset by new trip generation (77). The study found that the rebound effect in a household's travel (additional telecommuting-induced non-commute trips) becomes larger than the average one-way commute distance as telecommuting becomes more regular for that household. Similar findings were reported in (78, 79), where the authors found that telecommuting increases one-way commute distance, daily total work trips, and daily total non-commute trips. Another study, analyzing whether telecommuting causes people to move farther from their work location, adopts a path analysis and finds that, rather than telecommuting being a determinant of home location choice, the decision to telecommute is determined by the workers' job location (80). The study also finds that telecommuting positively affects non-commute travel distance for telecommuters and their households, and argues that this increase in travel is not due to location change. Other work in England found that telecommuters make significantly more non-commute trips than non-telecommuters, but the study did not look at travel distance (81). Other studies, also in England and using cross-sectional data from the English National Travel Survey from 2005-2012, find that teleworking is associated with additional travel, owing to both longer commute distance for teleworkers relative to non-teleworkers and additional non-commute trips by teleworkers (82, 83). A more recent study builds on the previous two studies by extending their dataset until 2019, and finds that teleworkers travel father each week than non-teleworkers, that the total household weekly travel is larger in households with one member teleworking, and that teleworkers travel farther for business compared to non-teleworkers (84).

The literature summarized above, in addition to many other studies, suffer from significant limitations. Indeed, the literature is largely inconclusive, with studies often offering conflicting conclusions on the impact of telecommuting on travel. A recent systematic review of the energy and climate impacts of teleworking found that only five out of the 15 reviewed telecommuting studies concluded that teleworking increases non-commute travel. The conflicting conclusions are likely largely due to unobserved confounding that is hard to control for in cross-sectional observational analyses, an issue that multiple studies acknowledge as a key limitation (76, 80, 82, 84). Nevertheless, some studies have attempted to use econometric and quasi-experimental techniques to control for unobserved confounding when looking at the relationship between telecommuting and travel outcomes. For example, (79) uses an instrumental variables approach to address the endogeneity of commute distance and individual's telecommuting decision, using "internet use at home" and "total number of phones" as instruments for telecommuting decision. The choice of instruments, however, is questionable and likely violates the necessary conditions for valid instrumental variable designs. For example, the "internet use at home" and "total number of phones" instruments may themselves be endogenous to commute distance in the same way the decision to telecommute is. Additionally, even if we ignore the issue of endogeneity of the instruments and the outcome, there are also endogeneity issues between the instruments and the treatment: the direction of causality between internet use at home and telecommuting is more likely to be from telecommuting to internet use at home: telecommuting causes the increase in internet use, not the other way around. (82) attempts to overcome the selection bias when investigating the effect of telecommuting on commute distance by adopting a propensity score matching approach, but such an approach can only adjust for selection bias on the observable characteristics used in the propensity score model, and does not solve the unobserved confounding problem. Our work addresses these gaps in the literature and overcomes the limitation of unobserved confounding by using a panel dataset and controlling for fixed and time-variant confounders.

Another contribution of our work is leveraging the natural experiment that occurred due to the COVID-19 pandemic, which caused telecommuting rates across many sectors to increase dramatically. This reduced the systematic differences between telecommuters and non-telecommuters present in pre-pandemic cross-sectional studies, when telecommuting rates were significantly lower. Research exploring the effects of telecommuting during the pandemic mostly focused on the environmental impacts of reduced peak-hour congestion, the shifting in energy load patterns due to working from home, or hypothesized about the long-term effects of the increased telecommuting rates on future residential location ((85), (86), (87)), but did not try to quantify the effect of telecommuting on non-commute trips or distance traveled. We fill this gap and leverage the high telecommuting rates to provide more robust and generalizable results on the transport impacts of telecommuting.

3. Data and Methods

3.1. Data Collection

We used a combination of data collection methods to develop a database that enables a broad understanding of how COVID-19 has affected people's travel behavior. This analysis uses a panel dataset consisting of five waves of surveys and a total of 1,962 unique respondents. To construct our sample, we began with a geographically representative panel of approximately 100,000 smartphone users tracked by Embee Mobile/Similarweb. We recruited from a random sample stratified by U.S. region of approximately 15,000 members of this panel to conduct five waves of surveys (August 2020, September 2020, December 2020, April 2021, July 2021). Our surveys included questions regarding respondents' employment status, travel and telecommuting behavior, vaccination status, demographic characteristics, and ideological beliefs. In the first wave, 1,321 individuals responded to the survey — a response rate of about 9%. We then targeted those same respondents in the subsequent survey waves, and supplemented our sample from the original pool of 15,000 members to account for drop-outs while maintaining a sample size of approximately 1,000 participants in each survey wave.

A key advantage of our dataset is the ability to link the individual survey responses with individual POI tracking data. This allowed us to supplement the stated responses of the study participants with their observed travel behavior, and to compute mobility metrics at different temporal aggregation levels (e.g., number of daily and weekly trips, total daily and weekly distance traveled). We also tracked stay-at-home and shelter-in-place policies across the states and counties of the study participants. An overview of the study timeline is shown in Figure III-1. The POI data is available from January 2020. The survey data collection started in early August 2020 and collected data in five total survey waves (August 2020, October 2020, December 2020, March 2021, June 2021).

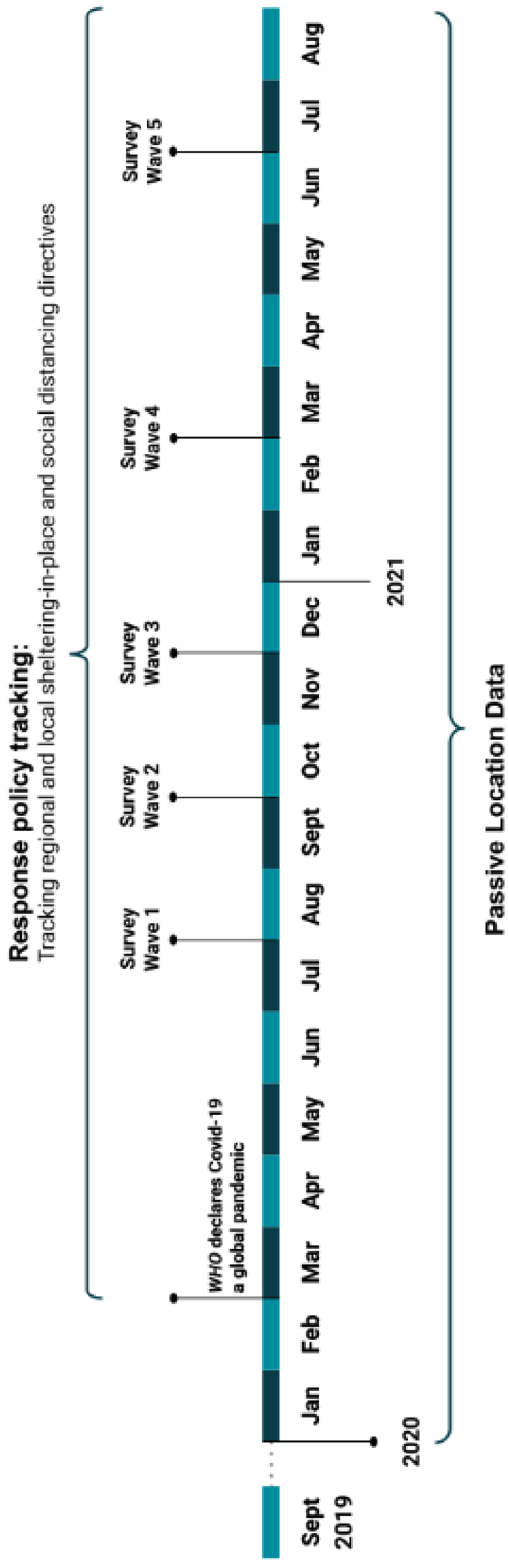


Fig. 2: Study Timeline

Figure III-1: Study and data timeline

3.2. Demographic summary

Table III-1 summarizes the demographic characteristics of full time workers from our survey respondents and compares those characteristics to those of full-time workers in the U.S., using the 2020 U.S. census data. In terms of gender, our respondents oversample females and undersample males. Our data also overrepresents individuals within the 25 to 59 age range compared to the U.S. census. In terms of income, our sample over-represents individuals with household incomes between $25,000 and $99,000 compared to the rest of the U.S. population, and under-represents high-income households with incomes above $100,000. Additionally, we over-represent racial and ethnic minorities and under-represent Caucasian workers. We also over-represent workers with a university or college degree, and under-represent those with less than a high-school education. On the other hand, statistics on household vehicle ownership closely track those of the U.S. population.

Despite some statistically significant differences between our sample means and the population means, the magnitudes of the differences are modest, and we do not have demographic categories that are missing completely. The biggest difference is in household income, which we control for in our models. This similarity in the characteristics of workers between our sample and workers in the U.S. suggests that our estimates of the travel impact of telecommuting are likely to generalize well to the U.S. population.

Table III-1 - Demographic Characteristics of Study Participants Compared to the U.S. Census

Category	**Sample (%) (N=809)**	**Population (%)**
Gender		
Male	47.5	52.66
Female	52.5	47.34
Age		
19 years and under	0.9	3.39
20 to 24 years	6.4	9.28
25 to 34 years	27.3	22.68
35 to 44 years	30.8	21.01
45 to 54 years	22.6^	20.77
55 to 59 years	5.7	9.75
60 to 64 years	4.0	7.25
65 years and over	1.4	5.89
Household Income		
$0 - $24,999 USD	13.3	7.20
$25,000 - $49,999 USD	34.5	15.54

$50,000 - $99,999 USD	37.7	33.47
$100,000 - $149,999 USD	10.3	21.55
$150,000 - $199,999 USD	1.9	10.70
$200,000 USD or more	2.3	11.54
Race/Ethnicity		
White/Caucasian	50.7	72.11
Black/African American	19.3	11.65
Native American/Alaskan Native	2.2	0.69
Asian or Pacific Islander	11.0	6.21
Other	6.6^	5.06
Mixed race	7.3	4.27
Hispanic Status		
Hispanic or Latino	20.0^	17.34
Not Hispanic or Latino	76.0	82.66
Education		
Less than High School	1.5	8.88
High School	38.3	46.31
University/College	51.2	31.56
Postgraduate Education	9	13.25
Household Vehicles		
0	7.3	8.6
1	38.2	32.7
2	33.5	37.2
3 or More	20.7^	21.4

^indicates the sample mean is *not* statistically different from the population mean at the 5% significance level.

3.3. Modeling framework and quasi-experimental designs

Our research objectives are twofold. First, we want to determine whether workers engage in more non-commute travel (in terms of number trips and travel distance) on telecommuting days, relative to commute days. Second, we want to determine whether this difference in non-commute travel activity (if it exists) between telecommuting and commute days is additive or substitutional; in other words, are those daily differences the result of newly generated trips, or of people shifting their existing travel from other days of the week to telecommuting days.

While similar questions have been studied before (see Section 2), our quasi-experimental designs and rich panel dataset allow us to overcome key limitations of selection bias and unobserved confounding to which previous studies in the literature have been prone. With repeated observations of the same individuals, often over the course of many weeks, we can control for individual unobserved differences using fixed effects models. Effectively, this means that we sweep out all cross-sectional variation in outcomes and telecommuting status, removing any factors that vary across individuals but are fixed over time, when testing hypotheses about the effect of telecommuting on travel. This stands in contrast to the vast majority of studies in the literature that leverage variation in telecommuting and outcomes across individuals. Our methodological approach and causal designs have been acknowledged and recommended in the transportation literature as being advantageous for estimating causal effects (1, 2).

To determine how non-commute travel changes on telecommute days — the first research objective — we look at full-time employees in our sample and use the passively collected GPS data to compute the following metrics for each individual on every day that they were observed in the data: their total number of daily non-commute trips, their total daily distance traveled, and whether or not they commuted to work on that day. We then regress the number of daily non-commute trips and the total daily distance traveled by a person on an indicator variable denoting whether or not the person commuted to work on that day, as well as a list of control variables. Crucially, we control for individual fixed effects, which allows us to control for unobserved but fixed confounders that affect the mobility outcomes for each individual. We also control for time-variant unobserved individual confounders by interacting the individual fixed effects with the monthly fixed effects; these controls eliminate any individual-specific factors that are trending over time. Finally, we control for various demographic variables (like household income, age, gender, race, and education level), and time fixed effects (for week of year, and day of week). The model specification is explicitly shown in Section 4.2.1.

To test whether non-commute trips on telecommute days are additive or simply shifted from other days — the second research objective — we use the survey data to determine the number of days that each full time employee has worked from home on the week that they took the survey. We use the survey data to infer this information, rather than the passive POI data, because the survey data allow us to explicitly distinguish between days worked at home and vacation or sick days. The downside of the survey data is that we only have this information for one week per wave, or a total of five weeks for individuals who took every survey wave, so the sample size is significantly smaller than the one used to investigate the first research objective.

To quantify the effect of one additional day of telecommuting on the total number of weekly trips, we again use the within-individual variation in the number of days worked from home across successive waves, and correlate that with the change in their total number of weekly trips on those same weeks. Specifically, we implement a first-differences design where we regress the individual differences in the total number of weekly trips taken on weeks corresponding to successive survey waves on the same individual differences in the number of days worked from home. The benefit of this approach is that we only assume that the individual fixed effects remain constant between successive waves, which are roughly three months apart on average.

Alternatively, a possible approach could be to include individual differences on non-successive waves in the regression. The downside of the latter approach is that it makes a stricter assumption that the individual fixed effects remain the same across all waves (from August 2020 till August 2021). By adopting the first-differences approach, we make less strict assumptions but sacrifice some statistical power due to excluding data points from non-successive waves. In Section 4.2.2, we show that the confidence intervals on the estimated model coefficients are sufficiently tight, so the loss of power due to adopting the first-differences approach is not problematic.

4. Results

4.1. Telecommuting during the COVID-19 pandemic

While our data show that the share of full-time workers in our sample who are back to commuting fully to their offices has increased in 2021 (Figure III-2), from a low of 51% in December 2020 to a high of 58% in July 2021 (p-value < 0.001), the fraction of people who telecommute at least once a week remains at unprecedented levels. These numbers suggest that partial telecommuting continues to be an option for a significant portion of workers, and this could remain the norm in a post-pandemic world. This trend will have significant implications on congestion and traffic during what was traditionally thought of as peak hour commute. A recent study by the Pew Research Center supports these findings and gives more insights into the factors affecting the decision to work from home (88). Interestingly, safety does not appear to be the main reason affecting people's decision to telecommute [cite our Vaccine paper], which further suggests that high telecommuting rates may persist in the long term, even as safety concerns continue to decrease in the general public.

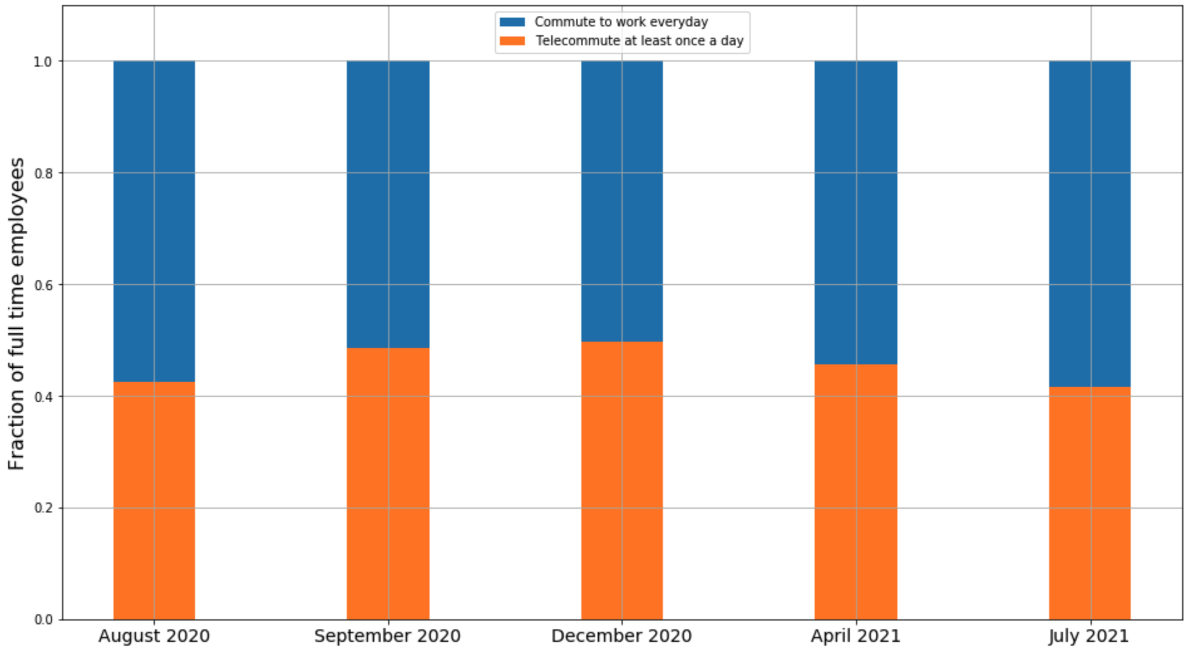


Figure III-2. The frequency of commuting and telecommuting to work for fully employed panelists. We define fully employed as those who report working at least 5 days a week in the past week. The sample size is between 402 and 477 individuals depending on the wave, except for the third wave in December 2021, when the number drops to 328 (likely due to vacation time).

Hybrid telecommuting schedule coordination could have important implications on the degree to which peak hour congestion decreases. For example, if all telecommuters choose to commute to work on Wednesday, then Wednesday peak hour congestion and emissions will be significantly higher than on other days of the week. On the other hand, if telecommuters hypothetically had perfect information about the choices of other telecommuters and were indifferent about which days they telecommuted, then an “equilibrium” could emerge in which the distribution of peak hour travel time is uniform over the days of the week.

Interestingly, our data show that in August 2021 (our most recent wave of survey data), Wednesdays are the most popular commute days within telecommuters (about 18% of telecommuters’ commute days). On the other hand, Mondays and Fridays are the least popular weekdays for commuting to work (13% and 14.5% of all telecommuters’ commute days), and about 10% of commute days happen on Saturday and Sunday. These findings suggest that there will be significant differences in peak hour traffic on different weekdays, and that the differences will be larger as more people telecommute.

Aside from schedule coordination, in the next section, we investigate other transport impacts of telecommuting by quantifying the net effects of telecommuting on trip generation and distance traveled.

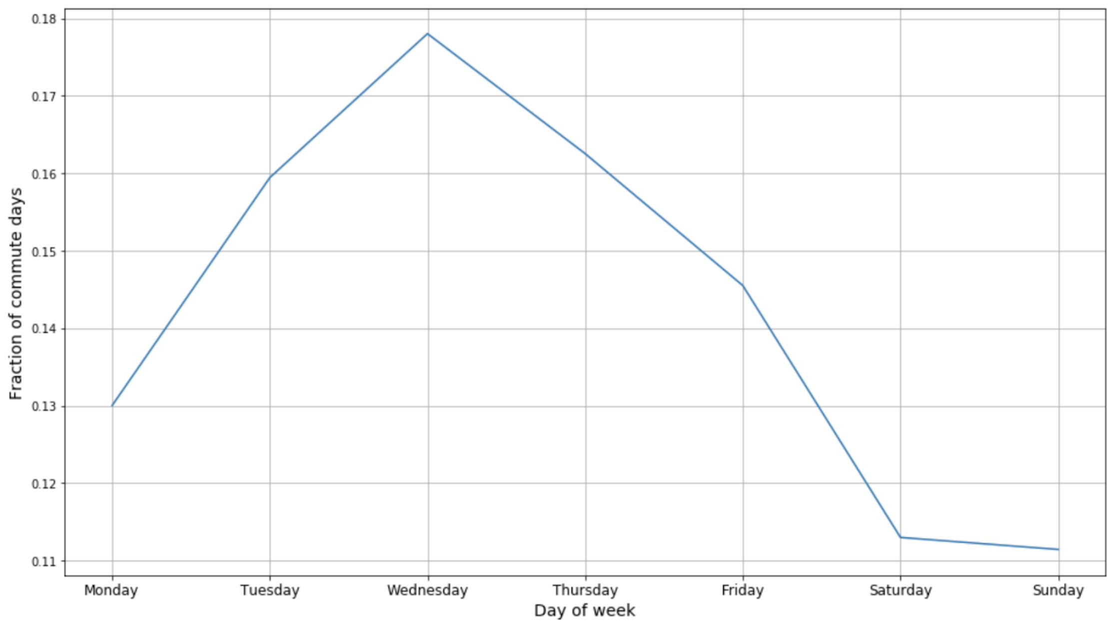


Figure III-3. Weekly distribution of commute days for telecommuters in August 2021.

4.2. Transport impacts of telecommuting

With hybrid work models persisting, we next study the effects of telecommuting on key transport outcomes. This section is split into two subsections: 1) The effect of telecommuting on the number of daily non-commute trips and the total daily distance traveled, and 2) The effect of telecommuting on the number of weekly non-commute trips and the total weekly distance traveled. We implement two different causal empirical designs to answer those questions, the details of which are presented below.

4.2.1. Effect of telecommuting on daily mobility

In this subsection we quantify the effect of telecommuting on the number of daily non-commute trips and the total daily distance traveled on telecommuting days relative to regular commute days. Section 3.3 describes the modeling approach and identifying assumptions in greater detail. Our regression model is shown in equation 1:

$$
\begin{aligned}
Y_{it} = \alpha_i \\
&+ \zeta_t \\
&+ \gamma_{it} \\
&+ \beta * I(t = telecommute\ day) \qquad \text{(eq. 1)}
\end{aligned}
$$

In other words, we assume that the outcome Y_{it} (i.e.: the number of daily non-commute trips, and the total daily distance traveled) for individual *i* on day *t* is a function of individual fixed effects (α_i), time fixed effects (ζ_t), individual effects that vary with time (γ_{it}), and most importantly the effect of telecommuting (β), which is our parameter of interest. To control for time and seasonal trends, we use year and week of year indicator variables. To control for individual fixed effects, we use individual indicators, and to control for individual effects that vary with time, we interact the individual indicator with year and month of year indicators. In other words, we control for fixed unobserved confounders as well as unobserved confounders that could change monthly. This allows for our estimate of the effect of telecommuting (β) to be more robust than others in the literature that use cross-sectional datasets. For the estimation technique, the final model estimated and presented below is a linear regression model with pooled standard errors to account for the correlation structure among observations coming from the same individual. Since the outcome is not continuous and is strictly positive, we also estimated negative binomial regressions and found that the results for the treatment effects (after performing the necessary transformations on the coefficients) were very close to, and not significantly different from, the ones estimated via linear regression. Because of this similarity, the final results presented below are the results of the linear regression since they are more intuitively interpreted than the exponentiated negative binomial coefficients.

Our null hypothesis is that full-time workers make the same non-commute trips on telecommute and commute days. Under this null hypothesis, our estimates of the effect of telecommuting (β) on the mobility outcomes of interest should be equal to zero. While our estimates should be interpreted as the effect of telecommuting for telecommuters, due to the historically high rates of telecommuting during the pandemic and the similarity between the characteristics of our sample and the broader full-time workers population in the U.S. (see Section 3.2.), it is likely that our estimates will generalize to most workers, since workers of various backgrounds, industries, and characteristics have been forced to telecommute during the pandemic, reducing the systematic differences between telecommuters and non-telecommuters that have existed in pre-pandemic datasets.

We estimate the models described in this section for four different time periods: Pre-pandemic (January and February 2020), Pandemic pre-vaccine (March 2020 to December 2020), Early vaccination period (January 2021 to May 2021), and Widespread vaccination period (June 2021 to December 2021). We conduct our estimation in this manner to investigate whether our estimates are: 1) consistent, and 2) have stabilized since the onset of the pandemic. Indeed, during the early phases of the pandemic when lockdowns and restrictions were prevalent, one can reasonably assume that effects on telecommuting-induced non-commute trips are attenuated.

Tables III-3 and III-4 present the estimates of the effect of telecommuting on non-commute trips and total distance traveled, respectively. The results show clearly that telecommuting has a positive relationship with the number of non-commute trips that employees make, with an average of about one additional non-commute trip in the second half of 2021. When it comes to total distance traveled, our estimates show that people travel between 0.12 and 9.8 kilometers more on commute days relative to telecommute days, depending on the analyzed time period. In the pre-pandemic period and the late 2021 pandemic period, the reduction in total daily distance traveled is insignificant, suggesting that the additional non-commute trip's distance is close to the commute distance. During the pandemic, however, the reduction in total distance traveled on telecommuting days is significantly different than zero, suggesting that the additional non-commute trip on telecommute days is shorter than the displaced commute trips, resulting in a net reduction in total distance traveled. This suggests that as the pandemic restrictions started easing and vaccines became more widely available, people started returning to their previous travel behaviors on telecommute days.

Table III-3: Effect of telecommuting on total daily non-commute trips

Time Period	Estimate for β (in trips)	Standard error	p-value
Pre-pandemic	0.90	0.042	<0.001
Pandemic pre-vaccine (Mar-Dec 2020)	0.85	0.024	<0.001
Early vaccination period (Jan-Jun 2021)	0.97	0.028	<0.001
Widespread vaccination period (Jun-Dec 2021)	1.03	0.051	<0.001

Table III-4: Effect of telecommuting on total daily distance traveled

Time Period	Estimate for β (in km)	Standard error	p-value
Pre-pandemic	-0.12	2.8	0.89
Pandemic pre-vaccine (Mar-Dec 2020)	-9.8	0.9	<0.001
Early vaccination period (Jan-Jun 2021)	-6.4	1.4	<0.001
Widespread vaccination period (Jun-Dec 2021)	-3.5	2.7	0.20

4.2.2. Effect of telecommuting on weekly mobility

To investigate whether the increase in the number of non-commute trips on telecommute days is the result of employees shifting existing trips from other days of the week to telecommuting days, in this section we quantify the effect of one additional day of telecommuting per week on the total number of weekly non-commute trips. If employees are only shifting their non-commute trips to telecommuting days instead of engaging in additional activities, then the number of weekly non-commute trips should not change in response to an increase in weekly telecommuting days. Consequently, our null hypothesis is that one day of extra telecommuting results in no additional weekly non-commute trips.

Equation 2 presents our model to estimate the effect of an additional day of telecommuting per week on the number of weekly trips (refer to Section 3.3 for more details on the identification strategy):

$$\begin{aligned} Y_{it} = \alpha_i & \\ + \gamma_t & \\ + \beta_e * \ \# telecommute\ days_{it} & \qquad (eq.\ 2) \end{aligned}$$

Where Y_{it} is the outcome (the number of weekly trips or the total weekly distance traveled for employee i at survey wave t (t = 1, 2, 3, 4, 5)), α_i is the individual fixed effect for individual *i*, γ_t is the time fixed effect at wave *t*, and β_e is the effect of one additional day of telecommuting. Stated differently, we assume that the number of trips that an employee makes on a given week and their total weekly distance traveled are a function of fixed individual specific characteristics that do not vary with time, fixed time effects that do not vary across individuals, and the individual's number of weekly telecommute days. As mentioned in Section 3.3, one way to estimate β_e is through a regression using the full set of observations, i.e. using all employees' data across all survey waves. However, this approach makes the strong assumption that individual fixed effects (β_i) are unchanged across all survey waves, from August 2020 through August 2021. To avoid making this assumption, we instead take the first differences of the regressor and outcome of interest (i.e. the differences in Y_{it} and $\#\, telecommute\ days_{it}$) for successive waves only, and regress the first differences in Y_{it} on the first differences in $\#\, telecommute\ days_{it}$, while controlling for time fixed effects. By doing so, we instead assume that the individual fixed effects remain unchanged only between successive waves, but can vary between non-successive waves. The simplified regression equation that we estimate is shown in equation 3:

$$\begin{aligned}
\Delta Y_{it} = \beta_0 &* I(Wave\ 2 < t < Wave\ 1) \\
&+ \beta_1 * I(Wave\ 3 < t < Wave\ 2) \\
&+ \beta_2 * I(Wave\ 4 < t < Wave\ 3) \\
&+ \beta_3 * I(Wave\ 5 < t < Wave\ 4) \\
&+ \beta_e * \Delta\, \#\, telecommute\ days_{it} \qquad \text{(eq. 3)}
\end{aligned}$$

where ΔY_{it} is the difference in the number of weekly trips taken (or the total weekly distance traveled) by individual *i* between two successive waves. Note that the individual fixed effects cancel out, and only four time effects are estimated, corresponding to the first differences in the five survey waves. Table III-5 presents the results of the model estimation for β_e, which quantifies the effect of an additional day of telecommuting per week on the number of weekly non-commute trips or the weekly distance traveled. The results show that each additional day of telecommuting is associated with an average increase of 1.07 weekly trips. More importantly, the estimate is precise enough to rule out the possibility that telecommuting induces no additional trips: the 95% confidence interval on the coefficient excludes zero, leading to a rejection of our null hypothesis. These results confirm that the additional trips on telecommuting days are not trips the individuals would have made regardless of whether or not they telecommuted, and instead are newly generated trips.

The results also show that an additional day of telecommuting results in an average reduction of 15.5 kilometers in total weekly traveled distance. While our estimate's confidence interval is wide, the estimate is significantly different from zero, which suggests that the additional telecommuting induced trip is shorter than the commute trip, and the added non-commute travel distance does not fully offset the reduction in commute distance associated with telecommuting.

Table III-5: Effect of one day of telecommuting on total weekly non-commute trips and total weekly distance traveled

Parameter	Estimate	Standard error	95% Confidence interval
β_e (for number of weekly trips)	1.07	0.33	[0.42, 1.73]
β_e (for total weekly distance traveled)	-15.50	4.90	[-25.1, -5.9]

5. Conclusion

In this work, we contribute to the literature on the transport impacts of telecommuting and provide insights on their effects on inducing travel. We leveraged a unique dataset of passive POI data and five waves of actively collected stated preference survey data on a panel of individuals in the U.S., and implemented empirical quasi-experimental designs to quantify the causal effects of telecommuting on daily and weekly trips and travel distance. Our data and designs allow us to control for individual unobserved fixed and time-variant confounders by leveraging within-individual variation in telecommuting and trip-making frequency over several weeks in 2020 and 2021. Our results can be summarized by the following key takeaways:

- Individuals who telecommute make an average of about one non-commute trip on telecommuting days. The additional distance traveled on this trip is however shorter than the two-way commute distance, as individuals travel significantly shorter distances on telecommuting days relative to commute days.
- The additional non-commute trip that individuals make on telecommuting days is a newly generated trip, not a trip that has been shifted from other days of the week. The net effect of one additional day of telecommuting per week on weekly distance traveled is also negative, confirming that the newly generated non-commute travel distance does not fully offset the reduction in the two-way commute distance.

Our work confirms previous findings that telecommuting-induced trip generation partially offsets the reduction in commute trips (76, 77), and that the travel reduction effects of telecommuting will be overestimated if the new trip-generation is not taken into account. Nevertheless, the net effect of telecommuting on total distance traveled remains negative, meaning that the additional distance traveled during the newly generated trip is shorter than the two-way commute distance. Our findings are more robust and provide stronger evidence on the travel-inducing impacts of telecommuting than previous studies in the literature, by overcoming the unobserved confounding limitations that are often present in cross-sectional observational studies, and which have been acknowledged by previous studies (84).

In future extensions, researchers should further investigate any modal differences between telecommuting-induced non-commute trips and commute trips in order to more accurately determine the environmental impacts of telecommuting. Comparing the different timing of those trips will also prove important in determining their environmental footprints: trips made during peak congestion hours are more environmentally costly than off-peak trips. Finally, researchers should extend the existing literature on the relationship between telecommuting and home-to-work distance. The COVID-19 pandemic has resulted in a natural experiment where telecommuting rates have increased across the board: this presents an opportunity to deal with the endogeneity between telecommuting decisions and home location / distance from work. Researchers can take advantage of this emerging context to shed light on this long-standing question in transportation research.

6. Acknowledgement

I would like to thank my co-authors on this paper, Mohammed Amine Bouzaghrane, and professors Joan Walker and Michael Anderson. I would also like to thank the Embee Mobile team for providing the data that made this research possible.

Chapter 3: Early pandemic behaviors and the role of vaccines in reversing pandemic mobility trends: Evidence from a U.S. panel

Executive Summary

The COVID-19 pandemic has disrupted travel behavior and resulted in the emergence of new mobility trends. In this paper, we study the degree to which vaccines played a role in reversing pandemic-induced travel behaviors and contributing to a "return-to-normal". Using five waves of original U.S.-based survey data combined with passive smartphone tracking data, we show that in the early phases of the pandemic, the behavioral response of people in the U.S. was heterogeneous: individuals with low levels of concern about being infected with COVID-19 engaged in riskier behaviors than those with higher levels of concern, such as traveling more, eschewing masks, attending large gatherings, and using public transportation. These differences in behaviors persisted throughout 2020, prior to widespread vaccine availability. Once vaccines became available in early 2021, many of the behaviors changed. We find that getting vaccinated significantly reduced concern about contracting and spreading the disease, and significantly increased mobility and travel frequency. We also show that getting at least partially vaccinated increases the frequency of using public transportation. When it comes to working from home, we find that telecommuting at least once a week remained high in 2021, despite the significant decrease in the fraction of full-time employees that work from home every day compared to 2020 during the peak of the pandemic. Unlike our findings for mobility and transit, getting vaccinated did not have a significant effect on employees' decisions to return to office, which suggests that the decision to return to in-person work is not primarily driven by employees' safety concerns, and is instead a function of employers' expectations and their decisions to reopen their offices. We discuss the implications of our findings on understanding travel behavior during the pandemic impact and recovery periods.

1. Introduction

The COVID-19 pandemic has caused unprecedented disruptions to many aspects of everyday life. Prior to the development of vaccines, government and public officials relied on non-pharmaceutical interventions (NPIs) and health recommendations to control the disease’s spread, including recommending or mandating mask-wearing, avoiding large gatherings, reducing travel, and sheltering in place. Along with widespread office closures and businesses transitioning to working from home, these measures resulted in significant impacts on the transportation system. Despite the extensive literature on the changes in travel behavior during the pandemic, research on whether vaccination changes pandemic-induced mobility behaviors has been limited in scope. Our work fills this gap in scholarship and makes several contributions.

First, we examine the question of whether travel behavior and compliance with public health recommendations in the early phases of the pandemic are heterogeneous. Using five waves of original U.S.-based survey data, we analyze individuals’ reported degree of compliance with mask-wearing recommendations and avoidance of large gatherings in the early phases of the pandemic preceding vaccine availability. We also use passively collected Point of Interest (POI) data of our panelists and examine whether compliance with public health measures is correlated with the frequency of travel. We then examine whether concern about contracting COVID-19 determines compliance with public health measures and reduces travel activity. Next, using a series of difference-in-difference designs, we examine the role of vaccines in reducing concern about contracting COVID-19 and reversing pandemic-induced travel behaviors. We focus on three main questions: whether getting vaccinated increases individuals' travel frequency, increases public transportation use, and encourages a return to in-person work. Our analysis contributes important insights on understanding human behavior and mobility during pandemic impact and recovery periods.

The rest of this manuscript is organized as follows: Section 2 contains a review of the literature to which we contribute. Section 3 introduces the data we used in this analysis. We present our findings in Section 4, which contains two subsections: 1) early pandemic compliance behaviors, and 2) the effects of vaccination on personal mobility, transit use, and telecommuting. We finish with a conclusion and discussion of the policy implications of our findings in Section 5.

2. Mobility and human behavior during the COVID-19 pandemic

Prior to the development and availability of vaccines, public health officials focused on non-pharmaceutical strategies to manage the spread of COVID-19. Multiple studies have examined the effectiveness of NPIs in reducing the risk of contracting COVID-19 and the severity of symptoms. Early evidence from China showed that NPIs effectively reduced the serial interval of SARS-CoV-2, the virus that causes COVID-19, allowing for quicker identification and isolation of cases(40). Another study attempted to rank the effectiveness of NPIs and found that banning public gatherings, closing entertainment centers, and shutting down citywide public transport were the most effective at slowing the early spread in China(41). Other work found that reducing the number of daily contacts from 14-20 down to 2 contacts per person effectively reduced the reproduction number below the epidemic threshold(42), and that adherence to NPIs and behavioral mitigation measures strongly predicts mortality(43). Despite this evidence, public opinion remained divided on the appropriateness of public health recommendations. In the U.S., agreement with and support of public health measures has been highly politicized, with multiple studies showing that political ideology is a strong predictor of policy preferences related to NPIs(44, 45). This has led to varying levels of compliance and adherence to public health recommendations(46). Our work joins the literature on the heterogeneity of compliance with NPIs and examines its implications on the transportation system.

We also contribute to the literature on the travel behavior impacts of the pandemic and the role that vaccines played in reversing some of the pandemic-induced mobility trends. Multiple studies have examined the initial impacts of COVID-19 on travel behavior(47–50). Other work proposed and illustrated an approach for policy making in a post-pandemic world that accounts for future public health threats(51). Some research, relying upon individuals' reported expectations about their post-pandemic behavior, has argued that the increase in telecommuting, decrease in public transit use, and decrease in travel are likely to persist in a post-pandemic world(52).

On the telecommuting front, researchers found that workers would like to maintain their ability to telecommute after the pandemic is no longer a threat. Research conducted by the Pew Research Center shows that, for individuals whose work responsibilities can be completed from home, the fraction of telecommuters increased from 20% pre-pandemic to 71% during the pandemic. Perhaps more interestingly, half of those individuals expressed their desire to continue telecommuting post-pandemic, suggesting a continuation of the pandemic-induced telecommuting trend(53). Another study showed that the fraction of workers who expected to telecommute at least a few times a week post-pandemic was 26%, double that of pre-pandemic levels(54). To the best of our knowledge, no study has attempted to investigate whether getting vaccinated had an effect on encouraging employees to return to in-person work. We fill this gap.

Like telecommuting, public transportation has been highly affected by COVID-19 and is the most heavily impacted transport mode during the pandemic. For example, public transit ridership plummeted by more than 90% in New York City in March 2020(55), and similar declines have been observed throughout the U.S. and globally. A study in Sweden found that transit ridership decreased by 40-60% even when transit service reductions were minor(56). In Japan, a study found larger decreases in public transit mode use relative to car travel in three out of four studied cities(57). Some studies found that the higher fear of infection and overcrowding on public transit are key drivers of the reduction in demand, with those who perceive public transit as having a higher infection risk being less likely to use it(58, 59). One study suggested that the impact on the public transportation system is likely to persist until confidence in those modes is restored(60).

Despite the extensive literature on pandemic mobility, research on whether vaccination changes pandemic-induced mobility behaviors remains limited in scope. A study in Israel found that for tourism related travel, vaccines did not necessarily affect individuals' desire to travel(61). The study has limitations, however: it relies on stated intentions to travel in the future, and used three waves of independent cross-sectional data, meaning that participants varied significantly between different waves. Another study focusing on the tourism industry found that getting vaccinated increases the probability of taking a holiday trip by 8.3 percentage points in the general population and 11.3 percentage points for vaccinated individuals(62). While the study attempted to provide causal evidence on the effect of vaccination, its use of cross-sectional data raises concerns about confounding, even when controlling for observable differences between the vaccinated and unvaccinated groups using regression adjustment and propensity score weighting. Our work expands this literature by estimating quasi-experimental designs and quantifying the effect of vaccines on weekly mobility, transit use, and telecommuting.

3. Data

3.1. Data Collection

We used a combination of data collection methods to develop a database that enables a broad understanding of how COVID-19 has affected people's travel behavior. The analysis uses a panel dataset consisting of five waves of surveys and a total of 1,962 unique respondents. To construct our sample, we began with a geographically representative panel of approximately 100,000 smartphone users tracked by Embee Mobile. We recruited from a random sample stratified by U.S. region of approximately 15,000 members of this panel to conduct five waves of surveys (August 2020, October 2020, December 2020, April 2021, July 2021). In the first wave, 1,321 individuals responded to the survey — a response rate of about 9%. We then targeted those same respondents in the subsequent survey waves, and supplemented our sample from the original pool of 15,000 members to account for drop-outs while maintaining a sample size of approximately 1,000 participants in each wave. Our survey included questions regarding respondents' employment status, travel and telecommuting behavior, vaccination status, demographic characteristics, and ideological beliefs.

A key advantage of our dataset is the ability to link the survey responses to passive location tracking POI data. This allowed us to supplement the stated responses of the survey respondents with their observed travel behavior, and to compute mobility metrics at different temporal aggregation levels (e.g., number of daily and weekly trips), which we analyzed in conjunction with the survey responses to understand how different demographic and ideological groups traveled during the pandemic. An overview of the study timeline is shown in Figure II-1. The passive data is available from January 2020. The survey data collection started in early August 2020 and collected data in five total survey waves, ending in July 2021.

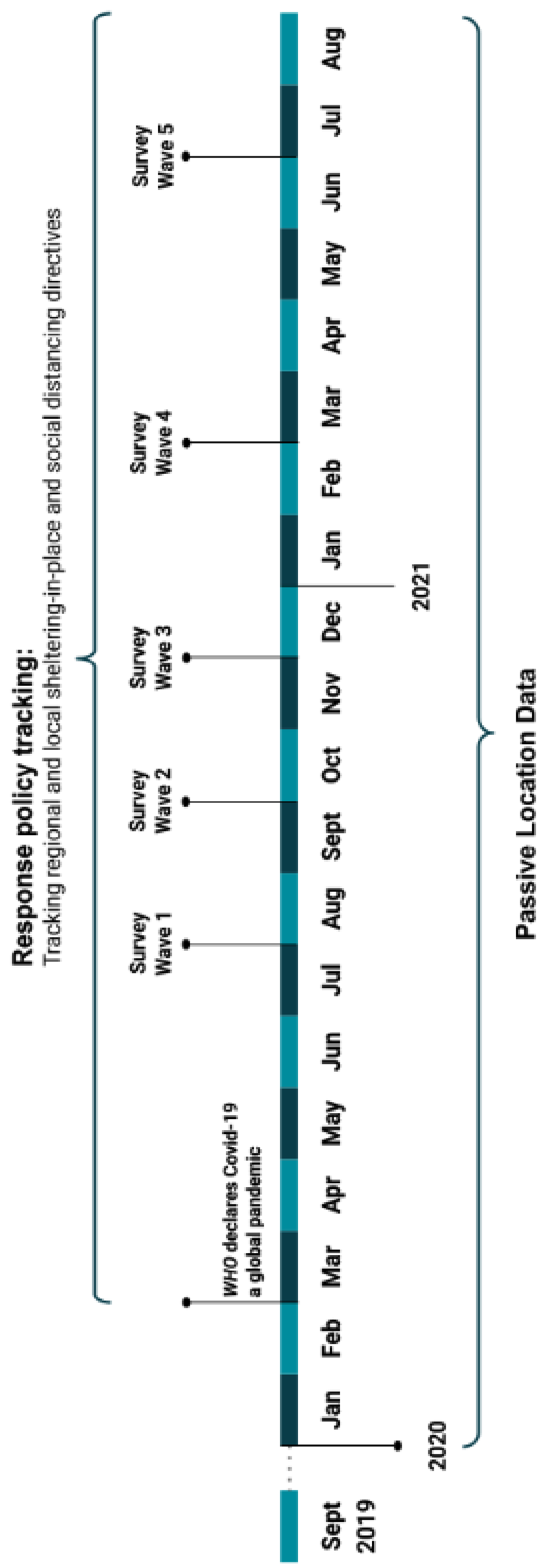


Figure II-1: Data collection timeline for passive and survey data.

3.2. Demographic Summary

Table II-1 summarizes the demographic characteristics of our survey respondents and compares those characteristics to national statistics from the U.S. census. When compared to the U.S. population, our data oversamples men and undersamples women. Our data also significantly oversamples young adults aged between 25 and 54 years old: this age group represented about 69 to 73% of survey participants (depending on the wave), compared to 39% in the U.S. population. When it comes to household income, our sample overrepresents households with low- and medium-income levels and underrepresents high-income households: about 59 to 65% of the survey respondents earned $50,000 or less a year, compared to only 40% in the U.S. population. Similarly, about 10 to 11% of our respondents belong to households that earned $100,000 or more a year, compared to 31% of households in the U.S. population. Our sample also underrepresents white Americans and overrepresents other groups. Caucasians represent approximately 72% of the U.S. population compared to approximately 52-56% of our respondents across waves. Black respondents represent about 13% of the U.S. population but comprise 18% of our respondents, across all five waves. Our sample underrepresents non-Hispanics/Latinos when compared to the U.S. population, with 75-77% of our sample identifying as a non-Hispanic/Latino compared to 82% of individuals in the U.S.

Despite the non-representativeness of some characteristics in our sample, we do not attempt to re-weight observations to match the U.S. population, since our analysis is comparative in nature: we investigate the behavior of specific groups relative to other groups, and we do not attempt to present aggregate statistics that are nationally representative.

Table II-1. Demographic characteristics of study participants compared to the U.S. Census

Category	Wave 1 (%) (Aug. 2020)	Wave 2 (%) (Oct. 2020)	Wave 3 (%) (Dec. 2020)	Wave 4 (%) (Apr. 2021)	Wave 5 (%) (Jul. 2021)	Population (%)
Sample Size	1,333	1,110	810	983	842	-
Gender						
Male	40.8	40.4	40.5	43.9	45.2	49.2
Female	59.2	59.6	59.5	56.1	54.8	50.8
Age						
19 years and under	3.8	3.8	3.5	4.3	3.9	25.3
20 to 24 years	9.7	9.0	8.1*	10.2	9.4	6.8
25 to 34 years	23.7	22.4	23.0	23.7	24.1	13.9
35 to 44 years	27.4	26.1	26.8	27.2	27.0	12.6
45 to 54 years	20.2	21.2	22.0	20.4	21.3	13.0
55 to 59 years	6.7*	7.3*	6.7*	5.7*	5.6*	6.7
60 to 64 years	4.4	4.7	4.7	4.0	3.6	6.2
65 years and over	4.3	5.5	5.5	4.6	5.2	15.6
Household Income						
$0 - $24,999	35.1	31.8	30.4	30.0	27.8	19.3

$25,000 - $49,999	30.1	32.5	30.8	28.7	29.1	21.2
$50,000 - $99,999	24.5	25.4	27.3*	30.6*	31.1*	29.9
$100,000 - $149,999	6.9	6.4	7.6	6.7	7.6	15.1
$150,000 - $199,999	1.4	2.1	2.1	2.0	2.1	6.8
$200,000 or more	1.8	1.7	1.8	1.9	2.1	7.7
Race						
Asian or Pacific Islander	7.4*	7.7	8.4	10.7	11.0	5.7
Black/African American	18.1	18.9	17.4	18.4	18.1	12.7
Mixed Race	6.5	6.4	5.7	7.1	6.8	3.3
Native American/Alaskan Native	2.5	2.8	3.1	1.9	1.9	0.9
White/Caucasian	54.5	54.7	56.5	52.2	53.3	72.5
Other	8.0	7.0	6.5*	6.3*	6.9	4.9
Hispanic Status						
Hispanic or Latino	20.3*	18.9*	18.8*	18.8*	19.5*	18.0
Not Hispanic or Latino	75.8	77.9	77.5	77.0	77.3	82.0
Education Level						
Less than High School	3.5	3.1	2.8	3.8	2.9	10.1
High School	46.4	46.2	45.8	44.8	42.8	51.5
University/College	43.8	44.4	44.7	44.9	47.3	27.5
Post-graduate Education	6.1	6.2	6.5	6.4	7.1	11.0
Household Size						
1	15.1	16.3	14.2	15.4	14.8	28.0
2	25.2	25.8	26.3	26.3	26.3	34.0
3	20.8	21.5	24.1	20.4	23.1	15.6
4	18.3	16.2	16.8	19.6	18.7	13.0
5	10.6	11.2	9.6	8.4	7.7*	6.0
6+	9.8	8.4	8.4	9.2	8.4	2.3

*indicates a statistic representative of the U.S. population at the 5% significance level

4. Results

4.1. Early pandemic behavioral compliance and mobility

Public opinion has been divided about the severity of the virus and the need for the implemented NPIs such as wearing protective masks and avoiding large gatherings ever since COVID-19 was declared a global pandemic(44). Our data show a clear difference in compliance with NPIs among our study participants, with a small fraction of individuals not complying with public health recommendations. The risk of contracting and spreading the disease by this minority is further compounded by their engagement in multiple risky behaviors at the same time, as we show in this section.

Firstly, when it comes to mask-wearing, our data show that the majority of respondents followed public health recommendations. Data from the first survey wave (August 2020) reveal that, when indoors (excluding one's own home), around 75% of participants reported always wearing masks, while 11% reported only sometimes or never wearing masks. When socializing with others, 49% of participants reported always wearing masks, while 30% reported only sometimes or never wearing masks. When traveling not in one's own private vehicle, approximately 52% of respondents reported always wearing masks, while 13% reported only sometimes or never wearing masks. A similar pattern appears with the attendance of large gatherings; our first survey wave data show that while the majority of respondents avoided large gatherings as recommended by public health officials, approximately 18% reported attending a gathering of 10 or more people in the previous two weeks.

More importantly, we find that those who do not follow one public health recommendation were also less likely to follow others. For example, in August 2020, individuals who attended large gatherings with more than 10 people were five times more likely to never wear masks while socializing with other people, and about four times more likely to never wear masks while traveling in public (p-value < 0.01 for both). These individuals also made more trips, averaging 17.7 weekly trips in August 2020, compared to 14.8 trips for those who did not attend large gatherings. This trend holds in the October 2020 and December 2020 survey waves (Figure II-2).

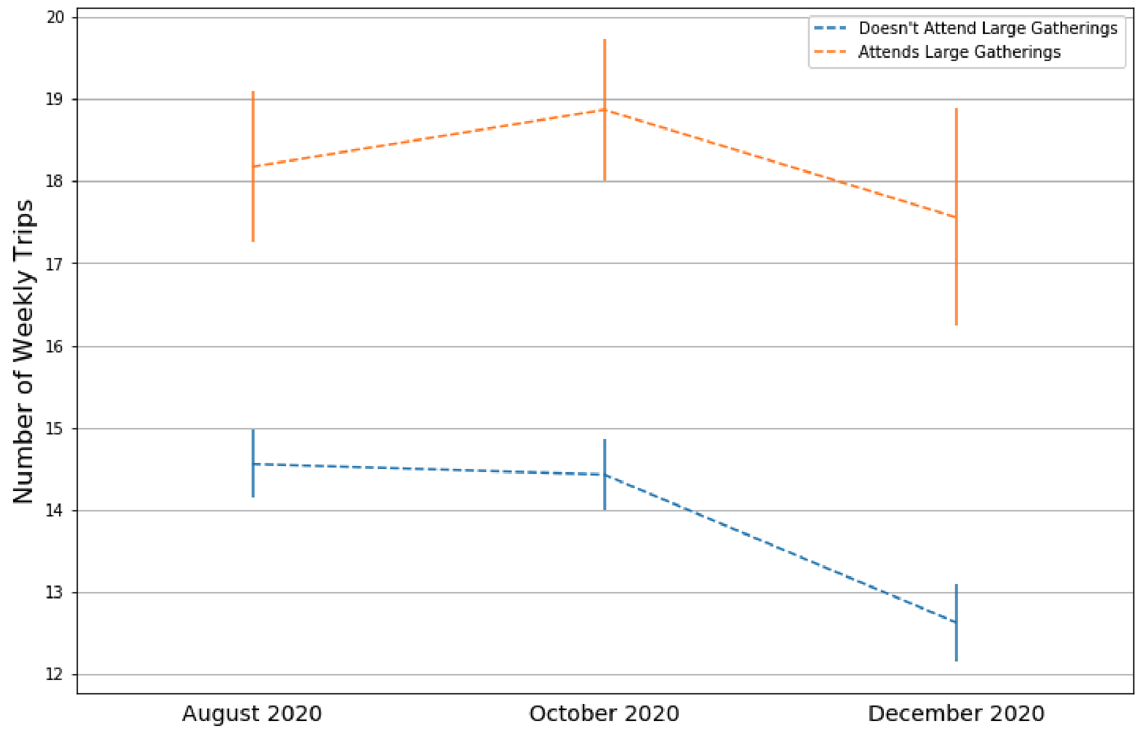


Figure II-2: The number of weekly trips (with 95% confidence intervals) exhibited by people who do and do not attend gatherings larger than 10 people.

These findings suggest that, similar to the spread of the disease, the behavioral risk during the initial phase of the pandemic was also overdispersed: a small share of individuals engaged in multiple risky behaviors and exhibited more mobility compared to the rest of the population. This suggests that human behavior can explain, at least in part, the overdispersion in the spread of the disease. This insight has important policy implications, specifically around the effectiveness of behavioral interventions, that we discuss further in Section 5.

Similarly to NPIs, we find that opinions about the effectiveness, safety, and willingness to get vaccinated varied across segments of the population. When asked about their plans and willingness to get vaccinated, approximately 29% of participants reported that they intended to receive the COVID-19 vaccine once available, while another 28% of respondents reported being unwilling or unlikely to get vaccinated (the remaining participants either declined to answer or were unsure about their decision). We also found an association between vaccine hesitancy and engaging in risky behaviors. The fraction of individuals unwilling to get vaccinated was significantly larger among those who did not follow public health measures: 35% of those unwilling or unlikely to get vaccinated reported either not wearing masks where recommended or attending gatherings larger than 10 people. Furthermore, 50% of those who attended gatherings of 10 people or more and did not wear masks reported an unwillingness to get vaccinated. In short, those who followed public health measures and significantly reduced their mobility are also significantly more likely to get vaccinated.

An important predictor of people's compliance with public health measures, mobility levels, and willingness to get vaccinated in our data was their level of concern about contracting COVID-19 and the severity of the disease. We find that individuals with lower levels of concern about contracting the virus were significantly more likely to engage in riskier behaviors, such as less frequent mass wearing where recommended by public health officials and more frequent attendance of large gatherings. Additionally, these individuals exhibited significantly higher mobility relative to the rest of the population. Figure II-3 illustrates these patterns. These findings add to the literature that found that risk perception and functional fear of COVID-19 are important predictors of behavioral compliance with COVID-19 safety measures(63, 64). Our work has significantly more power due to larger sample sizes relative to the two previous studies in the literature.

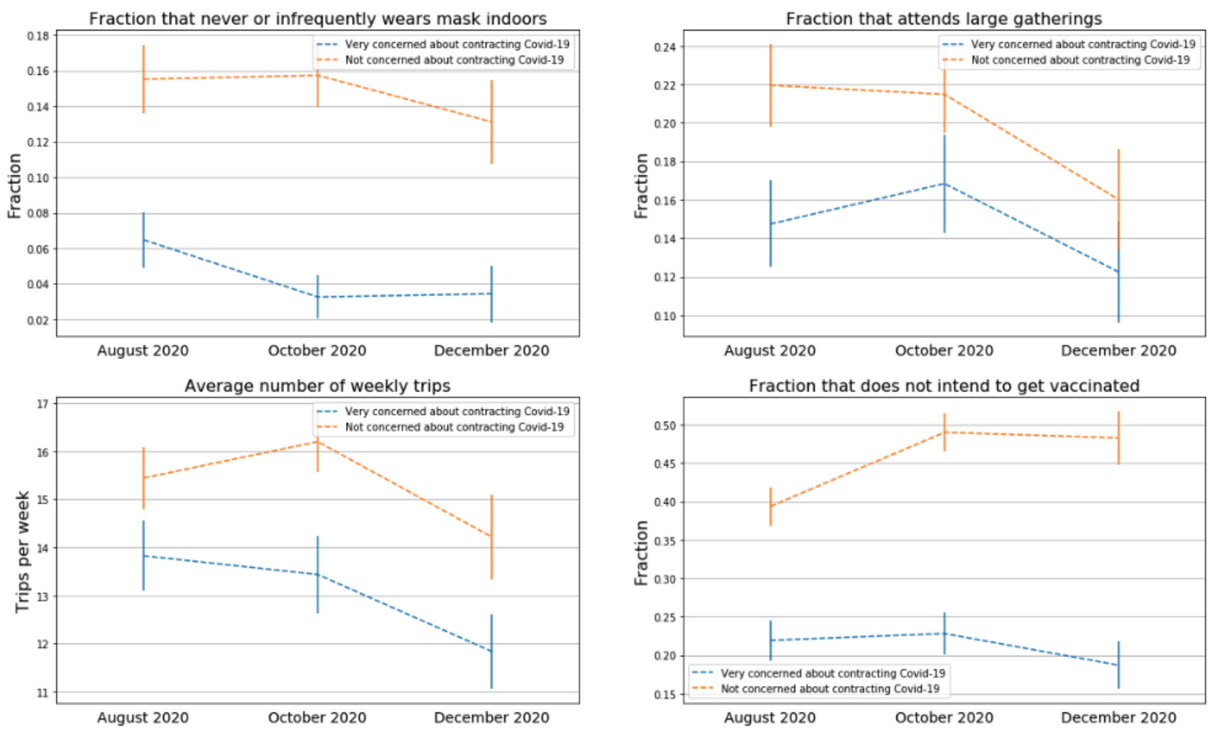


Figure II-3: Comparing the behaviors of participants who are concerned (in blue) vs. not concerned (in orange) about contracting COVID-19. The behaviors compared are (from top left to bottom right): Not wearing masks indoors, attending gatherings with more than 10 people, number of weekly trips taken, and unwillingness to get vaccinated (asked before vaccines became available). Lines at the end of each bar are 95% confidence intervals.

4.2. Effects of vaccination

4.2.1. On mobility

Our initial survey waves establish that following public health recommendations and mobility were, in the early phases of the pandemic, strongly correlated with individuals' level of concern about contracting COVID-19. As vaccines became the primary tool to combat the pandemic in 2021, a natural question that arises is how getting vaccinated affected the degree to which an individual worried about contracting the disease. And, if levels of concern change, how did these changes affect travel behavior?

To answer those questions we examined the responses of individuals who got vaccinated by the end of the April 2021 wave of our survey, in which roughly 20% of survey respondents were vaccinated. We limited the analysis to those who got vaccinated by April 2021 in order to have a more well-defined treatment period (including those who were vaccinated in the August 2021 wave would make the treatment period eight months long, with no precise information on when the treatment occurred). Our results show a significant increase in the share of individuals who were not concerned about being infected with COVID-19 among those who received a vaccine (Figure II-4), from 25% before getting vaccinated (in the December 2020 survey wave) to 41% after getting fully vaccinated, an increase of about 64% (p-value < 0.01). In other words, getting vaccinated has equalized the level of concern about contracting COVID-19 between vaccinated and unvaccinated individuals. Also worth noting is the decrease in concern among those who were not fully vaccinated by the April 2021 wave, although the decrease was much smaller in magnitude compared to vaccinated individuals (Figure II-4). This reduction in concern for unvaccinated individuals likely followed the change in public narrative regarding the pandemic that happened around the time vaccines became available, when government officials adopted a noticeably more optimistic tone(65).

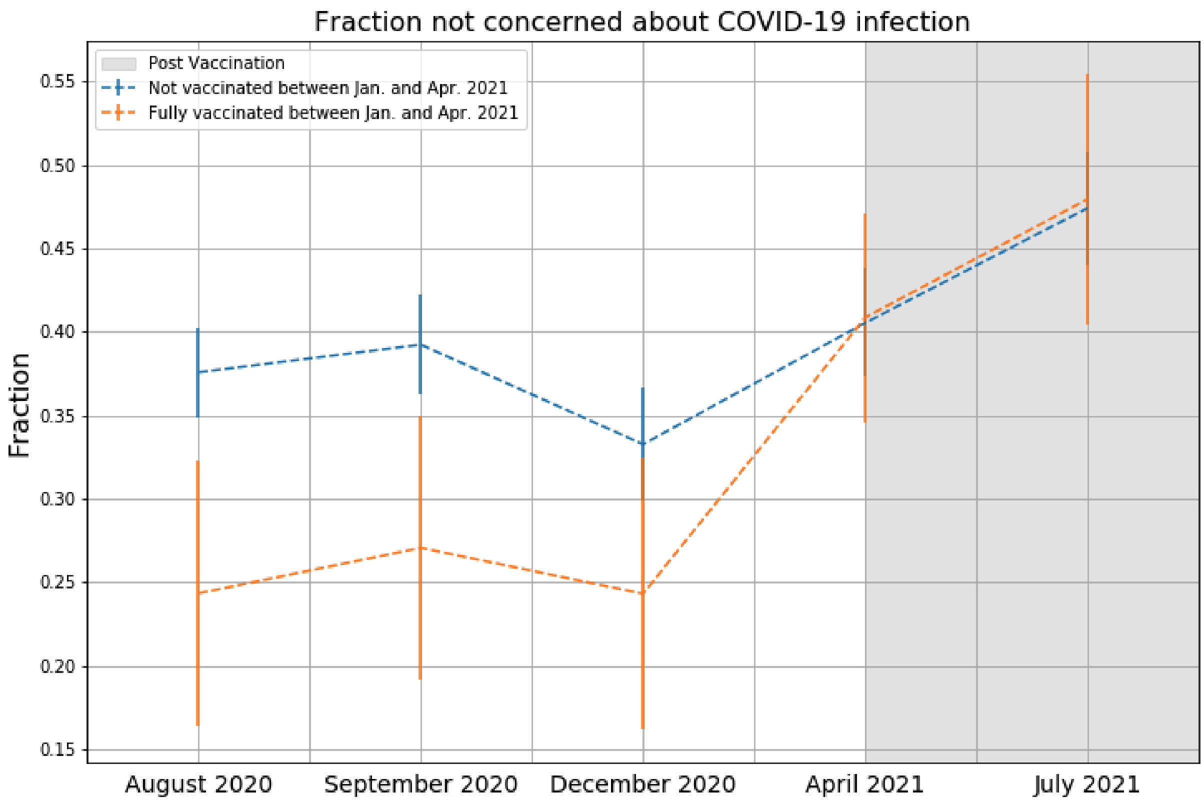


Figure II-4. The evolution of the level of concern for those early vaccinated and unvaccinated individuals. Sample sizes: Vaccinated = ~190 panelists, Not (or partially) vaccinated = ~ 800 panelists.

In light of this reduced concern about contracting COVID-19 and the previously shown association of this concern with mobility levels, we investigated the effects of vaccines on personal mobility by focusing on survey participants who were fully vaccinated as of the April 2021 survey wave. This corresponds to the first wave of surveys following vaccine availability in the U.S. Figure II-5 shows the evolution of the number of weekly trips taken by those who got fully vaccinated before May 2021, relative to the rest of the population. Before the onset of the pandemic, in February 2020, the two groups of individuals exhibited similar mobility levels. The Figure II-shows that participants who got vaccinated before May 2021 reduced their mobility more in the early phases of the pandemic. Once vaccinated, however, these participants increased their weekly trips at a significantly faster rate compared to the unvaccinated.

To quantify the net effect of getting vaccinated on weekly mobility, we estimated a difference-in-differences regression where the treatment is the individual's vaccination status as of the April 2021 wave of our survey and the outcome is the number of weekly trips carried out by an individual, as inferred from the passive mobility data. Our key identification assumption is the parallel trends assumption, i.e. that travel activity would have evolved similarly in both groups absent the vaccine rollout. Figure II-5 suggests that the parallel trends assumption holds, as the two groups' average weekly trips track closely prior to January 2021. The grayed-out area is the treatment period between

January 1st 2021 and April 30th 2021, corresponding to the early vaccination period in the U.S. The posttreatment period is the four weeks after the end of the treatment period (May 2021), and the pretreatment period is the four weeks before the treatment period (December 2020). Equation 1 shows our structural equation specification: the outcome, , is the average number of weekly trips for group *g* at time *t*, where *g* is the treatment status (vaccinated in wave 4 vs. unvaccinated in wave 4), and *t* is the treatment period (pre-treatment vs. post-treatment). Table II-2 summarizes the results of the regression. We find individuals who got vaccinated increased travel by 4.8 trips per week (β_2+β_3) during the post-treatment period, on average, relative to the pretreatment period, while the unvaccinated increased travel by only 1.8 trips per week (β_2) during the same period. The difference-in-differences estimate (β_3) is 3 trips per week (p-value < 0.001), or 170% of the increase for the unvaccinated. This result provides strong evidence that people significantly changed their travel habits after getting vaccinated, suggesting that vaccines play a significant role in the return to normal for vaccinated individuals.

$$
\begin{aligned}
Y_{gt} = \beta_0 & \\
& + \beta_1 * I(g = Vaccinated\ in\ Wave\ 4) \\
& + \beta_2 * I(t = Post\ Vaccination) \\
& + \beta_3 * I(g = Vaccinated\ in\ Wave\ 4) * I(t = Post\ Vaccination) \qquad \textit{(eq. 1)}
\end{aligned}
$$

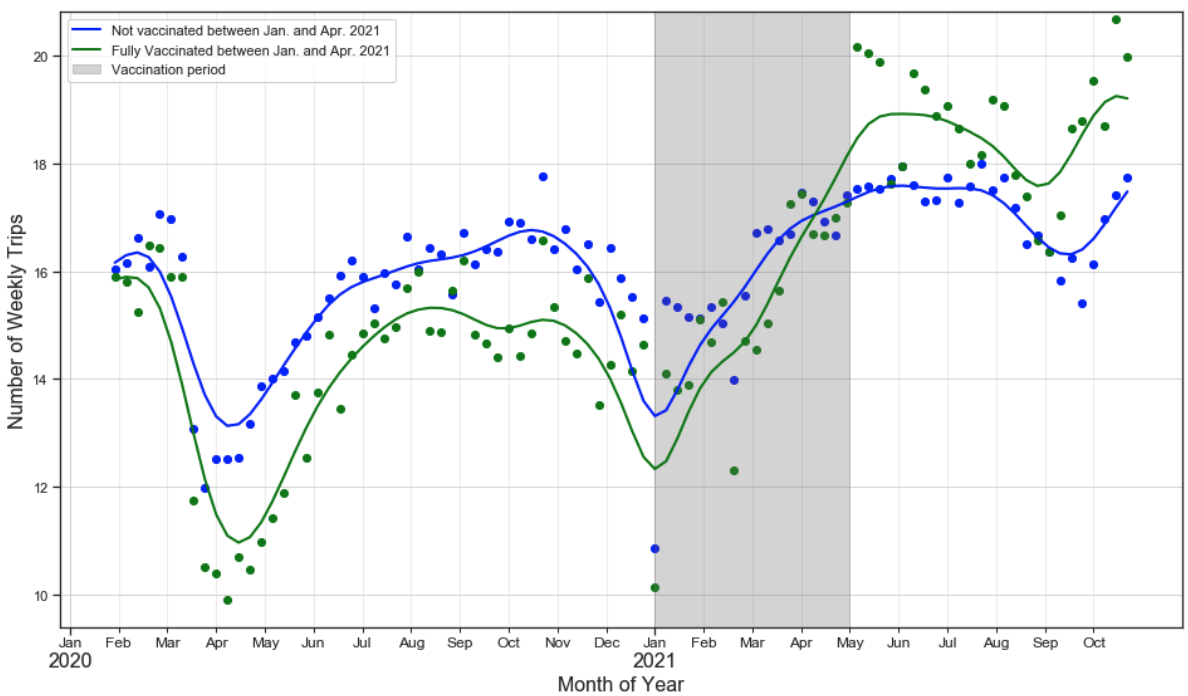


Fig. 5. The evolution of the number of weekly trips for vaccinated and unvaccinated individuals. Sample sizes: Vaccinated = ~190 panelists, not (or partially) vaccinated = ~ 800 panelists.

Table II-2. Estimated coefficients for the difference-in-differences regression in equation 1.

Parameter	Estimate	Standard error	p-value
β_0	15.7	0.208	< 0.001
β_1	-1.2	0.588	0.045
β_2	1.8	0.319	< 0.001
β_3	3.0	0.849	< 0.001

Aside from the vaccine effect, Figure II-5 also shows that mobility has significantly increased for non-vaccinated individuals, albeit at a slower rate than vaccinated individuals (as shown by the positive and statistically significant). This coincides with the significant reduction in concern among unvaccinated individuals shown in Figure II-4. The magnitude of the change in concern highly correlates with the change in mobility as well, with a smaller decrease in concern reflected in a smaller increase in the number of weekly trips. This presents a challenge for public health officials, as it is harder to motivate people to avoid risky behaviors and engage in preventative measures when perceived risk drops.

4.2.2. On transit use

In our first survey wave (August 2020), when asked about factors that would encourage their increased use of transit, about 39% of pre-pandemic transit riders indicated that the development of an effective COVID-19 treatment or vaccine would increase their use of public transportation. Data from our post-vaccine availability survey waves confirm that vaccines indeed play an important role in reducing the safety concerns around public transportation and increasing public transportation use, offering a more optimistic perspective on the future of public transportation. As described in section 4.1, worries of getting infected by COVID-19 decreased significantly in 2021 relative to 2020, particularly among vaccinated individuals. This decrease in concern is associated with an increase in public transportation use among vaccinated individuals, where safety has been shown to be one of the main reasons for the reduction in ridership(58, 66). In the April 2021 and July 2021 survey waves, about 18% of individuals who got at least one shot of a COVID-19 vaccine before May 2021 reported that they had used public transportation at least once in the preceding two weeks. Only 12% of these individuals reported using public transportation in the first three waves of the survey (August 2020, October 2020, December 2020), representing a 50% increase from 2020 to 2021 (p-value < 0.05). There was no significant difference in the frequency of using public transit among those who did not receive any vaccine, with about 15-16% of this group reporting using transit in both pre-vaccine and post-vaccine availability survey waves (Figure II-6). The difference-in-differences estimator

When it comes to public transportation as a primary mode of commuting to work, we find a similar trend between vaccinated and unvaccinated individuals, albeit among a much smaller sample. Our data show that between 6-7% of all commuters use transit to commute to work, a fraction comparable to the national average of about 5%. This share, and its evolution, differ across the groups of vaccinated and unvaccinated individuals. Of those vaccinated between January and April 2021, only 3% used transit as a primary mode of commute in 2020 (before getting vaccinated), compared to 5.5% after getting vaccinated in 2021, an 80% increase. Among unvaccinated individuals, however, the numbers are roughly the same before and after vaccination (about 7%).

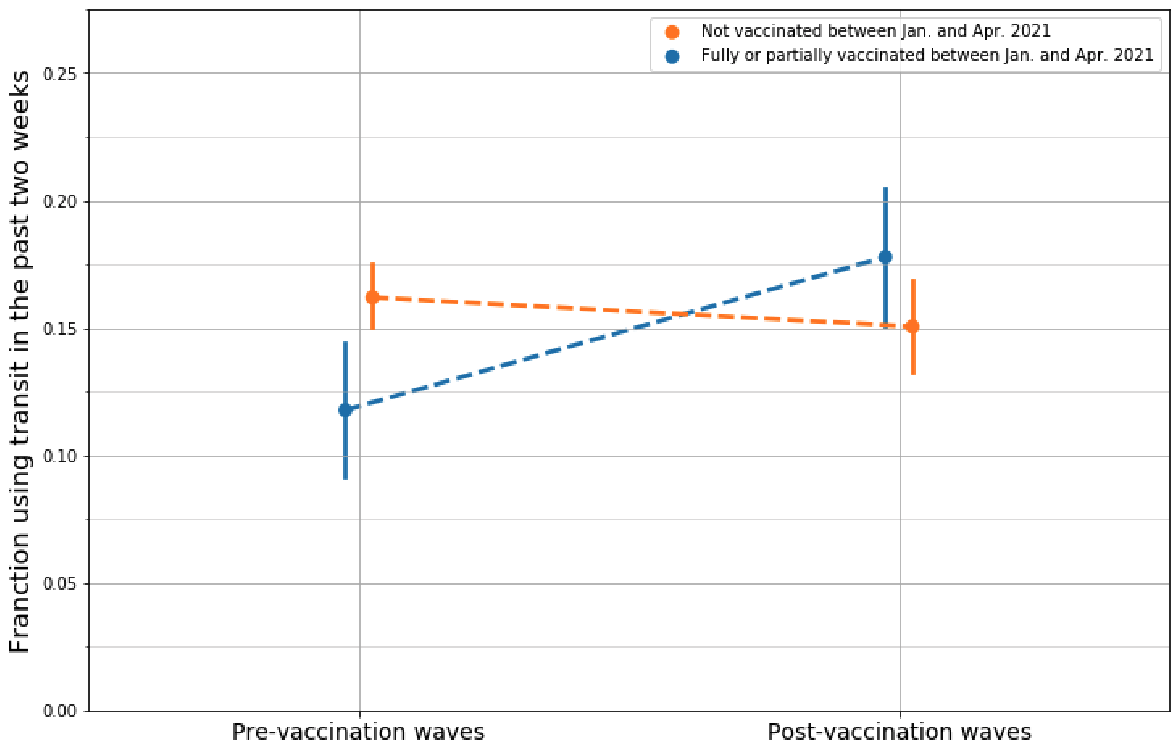


Figure II-6. The change in the frequency of using transit for vaccinated and unvaccinated individuals.

This increase in the frequency of using transit among vaccinated individuals is likely due to the effect of vaccines on reducing the concern about being infected by COVID-19, as shown in section 4.2.1. Indeed, our data show that concern about being infected by COVID-19 is strongly correlated with the frequency of using public transit. Across all survey waves, panelists who say they are not concerned about being infected by COVID-19 are 26% more likely to report using public transit in the past two weeks (p-value < 0.05) — 19% of those who say they're not concerned about being infected by COVID-19 report using transit at least once in the past two weeks, compared to 15% of those who are concerned.

Our results and data further suggest that addressing safety concerns on public transportation modes can significantly reduce the impact of potential future pandemics on transit ridership, even when transit service changes and reductions are inevitable. In our first wave of surveys, preceding vaccine availability, factors such as the widespread use of face masks and increased sanitation/cleaning on public transit were more frequently reported as measures that would increase transit use (about 45% of transit riders) than factors related to service changes (about 22% of transit riders)(66). The observed increase in transit use after vaccination provides further evidence that addressing safety concerns are key to protecting transit ridership in future pandemics and disease outbreaks.

4.2.3. On telecommuting

While telecommuting at least once a week remains the norm for the majority of workers in 2021, our data show a decrease in the share of individuals who fully telecommute (Figure II-7), from a high of 39% in December 2020 to a low of 28% in July 2021 (p-value < 0.001). On the other hand, the number of people commuting to work on every workday increased by approximately 16%, from a low of 50% in December 2020 to just above 58% in July 2021 (p-value < 0.01). These results suggest that while partial telecommuting continues to be an option for a significant portion of workers, full remote work will not remain the norm in a post-pandemic world. Hybrid work models seem to be the most likely outcome, and this trend will have significant implications on congestion and traffic patterns.

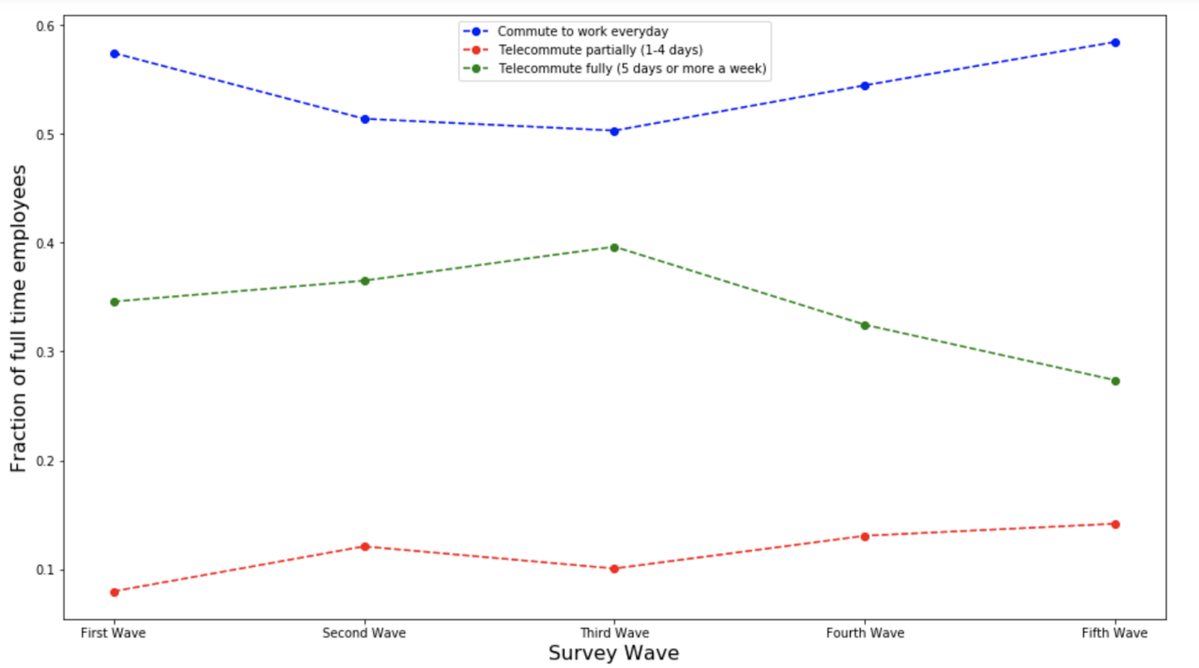


Figure II-7. The frequency of commuting and telecommuting to work for fully employed panelists. We define fully employed as those who report working at least 5 days a week in the past week. The number ranges from 402 to 477 depending on the wave, except for the third wave in December 2021 when the number drops to 328 (likely due to vacation time).

In contrast to our general travel behavior results, the observed decrease in the frequency of telecommuting was not significantly different between vaccinated and unvaccinated individuals. To measure the impact of vaccination status on telecommute frequency, we estimate a similar difference-in-differences design for two cohorts: Fully employed participants who got vaccinated between January 2021 and April 2021 compared to those who did not (hereafter referred to as the fourth wave cohort), and fully employed panelists who got vaccinated between April 2021 and July 2021 compared to those who did not (hereafter referred to as the fifth wave cohort). The sample sizes for those two cohorts are shown in Table II-3. The models' specification is shown in Equation 2: the outcome, , is the fraction of people working fully remotely in group *g* at time *t*, where *g* is the treatment status (vaccinated vs. unvaccinated) during the wave of interest (wave 4 or wave 5, depending on the cohort), and *t* is the treatment period (pre- vs post-vaccination). Table II-4 shows the models' estimated parameters for both cohorts.

$$
\begin{aligned}
Y_{gt} = \beta_0 \\
&+ \beta_1 * I(g = Vaccinated) \\
&+ \beta_2 * I(t = Post\ vaccination) \\
&+ \beta_3 * I(g = Vaccinated) * I(t = Post\ vaccination) \qquad (eq.\ 2)
\end{aligned}
$$

Table II-3: Cohort definitions for the difference-in-difference designs to quantify the effect of vaccination on the likelihood of working fully remotely.

Vaccination status	Fourth wave cohort	Fifth wave cohort
Vaccinated	143	152
Unvaccinated	316	156

Table II-4. Estimated coefficients for the difference-in-differences regression in equation 2.

	Fourth wave cohort		Fifth wave cohort	
Parameter	Estimate	p-value	Estimate	p-value
β_0	0.43	< 0.001	0.29	< 0.001
β_1	-0.089	-0.58	0.16	0.020
β_2	-0.083	0.035	-0.059	0.19
β_3	0.0035	0.96	-0.050	0.44

Our estimates for the effect of getting vaccinated were insignificant in both cohorts. The similar drop in telecommuting rates for both the vaccinated and unvaccinated, as shown in the small and statistically insignificant parameter in both regressions, suggests that while vaccines increase mobility by reducing the concern about contracting the disease, the decision to return to in-person work is not primarily driven by the safety and drop in fear of infection provided by the vaccine. Rather, it is likely determined by other external factors including but not limited to businesses deciding to return to in-person modes of operation. A recent study by the Pew Research Center gives more insight into some of the reasons some workers are returning to their offices, citing increased productivity in the workplace, not having the proper space and resources at home, having more opportunities for advancements in the workplace, or feeling pressure from coworkers or supervisors to work in-person(67). These findings, along with ours, suggest that safety is not the main concern for working in person, and that equilibrium telecommuting rates will ultimately be a function of both employer expectations and decisions to re-open office spaces, as well as employee preferences.

5. Conclusion and policy implications

This work contributes to the understanding of human and travel behavior during the pandemic impact and recovery periods. First, we showed that the changes in the frequency of travel, as well as compliance with public health recommendations more broadly, have been different for different groups of people. While the majority of the participants in our study complied with public health recommendations and significantly reduced their travel, a minority did not. These two groups have significantly different levels of concern about contracting and spreading the disease, and we find that higher concern is strongly associated with higher compliance and reduced travel. Next, we showed that vaccination plays a significant role in reducing concern among vaccinated individuals, which in turn affects their travel behavior. We estimated difference-in-differences regressions leveraging the longitudinal aspect of our dataset and found that vaccines have played a significant role in reversing some of the pandemic-induced mobility trends. We showed that vaccines led to a significantly higher increase in mobility for vaccinated individuals relative to the unvaccinated. We also showed that the frequency of public transportation use significantly increased for vaccinated individuals in 2021 relative to 2020, while it remained the same for the unvaccinated. With telecommuting, we see a significant drop in the fraction of full-time employees that worked from home every day in 2021 relative to 2020, but the fraction of employees telecommuting at least once a week remained high at 42% in August 2021. The drop in fully remote work was not significantly different between vaccinated and unvaccinated individuals, suggesting that decisions to return to in-person work are not driven by individual safety concerns and instead are likely a function of employer expectations, and that the high telecommuting rates may well persist in a post-pandemic world.

Our results have several implications on policy-making during pandemic times:

Insight #1: We showed that in the early phases of the pandemic preceding vaccine availability, while the majority of individuals in our panel complied with public health recommendations, a minority did not. This minority carried significantly more risk of contracting and spreading the disease by engaging in multiple non-compliant behaviors at the same time. For example, while only 18% of our panelists reported attending gatherings larger than 10 people in the August 2020 wave of the survey, these individuals were five times more likely to not wear masks while socializing with others, and four times more likely to not wear masks while traveling in public (not using their private vehicle). We also find that these individuals traveled significantly more than others in the early phases of the pandemic, increasing their risk of spreading the virus and highlighting the role of the transportation system in propagating the disease.

The concentration of “behavioral” risk – risk associated with non-compliant behaviors like eschewing masks, attending large gatherings, traveling more, etc. – among a small subset of individuals draws parallels to the literature on superspreading: similar to how a small number of individuals has been shown to cause the majority of cases in the early outbreak of the disease, a small number of individuals carry disproportionately higher risk of contracting and spreading the disease based purely on their behavior and non-compliance with multiple recommended safety protocols. From a policy standpoint, finding that superspreading can be, at least partially, explained by human behavior, rather than being a fully biological phenomena, is positive. It suggests that policies aimed at managing human behavior can be effective at controlling the spread of the disease absent any vaccines or pharmaceutical interventions. A key to the success of those policies, however, is ensuring high compliance from the public. Some work suggested that the overdispersed nature of the spread is its Achilles' heel(68), since this allows interventions to focus on key behaviors that feature a high number of interactions like indoor events and large gatherings. Our work suggests that this may be an oversimplification of the problem, however, since people who engage in those activities are also the ones who are resistant to behavioral interventions. For example, our August 2020 survey data show that those who do not wear masks where recommended by public health officials are 181% more likely to believe that COVID-19 related restrictions are too strict (p-value < 0.001). This implies that simply mandating public health recommendations would likely not be effective on the people whose compliance is needed most. Previous work studied how political ideology and demographic factors can predict the perception of threat of COVID-19(44, 45), which we show predicts compliance. More work is needed to develop communication strategies that speak to different demographic and ideological groups and are effective at engaging those with low safety concerns to improve their compliance with behavioral prevention measures.

Insight #2: The second important implication of this work relates to the role of vaccines in reversing safety concerns and inducing a return to pre-pandemic mobility levels. The literature on the effect of vaccination on pandemic mobility is limited, with studies providing conflicting findings. Our data clearly shows the positive effect of vaccination on weekly mobility for the vaccinated early in 2021. Our data also suggests that the mediator through which vaccination affects mobility is likely the level of concern about contracting the disease: concern about contracting COVID-19 is highly correlated with mobility levels (Figure II-3), and vaccines are very effective at reducing concern among vaccinated individuals (Figure II-4). The effect of vaccines on easing concerns and increasing mobility has not been limited to vaccinated individuals: our data importantly show that both vaccinated and unvaccinated individuals significantly increased their mobility right after vaccines became available in early 2021 (Figure II-2). The ease in concern among unvaccinated individuals coincides with the significantly more optimistic rhetoric adopted by government and public health officials in early 2021. This reduction in concern and increase in mobility early in 2021, especially among unvaccinated individuals, poses problems for public health and government officials. Multiple studies have found that reduced mobility can have a similar effect on the number of infections and hospitalizations(69, 70) as vaccination, with one study suggesting that the effect of mobility reduction is similar to that of a 30% vaccination rate with 70-90% vaccine efficacy(69). This suggests that, at least in the early days of vaccine availability when vaccination rates are still low, the net effect of vaccine availability on transmission and hospitalization numbers can be negative if safety protocols (e.g. mask wearing) and government restrictions are eased prematurely. Political and public health messaging should be careful to avoid narratives that suggest the pandemic is over, as this can result in a premature sense of safety that could decrease the motivation to comply with behavioral safety measures.

When it comes to public transit, we find a significant increase in the frequency of using public transport among those who got vaccinated early. We also find a high correlation between the concern about contracting COVID-19 and using public transit across all our survey waves, with 19% of non-concerned individuals using public transit in the past two weeks, compared to only 15% of the concerned individuals. Our results collectively show that safety concerns are an important reason for the public transit ridership declines during the pandemic, and consequently, dealing with safety concerns through improved sanitization, adding capacity constraints, and enforcing NPIs like requiring mask-wearing can soften the magnitude of ridership declines in future pandemics and public health outbreaks.

Insight #3: The third implication of this work relates to our finding that telecommuting will remain high after the pandemic is over. Our data also gives insight into the reasons why people choose to return to work, showing that safety is not an important factor in that decision. We show this by demonstrating that getting vaccinated does not affect the magnitude of the reduction in telecommuting rates, despite significantly affecting the level of concern about contracting the disease. In addition to the lack of vaccination effect, our data also suggest that working fully remotely is not going to remain the norm for many employees, as evident in the sharp decrease in the fraction of people working fully remotely in 2021. Hybrid telecommuting models seem likely to persist, and travel demand models need to account for this new trend which has several implications on congestion, peak hour commute, and trip making in general. The literature on the transport impacts of telecommuting remains conflicted, with some studies suggesting that telecommuting results in fewer trips and vehicle miles traveled, while others suggest the reduction in commute trips is largely offset by the generation of new trips. With high telecommuting rates likely to persist post-pandemic, more comprehensive analyses on the transport impact of telecommuting are needed to appropriately account for its impact in travel demand models, environmental impact reports, and emission analyses.

6. Acknowledgement

I would like to thank my co-authors on this paper for their invaluable input: Mohammed Amine Bouzaghrane, Meiqing Li, Drake Hayes, Madeleine Parker, and professors Joan Walker, Michael Anderson, Daniel Rodriguez, Dan Chatman, Karen Frick, and Raja Sengupta. I would also like to thank the Embee Mobile team for providing the data that made this research possible.

Chapter 4: Learning and Optimizing Charging Behavior at PEV Charging Stations: Randomized Pricing Experiments, and Joint Power and Price Optimization

Executive Summary

In this chapter, we introduce, implement, and assess a framework for jointly optimizing the pricing policy and the charging schedule of electric vehicles (EVs) by learning and nudging human behavior with pricing. The proposed methodology uses time-based pricing to nudge user behavior at charging stations towards actions that achieve the station operator's objectives. The optimization framework incorporates endogenous human behaviors by explicitly accounting for the willingness to delay charging, as well as the plug-in duration of each session, as a function of the hourly prices. The approach also addresses the issue of overstay at EV charging stations by casting the problem as a trade-off between occupying resources and giving more flexibility to the station operator. We discuss the design and analysis of the behavioral experiments used to model the charging behavior of participants at the charging stations, and demonstrate the effectiveness of those learned behavioral models in an optimization framework aimed at minimizing the total cost of providing the charging service without sacrificing the user experience. Our simulations show that our framework increases total revenue, reduces utility cost, and increases net profit for the station operator, while maintaining a high level of service and consumer utility.

1. Introduction

Plug-in electric vehicle (PEV) adoption has been on the rise, and the transportation system is steadily moving towards more electrification. Recent forecasts anticipate that the number of PEVs will seize one third of the entire vehicle sales market by 2025, and that more than one half of the new vehicles sold will be electrified by 2030 (89). This fast market penetration has been further accelerated by local and regional policies. For example, California recently set targets of five million zero-emission vehicles by 2030 and 100% clean electricity by 2045 (90). At the center of these policies are PEVs, which reduce greenhouse gas emissions by transitioning oil consumption to clean electricity.

The environmental benefits of replacing conventional vehicles with electric ones, however, are threatened by the increased demand on the electric grid and the potential increase in peak loads if the charging is left unmanaged. Currently, most EV owners charge their vehicles at home during the evening hours (91, 92). This simultaneous use of the charging stations in the future may undermine their potential by exacerbating existing peaks, and potentially exceeding the capacity of the power grid, requiring costly infrastructure upgrades (93, 94). In addition, charging in the evening and overnight can be temporarily mismatched with clean renewable electric energy sources, like solar in California, which is much more abundant during the day.

Owing to the reasons above, a growing body of research has been studying workplace charging as a better alternative to charging at home (95–97). Despite the growing interest, workplace charging still only accounts for about 15-20% of all PEV charging events (95). One important reason for the small adoption of workplace charging is the challenge of running an economically viable service, with most employers currently providing this service for free for their employees (97). In addition, charging at the workplace is often less convenient for PEV users, since it usually requires them to unplug their vehicle after a certain plug-in duration is reached or when their vehicle is fully charged. This makes competing with charging at home even more challenging, and puts tight constraints on the pricing strategy of the charging service (98).

Nevertheless, workplace charging still has the potential to shift a significant amount of the charging demand away from charging at home, and in doing so, reduce the peak loads on the electric grid. Workplace charging can also reduce the carbon footprint of the PEV charging service by aligning charging load with solar electricity generation. To deal with the aforementioned challenges, however, a deeper understanding of user charging behavior, their preferences, and price thresholds is imperative to successfully design smart charging stations that are economically viable (i.e., achieves the station operator's financial objectives), as well as attractive to EV users who have the option to charge elsewhere. Despite its importance, research on consumer behavior in the PEV charging context is limited, data is lacking, and consumer preferences, especially at workplace charging stations, remain poorly understood. We address this gap by implementing SlrpEV (Smart LeaRning Pilot for Electric Vehicles). Our research objective is to understand users' behavior and preferences when charging their vehicles at work in order to design charging stations that can be optimized to achieve the station operator's objectives.

1.1. Introducing SlrpEV

SlrpEV is a workplace charging program that provides plug-in charging service to PEV owners at two campuses within the University of California: Berkeley and San Diego. SlrpEV can be seen as a cyber-physical and human system (Figure IV-1). The physical component comprises the two charging sites with a total of eight Electric Vehicle Supply Equipments (EVSEs). The cyber component refers to the backend software and communication system for charging operation, namely the pricing optimization and energy scheduling algorithms. The human component, which is the main focus of this paper, refers to behavior of PEV users and their charging decisions, including their energy demand, willingness to delay charging, plug-in duration, etc.

The main goal of the SlrpEV project and the focus of this paper is to understand PEV drivers' charging habits and preferences at the workplace in order to improve the performance of the charging stations. Specifically, the project seeks to estimate the price thresholds and incentives at which users would accept delaying the charging of their vehicles instead of immediately charging it, giving flexibility to the charging operator to optimally deliver the required energy. We also model the relationship between how long users plug-in their vehicles and the hourly price they are paying. This knowledge is then used to design “smart” chargers that jointly optimize the pricing strategy and the energy delivery schedule while taking into account the needs of the consumers (PEV owners) as well as the suppliers (charging facility operators). More specifically, our objective is to jointly optimize pricing and power to increase net profit and minimize the cost of providing the charging service to the station operator. To achieve this objective, SlrpEV offers PEV users two choices for charging their vehicle: Regular charging, and Scheduled charging. Under the Regular option, the charging process starts as soon as the vehicle is plugged in, similar to how other standard charging stations currently operate. Scheduled charging, on the other hand, requires the user to input their requested energy (via miles of added range) and desired departure time before they start the charging session. The user is then guaranteed to receive their requested charge by their specified deadline, but not sooner. This provides the station operator with flexibility that can be used to optimally allocate the energy delivery schedule to achieve various possible objectives: minimizing emissions, maximizing revenue and profit, minimizing power variability, etc.

Another feature of SlrpEV that differs from many of the existing charging options is that pricing is done by hour, as opposed to by kWh. This automatically gives the operator the ability to influence the plug-in duration by leveraging the sensitivity of the users to the hourly prices. We exploit this relationship between charging duration and hourly prices to set optimal pricing policies that allow us to significantly reduce energy costs and increase the profits of the station operator, without sacrificing the level of service for the users.

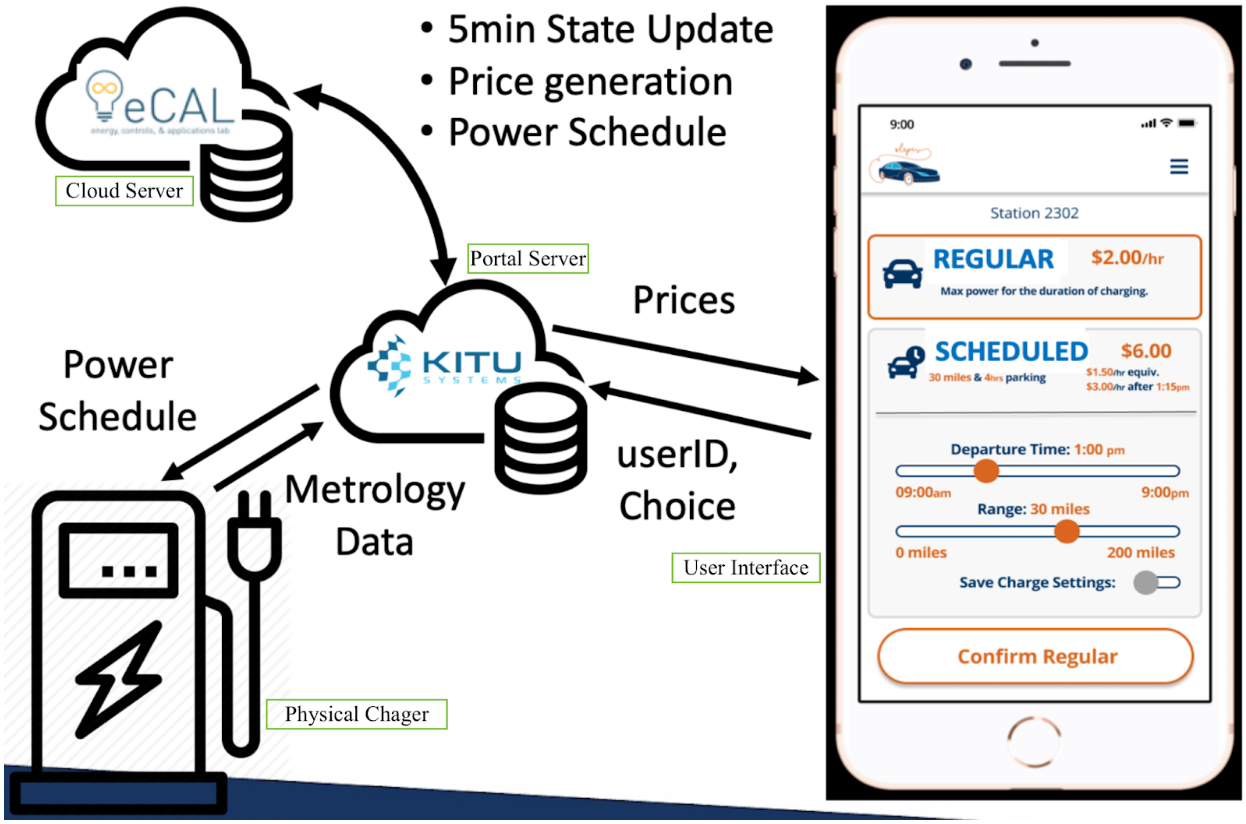


Figure IV-1: Schematic of the SlrpEV ecosystem

The rest of this manuscript is organized as follows: Section 2 reviews the existing literature on PEV charging station operation and planning that is relevant to our work. Section 3 formulates the price and power optimization problem, which motivates the behavioral modeling section. Section 4 summarizes the data we used in this project. Section 5 analyzes the results in three subsections: 1) Modeling the choice behavior between Regular and Scheduled charging, 2) Modeling plug-in duration, 3) and optimizing the choice and duration of charging to minimize station operator costs. We conclude and provide future directions in Section 6.

2. Literature review

Research on PEV charging station operation has been broad and diverse. Previous work ranged from modeling and analyzing network level interactions between the power and transportation systems, charging station interactions with renewable and green energies, power management and efficient scheduling of energy delivery, all the way to efficient energy storage systems and battery technology. Despite the breadth of topics, the literature remains relatively lacking when it comes to an accurate and reliable understanding of the behavior and decision making process of PEV users. This understanding is crucial to unlocking the full environmental benefits of electrification, especially as the optimal scheduling of energy delivery to PEV users will be essential as more and more vehicles become electric, in order not to overwhelm the electric grid infrastructure (99–103). A successful scheduling strategy should consider how PEV users make active decisions about their charging behavior (e.g. by delaying their charging, charging at different times of the day, etc.) in order to optimize operations. Recently, more studies have considered the demand side of PEV charging operations and focusing on understanding the consumer. In the OptimizEV project (100, 104), the authors proposed deadline differentiated pricing, which explicitly attempts to exploit the latent flexibility in demand to reduce power demand peaks. They showed how by incentivizing longer deadlines through reduced energy prices, they were able to optimally schedule demand and load peaks. The choice of the price thresholds, however, was arbitrary, and the work does not attempt to estimate the relationship between prices and flexibility. In (105), Daziano attempts to better understand customer behavior and estimate economic quantities related to the same OptimizEV project by conducting a stated-preference choice experiment to analyze the incentives needed for customers to accept delayed charging. The author found that customers were willing to delay their charging if they received a discount on the annual fee of their charging program of $2.66 per hour of control given to the operator. This means that customers negatively perceived the loss of control over their charging schedule. Thus an incentive would be required to nudge them to delay charging. The limitations of this study, however, stem from the fact that it relies solely on a stated-preference experiment. Despite the convenience and prevalence of stated preference surveys in research, humans often behave differently in real life than when presented with hypothetical choice scenarios. The study also requires participants to make unintuitive trade-offs that are complex and abstract, like valuing an hour of delay on their charging schedule in terms of yearly discounts, which may result in inaccurate or misguided responses.

Our work also contributes to the literature on understanding the overstay issue at charging stations. Overstay is an important problem that affects the operational efficiency and level of service at charging stations. At its worst, overstay can be perceived negatively as it reduces station utilization and can result in unsatisfied demand, especially as the charging infrastructure remains relatively sparse, leading to reduced social welfare. Common ways to address this problem are by setting an overstay penalty (106). For example, Tesla’s “idle fee” is one solution that has been implemented in practice to encourage people to pick up their car right when it is charged (107). An alternative view of overstay, however, is as an opportunity to shift power demand and coordinate energy delivery to minimize the environmental impacts and avoid power peaks. The latter view will become more important as PEV penetration increases and the charging infrastructure expands to accommodate increased demand. Deadline differentiated pricing like in (100) is one example that tries to exploit overstay to optimally control power demand for PEV charging in residential settings. Namely, a menu of deadline differentiated prices give discounts for longer plug-in duration. None of the studies that we are aware of attempt to actually model plug-in duration and develop a quantitative understanding of the relationship between charging duration and charging price. Our work fills this gap.

Finally, our work also contributes to the growing literature on workplace PEV charging. Workplace charging is the main charging means for about 40% of current EV owners who lack access to a home charging facility. While charging during the day is environmentally preferred to overnight charging for grids with large penetration of solar power, workplace charging comes with its own set of challenges. Specifically, workplace charging is very likely to be synchronized due to similar arrival and departure patterns dictated by working hours (108). In (108), the authors propose an integrated energy management scheme where employee's PEVs can be used as an energy supply source to provide the building with energy, leveraging the stable arrival-departure patterns of employees and their willingness to accept scheduled charging due to their flexibility. The user behavior (e.g., willingness to delay charging, willingness to accept using the PEV batteries as energy storage), and the sensitivity of those behaviors to the price of charging, are not studied in the aforementioned paper, neither are the incentives required by the drivers to accept scheduled charging. In (109), the authors analyze the demand side more closely, studying the effect of social norms on charging etiquette in the workplace. Their work focuses solely on the problem of overstay, however, and attempts to study the effectiveness of price and non-price interventions that encourage better and more efficient resource sharing etiquette. However, this study does not attempt to study the effectiveness of incentives intended to delay the plug-in duration to leverage scheduled charging. In other words, their work focuses on how to maximize throughput, instead of how to minimize the cost of charging or optimize energy delivery. One potential reason for why user charging behavior at the workplace is poorly understood is the lack of experimental data where prices can be randomly perturbed to learn user behavior. Existing research and datasets on user behavior in the workplace are mostly based on rigid pricing structures where price does not vary randomly across sessions and individuals (110–112). This significantly limits the researchers' ability to unbiasedly and robustly estimate key behavioral quantities that vary with price, like plug-in duration, choice between different charging alternatives, energy demand, incentive structures, etc. In this work, we fill this gap by conducting randomized experiments to estimate the impact of different pricing strategies on key behavioral outcomes of EV users' charging behavior.

3. Problem Formulation

Our objective is to jointly optimize the hourly prices and energy delivery schedule to PEV users in order to minimize the total utility costs and maximize net revenue for the station operator, given a set time-of-use (TOU) electricity rates. Conceptually, this framework allows us to achieve other objectives as well, like minimizing emissions, or shifting the electricity demand curve, since TOU prices can be used as a proxy for those objectives. For example, if the objective is to shift charging electricity demand from the afternoon peak, then setting the TOU electricity rates to be significantly more expensive during that peak would incentivize the station operator to try and shift the demand curve away from this peak period. Similarly, if the objective is to minimize emissions, then the TOU rates can be set to negatively correlate with high emission periods to disincentivize charging during those periods.

Consider a user, *i*, arriving at a SlrpEV station to start a charging session. This user is faced with three options: Regular charging, Scheduled charging, or leaving. Each of the charging options (excluding the leave option) is associated with its own cost of energy for the station operator.

Under the Regular option, a session *i* arriving at time t_o and leaving at time t_{reg} has a cost (from the station operator's perspective) that can be expressed as follows:

$$C_{Reg} = \sum_{t=t_o}^{t_c} c_t P_{max_i} - \sum_{t=t_o}^{t_{reg}} r_{reg_i} \qquad \text{(eq. 1)}$$

where t_c is the minimum time needed to fully charge the car at maximum power (if the user departs before their car is fully charged, then t_c = t_{reg}). Symbol c_t is the cost of electricity in $/kWh at time *t*, P_{max_i} is the maximum charging power that user *i*'s car can accept, and r_{reg_i} is the hourly rate paid by the user (i.e. the operator's revenue) if they choose the Regular option.

On the other hand, if the user chooses the Scheduled option, they are asked to indicate their desired energy required (E_{req}) and a fixed departure time (t_{sch}) for when they intend to pick up their car. Their cost to the operator can then be expressed as follows:

$$C_{Sch} = \sum_{t=t_o}^{t_{sch}} (c_t P_t - r_{sch_i}) \qquad \text{(eq. 2)}$$

where P_t is the charging power at time *t* (controlled by the station operator), and r_{sch_i} is the hourly rate paid by the user (i.e. revenue) if they choose the Scheduled option in session *i*.

Given the costs associated with each alternative, the total expected cost to the operator can be expressed as follows:

$$C_{Total} = p_{Reg} * C_{Reg} + p_{sch} * C_{Sch} + p_{Leave} * C_{Leave} \qquad \text{(eq. 3)}$$

where p_{Reg}, p_{sch}, and p_{Leave} are the probabilities that the user chooses the Regular option, the Scheduled option, or the Leave option in session *i*, respectively. Note that p_{Reg}, p_{sch}, and p_{Leave} must sum to one. Since our definition of the cost refers to the total cost of providing the charging service and does not include any opportunity cost, the cost of leave, C_{Leave}, is assumed to be zero.

The objective of the station operator is to minimize the total cost shown in *Equation 3*. To achieve this objective, there are two levers that the operator can control: the hourly prices for the Regular and Scheduled options, r_{reg} and r_{sch}, respectively, and the charging power at time *t*, P_t , if a user chooses the Scheduled option. On the other hand, there are constraints on those decision variables that the station operator has to abide by, namely:

- The charging power at time t, P_t , must be non-negative and upper-bounded by the maximum power that user *i*'s vehicle can accept:
  - $0 \leq P_t \leq P_{max_i}$
- The total energy delivered to the user in session *i* if they choose scheduled has to be equal to their total energy requested (E_{req}) at the user specified deadline, t_{sch}
  - $\sum_{t=t_o}^{t_{sch}} \Delta t \cdot P_t = E_{req}$, where Δt is the time step of the system.
- The hourly prices of Regular and Scheduled, r_{reg} and r_{sch} respectively, are non-negative.
  - $r_{reg} \geq 0$ and $r_{sch} \geq 0$

There are a few key details worth highlighting in the above equations and definitions; many parameters in the problem formulation are dependent on user behavior and are endogenous to the prices that the operator controls. Namely, the probabilities of choosing each alternative, p_{Reg}, p_{sch}, and p_{Leave}, all depend on the hourly prices of the Regular and Scheduled options, r_{reg} and r_{sch} (see Section 5.1 for details on how those probabilities were modeled). In addition, the plug-in duration, t_{reg} and t_{sch} for Regular and Scheduled, respectively, will also depend on the hourly price that the user is paying. Intuitively, a discounted hourly rate will result in a longer plug-in duration, since the user can leave their vehicle parked for longer while still maintaining a lower total session cost. Our work importantly focuses on learning those behavioral responses to the hourly prices, where the literature is particularly lacking. We conduct randomized experiments and use methods from the econometrics literature and random utility theory to model the choice making behavior of users when choosing between the charging options at SlrpEV. We also estimate the relationship between session duration and price using univariate Kernel regressions. Those learned behavioral models are then used in the optimization to inform price-setting, as shown in the simulations in Section 5.3.

4. Data

The main data used in this study is from the revealed preference experiment, which was conducted on two University of California (UC) campuses: San Diego and Berkeley. The experiment is a real life choice experiment where individuals choose between the two charging options described in Section 3.1: Regular and Scheduled, and pay accordingly. The key advantage of those experiments is the ability to create random variations in the prices and observe the resulting behavioral responses of the participants, namely their choice of alternative and their plug-in duration. Most of the existing studies in the literature lack this ability and either use passive observational data or have rigid pricing policies that do not allow randomization and experimentation (100, 112), which limits their ability to learn, and consequently control, user behavior.

Throughout these experiments, we present users with a randomly chosen hourly rate for each of the Regular and Scheduled alternatives. The distribution of the randomized prices for Regular and Scheduled, as well as the distribution of the differences in their hourly rates, are shown in Figure IV-2. As the distributions show, most of the randomized prices are below $2/hr, in order to make SlrpEV price competitive with other nearby available charging options and to avoid user churn; two public stations within 0.5 miles offered $1.50/hr during the experimental period. Nevertheless, higher prices were presented occasionally in order to better analyze and model the relationship between hourly prices and plug-in duration.

Figure IV-2 also shows that the distribution of the Regular prices has more mass over higher hourly rates than that of the Scheduled prices. This is because, everything being equal, the Regular option is more convenient for the users, since users will get maximum charging immediately without having to worry about the overstay fee present in the Scheduled option. Indeed, if prices were equal, then by picking the Scheduled option, a participant would sacrifice the benefits of immediate charging, and risk an additional overstay fee without any added benefits. Thus, if the difference in the hourly rates was randomly and symmetrically distributed around zero, then the final sample of choices will likely be heavily skewed towards the Regular option. To avoid this situation, a discount should be applied to make the Scheduled option competitive with the Regular and have a roughly balanced sample of choices to analyze.

In order to develop a prior on what the average discount for the Scheduled option should be, we conducted a stated preference experiment where participants were asked to choose between two hypothetical charging scenarios that are similar to the real Scheduled and Regular options. We found the willingness to pay for the convenience of the Regular option was, on average, around 30 cents/hour. We used this value as a prior when generating the hourly rates for the revealed preference experiment. As Figure IV-2 shows, the average difference in the randomized prices for the two charging options is about 26 cents/hour, but the differences can be up to $2/hour in order to broadly understand the behavioral responses at different prices. The resulting choice split between the Regular and Scheduled options is presented in Figure IV-3, which shows that the samples are reasonably balanced with very modest discrepancies.

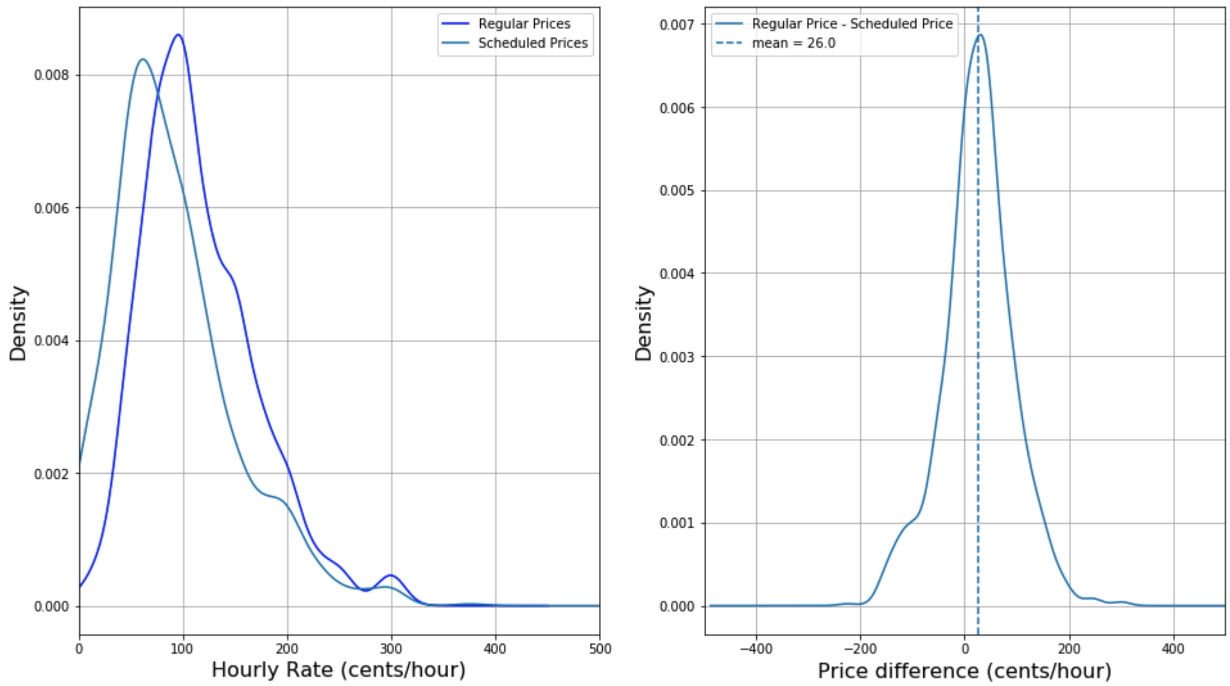


Figure IV-2: Sampling distributions of the hourly rates and hourly rate differences of Regular and Scheduled

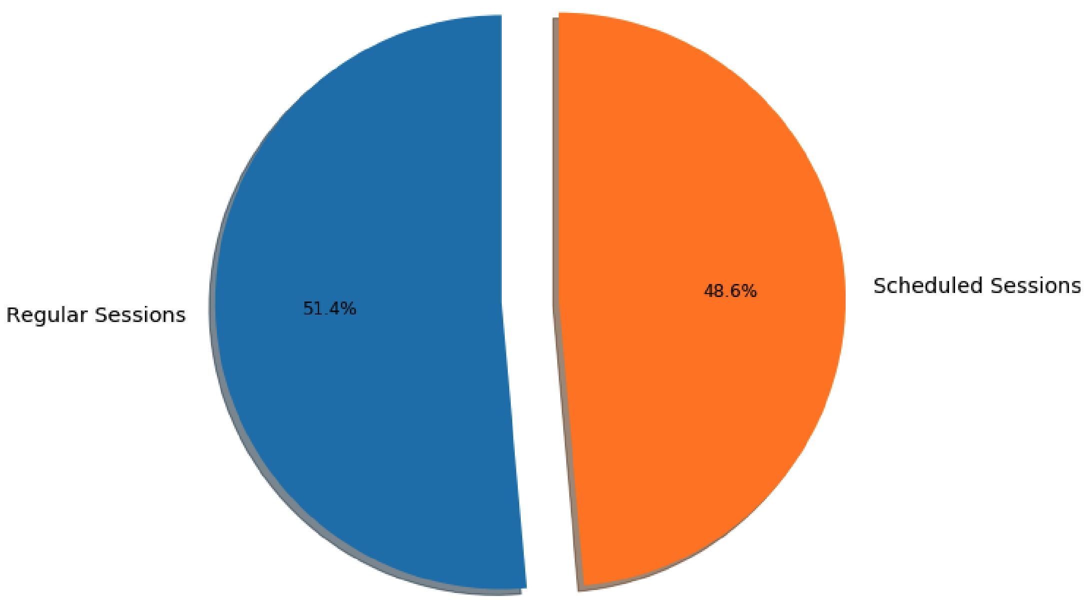


Figure IV-3: The split of the Regular and Scheduled choices in the experimental data

5. Results and Discussion

5.1. Choice Modeling

5.1.1. Estimation

A key quantity to estimate in order to optimally set the pricing policy is the level of discount needed for the Scheduled option to become competitive with the Regular option. In order to estimate this quantity, we model the charging choice making behavior with a discrete choice model. Understanding this behavior will allow the station operator to incentivize the alternative that achieves a certain desired outcome, whether that is gaining power delivery flexibility (Scheduled), or maximizing throughput (Regular). One of the key advantages of this work is our ability to conduct real-time randomized experiments where we exogeneously vary the hourly rates (and hourly rate differences) for the Regular and Scheduled options and observe the corresponding changes in the participants' choices. The ability to randomize the prices makes the modeling process much more robust and less reliant on strong assumptions like unconfoundedness, no-selection bias, and functional forms, since we do not need to make assumptions about the data generating process.

We adopt a random utility theory approach and use a binary logit model to capture the decision making process of participants choosing Regular or Scheduled. Random utility models, like binary logit, are widely used in the literature for modeling individual choice behaviors in the context of EVs (105, 113–115). These models are most frequently developed based on a utility maximization framework - each alternative choice available has a corresponding utility for each individual, and individuals are assumed to be rational agents, i.e. they choose the alternative that maximizes their utility. We adopt a similar framework for developing and estimating our models.

For this work, we are mainly interested in the effect of changing the hourly prices and hourly price differentials on the probability of choosing each alternative. We are also interested in whether those choice probabilities vary depending on the time of the day. Indeed, one may reasonably expect that late afternoon arrivals may be more inclined to choose Regular to minimize their charging time, compared to early morning sessions. Note that we do not include individual characteristics and demographic information in the final model specification. This is because the model will be later used to optimize the prices in order to achieve a certain probability split between Regular and Scheduled. Including individual level covariates in the model would lead to price discrimination based on individual factors like income, gender, etc, which must be avoided. Our final model thus only included price and temporal features, as these are the inputs that would be used in the optimization framework.

More specifically, we specify the utility equations of each alternative for the user in session *i* as shown in equation 1:

$$U_{sch_i} = \boldsymbol{\beta}_{\Delta p} * (r_{sch_i} - r_{reg_i})$$
$$+ \boldsymbol{\beta}_{pm} * I(2pm < t_0 < 8pm)$$
$$+ \boldsymbol{\varepsilon}_{sch_i}$$

$$U_{reg_i} = ASC_{reg} + \boldsymbol{\varepsilon}_{reg_i}$$

In other words, we assume that the utility of the scheduled choice is a linear function of the difference between the hourly price of regular (r_{reg}) and the hourly price of Scheduled (r_{sch}), while allowing for different intercepts depending on the arrival time. The utility of Regular, on the other hand, only includes an alternative specific constant (ASC_{reg}), and the alternative specific constant for Scheduled is fixed to zero in order to uniquely identify the model parameters. Symbols $\boldsymbol{\varepsilon}_{sch_i}$ and $\boldsymbol{\varepsilon}_{reg_i}$ are individual random error terms, which we assume are independently and identically distributed (i.i.d) over alternatives and individuals as Extreme Value Type I with a location parameter of 0 and a scale parameter of 1.

When making a choice, the user is assumed to choose alternative n instead of m if the utility of n exceeds that of m. Denote $V_{n_i} = U_{n_i} - \boldsymbol{\varepsilon}_{n_i}$ be the systematic utility of alternative n for session *i*. The probability of user choosing alternative *n* in session *i,* $P(y_{n_i})$, can be expressed mathematically as follows:

$$P(y_{n_i}) = P(U_{n_i} > U_{m_i})$$
$$= P(V_{n_i} - V_{m_i} > \boldsymbol{\varepsilon}_{m_i} - \boldsymbol{\varepsilon}_{n_i})$$
$$= 1 - \frac{e^{V_{n_i} - V_{m_i}}}{1+e^{V_{n_i} - V_{m_i}}}$$
$$= \frac{1}{1+e^{V_{n_i} - V_{m_i}}}$$

Since the individual disturbances, $\boldsymbol{\varepsilon}_{in}$ and $\boldsymbol{\varepsilon}_{ij}$, are i.i.d Extreme Value Type I, then their difference follows the logistic distribution.

Given the expression for the individual probability of choosing each alternative, we can derive the likelihood equation for the model and estimate the model parameters using maximum likelihood estimation. The likelihood of the model, L, is the model’s probability that all individuals make the choices that were observed in real life. Since we assume that the individual disturbances, $\boldsymbol{\varepsilon}_{in}$ and $\boldsymbol{\varepsilon}_{ij}$, are independent across individuals and choice situations, this likelihood is the product of the individual probabilities across all observations:

$$L = \Pi_{n=1}^{N} P(y_i)$$

where y_n is the observed choice of individual n. The model parameters $\boldsymbol{\beta}_{\Delta p}$, $\boldsymbol{\beta}_{pm}$, ASC_{reg} are estimated by maximizing the log-likelihood of the model:

$$max\, \Sigma_{n=1}^{N} log\, P(y_i)$$

5.1.2. Results and Validation

The estimated model coefficients are shown in Table IV-1.

Table IV-1: Estimated parameters of the discrete choice model

Parameter	Estimated coefficient	p-value
ASC_{REG}	0.31	<0.005
$\boldsymbol{\beta}_{\Delta p}$	-0.02	<0.001
$\boldsymbol{\beta}_{pm}$	-1.14	<0.005

The model allows us to estimate key quantities of interest and draw several conclusions about user behavior at the charging station. First, the estimates show that for a price difference of zero (i.e., when both alternatives have the same price), the Regular option is significantly preferred over Scheduled. In fact, the model allows us to compute the discount amount that will make Scheduled competitive with Regular, where we define "competitive" as having equal probability. To do that, we note that in order for the two alternatives to have equal probability, their systematic utilities to the user must be equal:

$$P(y_i = Regular) = P(y_i = Scheduled) \Leftrightarrow V_{reg_i} = V_{sch_i}$$

$$\Leftrightarrow ASC_{reg} = \beta_{\Delta_p} * (r_{reg_i} - r_{sch_i}) + \beta_{pm} * I(2pm < t_0 < 8pm)$$

$$\Leftrightarrow (r_{reg_i} - r_{sch_i}) = \frac{ASC_{reg} - \beta_{pm} * I(2pm < t_0 < 8pm)}{\beta_{\Delta_p}} \qquad \text{(eq 4)}$$

After substituting the parameter values from Table IV-1, Equation 4 implies the station operator will need to discount Scheduled relative to Regular by 15 cents/hour during the day, to as high as 72.5 cents/hour in the afternoon (between 2pm and 8pm). This makes sense, since if the Scheduled price was not discounted, the user would only be risking overstay penalties if they choose Scheduled, without gaining anything for the additional risk, so Regular is the rational choice. The higher discount required in the afternoon is intuitive too, as it reflects the fact that users have less flexibility around their pick-up time, towards the end of the workday. That is, they prefer getting their vehicle charged as soon as possible. Interestingly, those values are very close to the willingness to pay estimated from the stated preference experiment prior to running the revealed preference experiment. This shows that stated preference surveys can be accurate tools to infer trade-offs and economic quantities using discrete choice models under the right conditions.

The ability to relate the probability of choosing Scheduled at each discount level is the main contribution of this model. It allows us to set the prices to achieve a desired distribution of Scheduled and Regular options throughout the day. To validate whether the model can achieve this objective, we conducted multiple experiments where we used the model to set the prices that achieve a certain probability split between Regular and Scheduled. The results of those experiments are combined in Figure IV-4. The results show that the split between Regular and Scheduled is very close to the real-world observed split when the corresponding prices are applied, suggesting the model can be used reliably to set the prices and achieve a desired probability distribution of the choices. Note this price modeling framework does not force a particular choice on a particular user, but achieves a desired distribution in aggregate.

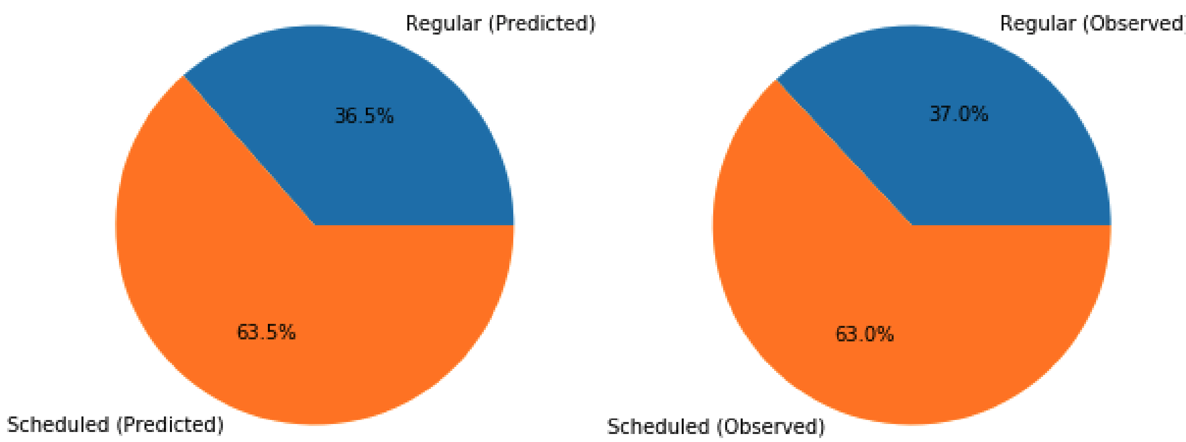


Figure IV-4: Performance of the discrete choice model on a test dataset.

5.2. Plug-in duration vs. Hourly rate

5.2.1. Estimation

As mentioned in Section 3, another important feature to estimate is the relationship between plug-in duration and hourly prices, since this relationship is key to the optimization objective. Figure IV-5 shows how the expected session duration decreases as the session's hourly price increases, from the real-world charging session data. This means that, if the station operator seeks to extend plug-in durations so they have more flexibility to optimize charging power delivery, then the station operator must provide a discounted hourly rate. These empirical results corroborate the deadline differentiated pricing scheme in (100, 104). Similarly, if the station operator's objective is to maximize utilization, increase throughput, and minimize charging time at the station, then the hourly rate should be increased to discourage long charging sessions and nudge the users to quickly pick-up their vehicle.

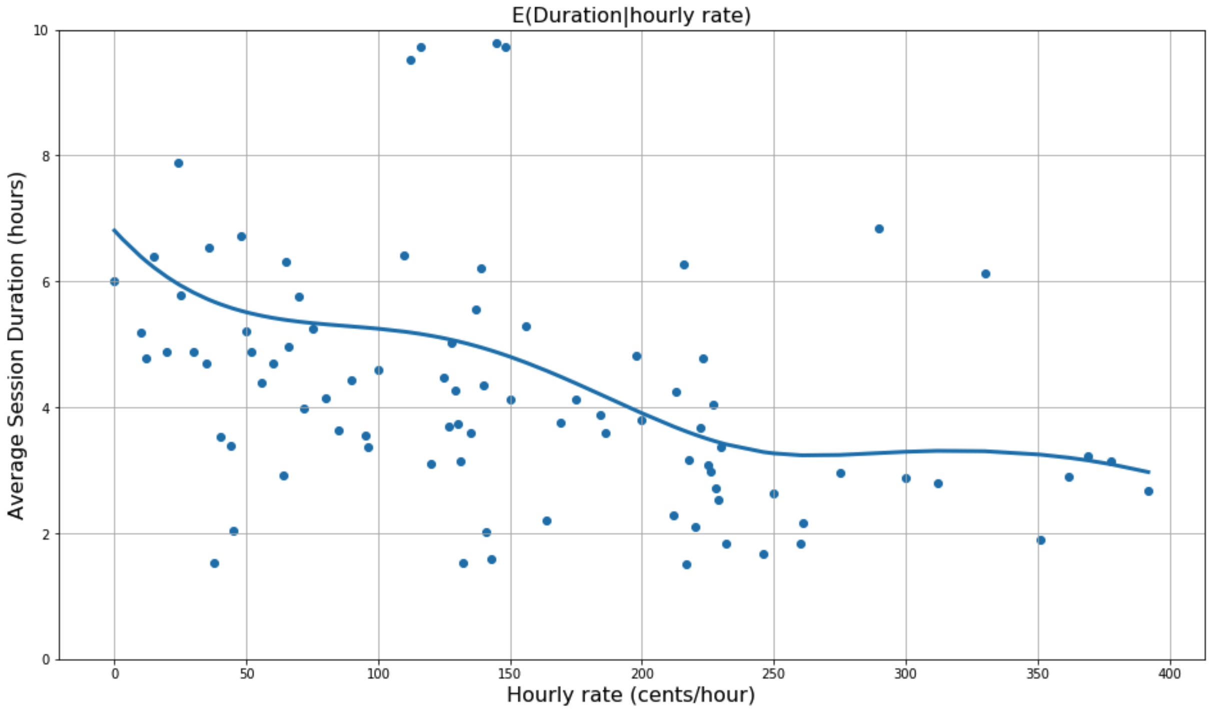


Figure IV-5: Expected value of the session duration as a function of the hourly price.

We fit a Kernel Regression model to the data in Figure IV-5 to predict the expected session duration given the price. A kernel regression is useful for capturing the non-linearity compared to a linear regression, and at the same time maintain monotonicity and better interpretability properties compared to tree-based and other machine learning methods. The fitted Kernel Regression is shown by the blue line in Figure IV-3.

5.2.2. Validation

To check whether the sessions' plug-in durations, and subsequently the energy demand, are responsive to controlled shifts in price, we conducted an experiment aimed at reducing electricity consumption in the afternoon (between 3pm and 8pm). This was motivated by the California Independent System Operator (CAISO) encouraging energy conservation during the summer months to ease demand peaks - a voluntary demand response. The experiment did not attempt to optimize power and energy delivery, and instead the objective was to shift sessions away from the 3pm - 8pm period strictly using the hourly rate of the sessions. More specifically, we progressively increased the prices each hour of the day in order to observe how the duration would respond. We compare the power demand profile during the experiment to a baseline period, which is a three week period of random price generation that occurred in September 2021.

Figure IV-6 shows the inverse relationship between hourly rate and plug-in duration during the experiment. Note that the prices shown in the Figure IV-are fixed for each arrival hour, and the blue dots represent the averaged session duration across all sessions that arrived at that hour. Figure IV-7 shows that energy demand dropped by a significant 40% between 3pm to 8pm compared to the baseline due to the higher hourly rates (from an average of 4.9 kWh per session in the baseline to an average of 2.9 kWh per session). This shows that session duration, and consequently energy demand, are responsive to controlled shifts in the hourly rates, which allows the station operator to influence those quantities via model-based optimization. Indeed, more or less energy can be shifted away by more or less aggressive pricing. Nevertheless, this experiment confirms that session duration and energy delivered can be shifted by pricing only.

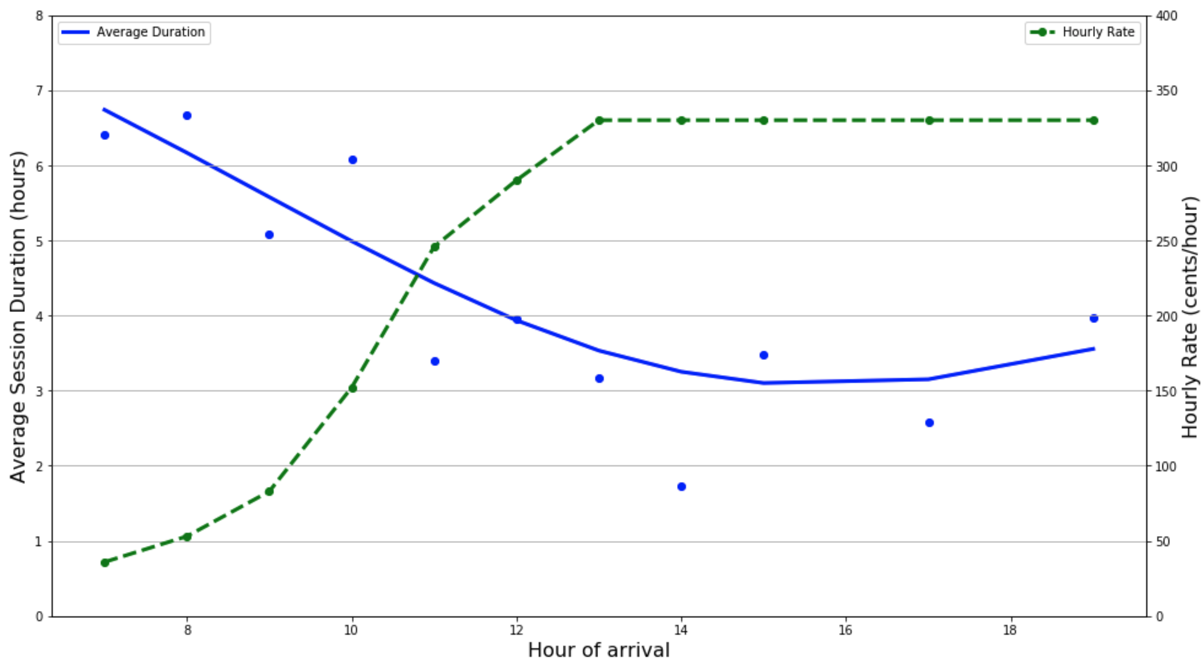


Figure IV-6: The hourly rates and the corresponding plug-in duration during the experiment, by hour of arrival. Blue dots represent the average duration of all sessions arriving at a given hour during the experiment.

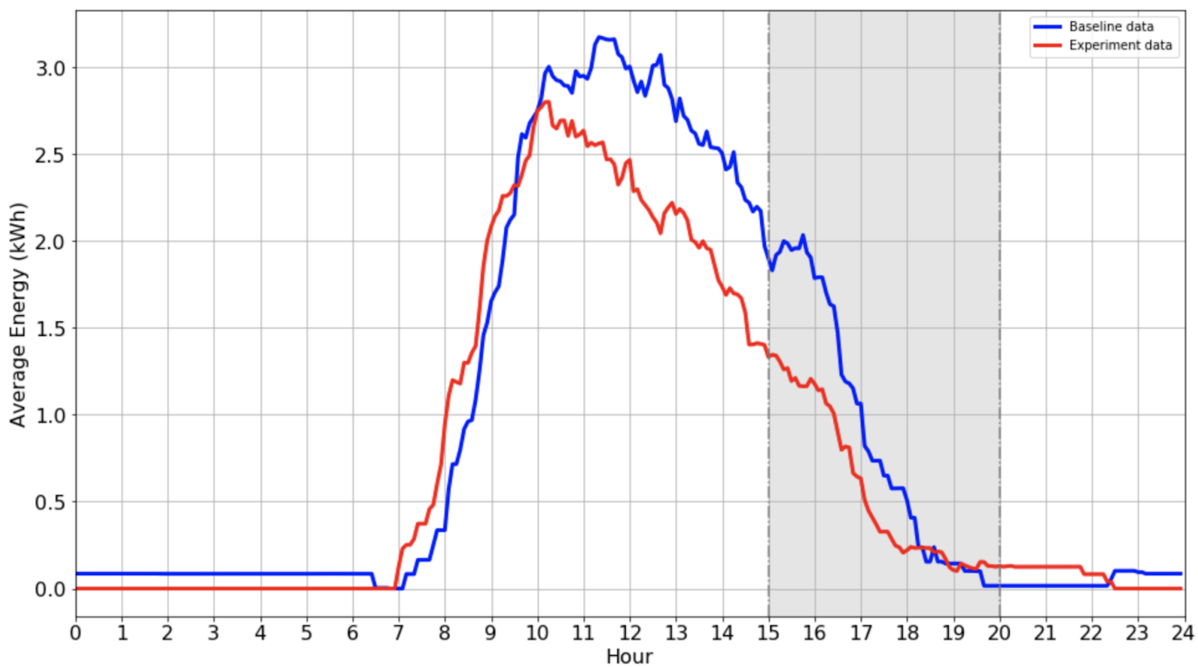


Figure IV-7: The daily energy demand profile during and before the experiment. The baseline data is a three week period of random price generation that occurred in September 2021. Prices during the baseline were randomized between \$1/hr and \$2/hr.

5.3. Operational Cost Optimization via Simulation

Now that we have shown and estimated the relationships between hourly prices and the choice probabilities, as well as between hourly prices and plug-in durations, those relationships can be substituted in the optimization framework shown in Section 3. We assess the framework through simulations: we first collect a period of real-life, uncontrolled baseline data where sessions and power were left unmanaged. We then use the same pattern of arrival and energy demand as the baseline, compute their optimized prices using our optimization framework defined in Section 3, and simulate the choice behavior of the users using our validated behavioral models. We assess the performance of our optimization framework compared to the baseline based on three key performance indicators:

1. Total revenue, cost, and profit generated by the station operator
2. The distribution of total energy delivered in peak and off-peak periods
3. The average hourly price paid by the consumer

5.3.1. Baseline data

The baseline data is a period of passive data collection where we fixed the prices of the two charging options and observed the corresponding charging behavior of the users. This data was collected between February 14, 2022 and March 8th, 2022, and included a total of 104 usable sessions. Prices were fixed for all the sessions at \$1.5/hr for Regular and \$2.5/hr for Scheduled. The fee for Regular was set to be comparable to two other public charging stations within 0.5 miles , where the charging cost was also \$1.5/hr at the time of the experiment. As expected, the vast majority of sessions selected Regular due to the large discount relative to Scheduled, with the exception of two sessions. We discarded those sessions and omitted them from the final analysis.

Table IV-2 summarizes the key characteristics of the baseline data and Figure IV-8 summarizes the distribution of arrival and departure times. Note that most of the arrivals happen in the morning between 6am and 10am.

Table IV-2: Session characteristics of the baseline data

Total (realized) sessions	104
Session hourly price	\$1.5/hr
Average session duration	3.62 hours
Average energy delivered per session	13.23 kWh

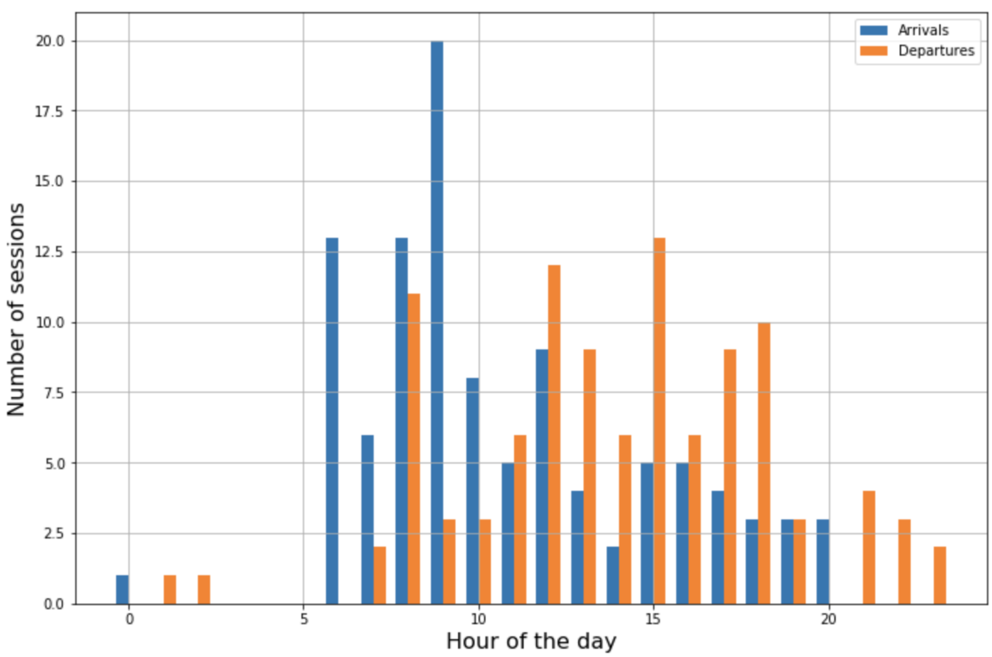


Figure IV-8: Arrival pattern of baseline sessions

5.3.2. Simulation workflow

Given the baseline data, the purpose of the simulation exercise is to simulate the behavior of those same baseline sessions if they had been subjected to our optimization framework. Then we compare the simulated behavior to the baseline using key performance indicators (KPIs) in order to assess the performance of our optimization framework and quantify its potential benefits. The simulation workflow is shown in Figure IV-9. The simulation inputs are the baseline sessions' individual information (time of arrival, energy required, and maximum power capacity), along with assumed hourly TOU electricity cost. Next, we compute the optimized prices using the framework described in Section 3. Given those prices, we simulate the expected behaviors – specifically, the choice of alternative, and the duration of stay – using the behavioral models estimated in Sections 5.1 and 5.2, respectively. Since those models were validated and have shown good predictive power in a similar setting, it is reasonable to assume that the behavioral responses predicted by the models would reasonably replicate what would happen if the framework were implemented in real life.

To simulate the choice of alternatives between Regular and Scheduled, we first compute the probability of each alternative (including leave) given their optimized prices using the discrete choice model. We then use Monte Carlo simulations to simulate a choice from the multinomial distribution parameterized by those probabilities.

As for the plug-in duration, we simply compute the expected duration given the chosen alternative's hourly price using the kernel regression estimated in Section 5.1. Note that this results in expected durations that have artificially smaller variance than the observed durations, since we are using the deterministic conditional expectation function to compute the expected duration. An alternative would be to model the duration as a stochastic outcome given the prices, which will render the problem as a stochastic optimization problem and significantly complicate the solution algorithm, since the session duration directly affects power optimization and energy delivery schedule for the Scheduled option. This is beyond the scope of this paper, but is one way to improve the simulations and is a welcome future extension of this work.

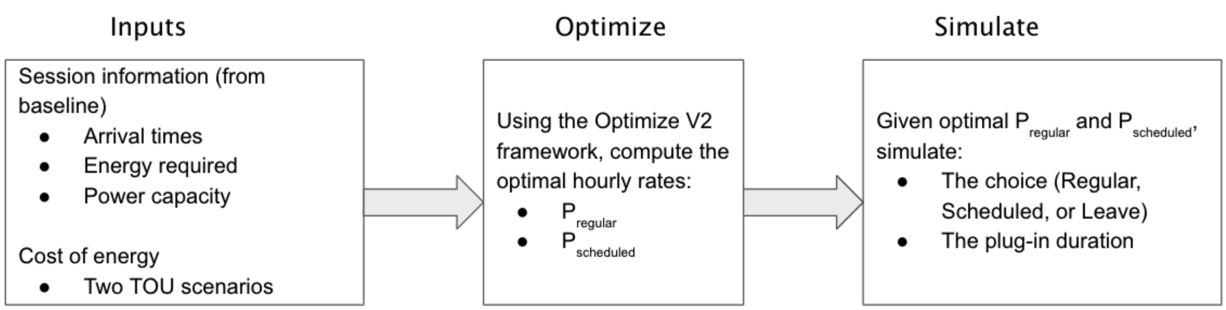


Figure IV-9: Simulation workflow

5.3.3. Simulation results

In this section, we present the results of the simulation exercise where we apply the framework outlined in Section 3 to a time-of-use (TOU) electricity rate scenario.

The first performance metric is explicitly optimized for: as Section 3 shows, the optimization objective is to minimize total net cost, or alternatively seen as maximize net revenue for the station operator. The second and third metrics are intended to give more insight into the mechanism through which net revenue is maximized. Indeed, net revenue can be increased by either increasing total revenue, decreasing total electric utility costs, or both. Ideally, we want to see an increase in net revenue that is caused by shifting energy delivery to off-peak periods, without necessarily increasing the hourly prices paid by the users.

The TOU plan that we analyze is shown in Figure IV-10, which corresponds to PG&E's peak-pricing TOU rate plan. Under this plan, peak pricing occurs between 4pm and 9pm, as shown by the high cost during those times. From the station operator's perspective, this TOU plan would incentivize shifting the energy load away from 4pm-9pm when energy prices are highest. The TOU rates shown in Figure IV-11 are used as input to the simulation, along with the inputs reflecting the baseline session information described in Section 5.3.2.

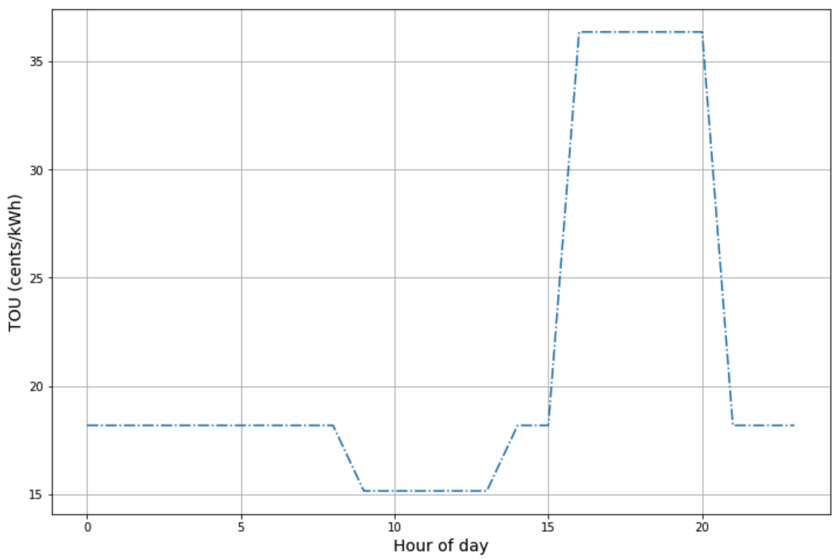


Figure IV-10: Electricity cost corresponding to PG&E's TOU rate plan

Figure IV-11 shows the performance of our framework using the first performance metric related to total revenue, cost, and profit generated by the station operator. The behavioral aware optimization framework led to statistically significant improvements: a reduction of 16.2% in total cost, an increase of 16.4% in total revenue, resulting in a significant 50.2% increase in total profit.

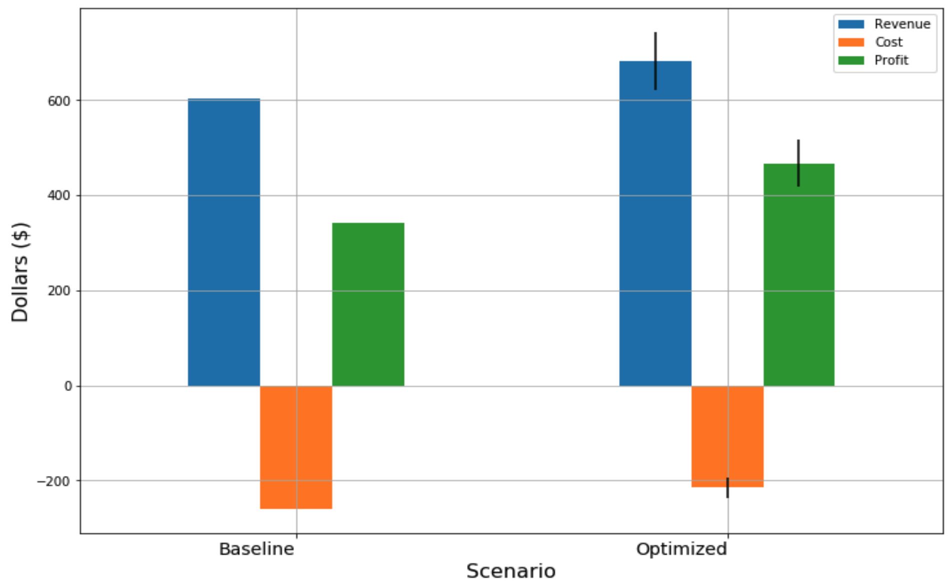


Figure IV-11: Revenue, cost, and profit comparisons between the baseline and the optimization framework

The reduction in cost and increase in profit are the result of a combination of factors: shifting energy delivery from the afternoon peak to cheaper off-peak periods, which is made possible by longer session durations and more flexibility to the station operator. Figure IV-12 shows the distribution of total hourly energy delivered between the baseline and the simulation. As shown in the figures, there is significantly more energy delivered in the super off-peak period in the optimized framework relative to the baseline (from 46% to 57%), and significantly less energy delivered during the peak (from 17% to 9%).

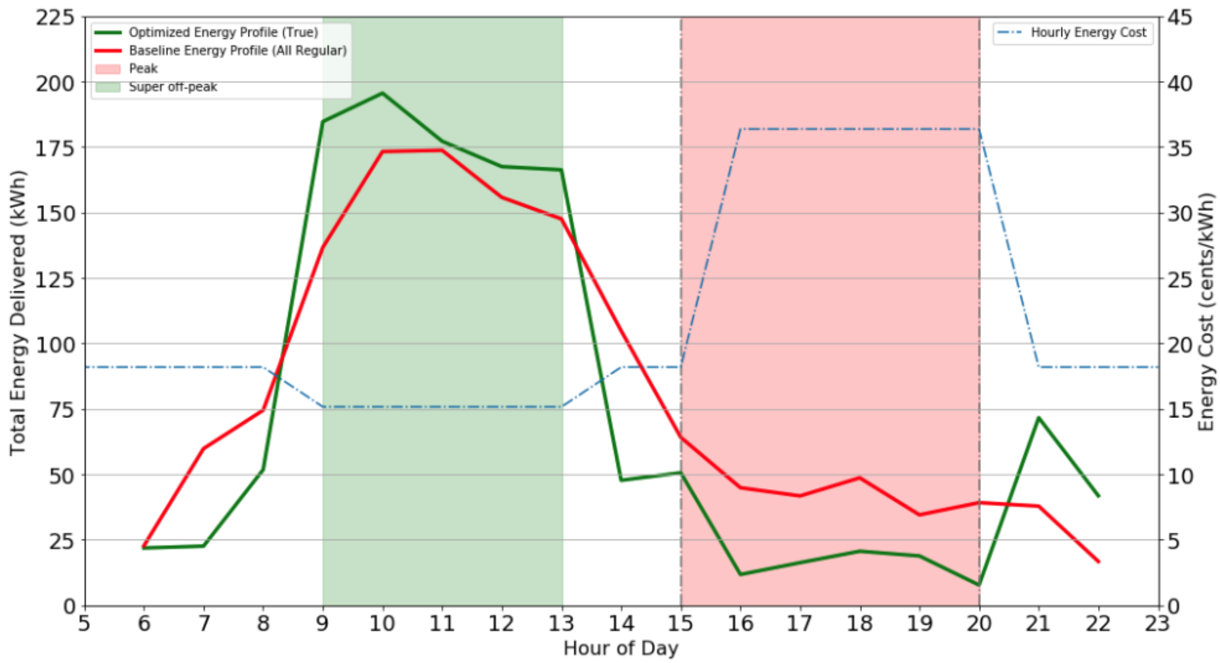


Figure IV-12: Distribution of energy delivered over the hours of the day - median simulation

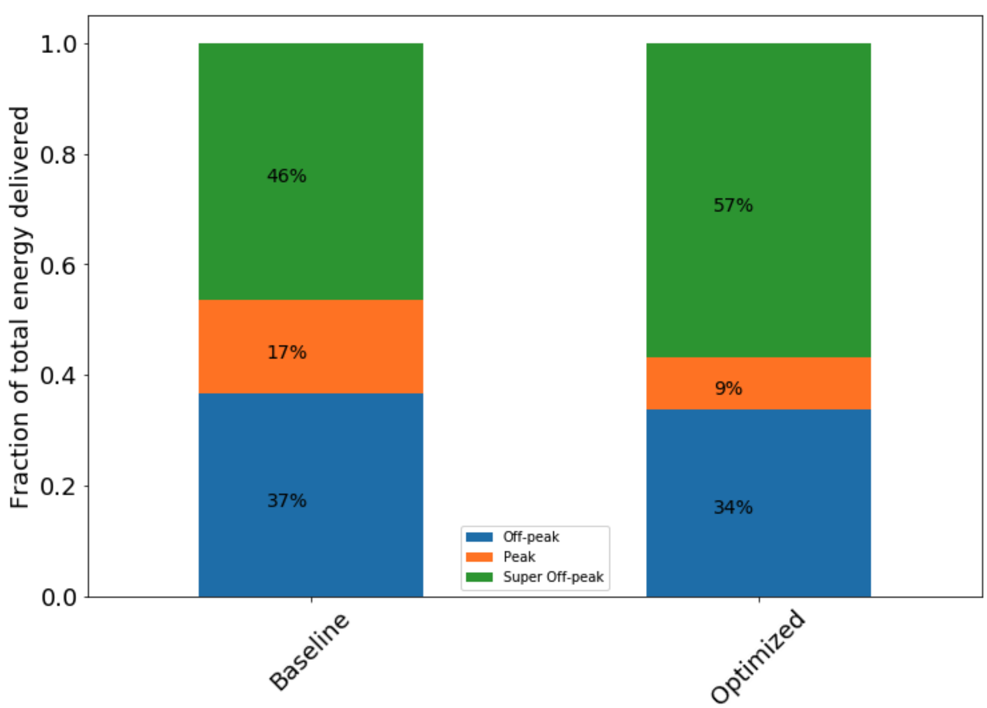


Figure IV-13: Fraction of total energy delivered during peak and off-peak periods - median simulation

On the user behavior side, an important outcome is that the increase in profit has not been the result of increasing the hourly rates that users are subjected to. In fact, the average session hourly rate in the simulation is $1.41/hr, a 6% discount over the baseline hourly rates. The increase in revenue is instead a result of a significant increase in the average plug-in duration, from 3.62 hours in the baseline to 5.3 hours under optimized pricing, as well as a decrease in the leave rate due to more discounted prices. Table IV-3 summarizes the session information for the median simulation in more detail. Overall, these results demonstrate that it is possible to simultaneously increase operator net revenue, decrease user costs, and concoct demand flexibility by optimizing both price and power using behavioral models.

Table IV-3: Session characteristics under the optimization framework - median scenario

Total Regular Sessions	23
Total Scheduled Sessions	87
Total Leave Session	22
Average Regular Price	\$2.24/hr
Average Scheduled Price	\$1.26/hr
Average Price paid (excluding Leave sessions)	\$1.41/hr
Average Session Duration	5.26 hours

6. Conclusion

In this paper, we introduced and assessed a framework to optimally operate an PEV charging station by exploiting inefficiencies that are present in the majority of existing stations. The framework jointly optimizes the pricing strategy and power at charging stations, and explicitly incorporates the key role that human behavior and decision making play in the successful operation of those stations by explicitly modeling choice behaviors and incorporating them into the optimization program. We describe the experiments used to estimate and validate those behavioral models, namely a discrete choice model of the choice between two charging options and a Kernel Regression model for the plug-in duration as a function of the hourly price. We then show, through simulations, how the framework improves the operation of the charging station by increasing net revenue and decreasing utility costs for the station operator without impacting the user experience or increasing the average cost paid by the EV owners.

Our work overcomes important limitations of the existing literature on optimizing PEV charging station operations. Specifically, there is a lack of a rigorous understanding of consumer behavior and their preferences in the PEV charging context, owing largely to the lack of good experimental data. This is particularly true for charging at the workplace. Our work fills this gap by designing and implementing randomized experiments to estimate key behaviors and relationships like the choice behavior between two charging options, and the relationship between plug-in duration and price.

A natural extension of this work is to implement this framework in real-time and compare the results to the simulations. This was not possible in this work due to technical limitations related to the charging infrastructure, which did not allow us to optimize the prices based on real-time inputs, like energy required and hour of the day. Such limitations can be overcome with a more flexible control mechanism. Another methodological contribution that could extend this work is to compare the experimental results of the price sensitivity of plug-in duration and willingness-to-delay charging to ones computed using observational data. Charging station and behavior data are becoming more available with the increased penetration of electric vehicles, and comparisons of this sort will either validate the use of observational data as a way to model charging behavior if the results are comparable to the experimental benchmarks, or emphasize the need for randomized experiments if they are not.

7. Acknowledgement

I would like to thank my collaborators on this paper for their invaluable inputs: Ayse Tugba Ozturk, Wente Zeng, and Professor Scott Moura. I would also like to thank Teng Zeng for his insights in the early phases of this research, and Total energies for funding support.

Chapter 5: Conclusion and Future Directions

5.1. Summary of Contributions

In this dissertation, I address a gap that I see in theory and practice between the fields of travel behavior and causal inference. I discuss a disconnect that manifests in misinformed statements about causality, erroneous modeling specifications that undermine the causal interpretation of modeling coefficients, and underuse of causal identification strategies and other advances in the causal inference literature by travel behavior researchers.

The dissertation comprised two parts: a conceptual part (Chapter 1), and an empirical part (Chapters 2, 3, 4). The chapters contribute the following to the transportation literature:

- The discussion presented in Chapter 1 of the disconnect between the causal inference and transportation literatures highlights important misconceptions and erroneous practices in transportation research when estimating and interpreting modeling parameters causally. Specifically, I distinguish between mediators and confounders in discussions of endogeneity in travel demand models, and show how controlling for mediators undermines the causal interpretation of the model parameters. I then introduce DAGs as a tool that transportation researchers can leverage to reason about their modeling specifications and avoid the problem of "bad controls" when the purpose is to interpret the model parameters causally.
- In Chapter 2, I overcome many of the methodological and sample size limitations of the existing literature on the transport impacts of telecommuting that predominantly either relies on large cross-sectional data, or small self-reported travel diaries data. The longitudinal aspect of the data used in this chapter coupled with the rigorous causal identification strategies I implemented allow me to control for many observed and unobserved confounders, resulting in more robust and defensible conclusions. I find overwhelming evidence that telecommuting results in the generation of about 1 non-commute trip, but that the net effect of telecommuting on distance traveled is negative, with one day of telecommuting resulting in a reduction of about 10 km in total daily distance traveled.
- The question answered in Chapter 3 results in a deeper understanding of the factors affecting the heterogeneity in the behavioral response to the pandemic among the U.S. population, as well as the role of vaccines in reversing pandemic induced mobility trends. The findings in this chapter have important implications on understanding and predicting human responses during future pandemic impact and recovery periods.
- In Chapter 4, I design and implement randomized pricing experiments, the gold standard of causal inference, in the context of user behavior at PEV charging stations. Research on consumer behavior in the PEV charging context is limited, data is lacking, and consumer preferences, especially at workplace charging stations, remain poorly understood. I address this gap, and quantify key behavioral quantities like the willingness and required incentives to delay

charging, and the relationship between plug-in duration and hourly prices. Then, I propose a novel optimization framework that incorporates those learned behaviors into the optimization objective and significantly improves the operational efficiency of PEV charging stations. My analysis shows that incorporating behavioral theory in the optimization framework results in significantly lower operational costs (up to 17%) and higher net revenues (up to 50%) for the charging station operator compared to the uncontrolled baseline, without sacrificing the user experience.

5.2. Extensions and Future Directions

This dissertation included conceptual / methodological chapters, as well as domain-specific empirical chapters. As such, the future directions of the topics explored in this dissertation will similarly be split into conceptual and domain-specific avenues.

First, on the methodological disconnect between the causal inference and travel demand modeling literature, I highlighted in this dissertation the misconceptions about causality in travel demand modeling discussions, the common issue of conditioning on post-treatment variables, also known as bad controls, distinguishing between confounders and mediators in modeling specifications, and the lack of rigor in explicitly specifying the causal design and identification assumptions required for producing valid causal inferences. An important and highly impactful extension would be to perform a comprehensive literature review of existing papers that estimate transportation choice models using observational data, extract the “implied” causal identification assumptions (ideally in the form of a causal graph) under which the conclusions drawn by the authors are valid causally, and discuss the validity of such assumptions and graphs. If those assumptions are not valid, e.g. the graph makes independence assumptions that are unlikely to hold, the next step is then to write down an alternative, valid causal graph, estimate the corresponding model based on the causal graph with causality as the goal, and then look at the change in the causal parameters of interest and the impact of this change on the research conclusions. This would be a highly impactful research area, both from a practical and conceptual standpoint. Currently, researchers in transportation demand modeling seem resistant to critiques (based on personal conversations with scholars in the field) related to the lack of causality in their frameworks. Performing such a comprehensive review and being able to quantify the consequences of not adhering to a rigorous causal framework when developing transportation demand models will force a more serious discussion in the field on the topic of causality in transportation and the lack thereof.

Second, on the effect of telecommuting on travel, I showed in Chapter 2 that telecommuting results in the generation of about 1 non-commute trip for every day of telecommuting, and that this trip is on average shorter than the two-way commute trip, meaning that the net effect of telecommuting on total distance traveled is negative. The chapter also highlighted the power of using a rigorous causal framework in providing robust and defensible estimates of the causal relationship between telecommuting and travel. In future extensions, researchers should further investigate any modal differences between telecommuting-induced non-commute trips and commute trips in order to more accurately determine the environmental impacts of telecommuting. Comparing the different timing of those trips will also prove important in determining their environmental footprints: trips made during peak congestion hours are more environmentally costly than off-peak trips. Finally, researchers should extend the existing literature on the relationship between telecommuting and home-to-work distance. The COVID-19 pandemic has resulted in a natural experiment where telecommuting rates have increased across the board: this presents an opportunity to deal with the endogeneity between telecommuting decisions and home location / distance from work. Researchers can take advantage of this emerging context to shed light on this long-standing question in transportation research.

Finally, on the topic of learning and optimizing human behavior at charging stations, I have shown in this study how one may learn accurate models for choice making between different charging options, as well as the relationship between plug-in duration and price using randomized experiments. I also showed, using simulations, how one may formulate an optimization framework that incorporates those learned behaviors to significantly improve station efficiency. A natural extension of this work is to implement this framework in real-time and compare the results to the simulations. This was not possible in this work due to technical limitations related to the charging infrastructure, but such limitations can be overcome with a more flexible control mechanism. Another methodological contribution that could extend this work is to compare the experimental results of the price sensitivity of plug-in duration and willingness-to-delay charging to ones computed using observational data. Charging station and behavior data is becoming more available with the increased penetration of electric vehicles, and comparisons of this sort will either validate the use of observational data as a way to model charging behavior if the results are comparable to the experimental benchmarks, or emphasize the need for randomized experiments if they are not.

5.3. Conclusion

The field of travel demand and behavioral modeling in transportation is often motivated by a causal objective. The models developed in this field are often used to support policy evaluations, which are causal evaluations by definition: we are interested in how the system reacts to external interventions. Despite the prevalence of those causal applications, the literature on travel demand modeling as a whole still lags behind when it comes to leveraging the advancements in the causal inference literature, and many causal inference techniques have yet to enter the field.

My aim in this dissertation was to shed some light on this disconnect, highlight examples where misconceptions of causality have been made, and give a high level overview of key causal inference concepts and identification strategies. I also aimed to practice what I preach, by answering three empirical causal questions where the focus was explicitly on estimating causal parameters, rather than on fitting predictive models.

My empirical chapters emphasized the importance of a good causal design and appropriate data collection, rather than the more commonly observed focus on modeling complexity. With weak causal designs and limited data, no degree of modeling complexity can overcome the fundamental problem of causal inference, unless the modeler is willing to resort to strong and likely implausible assumptions.

The hope is that this dissertation will spark some interest in causal inference as a field of study within the transportation research community, and will result in serious discussions of causality in travel demand modeling. I am very confident that the coupling of the two fields is bound to be beneficial for both: indeed, transportation demand modelers have made very important contributions to the topic of choice modeling given its relevance to us, and there is nothing stopping us from making the same level of contributions to the field of causal inference. Other applied disciplines like epidemiology, psychology, and political science have already done so, and given the relevance of our applications in travel demand modeling, we should join and be at the forefront of this interdisciplinary effort to tackle causality.

Bibliography

1. X. Cao, P. L. Mokhtarian, S. L. Handy, Examining the Impacts of Residential Self-Selection on Travel Behaviour: A Focus on Empirical Findings. *Transp. Rev.* **29**, 359–395 (2009).

2. P. L. Mokhtarian, X. Cao, Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Trans. Res. Part B: Methodol.* **42**, 204–228 (2008).

3. D. McFadden, The Choice Theory Approach to Market Research. *Marketing Science* **5**, 275–297 (1986).

4. K. E. Train, D. L. McFadden, A. A. Goett, Consumer Attitudes and Voluntary Rate Schedules for Public Utilities. *Rev. Econ. Stat.* **69**, 383–391 (1987).

5. M. Ben-Akiva, *et al.*, Integration of choice and latent variable models. *Perpetual motion: Travel behaviour research opportunities and application challenges* **2002**, 431–470 (2002).

6. C. R. Bhat, J. Y. Guo, A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Trans. Res. Part B: Methodol.* **41**, 506–526 (2007).

7. A. R. Pinjari, R. M. Pendyala, C. R. Bhat, P. A. Waddell, Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. *Transportation* **34**, 557–573 (2007).

8. J. L. Walker, E. Ehlers, I. Banerjee, E. R. Dugundji, Correcting for endogeneity in behavioral choice models with social influence variables. *Transp. Res. Part A: Policy Pract.* **45**, 362–374 (2011).

9. C. A. Guevara, D. Polanco, Correcting for endogeneity due to omitted attributes in discrete-choice models: the multiple indicator solution. *Transportmetrica A: Transport Science* **12**, 458–478 (2016).

10. C. A. Guevara, Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transp. Res. Part A: Policy Pract.* **82**, 240–254 (2015).

11. C. A. Guevara, M. Ben-Akiva, Endogeneity in Residential Location Choice Models. *Transp. Res. Rec.* **1977**, 60–66 (2006).

12. M. Ben-Akiva, STRUCTURE OF PASSENGER TRAVEL DEMAND MODELS in *53rd Annual Meeting of the Highway Research Board*, (1974) (August 31, 2022).

13. T. Brathwaite, J. L. Walker, Causal inference in travel demand modeling (and the lack thereof). *Journal of Choice Modelling* **26**, 1–18 (2018).

14. S. van Cranenburgh, S. Wang, A. Vij, F. Pereira, J. Walker, Choice modelling in the age of machine learning - Discussion paper. *Journal of Choice Modelling* **42**, 100340 (2022).

15. V. Chernozhukov, *et al.*, Double/Debiased/Neyman Machine Learning of Treatment Effects. *Am. Econ. Rev.* **107**, 261–265 (2017).

16. S. Wager, S. Athey, Estimation and Inference of Heterogeneous Treatment Effects using

Random Forests. *J. Am. Stat. Assoc.* **113**, 1228–1242 (2018).

17. S. Athey, J. Tibshirani, S. Wager, Generalized random forests. *aos* **47**, 1148–1178 (2019).

18. S. Athey, G. Imbens, Recursive partitioning for heterogeneous causal effects. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 7353–7360 (2016).

19. S. Athey, G. W. Imbens, Machine Learning Methods That Economists Should Know About. *Annu. Rev. Econom.* **11**, 685–725 (2019).

20. S. Athey, Beyond prediction: Using big data for policy problems. *Science* **355**, 483–485 (2017).

21. X. Nie, S. Wager, Quasi-oracle estimation of heterogeneous treatment effects. *Biometrika* **108**, 299–319 (2020).

22. J. M. Wooldridge, *Introductory Econometrics: A Modern Approach* (Cengage Learning, 2015).

23. H. M. A. Aziz, *et al.*, Exploring the impact of walk–bike infrastructure, safety perception, and built-environment on active transportation mode choice: a random parameter model using New York City commuter data. *Transportation* **45**, 1207–1229 (2018).

24. P. R. Rosenbaum, The Consequences of Adjustment for a Concomitant Variable That Has Been Affected by the Treatment. *J. R. Stat. Soc. Ser. A* **147**, 656–666 (1984).

25. C. F. Manski, *Identification Problems in the Social Sciences* (Harvard University Press, 1995).

26. E. L. Lehmann, G. Casella, *Theory of Point Estimation* (Springer New York) (August 22, 2022).

27. K. Train, *Discrete Choice Methods with Simulation* (Cambridge University Press, 2009).

28. J. Walker, M. Ben-Akiva, Generalized random utility model. *Math. Soc. Sci.* **43**, 303–343 (2002).

29. J. L. Walker, “Extended discrete choice models : integrated framework, flexible error structures, and latent variables,” Massachusetts Institute of Technology. (2001) (September 4, 2022).

30. D. B. Rubin, Causal Inference Using Potential Outcomes: Design, Modeling, Decisions. *J. Am. Stat. Assoc.* **100**, 322–331 (2005).

31. L. Keele, The Statistics of Causal Inference: A View from Political Methodology. *Polit. Anal.* **23**, 313–335 (2015).

32. A. S. Gerber, D. P. Green, *Field experiments: Design, analysis, and interpretation* (WW Norton, 2012).

33. A. Abadie, Semiparametric Difference-in-Differences Estimators. *Rev. Econ. Stud.* **72**, 1–19 (2005).

34. S. Athey, G. W. Imbens, Identification and Inference in Nonlinear Difference-in-Differences Models. *Econometrica* **74**, 431–497 (2006).

35. F. Richard Guo, A. R. Lundborg, Q. Zhao, Confounder Selection: Objectives and Approaches. *arXiv [stat.ME]* (2022).

36. C. Cinelli, A. Forney, J. Pearl, A Crash Course in Good and Bad Controls. *Sociol. Methods Res.*, 00491241221099552 (2022).

37. J. Pearl, Causal diagrams for empirical research. *Biometrika* **82**, 669–688 (1995).

38. G. W. Imbens, Potential Outcome and Directed Acyclic Graph Approaches to Causality: Relevance for Empirical Practice in Economics. *J. Econ. Lit.* **58**, 1129–1179 (2020).

39. M. A. Stinson, C. R. Bhat, Commuter Bicyclist Route Choice: Analysis Using a Stated Preference Survey. *Transp. Res. Rec.* **1828**, 107–115 (2003).

40. S. T. Ali, *et al.*, Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. *Science* **369**, 1106–1109 (2020).

41. H. Tian, *et al.*, An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).

42. J. Zhang, *et al.*, Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486 (2020).

43. J. Margraf, J. Brailovskaia, S. Schneider, Adherence to behavioral Covid-19 mitigation measures strongly predicts mortality. *PLoS One* **16**, e0249392 (2021).

44. W. Bruine de Bruin, H.-W. Saw, D. P. Goldman, Political polarization in US residents' COVID-19 risk perceptions, policy preferences, and protective behaviors. *J. Risk Uncertain.* **61**, 177–194 (2020).

45. D. P. Calvillo, B. J. Ross, R. J. B. Garcia, T. J. Smelter, A. M. Rutchick, Political Ideology Predicts Perceptions of the Threat of COVID-19 (and Susceptibility to Fake News About It). *Soc. Psychol. Personal. Sci.* **11**, 1119–1128 (2020).

46. ,Most Americans say they regularly wore a mask in stores in the past month; fewer see others doing it. *Pew Research Center* (2020) (March 28, 2022).

47. A. Tirachini, O. Cats, COVID-19 and Public Transportation: Current Assessment, Prospects, and Research Needs. *J. Public Transp.* **22** (2020).

48. J. De Vos, The effect of COVID-19 and subsequent social distancing on travel behavior. *Transp Res Interdiscip Perspect* **5**, 100121 (2020).

49. E. Jenelius, M. Cebecauer, Impacts of COVID-19 on public transport ridership in Sweden: Analysis of ticket validations, sales and passenger counts. *Transp Res Interdiscip Perspect* **8**, 100242 (2020).

50. A. Shamshiripour, E. Rahimi, R. Shabanpour, A. K. Mohammadian, How is COVID-19 reshaping activity-travel behavior? Evidence from a comprehensive survey in Chicago. *Transp Res Interdiscip Perspect* **7**, 100216 (2020).

51. J. Zhang, Transport policymaking that accounts for COVID-19 and future public health threats: A PASS approach. *Transp. Policy* **99**, 405–418 (2020).

52. D. Salon, *et al.*, The potential stickiness of pandemic-induced behavior changes in the United States. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021).

53. T. Mitchell, How Coronavirus has changed the way Americans work. *Pew Research Center's Social & Demographic Trends Project* (2020) (December 31, 2021).

54. D. Salon, *et al.*, The potential stickiness of pandemic-induced behavior changes in the United States. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021).

55. J. F. Teixeira, M. Lopes, The link between bike sharing and subway use during the COVID-19 pandemic: The case-study of New York's Citi Bike. *Transp Res Interdiscip Perspect* **6**, 100166 (2020).

56. E. Jenelius, M. Cebecauer, Impacts of COVID-19 on public transport ridership in Sweden: Analysis of ticket validations, sales and passenger counts. *Transp Res Interdiscip Perspect* **8**, 100242 (2020).

57. H. Morita, S. Nakamura, Y. Hayashi, Changes of Urban Activities and Behaviors Due to COVID-19 in Japan (2020) https:/doi.org/10.2139/ssrn.3594054 (December 31, 2021).

58. L. Tan, C. Ma, Choice behavior of commuters' rail transit mode during the COVID-19 pandemic based on logistic model. *Journal of Traffic and Transportation Engineering (English Edition)* **8**, 186–195 (2021).

59. S.-H. Cho, H.-C. Park, Exploring the Behaviour Change of Crowding Impedance on Public Transit due to COVID-19 Pandemic: Before and After Comparison. *Transportation Letters* **13**, 367–374 (2021).

60. W. Rothengatter, *et al.*, Pandemic waves and the time after Covid-19 - Consequences for the transport sector. *Transp. Policy* **110**, 225–237 (2021).

61. Y. Ram, N. Collins-Kreiner, E. Gozansky, G. Moscona, H. Okon-Singer, Is there a COVID-19 vaccination effect? A three-wave cross-sectional study. *Curr. Issues Tourism* **25**, 379–386 (2022).

62. D. Boto-García, J. Francisco Baños Pino, Propelled: evidence on the impact of vaccination against COVID-19 on travel propensity. *Curr. Issues Tourism*, 1–20 (2022).

63. C. A. Harper, L. P. Satchell, D. Fido, R. D. Latzman, Functional Fear Predicts Public Health Compliance in the COVID-19 Pandemic. *Int. J. Ment. Health Addict.* **19**, 1875–1888 (2021).

64. D. W. Snook, W. Kaczkowski, A. D. Fodeman, Mask On, Mask Off: Risk Perceptions for COVID-19 and Compliance with COVID-19 Safety Measures. *Behav. Med.*, 1–11 (2022).

65. S. G. Stolberg, Covid News: Biden's Team Sees Reasons for Optimism in Latest Data. *The New York Times* (2022) (April 19, 2022).

66. M. E. G. Parker, *et al.*, Public transit use in the United States in the era of COVID-19: Transit riders' travel behavior in the COVID-19 impact and recovery period. *Transp. Policy* **111**, 53–62 (2021).

67. T. Mitchell, COVID-19 pandemic continues to reshape work in America. *Pew Research Center's Social & Demographic Trends Project* (2022) (March 28, 2022).

68. K. Sneppen, B. F. Nielsen, R. J. Taylor, L. Simonsen, Overdispersion in COVID-19 increases the effectiveness of limiting nonrepetitive contacts for transmission control. *Proc. Natl. Acad. Sci. U. S. A.* **118** (2021).

69. F. Saldaña, J. X. Velasco-Hernández, The trade-off between mobility and vaccination for COVID-19 control: a metapopulation modelling approach. *R Soc Open Sci* **8**, 202240 (2021).

70. C. Spinella, A. M. Mio, Simulation of the impact of people mobility, vaccination rate, and virus variants on the evolution of Covid-19 outbreak in Italy. *Sci. Rep.* **11**, 23225 (2021).

71. D. Salon, M. G. Boarnet, S. Handy, S. Spears, G. Tal, How do local actions affect VMT? A critical review of the empirical evidence. *Transp. Res. Part D: Trans. Environ.* **17**, 495–508 (2012).

72. R. Kitamura, P. L. Mokhtarian, R. M. Pendyala, An Evaluation of Telecommuting As a Trip Reduction Measure (1991) (May 3, 2022).

73. D. K. Henderson, P. L. Mokhtarian, Impacts of center-based telecommuting on travel and emissions: Analysis of the Puget Sound Demonstration Project. *Transp. Res. Part D: Trans. Environ.* **1**, 29–45 (1996).

74. J. M. Nilles, Traffic reduction by telecommuting: A status review and selected bibliography. *Transportation Research Part A: General* **22**, 301–317 (1988).

75. P. L. Mokhtarian, A Synthetic Approach to Estimating the Impacts of Telecommuting on Travel. *Urban Stud.* **35**, 215–241 (1998).

76. S.-N. Kim, S. Choo, P. L. Mokhtarian, Home-based telecommuting and intra-household interactions in work and non-work travel: A seemingly unrelated censored regression approach. *Transp. Res. Part A: Policy Pract.* **80**, 197–214 (2015).

77. S.-N. Kim, Is telecommuting sustainable? An alternative approach to estimating the impact of home-based telecommuting on household travel. *International Journal of Sustainable Transportation* **11**, 72–85 (2017).

78. P. Zhu, L. Wang, Y. Jiang, J. Zhou, Metropolitan size and the impacts of telecommuting on personal travel. *Transportation* **45**, 385–414 (2018).

79. P. Zhu, Are telecommuting and personal travel complements or substitutes? *Ann. Reg. Sci.* **48**, 619–639 (2012).

80. S.-N. Kim, Two traditional questions on the relationships between telecommuting, job and residential location, and household travel: revisited using a path analysis. *Ann. Reg. Sci.* **56**, 537–563 (2016).

81. H. Budnitz, E. Tranos, L. Chapman, Telecommuting and other trips: an English case study. *J. Transp. Geogr.* **85**, 102713 (2020).

82. P. C. Melo, J. de Abreu e Silva, Home telework and household commuting patterns in Great

Britain. *Transp. Res. Part A: Policy Pract.* **103**, 1–24 (2017).

83. J. de Abreu e Silva, P. C. Melo, Does home-based telework reduce household total travel? A path analysis using single and two worker British households. *J. Transp. Geogr.* **73**, 148–162 (2018).

84. B. Caldarola, S. Sorrell, Do teleworkers travel less? Evidence from the English National Travel Survey. *Transp. Res. Part A: Policy Pract.* **159**, 282–303 (2022).

85. M. T. Jamal, N. A. Khan, Adoption of telecommuting and its impact on traffic congestion, energy consumption and emission levels in India. *Interdisciplinary Environmental Review* **21**, 249–268 (2021).

86. J. C. T. Bieser, B. Vaddadi, A. Kramers, M. Höjer, L. M. Hilty, Impacts of telecommuting on time use and travel: A case study of a neighborhood telecommuting center in Stockholm. *Travel Behaviour and Society* **23**, 157–165 (2021).

87. W. Marz, S. Sen, Does Telecommuting Reduce Commuting Emissions? (2021) https:/doi.org/10.2139/ssrn.3945279 (March 29, 2022).

88. K. Parker, J. M. Horowitz, R. Minkin, COVID-19 pandemic continues to reshape work in America. *Pew Research Center's Social & Demographic Trends Project* (2022) (June 25, 2022).

89. X. Mosquet, A. Arora, A. Xie, M. Renner, Who will drive electric cars to the tipping point? *United States - EN* (2020) (June 20, 2022).

90. State of California, Governor Newsom announces California will phase out gasoline-powered cars & drastically reduce demand for fossil fuel in California's fight against climate change. *California Governor* (2020) (June 20, 2022).

91. J. G. Smart, S. D. Salisbury, "Lessons learned about plug-in electric vehicle charging infrastructure from the EV Project and ChargePoint America" (Idaho National Laboratory (INL), 2015) (August 10, 2022).

92. ,Advanced Clean Cars midterm review (August 10, 2022).

93. J. Coignard, P. MacDougall, F. Stadtmueller, E. Vrettos, Will Electric Vehicles Drive Distribution Grid Upgrades?: The Case of California. *IEEE Electrification Magazine* **7**, 46–56 (2019).

94. D. Fischer, A. Harbrecht, A. Surmann, R. Mckenna, Electric vehicles' impacts on residential electric local profiles - A stochastic modelling approach considering socio-economic, behavioural and spatial factors. *Appl. Energy* (2018) https:/doi.org/10.1016/j.apenergy.2018.10.010 (August 9, 2022).

95. M. A. Nicholas, G. Tal, T. S. Turrentine, Advanced Plug-in Electric Vehicle Travel and Charging Behavior Interim Report (2017) (August 9, 2022).

96. M. A. Nicholas, G. Tal, Charging for Charging: The Paradox of Free Charging and its Detrimental Effect on the Use of Electric Vehicles (2013) (August 9, 2022).

97. M. Smith, J. Castellano, "Costs Associated With Non-Residential Electric Vehicle Supply

Equipment: Factors to consider in the implementation of electric vehicle charging stations" (2015) (August 10, 2022).

98. M. Nicholas, G. Tal, Dynamics of Workplace Charging for Plug-in Electric Vehicles: How Much is Needed and at What Speed? *World Electric Vehicle Journal* **6**, 819–828 (2013).

99. P. B. Andersen, T. Sousa, A. Thingvad, L. S. Berthou, M. Kulahci, Added Value of Individual Flexibility Profiles of Electric Vehicle Users For Ancillary Services in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, (2018), pp. 1–6.

100. E. Bitar, Y. Xu, Deadline Differentiated Pricing of Deferrable Electric Loads. *IEEE Trans. Smart Grid* **8**, 13–25 (2017).

101. A. I. Arif, *et al.*, Online scheduling of plug-in vehicles in dynamic pricing schemes. *Sustainable Energy, Grids and Networks* **7**, 25–36 (2016).

102. L. Calearo, A. Thingvad, K. Suzuki, M. Marinelli, Grid Loading Due to EV Charging Profiles Based on Pseudo-Real Driving Pattern and User Behavior. *IEEE Transactions on Transportation Electrification* **5**, 683–694 (2019).

103. A. González-Garrido, A. Thingvad, H. Gaztañaga, M. Marinelli, Full-scale electric vehicles penetration in the Danish Island of Bornholm—Optimal scheduling and battery degradation under driving constraints. *Journal of Energy Storage* **23**, 381–391 (2019).

104. P. Alexeenko, E. Bitar, Achieving Reliable Coordination of Residential Plug-in Electric Vehicle Charging: A Pilot Study. *arXiv [eess.SY]* (2021).

105. R. A. Daziano, Willingness to delay charging of electric vehicles. *Research in Transportation Economics*, 101177 (2022).

106. A. Biswas, R. Gopalakrishnan, P. Dutta, Digital Object Identifier System (June 19, 2022).

107. F. Lambert, Tesla increases Supercharger idle fees to decrease wait times. *Electrek* (2018) (June 19, 2022).

108. D. Wu, H. Zeng, C. Lu, B. Boulet, Two-Stage Energy Management for Office Buildings With Workplace EV Charging and Renewable Energy. *IEEE Transactions on Transportation Electrification* **3**, 225–237 (2017).

109. O. I. Asensio, C. Z. Apablaza, M. C. Lawson, S. E. Walsh, A field experiment on workplace norms and electric vehicle charging etiquette. *J. Ind. Ecol.* **26**, 183–196 (2022).

110. Z. J. Lee, T. Li, S. H. Low, ACN-Data: Analysis and Applications of an Open EV Charging Dataset in *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, e-Energy '19., (Association for Computing Machinery, 2019), pp. 139–149.

111. B. Ferguson, V. Nagaraj, E. C. Kara, M. Alizadeh, Optimal Planning of Workplace Electric Vehicle Charging Infrastructure with Smart Charging Opportunities in *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, (2018), pp. 1149–1154.

112. O. I. Asensio, M. C. Lawson, C. Z. Apablaza, Electric vehicle charging stations in the

workplace with high-resolution data from casual and habitual users. *Sci Data* **8**, 168 (2021).

113. Y. Ge, D. MacKenzie, “Modeling vehicle choices and charging behavior of plug-in electric vehicle owners jointly using dynamic discrete choice model” (2018).

114. M. Kajanova, P. Bracinik, P. Belány, Analysis of the discrete choice model representing the electric vehicle owners’ behavior in Slovakia. *Univ. Polytehnica Bucharest Sci. Bull. Ser. C: Electr. Eng.* **104**, 131–141 (2022).

115. E. Guerra, R. A. Daziano, Electric vehicles and residential parking in an urban environment: Results from a stated preference experiment. *Transp. Res. Part D: Trans. Environ.* **79**, 102222 (2020).