Title
Essays in Environmental and Public Economics

Author
Tarduno, Matthew Alexander

Publication Date
2022

Peer reviewed|Thesis/dissertation

Essays in Environmental and Public Economics

by

Matthew Tarduno

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor James Sallee, Chair
Associate Professor W. Reed Walker
Professor Michael Anderson

Spring 2022

Essays in Environmental and Public Economics

Abstract

Essays in Environmental and Public Economics

by

Matthew Tarduno

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Associate Professor James Sallee, Chair

Economists have long prescribed taxing externality-generating goods according to their marginal damages to internalize the social costs of economic activity. Real-world policies, however, often differ significantly from this prescription. In some settings, practical constraints render perfect Pigouvian taxation infeasible. In other settings, political systems favor imperfect externality regulation, even when first-best policies are possible. In this dissertation, I study externality regulation in the context of environmental pollution and traffic congestion, with an eye on designing optimal policies in the face of political and practical constraints on policymakers.

In the first chapter, I study *cordon zones*, regions in the center of cities where cars are charged a toll for entering during peak hours. These policies are blunt approximations for Pigouvian taxes: uniform prices cannot reflect all of the heterogeneity in congestion and pollution damages associated with urban driving, and the discrete spatial and temporal cutoffs commonly used in cordon zones invite externality leakage. To understand the implications of these imperfections for setting optimal cordon prices, I first extend existing models from public economics to characterize optimal cordon tolls. This exercise yields a set of parameters necessary for calculating optimal road prices. I then recover estimates of these parameters using a natural experiment where bridge tolls were adjusted in the San Francisco Bay Area.

Using these parameters, I estimate optimal cordon tolls for three US cities: New York, San Francisco, and Los Angeles, each of which is currently planning on implementing cordon pricing. Across these three cities, I find that because drivers are willing to substitute to unpriced alternative routes or travel times in response to tolls, optimal peak-hour cordon prices are *below* the average social damages associated with trips passing through the cordon

zone. Counterfactual simulations suggest that while cordon pricing would generate substantial welfare gains in each of the above cities, the imperfections in these policies mean that a majority of the theoretically possible welfare gains are left unrealized even under optimal peak-hour cordon pricing. To conclude this chapter, I discuss the prospects of improving the performance of cordon zones by allowing for granular time-of-day tolls, or by redesigning the cordon zones themselves.

In the second chapter, I investigate why voters support inefficient (standard-based) environmental policies over efficient (price-based) environmental policies. Using an information provision experiment conducted before and after voting on an environmental ballot initiative in Nevada, I estimate a model that maps voters' beliefs about the attributes of externality-regulating policies (effectiveness, cost, regressivity) to their support for these policies. I find that voting behavior is relatively unresponsive to perceived cost and perceived regressivity, but responsive to perceived policy effectiveness. Using this model, I decompose differences in support for a performance-based policy (Nevada's Renewable Portfolio Standard) and a hypothetical price-based policy (a carbon tax). Oaxaca-Blinder decompositions imply that differences in perceptions of policy attributes explain just 23% of the gap in support between renewable portfolio standards and carbon taxes, suggesting a significant role for "tax aversion." To the extent that misperceptions of policy attributes do explain differences in support for these two policies, the explained gap results from overly optimistic beliefs about the attributes of renewable portfolio standards. To conclude, I predict voting behavior under several counterfactual scenarios. I find that in this setting, targeting revenue toward "swing" voters is unlikely to significantly improve support for carbon taxes. Instead, the results of this pilot experiment highlight the importance of communicating to voters the efficacy of price-based policies.

In the third chapter, I leverage a natural experiment to study whether Transportation Network Companies (TNCs) like Uber and Lyft worsen traffic congestion, and discuss the policy implications for cities considering regulating or taxing these companies. Applying difference in differences and regression discontinuity specifications to high-frequency traffic data, I estimate that Uber and Lyft together decreased daytime traffic speeds in Austin by roughly 2.3%. Using Austin-specific measures of the value of travel time, I translate these slowdowns to estimates of citywide congestion costs that range from $33 to $52 million annually. Back of the envelope calculations imply that these costs are similar in magnitude to the consumer surplus provided by TNCs in Austin, meaning that the entrance of these companies can be thought of as a transfer of welfare from incumbent road users to TNC customers.

Because it is less costly for cities to track the movements of ridesharing vehicles than it is to build the infrastructure for cordon pricing, TNC taxes are often proposed as an alternative to conventional congestion pricing. Anecdotally, taxing Uber and Lyft also enjoys a political advantage over cordon-based pricing (for example, New York, Chicago, and San Francisco all passed the TNC taxes, but are each engaged in political debates over cordon pricing). The results in this paper suggest that even if TNC taxes or restrictions are politically expedient, they are unlikely to generate large net welfare gains through reduced congestion.

Contents

List of Figures

List of Tables

Acknowledgments

I owe a debt of gratitude to many people who helped me during my time as a graduate student. First, my advisors (James Sallee, Michael Anderson, and Reed Walker) have provided invaluable support and mentorship over the past five years. Jim's dedication to his students during the depths of the pandemic was truly exceptional. More than a few times, he was able to take my vague research ideas and refine them into something coherent, all over zoom, and while juggling a score of other advisees. What I appreciate most about Jim is that his feedback and advice always focused on long-term development. He took the time not just to help me over the next hurdle, but also to provide the perspective necessary to see the forest for the trees. He is a true role model, and I hope to emulate his mentoring and teaching style in my career.

Reed provided consistent feedback and a much-needed perspective on my research ideas, and always encouraged me to work smarter, rather than harder. His advice to broaden my circle of mentors beyond ARE has undoubtedly made me a more well-rounded economist, and more well-prepared for the market to boot.

I am particularly grateful for Michael's support during the second-year paper sequence, and for helping me through the subsequent submission and review process. As the profession marches unblinkingly toward a norm of 100-page research papers, Michael's ability to answer research questions both concisely and convincingly seems an increasingly rare skill, and the papers in this dissertation are the better for it.

Beyond my advisors, I have been fortunate to be surrounded by a spectacular group of mentors at Berkeley. I would like to thank Ernesto Dal Bó and Joseph Shapiro for providing insightful feedback as members of my orals committee, Meredith Fowlie and Sofia Villas Boas for elucidating the mechanics of discrete choice models, Ricardo Perez-Truglia for encouraging me to use information provision experiments in my work, Lucas Davis for his advice on both my second year and job market papers, and Severin Borenstein for many engaging lunch conversations.

I am deeply indebted to my fellow members of the 2017 ARE cohort. If not for their support and camaraderie, I'd probably be lost in some back room of Doe Library, still puzzling over Bellman equations. I could always count on Jesse Buchsbaum to keep things light and to help me strike the right balance between work and everything else that the Bay had to offer. Marshall Blundell helped all of us discover the hidden hikes and restaurants of Berkeley. And Jenya Kahn-Lang, Livia Alfonsi, and (honorary cohort member) Mia Smith's gatherings made Berkeley feel like a home. Outside of my cohort, I owe thanks to many others at ARE, but especially to Carmen Karahalios and Diana Lazo for supporting me on the market. ARE as it exists today would not be possible without their hard work.

I was also extremely fortunate to be a member of the Energy Institute at Haas. EI has been a wonderful academic community, and I'm grateful for the work done by the leadership and staff at EI to keep the lights on during the pandemic. I'd also like to thank my fellow EI GSRs — Susanna Berkouwer, Karl Dunkle Werner, and Stephen Jarvis for their friendship and guidance when I first arrived to EI, and Leila Safavi, Elif Tasar, and Sarah Johns for making Giannini less lonely as campus re-opened. I'm also grateful to Karen Notsund and Cristina Bentley, who saved me on more than a few occasions, and to Andy Campbell, who always brought curiosity and good spirits to lunches, not to mention fantastic salsa and coffee recommendations.

To my friends — especially Greg Kehne, Matt Rock, Greg Ferland, Aldis Inde, Kaleigh Kenny, Keelia and Chris Alder, Chris Owyang, Alice Murphy, and Henry Schmidt — thank you for keeping me grounded, sane, and laughing over the past five years. My sincere apologies for all of the times that I made your eyes glaze over with unsolicited descriptions of my research agenda, and for all such future infractions.

Anya Marchenko deserves a paragraph to herself: From Porto to Mendocino to the Rose Gardens of the East Bay, you've helped me find joy and adventure in spite of the pandemic. Your infectious curiosity helped me re-discover my love of the discipline, your questions keep me honest and provide perspective, and your confidence in me helped me through the low points of the market. I am beyond grateful, and I hope that I can return these favors as you start your own grad school journey.

I'd be remiss if I didn't also mention the role that the running community has played in my life and my studies. I am in many ways a reflection of my teammates and coaches at Brighton, Williams, and Berkeley. As the structure of college dissolved into the disorder of Ph.D. candidacy, the discipline I learned as a Baron, Eph, and Strawberry was oftentimes the only thing keeping me on track. Thank you for challenging and supporting me; *peace be your journey.*

I also owe thanks to the people who helped me find my way to Berkeley: Matthew Gibson was tremendously helpful in both preparing me for graduate school, but also in deciding whether it was the right path for me. The fact that I was one of his first advisees at Williams makes this all the more impressive. Marcella Alsan was also a crucial part of my path to Berkeley — from her I learned how to shoot high, as well as how to break lofty goals into manageable pieces. Her incredible papers and enthusiasm for those around her continue to inspire me today.

I'm not sure I'd be an economist if not for Michael Noto. I showed up in his economics elective at Brighton High School in the fall of 2011 after failing to test out of the school's economics requirement. Over the course of the following semester, through enthralling lectures about the great recession, NAFTA, and the lack of free lunches, he shook my confidence

in my plan to major physics in college. Two years later, after thoroughly burning out on Electricity and Magnetism as a college sophomore, economics was there waiting for me.

Finally, I'm thankful for the love and encouragement of my parents, John and Beth Tarduno, and to my sister Maria Tarduno, who is a constant source of inspiration and positive energy for those around her.

Chapter 1

For Whom the Bridge Tolls: Congestion, Air Pollution, and Second-Best Road Pricing

1.1. Introduction

Economists have long advocated for charging road users to address the negative externalities associated with urban driving (Vickrey, 1963; Johnson, 1964; Parry, 2002). Following early policy experiments in Singapore and London, a growing number of cities including New York, Los Angeles, and San Francisco, are considering implementing road pricing. Despite its history of advocating for road pricing, however, the economics literature offers little insight on how to implement road pricing in practice, especially given that real-world policy instruments differ significantly from the first-best policies prescribed by economists.

A first-best road pricing policy would charge drivers for the marginal social damages (the time cost imposed on others plus the social cost of pollution generated) associated with every vehicle trip. Practical constraints, however, render first-best road pricing infeasible in most settings. Implementing a first-best policy would require detailed information about each driver's routes and emissions, as well as real-time traffic data. It is typically too costly to collect this information through a passive sensor network, and proposals for GPS-based pricing schemes are often rejected on privacy grounds (Lehe, 2019; Giuliano, 1992). Consequently, city-wide road pricing often takes the form of *cordon zones* — regions in the center of a city where drivers are charged for entry. Real-world road pricing schemes therefore deviate from the first-best policy along two important dimensions: First, feasible cordon systems cannot account for all of the heterogeneity in congestion and pollution externalities across trips that all enter the cordon. Second, cordon zones leave nearby roads unpriced,

allowing for externality leakage. As a result, it is generally unclear how to set cordon prices even if policymakers have perfect information about the social damages associated with trips that pass through the city center (Parry, 2009).

In this paper, I adapt models from public finance to characterize optimal cordon prices in the face of these policy imperfections. I then generate empirical estimates of how drivers would respond to road pricing, and use these estimates together with formulas derived from the theoretical framework to calculate second-best cordon prices.

The second-best pricing framework I build stipulates a set of parameters necessary for calculating second-best road prices accounting for both leakage and imperfect pricing (i.e., many vehicle trips with different externalities are charged the same price). Calculating optimal prices requires information about (i) the heterogeneity in marginal trip-level externalities, (ii) the correlation between trip-level externalities and individual price-responsiveness, and (iii) the elasticity of substitution between priced and unpriced trips. Outside of road pricing, this framework can be applied to any setting where externality heterogeneity and leakage simultaneously prevent the implementation of a first-best corrective policy (e.g., electricity markets, or sin taxes).

In the empirical section of this paper, I use a natural experiment from the San Francisco Bay Area to recover estimates of each of the parameters necessary to calculate optimal cordon prices. In 2010, bridge tolls increased on all of the region's bridges, and peak-hour pricing was implemented on the region's busiest bridge. I use this variation in road prices together with administrative microdata from the region's electronic tolling system to estimate a discrete choice model of driving demand. The results from this exercise imply that the two policy imperfections — leakage and heterogeneity — create a tension in optimal cordon pricing. Trips associated with higher social damages are more elastic. Absent leakage, this heterogeneity would imply second-best optimal prices that are *above* average social damages (Diamond, 1973). The discrete spatial and temporal cutoffs in cordon pricing, however, incentivize some drivers to shift trips in time and space to avoid tolls. Absent heterogeneity, this leakage would imply optimal prices that are *below* average social damages (Green and Sheshinski, 1976). The structure of this discrete choice model simplifies the information required to apply the second-best tax framework (items (i)-(iii), above). Namely, it allows me to populate a substitution matrix between alternative driving times and routes based on a small number of parameters that describe driving choices.

I use this model of driving demand to calculate second-best optimal prices for the proposed cordon zones in San Francisco, Los Angeles, and New York. I find that the leakage effect strongly dominates the heterogeneity effect in each of these cities, resulting in second-best optimal prices that are below the average social damages associated with trips that enter the cordon. In San Francisco, for example, when cordon prices are constrained to

peak hours, the second-best optimal prices that account for both heterogeneity and leakage are $2 to $3. This is roughly half of the average social damages generated by trips that use the cordon during those periods ($4 to $6). Unsurprisingly, peak-hour cordon pricing performs poorly relative to the (infeasible) Pigouvian prescription. The second-best optimal road pricing scheme in San Francisco, for example, achieves only 28% of the total welfare gains, 30% of the congestion reductions, and 22% of pollution reductions relative to a policy where drivers are charged according to the marginal damages of each trip. Across the three cities that I examine, I find that optimal peak-hour cordon prices are more effective at internalizing congestion than they are at internalizing pollution. This reflects the fact that while congestion and pollution externalities are spatially correlated, average trip-level pollution damages do not exhibit the same within-day variation as congestion externalities, and are therefore poorly targeted by peak-hour congestion prices.

To conclude, I investigate the prospects for improving cordon pricing policies. Allowing a policymaker to set a fixed schedule of hourly prices between 6 a.m. and 7 p.m. generates sizable welfare gains relative to a cordon policy constrained to charge prices only during peak hours. I estimate that these welfare gains range from $146 million annually in San Francisco to $286 million annually in New York. In each city, however, a cordon zone with second-best hourly prices would realize less than half of the welfare gains possible under first-best prices.

This paper makes three primary contributions. First, this paper provides the first empirical estimates of optimal cordon prices that account for both pollution and congestion. I recover optimal peak-hour cordon prices that range from $2.20 in San Francisco to $7.92 in New York. While there are robust literatures documenting the reduced-form relationship between road pricing and traffic speeds (Yang, Purevjav, and Li, 2020; Gibson and Carnovale, 2015; Leape, 2006), as well as traffic and local air pollution (Currie and Walker, 2011a; Anderson, 2020; Gibson and Carnovale, 2015; Knittel, Miller, and Sanders, 2016; Tonne, Beevers, Armstrong, Kelly, and Wilkinson, 2008), these results have yet to be combined into optimal cordon prices that account for both of these externalities, as noted by Parry (2009). Importantly, the optimal road prices presented in the paper also account for imperfections in real-world policies. Both theoretical and empirical studies suggest that while price or quantity-based cordons can ameliorate pollution and congestion in some settings (Zhong, Cao, and Wang, 2017; Börjesson, Eliasson, Hugosson, and Brundell-Freij, 2012), policies designed without regard to agent re-optimization and heterogeneity may lead to poor or perverse policy outcomes (Davis, 2008, 2017; Zhang, Lawell, and Umanskaya, 2017; Hanna, Kreindler, and Olken, 2017; Green, Heywood, and Paniagua, 2020). Calculating optimal cordon prices through a second-best tax framework explicitly accounts for these considerations.

Second, this paper contributes to the literature on externality taxation by characterizing

second-best prices in the presence of both heterogeneous externalities *and* externality leakage. This framework combines two canonical models of second-best pricing: the "Diamond" model (Diamond, 1973), which shows that second-best uniform prices are a weighted average of heterogeneous externalities, and the "leakage" model, where second-best optimal prices reflect marginal damages, less a term that captures leakage (substitution) to other unpriced goods that also generate externalities (Green and Sheshinski, 1976, see also Davis and Sallee, 2020; Gibson, 2019; Holland, 2012). Specifically, I consider the setting where there are many externality-generating goods, the externalities vary across consumers and goods, and only a subset of the goods are taxable. I show that in the presence of both heterogeneity and substitution, the optimal second-best tax formula combines characteristics of the canonical Diamond and leakage models. Holding fixed all other taxes, the optimal tax on any *one* good is the Diamond-weighted marginal damages associated with the consumption of the good, less a term governed by the Diamond-weighted leakage to other goods. The optimal second-best tax vector solves a system of equations where terms in this system reflect individual externalities, own-price elasticities, and cross-price elasticities. This characterization is most closely related to Allcott, Lockwood, and Taubinsky (2019), who characterize the optimal vector of taxes on sugary drinks in the setting with welfare weights that reflect a planner's distaste for inequality. The optimal taxation problem in this paper also resembles the optimal collection of government revenue when taxes are distortionary (Ramsey, 1927). Namely, the solution to the road pricing problem involves a matrix of substitution elasticities, as does the general solution to the canonical Ramsey tax problem. In this setting, an untaxed good's idiosyncratic externality is analogous to each good's distortions in the many-good Ramsey problem.

This extension of optimal second-best pricing is applicable in settings outside of transportation. In energy markets, for example, the externalities associated with electricity generation differ based on the location of powerplants (urban or rural; upwind or downwind of population centers), and policies implemented by states or utilities may allow for externality leakage if electricity is imported from other jurisdictions. Sin taxes (e.g., alcohol taxes, cigarette taxes) similarly have heterogeneous impacts on consumers, and taxing any single product may induce consumers to substitute towards related (and undertaxed) sin products.

Lastly, this paper presents a new approach for estimating the willingness of commuters to shift the schedule of their trips. Scheduling costs are key parameters in the transportation economics literature (Vickrey (1963), Arnott, De Palma, and Lindsey (1990), Arnott, De Palma, and Lindsey (1993)) and an important driver of the theoretical welfare gains from congestion pricing (Kreindler, 2018). Adapting tools from the public finance literature on bunching (Saez, 2010; Kleven and Waseem, 2013), I develop an estimator that infers scheduling costs from the excess density of trips taken during times of day that fall just outside a

peak pricing window. Because peak pricing is used to alleviate congestion in bridges and tunnels in many cities, this estimation approach can be applied to understand scheduling in many other metro areas.

The rest of this paper is organized as follows: Section 1.2 characterizes the second-best optimal externality taxes in the presence of heterogeneity and leakage. Section 1.3 details the discrete choice model of driving demand that I use to back out the statistics necessary to estimate optimal prices. Section 1.4 outlines the setting and natural experiment that I use to estimate the model of driving demand, and Section 1.5 covers the data. In Section 1.6, I describe the empirical strategy that I use to estimate the model of driving demand. I present results in Sections 1.7 and 1.8, discuss these results in Section 1.9, and conclude in Section 1.10.

1.2. Theory: Externality Taxation Under Heterogeneity and Leakage

Public economics provides an unambiguous prescription for addressing market externalities: apply a (Pigouvian) tax equal to the marginal damages associated with consuming the externality-generating good. In practice, however, policy instruments typically lack the precision and coverage to execute this prescription. When corrective taxation cannot account for heterogeneous externalities or leakage (substitution) to other externality-generating goods, the second-best optimal tax on any given good may differ substantially from the tax instituted in the ideal Pigouvian policy. In this section, I outline canonical models for optimal taxation under each of these separate imperfections (heterogeneity and leakage), and then present a model that can be applied to instances where heterogeneity and leakage simultaneously prevent the implementation of the first-best.

1.2.1. Heterogeneity

For practical or legal reasons, policymakers are often constrained to apply a uniform corrective price to a good where the consumption externalities associated with that good are not uniform. In cordon zones, for example, drivers typically face a single charge for daytime trips, or a toll that charges one price for peak-hour trips, and a lower price for off-peak trips.

Under these pricing schemes, many trips that generate different externalities are charged the same price.[1] Sources of congestion heterogeneity include the total length of the trip, the

[1] Verhoef, Nijkamp, and Rietveld (1995) provide a theoretical overview of the Diamond model (which characterizes optimal corrective taxes under heterogeneous damages) as it applies to heterogeneous congestion and pollution externalities.

time that the trip is taken, and the specific roads used within and outside of the cordon. Sources of pollution heterogeneity include vehicle attributes, travel speed, and trip length.

Diamond (1973) characterizes the second-best optimal uniform tax on a good which generates heterogeneous externalities when consumed by different agents: The optimal tax is a weighted average of the individual externalities, where the weights (henceforth *Diamond weights*) are the individual own-price elasticities.

Formally, consider n consumers that derive utility from their consumption of an externality-generating good, α_i, and disutility from other's consumption of this good:

$$U^i = U(\alpha_1, ..., \alpha_i, ...\alpha_n) + \mu_i$$

The second-best optimal uniform tax in this setting is:

$$\tau^* = \frac{-\sum_h \sum_{h \neq i} \frac{\partial U^h}{\partial \alpha_i} \alpha_i'}{\sum_h \alpha_h'} \tag{1.1}$$

Where α_i' is the derivative of consumer i's demand for α with respect to the price of α, and $\frac{\partial U^h}{\partial \alpha_i}$ is the marginal external cost that consumer i imposes on consumer h by consuming α.

This expression captures an important principle in second-best corrective taxation: If individual elasticities are positively (negatively) correlated with idiosyncratic externalities, the second-best uniform tax on the externality-generating good will be larger (smaller) than the naive average of marginal damages. Intuitively, the role of corrective taxes is to move individuals to adjust their consumption of a product to the level where private marginal benefit equals the social marginal cost. If a given group is unresponsive to price, however, the second-best optimal tax described above will provide the correct incentive for the responsive group to consume at the level that balances private and social marginal costs.

1.2.2. Leakage

Just as legal or practical constraints prevent policymakers from perfectly targeting externalities, these constraints often also prevent policymakers from pricing all related externality-producing goods. Cordon prices, for example, price only trips that pass over the cordon's boundary, leaving trips that avoid the cordon unpriced.

Green and Sheshinski (1976) show that in the case of two externality-generating goods (one of which is taxable and one of which is not) and homogeneous marginal damages, the second-best prescription is to tax the taxable good at its marginal damages, less a term that

is increasing in the substitutability of the two goods, and increasing in the marginal damages of consuming the untaxable good.

Formally, consider two goods, x and y, with associated marginal external damages ϕ_x and ϕ_y. A representative consumer with an exogenous income derives utility from these two goods, and a quasilinear numeraire good, z:

$$U = U(x, y) + z$$

If a social planner is constrained to only tax x, the optimal tax is:

$$\tau_x^* = \phi_x + \frac{dy/dp_x}{dx/dp_x}\phi_y \tag{1.2}$$

The second-best optimal price balances the direct social damages associated with consumption of the taxable good (ϕ_x), with the leakage-associated social damages that result from an increase in the price of the taxable good ($\frac{dy/dp_x}{dx/dp_x}\phi_y$). In this paper, I will refer to the first term in this expression ($\frac{dy/dp_x}{dx/dp_x}$) as the *leakage share* between x and y.

As noted by Davis and Sallee (2020), equation 1.2 reflects separability insights from Kopczuk (2003): The second-best tax on x is the sum of a) direct damages associated with the consumption of this good, an b) a term that captures interactions between x and existing market distortions.

In the remainder of this section, I cover two extensions to the above models. In Section 2.3, I characterize optimal taxes for a general set of externality-generating goods, where only a subset of them can be taxed. In Section 2.4, I characterize optimal taxes for a general set of externality-generating goods, where only a subset of them can be taxed, *and* marginal damages are heterogeneous by consumer.

1.2.3. Leakage with Many Goods

Before characterizing second-best optimal taxes under both heterogeneity and leakage, I first extend the two-good model in Section 2.2 to the case of many (homogeneous) externality-generating goods, some of which are untaxable. This problem is a generalization of the two and three-good direct vs. indirect taxation problems presented in Green and Sheshinski (1976) and Sandmo (1978), and provides intuition useful for understanding the model with heterogeneity presented in Section 2.4.

Setup: A representative consumer chooses quantities of M goods, $(h_1, ..., h_M)$ and a numeraire, z. Each non-numeraire good has an associated (homogeneous) externality, ϕ_m

that is linear in the consumption of m. A policymaker can choose tax levels τ_j for goods $j \in \{1, ..., J\}$ where $J < M$. I assume goods $k \notin \{1, ..., J\}$ are un- or under-taxed.

In Appendix A.1, I show that under these constraints the optimal tax for good j holding fixed the taxes on all other taxable goods k is:

$$\tau_j = \phi_j + \frac{1}{\frac{\partial h_j}{\partial p_j}} \left(\sum_{k \neq j}^{J} \frac{\partial h_k}{\partial p_j} [\phi_k - \tau_k] + \sum_{l=J+1}^{M} \frac{\partial h_l}{\partial p_j} \phi_l \right) \tag{1.3}$$

This intermediate results is a generalization of the two-good case. Holding fixed all taxes other than τ_j, the optimal value for this final tax is its externality, ϕ_m, plus a term that captures the extent to which consumers switch to other goods, and the level of unpriced externality of those goods. Identifying the optimal tax vector requires simultaneously solving J equations in the form of Equation 1.3.

To do so, one can rewrite Equation 1.3 to separate the tax and externality terms:

$$\tau_j + \frac{1}{\frac{\partial h_j}{\partial p_j}} (\sum_{k \neq j}^{J} \frac{\partial h_k}{\partial p_j} \tau_k) = \phi_j + \frac{1}{\frac{\partial h_j}{\partial p_j}} \sum_{l=1}^{M} \frac{\partial h_l}{\partial p_j} \phi_l$$

This yields J equations, each linear in the J tax levels:

$$a_1^j \tau_1 + ... + a_l^j \tau_l + ... + a_J^j \tau_J = b_j \qquad \forall \quad j \in [1, J] \tag{1.4}$$

Where a_l^j and b_j are defined as:

$$a_l^j = \frac{\frac{\partial h_l}{\partial p_j}}{\frac{\partial h_j}{\partial p_j}} \tag{1.5}$$

$$b_j = \phi_j + \sum_{m=1}^{M} \frac{\frac{\partial h_m}{\partial p_j}}{\frac{\partial h_j}{\partial p_j}} \phi_l \tag{1.6}$$

The a and b terms have intuitive interpretations. a_l^j is the share of the reduction in overall consumption of good j that shifts to good l as a results of an increase in the price of good j. That is, each a term is a leakage share between two taxable goods. b_j is the overall reduction in externalities that results from the increase in the price of good j; this consists of a direct component, ϕ_j, plus the sum of leakage terms: $\sum_{m=1}^{M} \frac{\partial h_m}{\partial p_j} / \frac{\partial h_j}{\partial p_j} \phi_l$, which are negative if j is a normal good and m is a substitute for j.

This system can be written compactly as:

$$\begin{bmatrix} a_1^1 & \dots & a_J^1 \\ & \dots & \\ a_1^J & \dots & a_J^J \end{bmatrix} \begin{bmatrix} \tau_1^* \\ \dots \\ \tau_J^* \end{bmatrix} = \begin{bmatrix} b_1 \\ \dots \\ b_J \end{bmatrix}$$

$$\boldsymbol{A\tau} = \boldsymbol{b} \tag{1.7}$$

The optimal tax vector when there are J taxable goods out of M total externality-generating goods is:

$$\boldsymbol{\tau} = \boldsymbol{A}^{-1}\boldsymbol{b} \tag{1.8}$$

Equation 1.8 shows that solving for the second-best optimal vector of corrective taxes in a setting with incomplete tax coverage and substitution between many externality-generating goods requires a) the consumption externalities associated with each good, and b) the substitution matrix between all goods.[2]

1.2.4. Heterogeneity and Leakage

Finally, I characterize second-best taxes where a) there are many externality-generating products, b) policymakers can tax only a subset of these products, and c) externalities are heterogeneous in consumption of the products.

While I apply this model to urban driving externalities in this paper, many markets feature externalities and policy instruments that fit this description. Electricity generation, for example, produces environmental externalities that vary by location (Muller and Mendelsohn, 2007; Hernandez-Cortes and Meng, 2020), and local environmental policies may induce leakage if utilities import electricity across jurisdictional borders. Similarly, the consumption of "sin" goods may be associated with externalities or internalities that vary across consumers, and taxing any one product (e.g., cigarettes) may induce leakage towards other products (e.g., vape pens) that do not fall under a policymaker's purview (Herrnstadt, Parry, and Siikamäki, 2015).

Lastly, as I introduce heterogeneity, it is worth noting that I assume that the social planner acts to maximize aggregate welfare, as in Diamond (1973). The formulae that follow do not account for redistributive preferences — heterogeneity is included in the model to

[2]Note that this substitution matrix contains cross-price consumption *derivatives* and not cross-price consumption *elasticities*. $\boldsymbol{A}$ contains 1's along the diagonal; when all j goods are substitutes, the off-diagonal terms of $\boldsymbol{A}$ fall in the closed interval $[0, -1]$.

reflect the implications of differences in externalities, rather than understand how externality taxation interacts with inequality aversion.[3]

Setup: N heterogeneous consumers choose between M externality-generating goods and a numeraire, z. I denote individual i's consumption of good m as h_i^m. Each individual has an exogenous income μ_i. I assume that each consumer's utility is a function of their consumption of these M goods and a quasilinear numeraire, as well as other's consumption of these goods (which generate externalities and decrease i's utility): $U_i(h_1^1, ...h_1^M, ..., h_i^1, ...h_i^M, ...h_N^1, ...h_N^M) + z_i$.

As in Section 2.3, a policymaker can choose tax levels for goods $j \in \{1, ..., J\}$ where $J < M$. I assume goods $m \notin \{1, ..., J\}$ are un- or under-taxed. I denote τ^j as the tax on good j.

In Appendix A.1, I show that the optimal tax on τ_j as a function of the k other tax levels is:

$$\tau_j = \frac{\sum_{i=1}^{N}\sum_{g}^{N}(\frac{\partial U^i}{\partial h_g^1}\frac{\partial h_g^1}{\partial p_j} + ... + \frac{\partial U^i}{\partial h_g^M}\frac{\partial h_g^M}{\partial p_j})}{\sum_{i=1}^{N}\frac{\partial h_i^j}{\partial p_j}} + \frac{\sum_{k\neq j}^{J}\frac{\partial h_i^k}{\partial p_j}\tau_k}{\sum_{i=1}^{N}\frac{\partial h_i^j}{\partial p_j}} \tag{1.9}$$

This expression for the optimal level of a given tax is equivalent to the equation for substitutes with homogeneous damages (Equation 1.3) where each of the marginal damages have been replaced by Diamond-weighted externalities that account for heterogeneity in marginal damages across individuals. As in the case of many substitutes with homogeneous damages, the optimal tax vector solves a system of J equations:

$$\begin{bmatrix} a_1^1 & ... & a_J^1 \\ & ... & \\ a_1^J & ... & a_J^J \end{bmatrix} \begin{bmatrix} \tau_1^* \\ ... \\ \tau_J^* \end{bmatrix} = \begin{bmatrix} b_1 \\ ... \\ b_J \end{bmatrix}$$

$$\boldsymbol{A\tau} = \boldsymbol{b} \tag{1.10}$$

Where a_l^j and b_j are defined as:

$$a_l^j = \frac{\sum_{i=1}^{N}\frac{\partial h_i^l}{p^j}}{\sum_{i=1}^{N}\frac{\partial h_i^j}{\partial p_j}} \tag{1.11}$$

[3]See Allcott, Lockwood, and Taubinsky (2019) for a characterization optimal corrective taxation with incomplete instruments and social preferences for redistribution.

$$b_j = \underbrace{\frac{\sum_i^N \sum_{g \neq i}^N \frac{\partial U_i}{\partial h_g^j} \frac{\partial h_g^j}{\partial p_j}}{\sum_i^N \frac{\partial h_i^j}{\partial p_j}}}_{\text{Diamond-weighted externality of good } j} + \underbrace{\sum_{l \neq j}^M \frac{\sum_i^N \sum_{g \neq i}^N \frac{\partial U_i}{\partial h_g^l} \frac{\partial h_g^l}{\partial p_j}}{\sum_l^N \frac{\partial h_i^j}{\partial p_j}}}_{\text{Diamond-weighted leakage shares}} \tag{1.12}$$

Solving for the second-best optimal vector of corrective taxes therefore requires (i) the (heterogeneous) externalities associated with each good, (ii) the relationship between these heterogeneous externalities and individual price elasticities, and (iii) individual-level substitution matrices between goods.

These are considerable information requirements. In what follows, I demonstrate how to use the structure of discrete choice modeling to reduce the dimensionality of this problem. Specifically, rather than estimating how each driver substitutes between each possible trip, I use a discrete choice model over routes and times of day to populate the substitution matrix of options facing drivers based on the attributes of those trips.

1.3. A Discrete Choice Model of Driver Behavior

The theory outlined in Section 1.2 stipulates that calculating the second-best optimal cordon prices requires information about the heterogeneity in the price responsiveness of different types of trips that cross a cordon, as well as the rates of substitution between trips that can and trips that cannot be priced. To recover these parameters, I estimate a canonical "bottleneck" model of driving demand (Arnott, De Palma, and Lindsey, 1990, 1993).

Formally, imagine drivers i who choose between departure times h and a routes r to satisfy their demand for travel. Included in this choice set is the outside (no trip) option, which is normalized to zero utility. Each driver has an exogenous ideal arrival time, h_i^A. Drivers are atomistic and face travel times $T(h,r)$ and tolls $p(h,r)$ that may vary by route and time of day. A driver arriving before or after their ideal arrival time incurs disutilities γ_e and γ_l per minute, respectively. Drivers also incur disutility α from each minute spend commuting. Utility is thus:

$$u(h_i, r_i) = -\alpha T(h_i, r_i) - \gamma_e \underbrace{\lfloor h_i + T(h_i, r_i) - h_i^A \rfloor_-}_{\text{time early}} - \gamma_l \underbrace{\lfloor h_i + T(h_i, r_i) - h_i^A \rfloor_+}_{\text{time late}} - \beta p(h_i, r_i) \tag{1.13}$$

Each driver chooses the route (r_i) and time of day (h_i) that maximizes their expected utility:

$$\{h_i^*, r_i^*\} = arg\, max_{h_i, r_i} \{u(h_i, r_i))\} \tag{1.14}$$

To clarify the mapping between this discrete choice model and the optimal tax formula (Equation 1.10), a "good" (h^j in the notation used in Section 1.2) is a trip taken on a given *route* at a given *time of day*: $\{h, r\}$. Typical cordon zones have discrete spatial and temporal cutoffs.[4] The possibility of leakage reflects the ability of drivers to adjust trips in time (h) and space (r) to avoid tolls. Heterogeneity in externalities results from the fact that trips that enter a cordon zone during the same time of day are charged the same price, but differ in pollution externalities (a function of trip length, vehicle characteristics, and travel speed) as well as congestion externalities (a function of trip length and traffic density along the trip). To estimate the relationship between idiosyncratic externalities and price-responsiveness, I allow β (the coefficient on price) to vary across externality quantiles during estimation.

The value of estimating this discrete choice model is that it greatly reduces the number of parameter estimates required for applying the optimal tax formula outlined in Section 1.2. For any choice set (e.g., cordon vs. non-cordon routes at various times of day), equation 1.13 implies a matrix of own and cross-price elasticities between choices, which reflect model primitives (α_e, γ_e, γ_l, β) and trip attributes (T, p, *time late*, and *time early*). In Section 1.4 through 1.7, I use tolling microdata to recover estimates of each of these parameters using a mixed logit model. I also apply a bunching estimator to the introduction of peak-hour pricing in the Bay Area to produce separate estimates of scheduling parameters, γ_e and γ_l.

[4]The London Cordon Zone, for example, charges road users £15 between 7 a.m. and 10 p.m. for entering the city center. The Milan Cordon Zone charges users €2 to €5 based on vehicle type between 7:30 am and 7:30 pm. The proposed cordon zones in San Francisco would only charge drivers for trips during peak hours (6-10 am and 3-7 pm).

1.4. Natural Experiment: Traffic Tolling in the San Francisco Bay Area

I use administrative tolling data from the San Francisco Bay Area together with revisions to regional bridge tolls to estimate the model of driving demand outlined in Section 1.3.

1.4.1. Bay Area Bridge Tolls

FasTrak is an electronic tolling system used in California. Drivers are charged for using certain roads (bridges and high-occupancy toll lanes) via transponders mounted to the car's dash.[5] In the San Francisco Bay Area, tolls are collected on each of the region's trans-bay bridges (mapped in Figure 1.1) for westbound trips only.

FIGURE 1.1 — SAN FRANCISCO BAY AREA BRIDGES

Figure 1.1: This maps shows the four San Francisco Bay Area bridges used to estimate driver responses to toll prices in this paper. The *Richmond Bridge* connects Richmond and the eastern Bay Area to San Rafael and Marin County. The *Bay Bridge* connects Oakland to San Francisco. The *San Mateo Bridge* connects Hayward to San Mateo. The *Dumbarton Bridge* connects Fremont to Palo Alto. Each of these bridges charges drivers for westbound trips (as detailed in Figure 1.2).

[5]Drivers can pay with cash if they do not purchase a FasTrak device. Between 2010 and 2019, cash payers represented roughly 10% of all trips on Bay Area bridges.

1.4.2. Variation in Toll Prices

Bay Area FasTrak tolls vary by bridge, vehicle type, and time of day. I focus on passenger vehicles (as opposed to light and heavy-duty trucks), which constitute roughly 97% of vehicle trips on Bay Area bridges.[6] Currently, passenger vehicles are charged between $3 and $7 dollars, depending on the time of day, the number of occupants, and whether or not the vehicle is electric/hybrid.

In this paper, I leverage several changes in the tolling structure that occurred on July 1, 2010 to identify the parameters necessary to calculate optimal road prices. In 2009, the Bay Area Toll Authority (BATA) adopted Resolution 90, which increased the base prices for passenger vehicles from $4 to $5 beginning on July 1, 2010, and established peak-hour pricing on the Bay Bridge (detailed below). This intertemporal variation in toll prices is plotted in Figure 1.2.

[6] Between 2009 and 2019, the four major Bay Area bridges recorded roughly 285,000 FasTrak transactions daily for passenger vehicles, versus 7,000 daily transactions for vehicles with three of more axles.

Figure 1.2 — Variation in Passenger Vehicle Bridge Tolls

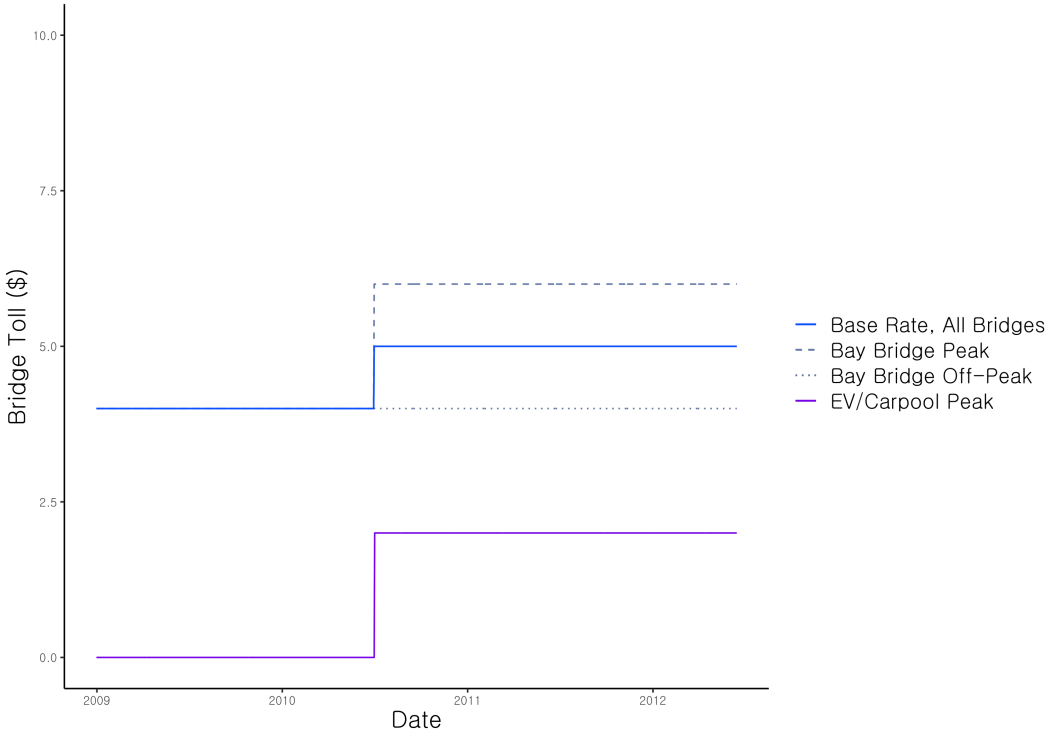


Figure 1.2: This figure shows Bay Area bridge tolls between 2009 and 2012 for passenger vehicles. Prices are uniform across bridges, with the exception of the Bay Bridge, which connects San Francisco and Oakland. Beginning in 2010, passenger vehicles crossing the Bay Bridge faced a two-dollar difference between peak and off-peak prices. The peak ($6) and off-peak ($4) prices are plotted above as dotted and dashed lines, respectively. EV and carpool trips were free on all bridges prior to 2010. Beginning in July of 2010, EV/carpool trips were charged the base rate ($5 on the San Mateo, Dumbarton, and Richmond Bridges; $4 on the Bay Bridge), except during peak hours, where they receive a discount ($2.5) on all bridges.

1.4.3. Peak-hour Pricing on the Bay Bridge

To address acute congestion on the region's busiest bridge, the Bay Area Toll Authority imposed peak hour pricing on the Bay Bridge (which connects San Francisco and Oakland) beginning on July 1, 2010. Passenger vehicles crossing westbound through the Bay Bridge toll plaza on weekdays between 5 a.m. and 10 a.m., or between 3 p.m. and 7 p.m. (henceforth *peak hours*) were charged $6. Tolls for all other hours (henceforth *off-peak*) remained at the pre-2010 price of $4.

Prior to July 1, 2010, passenger vehicles with two or more passengers, as well as eligible electric and hybrid electric vehicles were not subject to tolls on any Bay Area bridges. Starting in 2010, these vehicles were subject to the full toll value during off-peak hours, but retained a discount during peak hours: EV/carpool trips were charged $2.50 to use Bay Area bridges between July 1, 2010 and January 1, 2019.

Figure 1.3 — Peak Pricing on Bay Area Bridges

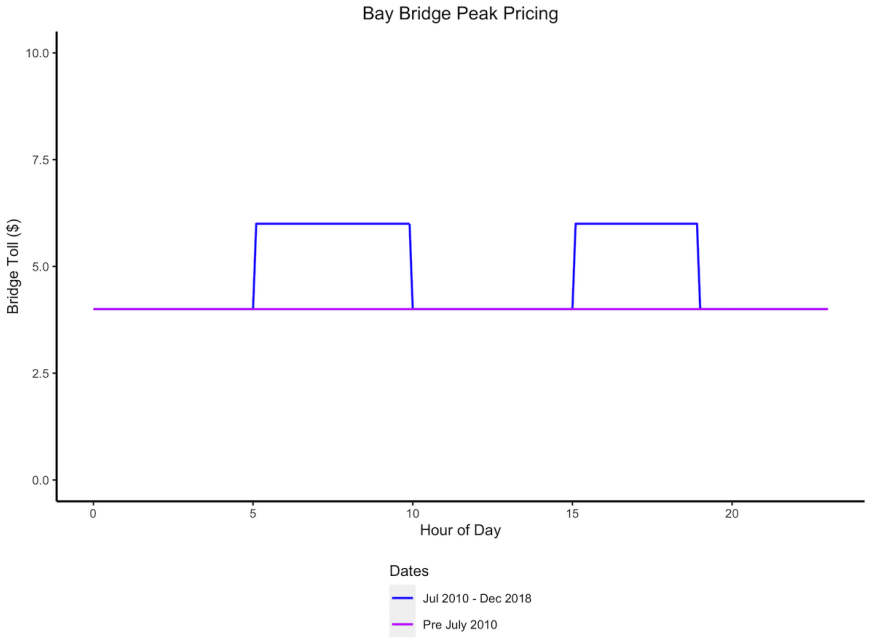


Figure 1.3: This figure displays peak-hour pricing schemes for passenger vehicles (vehicles with two axles) on California's Bay Bridge, which connects San Francisco and Oakland. Beginning on July 1, 2010, passenger vehicles crossing westbound on weekdays during peak hours (between 5 a.m. and 10 a.m., or between 3 p.m. and 7 p.m.) faced higher prices than vehicles crossing during off-peak hours. Peak-hour prices are displayed on large variable-message sign about the Bay Bridge toll plaza. Weekend trips on the Bay Bridge and trips on the other major Bay Area bridges are not subject to peak pricing, instead charging the base rate for passenger vehicles ($4 for pre-2010 and $5 for July 2010 - December 2018).

Foreman (2016) uses reduced-form approaches to provide valuable estimates of the responses of Bay Area drivers to this change in bridge prices. The number of vehicle trips during peak hours on the Bay Bridge decreased by 6 to 8% (400 to 550 vehicles per hour) following the imposition of peak hour pricing. Travel during off-peak hours on the Bay Bridge increased by 4 to 20% (225 to 400 vehicles per hour). Point estimates suggest the $1 increase on the San Mateo and Dumbarton bridges led to modest decreases in bridge use (15 to 48 vehicles per hour). Notably, crossings on the San Mateo Bridge *increased* by 100 to 200 vehicles (around a 5%) during peak hours, implying that some drivers switched from the Bay Bridge its closest substitute in response to the peak-hour price difference across routes.

To summarize this variation in road prices in this empirical setting, the 2010 revision to bridge tolls in the San Francisco Bay Area replaced uniform prices with prices that varied across bridges and times of day. Reduced-form analyses of this policy suggest that drivers responded to these pricing by reducing the overall number of trips, as well as shifting their

trips in time and space. In the following sections, I use this variation in prices together with microdata on driver choices to estimate the model of personal vehicle travel described in Section 1.3.

1.5. Data

1.5.1. Reconstructing Choice Sets

Estimating the model outlined in Section 1.3 requires individual-level data on travel choices, travel times, and road prices. To construct this choice set, I combine administrative microdata from the FasTrak tolling system with historic travel time data purchased from TomTom's *Historic Traffic Stats* database.

FasTrak Toll Data: I use administrative microdata from the FasTrak tolling system to create a panel of individual-level driving choices. These microdata record any electronic transactions that occurred on the four trans-bay bridges between January 1, 2009 and July 1, 2019. A single observation in this data set includes the date, time, and location of the vehicle crossing, as well as the vehicle class (axle number), the price paid, and an indicator for whether the vehicle used the EV/carpool lane. For vehicles with registered FasTrak devices (vehicles that did not pay cash) the microdata also include a unique FasTrak id number. Roughly 40% of observations that use a FasTrak device also list the home zip code associated with the FasTrak holder. These data contain hundreds of millions of trip records.

I restrict the dataset on several dimensions. First, I include only devices with a valid (Bay Area) zip code. Second, I drop devices with infrequent use (fewer than 50 weekday trips in the year prior to the 2010 price change), or users that take multiple trans-bay trips per day (greater than 500 weekday trips in the year to the 2010 price change). Lastly, for the purposes of estimation, I consider only trips taken in a narrow window (weekdays between June 15^{th} to July 15^{th}) before and after the 2010 change in toll prices. The resulting panel consists of 32,104 FasTrak devices and 1,078,044 bridge crossings.

These sample restrictions reflect the information requirements of the discrete choice model of driving demand. Recall that this model specifies driver utility as a function of trip attributes: travel time, time late or early, and price. Zip code information is necessary for assigning travel times to vehicle trips based on the distance between households and bridges. The restrictions based on the frequency of trips reflects the need to infer ideal arrival times for drivers. For FasTrak devices associated with daily commuters, ideal arrival times can be inferred based on bridge-crossing times prior to July 1, 2010 (detailed below). Drivers that infrequently use bridges, or that use bridges many times a day, are not well-described by

the discrete choice model I employ in this paper, as it is unclear how to assign these trips an ideal arrival time and trip termini. While imposing these sample restrictions comes at the cost comprehensiveness, estimating the discrete choice model provides a distinct benefit relative to a reduced-form approach: For any given choice set (e.g., driving options subject to cordon prices) the structure of the discrete choice model directly implies the substitution parameters required for calculating optimal prices.

Travel Time Data: Because the FasTrak microdata include only the device zip code and bridge used, I must infer trip travel times. I do so in two steps.

First, based on the zip code and travel behavior of a given vehicle, I use data from the 2012 California Household Travel Survey (CHTS) to infer a probability distribution over destinations for that vehicle. For example, if I observe a driver from Oakland traveling via the Bay Bridge, I enumerate the destination cities of all CHTS drivers from Oakland who reported using the Bay Bridge. I repeat this for all of the driver's trips, resulting in a probability distribution over endpoints for each FasTrak device.

Second, I use TomTom's Historic Traffic Stats data to reconstruct the travel time between an individual's home zip code and each of the possible destination endpoints. The FasTrak data provide hourly traffic speeds for major roads in the 12 months before and the 12 months after the July 2010 adjustment to Bay Area tolls. Importantly, I also use the TomTom data to estimate counterfactual travel times. The result is a reconstruction of each driver's choice set, namely the travel time and price for each trip that driver took, as well as the price and travel time if they had taken that same trip at a different time of day, or using a different bridge. This choice set construction is described in full detail in Appendix A.5.

Ideal Arrival Times: Ideal arrival times, h_i^A in equation 1.13, are not directly observed, and therefore must be inferred from each driver's activity. For each driver, I assign h_i^A as the modal bridge crossing time of each individual during weekdays between January 1, 2010 and July 1, 2010, plus the weighted average travel time between the bridge toll plaza and each of the possible endpoints for that driver.

For an illustrative example, consider a driver who exclusively uses the *Bay Bridge* during the pre-period, and who most commonly crosses this bridge at 9 in the morning. A trip taken by this individual that crosses the bridge at 9 a.m. would be assigned a value of zero for *time late* and *time early*. A trip taken by this individual that crosses the bridge at 10 a.m. would be assigned a value of *time late* of 1, plus any difference in expected after-bridge travel time between 9 a.m. and 10 a.m.

Lastly, it is worth noting that pre-period bridge crossing times may not indicate actual ideal crossing times if within-day traffic conditions provide sufficient incentive for drivers

to shift their trips in time to reduce overall commute times. The estimates of scheduling elasticities that I recover from responses to peak-hour pricing on the Bay Bridge, however, are inconsistent with this type of strategic scheduling. If Bay Area drivers have schedule costs low enough to induce them to strategically reschedule trips in the absence of peak-hour pricing, a much higher portion of drivers should have responded to the imposition of peak-hour pricing by rescheduling trips to just outside of the peak pricing window.

1.5.2. Externalities

Although data on trip-level externalities is not necessary for estimating a model of driving demand, second-best optimal road prices depend on the correlation between the price elasticity of demand for a given trip and the idiosyncratic externalities associated with that trip (see Section 1.2). I therefore estimate the externalities (congestion and pollution) associated with each FasTrak trip.

Note that I do not include accident externalities when calculating trip-level externalities. Although most estimates of per-mile externalities in the economics literature suggest that accident externalities constitute a significant portion of the overall social costs of driving (Parry and Small, 2005; Anderson and Auffhammer, 2014), empirical evidence suggests that the social benefits from reduced accidents in cordon zones are an order of magnitude smaller than the benefits associated with reduced congestion and air pollution (Green, Heywood, and Paniagua, 2020). Broadly, this empirical evidence reflects the fact that the type of driving curtailed by cordon pricing — slow, daytime trips in city centers — results in relatively few fatal traffic accidents. I provide further discussion of the relationship between accidents and optimal cordon prices in Appendix A.8.

Congestion Externalities: Congestion externalities vary significantly in space and time. The transportation economics literature canonically presents congestion externalities as a function of traffic *density*, measured in vehicles per lane-mile (Small, Verhoef, and Lindsey, 2007). To assign congestion externalities to trips in the FasTrak dataset, I use estimates from Yang, Purevjav, and Li (2020), who show that the marginal external (travel time) cost of traffic is convex in traffic density. That is, congestion externalities are negligible when there are few other vehicles on the road, but increase sharply with the number of vehicles per lane-mile. The congestion costs from this paper are reproduced in Figure 1.4.

Using a comprehensive network of traffic sensors on roadways in the Bay Area, I infer the density along the route for each FasTrak trip. These traffic sensors are mapped in Figure A.1. For each trip, I use HERE Technology's *Routes* API to identify the likely route between the zip code associated with the device and the bridge crossed. For each traffic sensor along

the driver's route, I use estimates from Yang, Purevjav, and Li (2020) to assign a marginal external congestion cost (in dollars per mile)[7] to this point based on the average traffic density at that sensor at the time of day when the trip was taken. A trip's total congestion externality is then the average of the external congestion costs (in dollars per mile) along the route, times the length of the trip.

As noted above, because one of the trip termini is missing from the FasTrak data, I impute the congestion externalities for the missing segment of the trip (between the bridge to the place of work) using the likely destination locations conditional on observable characteristics (home zip code, bridge used). Note that the majority of variation in externalities is driven by the choice of bridge and time of day, suggesting any noise in this imputation process should not meaningfully impact estimates of the relationship between idiosyncratic externalities and price responsiveness.

FIGURE 1.4 — CONGESTION COSTS, REPRODUCED FROM YANG ET AL. (2020)

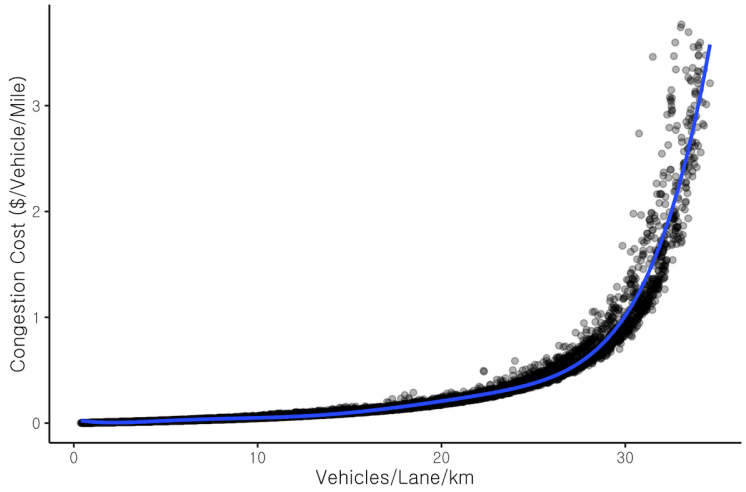


Figure 1.4: Congestion costs reproduced from Yang, Purevjav, and Li (2020), who exploit variation in traffic density generated by Beijing's driving restriction to estimate the relationship between traffic density and speed. The original results are presented in Yuan/Vehicle/km. I convert these values to dollars by a) converting currencies, and b) replacing the Beijing-specific value of time from (50% of the average wage rate in Beijing) with a $20 value of travel time, which reflects San Francisco-specific estimates from Goldszmidt et al. (2020).

Emissions Externalities: Fuel combustion and brake wear in passenger vehicles generates several air pollutants. These include "global" pollutants like CO_2 and methane, which contribute to climate change, as well as "local" pollutants like particulate matter (PM),

[7]The estimates from Yang, Purevjav, and Li (2020) are in yuan/vehicle/km. I convert these values to dollars by a) converting currencies, and b) replacing the Beijing-specific value of time from (50% of the average wage rate in Beijing) with a $20 value of travel time, which reflects research by Goldszmidt, List, Metcalfe, Muir, Smith, and Wang (2020).

nitrogen oxides (NO_x) and reactive organic compounds (ROCs), which negatively impact the health of nearby residents (Anderson, 2020; Currie and Walker, 2011a; Deryugina, Heutel, Miller, Molitor, and Reif, 2019). Vehicle emissions factors — the amount of a particular pollutant that a vehicle emits while traveling a mile — depends on a number of variables, including the type of fuel consumed, the fuel economy, the vehicle vintage[8], and vehicle speed.[9]

I estimate emissions for FasTrak trips using data from the California Air Resource Board's Emissions Factor Database (EMFAC). This database contains estimates of the average emissions rates of vehicles registered in each county as a function of vehicle speed. I then assign social costs to these trip-level emissions. For global pollutants, I use the EPA's 2021 social cost of carbon ($51 per ton) and methane ($1,500), respectively. Local pollutant damages reflect the cost of emitting each pollutant at ground level in San Francisco, according to the EASIUR model of local pollution damages. See Appendix A.3 for details on individual pollutant costs.

[8]Older vehicles have higher emissions factors for two reasons: They were subject to less stringent tailpipe emissions and fuel economy standards when they were built, and emissions abatement technologies (catalytic converters) depreciate over a vehicles lifetime.

[9]Vehicle speed impacts emissions through engine efficiency and the intensity of brake ware.

Figure 1.5 — Pollution Externalities at Various Speeds

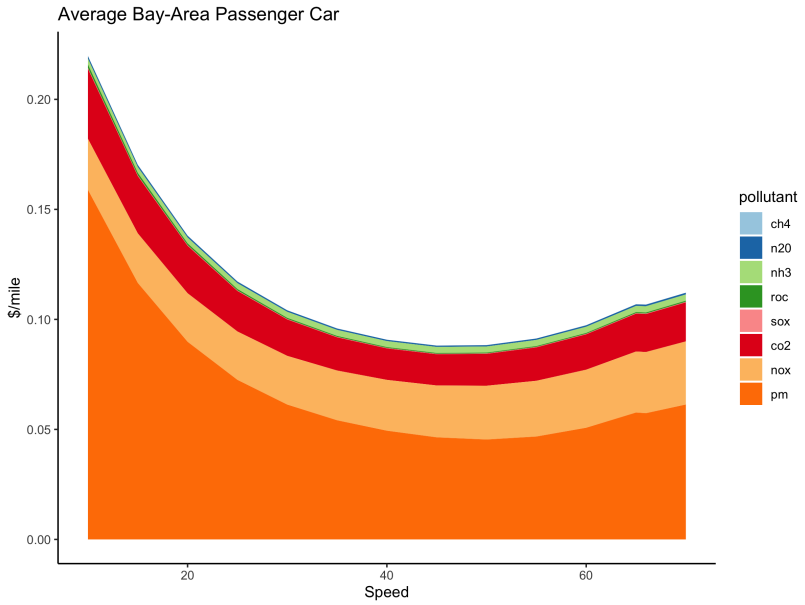


Figure 1.5: This figure plots per-mile pollution externalities at various speeds for an average passenger vehicle in the Bay Area. These costs reflect VMT-weighted average emissions factors (in grams/mile) of different pollutants at different speeds reported by California's Emissions Factor Model (EMFAC). The EMFAC emissions factor estimates reflect state DMV and smog check data. To convert these emissions factors to per-mile costs, I multiply the emission factor for each pollutant by the corresponding social cost of each pollutant. For local pollutants, the social cost is calculated using the Estimating Air pollution Social Impact Using Regression (EASIUR) Online Tool, calibrated with coordinates from San Francisco. For global pollutants, I use the EPA's 2021 social costs of \$51 per ton of CO_2 and \$1,500 per ton of CH_4, respectively. All values are in 2020 dollars.

Together, the data described in this section allow me to recreate the choices and choice set facing a sample of Bay Area drivers, augmented with estimates of the social costs associated with each trip choice.

1.6. Empirical Strategy

I use two strategies to recover the primitives that determine driving behavior. In my preferred specification, I use the variation in toll prices in 2010 together with the FasTrak microdata to estimate the parameters of Equation 1.13 using a mixed logit regression. As a check for the results from this first method, I apply a bunching estimator to the Bay Bridge's notched tolling schedule, producing a second set of empirical estimates of scheduling costs.

1.6.1. Multinomial and Mixed Logit Regressions

As described in Section 1.5, the FasTrak microdata and the TomTom historic traffic data allow me to reconstruct the attributes of elements in the choice set (routes and times of day) for each driver. I then use this reconstructed choice set to estimate the discrete choice model of driving demand outlined in Section 1.3.

$$\mathbb{1}(h_i = 1 \wedge r_i = 1) = -\alpha T(h_i, r_i) - \gamma_e \underbrace{|h_i + T(h_i, r_i) - h_i^A|_-}_{\text{time early}} - \gamma_l \underbrace{|h_i + T(h_i, r_i) - h_i^A|_+}_{\text{time late}} - \beta p(h, r) + \epsilon_{h,r,i} \tag{1.15}$$

Where $\mathbb{1}(h_i = 1 \wedge r_i = 1)$ is an indicator variable that takes a value of 1 if individual i crosses bridge r at time of day h, and zero otherwise. The routes available to a driver are each of the four Bay-Area bridges. Times of day are discretized at 12-minute intervals.

This estimation strategy leverages variation in trip-level attributes that reflect both the 2010 changes in toll prices, as well as differences in the attributes of trips available to drivers across routes or times of day.

The identifying variation in price, $p(h_i, r_i)$, comes from the revision to bridge tolls, which is detailed in Section 1.4. The identifying variation for travel time $T(h_i, r_i)$ comes from both within-day differences in travel time along a given route for each driver, as well as differences in travel times across routes (bridges) conditional on departure time. The variation in total travel times in response to the 2010 change in toll prices is negligible (Foreman, 2016). The schedule cost parameters reflect i) the tradeoff between travel time and late or early arrival induced by variation in travel times throughout the day, and ii) the tradeoff between early or late arrival and lower toll prices for peak-hour travelers on the Bay Bridge.

Peak-hour pricing on the Bay Bridge constitutes a potential threat to identification through reverse causality. Peak-hour pricing was imposed on the Bay Bridge in response to high demand for trips connecting Oakland to San Francisco during peak hours. If the high peak-hour demand on this bridge is completely explained by trip attributes — travel time and scheduling costs — then price will be uncorrelated with the error term $\epsilon_{h,r,i}$. If,

however, this high demand was the result of factors unobserved by the researcher that make peak-hour travel on the Bay Bridge attractive, then peak-hour pricing on the Bay Bridge would create a mechanical correlation between $p(h_i, r_i)$ and $\epsilon_{h,r,i}$. To address this threat to identification, I also estimate an instrumental variables regression where *post* acts as an instrument for *price*, which leverages only the level shift in prices to estimate the coefficient on price.

The estimated parameters of equation 1.15 imply a matrix of own and cross-price elasticities between routes and hours of day that I use to solve for second-best cordon prices in San Francisco. Formally, the own and cross-price elasticities from a multinomial logit regression used to estimate this model are:

$$\varepsilon_{\{h^j,r^k\},\{h^l,r^m\}} = \begin{cases} \beta p(h^l, r^m)(1 - s_{\{h^l,r^m\}}), & \text{if } i = l \wedge j = m \\ \beta p(h^l, r^m) s_{\{h^l,r^m\}}, & \text{otherwise} \end{cases} \tag{1.16}$$

Where $\{h^j, r^k\}$ denotes route r^k taken at time h^j, $s_{\{h^j,r^k\}}$ is the share of total trips taken via route r^k at time h^j, and β and $p(h, r)$ are defined as above. Importantly, ordinary logit models exhibit restrictive substitution parameters. Namely, the cross-price elasticities for a given good are constant across all alternatives, implying proportional substitution following a price increase of any one good. I relax this assumption in my preferred specification — a random coefficients ("mixed") logit regression. This regression estimates a joint distribution of coefficients (θ) which implies idiosyncratic pairwise substitution parameters between trip options:

$$\varepsilon_{\{h^j,r^k\},\{h^l,r^m\}} = \frac{p(h^l, r^m)}{s_{\{h^j,r^k\}}} \int \beta s_{\{h^j,r^k\}}(\theta) s_{\{h^l,r^m\}}(\theta) f(\theta) d\theta \tag{1.17}$$

1.6.2. Bunching Estimator

In this section, I outline how I use notches in the peak-hour tolling on San Francisco's Bay Bridge to recover the scheduling costs of drivers. This alternative empirical approach acts as a check for the results from the logit regressions.

Bunching estimators are used to infer structural parameters from the empirical density of choice variables around kinks or notches in a budget set (Chetty, Friedman, Olsen, and Pistaferri, 2011; Saez, 2010; Kleven and Waseem, 2013). While bunching estimators allow for the estimation of structural parameters using cross-sectional data, doing so often necessitates strong assumptions regarding the distribution of choice variables under a counterfactual (no-notch) budget set (Blomquist, Newey, Kumar, and Liang, 2021). The panel data in this setting allow me to directly compare the density of trips under notched (peak-hour) and non-notched pricing schemes, thereby circumventing distributional assumptions. Broadly, bunching estimators use changes in the density of choice variables to identify characteristics of a "marginal buncher" — an individual who is indifferent between two positions along a notched/kinked budget set. Before presenting the bunching estimator, it is therefore useful to characterize the marginal bunching individual in this setting.

Consider a group of drivers with homogeneous scheduling costs and perfect control over when they cross a bridge that charges different tolls during peak and off-peak hours. A "buncher" is a driver who would cross the bridge during peak hours in the absence of peak-hour pricing, but who would adjust their travel time to just avoid the extra toll in a world with peak-hour pricing. For the *marginal* buncher, the utility from the lower price is equal to the scheduling costs of adjusting their trip to cross outside of peak hours. Equation 1.18 shows this indifference condition in terms of structural parameters. For simplicity, I examine the case of a driver who faces a decision of whether or not to shift their trip earlier:

$$\underbrace{\beta \Delta p}_{\text{Benefit from shifting}} = \underbrace{\gamma_e \Delta h}_{\text{Cost of shifting}} \tag{1.18}$$

Following the notation from Equation 1.13, β is the marginal utility of a dollar (normalized to 1), Δp is the change in price at the notch, γ_e is the cost (in dollars/hour) of shifting a trip earlier, and Δh is the number of hours between the price notch and the time of day when the marginal buncher would have crossed the bridge in the absence of a price notch. The scheduling cost, γ_e, can then be written as a function of the size of the price notch, and the time that the marginal buncher would have to adjust their trip in order to cross the bridge before peak hours:

$$\gamma_e = \frac{\beta \Delta p}{\Delta h} \tag{1.19}$$

If travel times also differ significantly in the neighborhood of the price notch, this condition becomes:

$$\gamma_e = \frac{\beta \Delta p + \alpha \Delta T}{\Delta h} \tag{1.20}$$

Where ΔT is the difference between a driver's total travel time if they cross the bridge just before the beginning of peak hours, and a driver's total travel time if they cross the bridge at the time of day when the marginal buncher would have crossed the bridge in the absence of a price notch. The characterization of a marginal buncher is plotted in Figure 1.6.

Equations 1.19 and 1.20 imply that the relevant scheduling cost (either γ_e or γ_l) is inversely proportional to the width of the density trough on the relatively expensive side of the peak-hour price notch. Intuitively, the width of the density trough reflects how far the marginal buncher moves their trip in response to a price incentive. All else equal, decreasing scheduling costs makes drivers more willing to shift their trips further from their ideal travel time for a given level of compensation. A wider density gap therefore implies lower scheduling costs.

Because the peak-hour pricing on the Bay Bridge (see Figure 1.3) creates *notches* rather than *kinks* in the budget sets of drivers, the region immediately adjacent to the price notch is strictly dominated under any scheduling cost. The fact that there is still a positive density of crossings during this dominated period suggests frictions may prevent drivers from perfectly optimizing (Kleven, 2016). In this setting, these 'frictions' may reflect inattentiveness (Finkelstein (2009), for example, finds that toll prices are less salient for drivers with automatic toll tags) or the inability to perfectly time bridge crossings due to traffic shocks.

To account for these optimization frictions, as well as heterogeneity in scheduling costs, I use an estimator similar to Kleven and Waseem (2013). I first compare the density of trips in the dominated region before and after the imposition of peak pricing to identify the fraction of individuals with crossing times in the vicinity of the notch who are unresponsive to the price signal. I then estimate the excess trip mass on the relatively inexpensive side of the price notch:

$$B = \int_{\gamma_e} \int_{h*}^{h^*+\Delta h} (1-a) f_0(h) dh \simeq (1-a) f_0(h^*) E[\Delta h] \tag{1.21}$$

Where B is the excess bunching mass on the relatively inexpensive side of the notch, a is the fraction of drivers in the strictly dominated region, and $f_0(h)$ is the counterfactual (no-notch) density of vehicle crossings as a function of the time of day, h. $E[\Delta h]$ is the average

adjustment among drivers who bunch at the price notch. Solving Equation 1.21 for Δh and plugging into Equation 1.20 yields the bunching estimator:

$$\gamma_e = \frac{\beta \Delta p}{B/((1-a) f_0(h^*))} \tag{1.22}$$

Relaxing the assumption that travel times are relatively flat around the notch point is straightforward, but necessitates the value of travel time:

$$\gamma_e = \frac{\beta \Delta p + \alpha \Delta T}{B/((1-a) f_0(h^*))} \tag{1.23}$$

In all bunching estimates, I use a \$20 value of travel time, which reflects San Francisco specific findings from Goldszmidt et al. (2020). I also present estimates of scheduling parameters that ignore time savings (equation 1.22) in Appendix A.4.

Figure 1.6 — The Relationship Between Scheduling Costs and Bunching

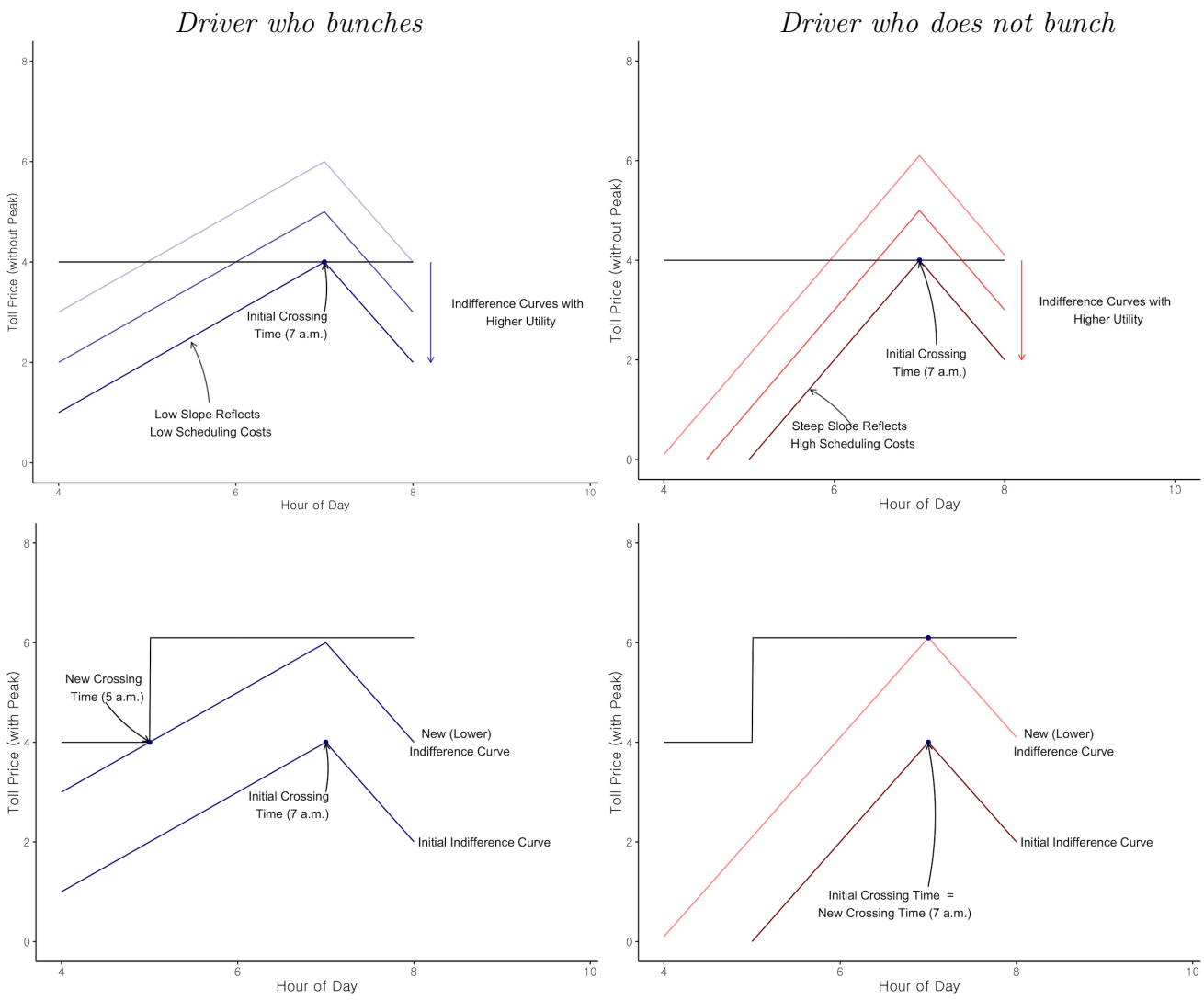


Figure 1.6: This figure illustrates the relationship between scheduling costs and bunching behavior in peak-hour toll schemes, as predicted by the discrete choice model outlined in Section 1.3. For expositional ease, this figure plots the case where travel times are constant throughout the day. The triangular shape of the indifference curves reflects the fact that the further a trip is from a given driver's ideal crossing time, the higher the compensation (via a lower toll price) required to maintain any given level of driver utility. In the right two panes, I plot indifference curves (red) of a driver with *high* scheduling costs, who does not shift their trip in response to peak pricing. In the left two panes, I plot indifference curves of a diver with *low* scheduling costs, who does shift their trip in response to peak pricing. All else equal, when scheduling costs are lower, drivers are more willing to adjust their travel times in response to peak pricing, implying a larger mass of trips around price notches.

1.7. Results

In this section, I present estimates of the parameters in the discrete choice model of driving demand (equation 1.13) estimated using the empirical approaches outlined in Section 1.6.

1.7.1. Logit Regression Results

Table 1.1 presents the results from equation 1.15, estimated via multinomial logistic regression. The point estimates from this regression imply that on average, drivers are indifferent between saving roughly $11 and saving an hour of travel time; they are indifferent between saving roughly $6 arriving an hour early, and they are indifferent between saving roughly $4 arriving an hour late.

Table A.2 and Figure A.2 show results of estimating equation 1.15 via mixed logit. Allowing for heterogeneity in the logit parameters produces results that are qualitatively similar to the results in Table 1.1. In Table 1.2, I allow price responsiveness to vary with road user's idiosyncratic externalities. To do so, I break FasTrak devices into quartiles based on the average estimated externality (both pollution and congestion) of each device's trips. The results in Table 1.2 suggest that price elasticity and idiosyncratic externalities are positively correlated: drivers that travel longer distances at more congested times and places are more price responsive on average than are drivers who take shorter, less-congested trips.

Table 1.1 — A Discrete Choice Model of Driving Demand

	Specification			
Variable	(1)	(2)	(3)	(4)
Travel Time ($/hr)	12.626	12.575	11.266	9.659
	(1.926)	(1.912)	(1.76)	(1.384)
Time Early ($/hr)	6.062	6.043	5.514	2.171
	(0.307)	(0.305)	(0.282)	(0.366)
Time Late ($/hr)	4.447	4.433	4.028	2.163
	(0.206)	(0.205)	(0.19)	(0.326)
Price	1.000	1.000	1.000	1.000
	(1.002)	(1.001)	(0.91)	(0.175)
Day of Week FE		Yes	Yes	Yes
Bridge FE			Yes	Yes

Table 1.1: Results from Equation 1.15, a discrete choice model where drivers choose over routes and times of day, estimated using the FasTrak tolling microdata described in Section 1.5. The dependent variable is whether an individual i elects to take a trip on route r at time of day h. *Travel time* is the travel time (in hours) that driver i would incur by traveling via route r at time h. *Time early* is the number of hours that that driver i would arrive before their ideal arrival time if they were to travel via route r at hour h. *Time late* is analogously defined. *Price* is the toll that driver i would incur by traveling via route r at hour h. As the coefficient on price is normalized to 1, the other coefficients can be interpreted as costs in dollars per hour. Columns (1) through (3) show results of a standard logit regression. Column (4) presents results of an instrumental variables regression where an indicator for post July 1, 2010 acts as an instrument for *price*. Standard errors in all regressions are clustered at the individual and zip code levels. All values are in 2010 dollars.

Table 1.2 — Mixed Logit with Price Elasticities by Externality Quartile

Variable	**Mean**	**Std Dev**
Time early (γ_e)	-1.247	0.488
	(0.113)	
Time late (γ_l)	-3.409	2.078
	(0.235)	
Travel time (α)	-5.371	1.57
	(0.53)	
Price (first externality quartile)	-0.016	
	(0.556)	
Price (second externality quartile)	-0.148	
	(0.105)	
Price (third externality quartile)	-0.285	
	(0.108)	
Price (fourth externality quartile)	-0.126	
	(0.115)	

Table 1.2: Results from Equation 1.15, a discrete choice model where drivers choose over routes and times of day, estimated using a random coefficients ("mixed") logit model. The dependent variable is whether an individual i elects to take a trip on route r at time of day h. *Travel time* is the travel time (in hours) that driver i would incur by traveling via route r at time h. *Time early* is the number of hours that that driver i would arrive before their ideal arrival time if they were to travel via route r at hour h. *Time late* is analogously defined. *Price* is the toll that driver i would incur by traveling via route r at hour h. I interact *price* with *externality quartile*, a categorical variable defined at the individual level that indicates the average intensity of externalities (both pollution and congestion) for trips taken by each device in the FasTrak dataset. Two-way standard errors are clustered at the individual and zip code levels.

1.7.2. Bunching Estimator Results

Applying a bunching estimator to notches in the pricing schedule on the Bay Bridge, I recover scheduling cost parameters (γ_e and γ_l in Equation 1.13) that range from \$6 to \$15 per hour.

Figure 1.7 plots the difference between the density of trips by time of day before vs. after the imposition of peak-hour pricing for the 5 a.m. price notch on San Francisco's Bay Bridge. The bunch in the density of trips prior to 5 a.m. (which is does not exist in the data prior to July 1, 2010) is consistent with a model of driving demand where drivers are willing to shift their trips in response to price incentives, but scheduling costs a) prevent all drivers from doing so, and b) lead drivers that do shift to adjust their travel time by the minimum amount necessary to receive the incentive.

Figures A.3 and A.4 plot the frequency of vehicle trips of before versus after the imposition of peak hour pricing for all hours of day. Qualitatively, the bunches appear to be most pronounced during the early morning (5 a.m.) and early afternoon (3 p.m.) price notches. Intuitively, this suggests that it is less costly to arrive early than arrive late for both morning and evening trips.

Using equation 1.23, I estimate that during morning commute hours, the marginal driver is roughly indifferent between saving \$6 being an hour early, and indifferent between saving \$15 and being an hour late. During evening commute hours, the marginal driver is roughly indifferent between saving \$9 being an hour early, and indifferent between saving \$13 and being an hour late. These estimates are summarized in Table 1.3.

Figure 1.7 — Bunching in Response to Peak-Hour Pricing

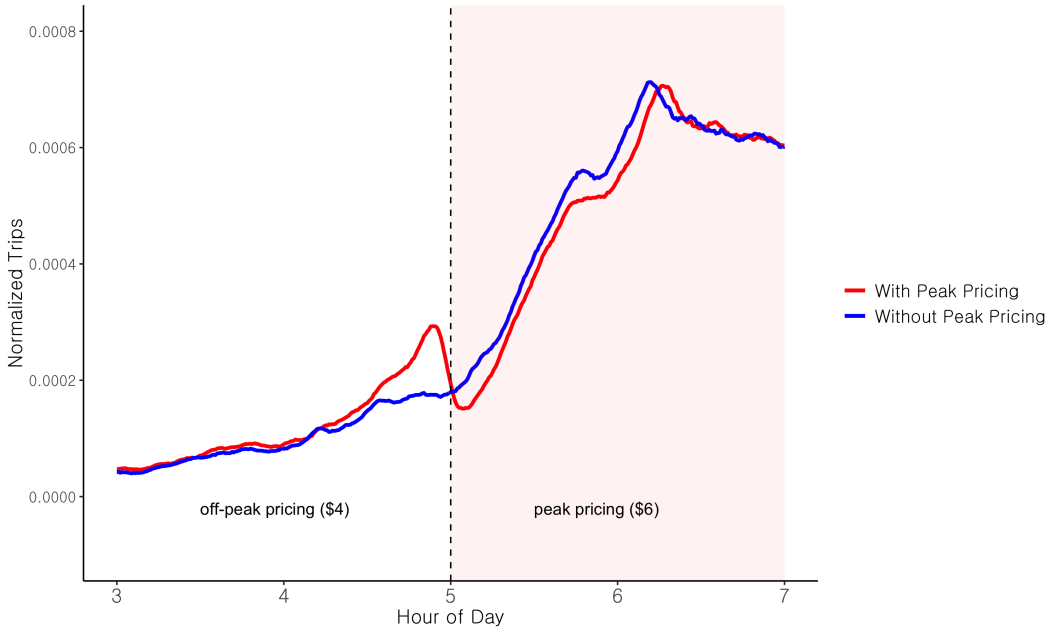


Figure 1.7: This figure plots the difference in the number of trips in the 6 months before (blue) vs the 6 months after (red) the imposition of peak-hour pricing on the Bay Bridge on July 1, 2010. To facilitate comparison, the number of trips at each time of day is normalized (divided by the total number of daily pre or post-period vehicle trips). The red shaded region demarcates times of day that were subject to peak-hour pricing after July 1, 2010. The vehicle trip counts reflect administrative tolling microdata collected by the Bay Area Toll Authority. Excluded from this graph are trip using the carpool/EV lane, which face a different pricing scheme. Figures A.3 and A.4 plot bunches for the other price notches (10 a.m., 3 p.m., and 7 p.m.) in the peak-hour pricing scheme on the Bay Bridge.

Table 1.3 — Estimating Scheduling Costs via Bunching

Parameter	**Estimate ($\hour)**
Time Early (5 a.m. notch)	6.195
	(0.419)
Time Early (3 p.m. notch)	9.744
	(1.545)
Time Late (10 a.m. notch)	15.498
	(2.593)
Time Late (7 p.m. notch)	13.759
	(1.306)

Table 1.3: This table shows estimates of the costs to drivers of scheduling trips earlier or later than the driver's ideal trip time (γ_e and γ_l in equation 1.13). I recover these estimates using equation 1.23, which relates scheduling costs to the number of additional vehicle trips observed in the period just outside of the peak-hour pricing period on San Francisco's Bay Bridge. In addition to the number of extra trips, equation 1.23 reflects scheduling frictions, as well as any time savings that result from drivers adjusting their trips to fall just outside of peak hours (assuming a $20 value of travel time for Bay-Area travelers, as estimated by Goldszmidt et al. (2020)). The additional bunching mass at price notches is estimated by comparing the number of trips in the neighborhood of the threshold time before vs. after the imposition of peak-hour pricing (see equation 1.21) using administrative tolling data from the Bay Area Toll Authority. Bootstrapped standard errors are in parentheses. All values are in 2010 dollars.

Appendix A.2 contains figures that examine the persistence of bunching behavior and the role of tax salience in determining bunching. Figure A.6 shows that the bunching behavior is more extreme for drivers who pay in cash than it is for drivers who pay electronic tolls, corroborating findings by Finkelstein (2009). The scheduling cost estimates in Table 1.3 and elsewhere in this paper reflect behavior of drivers using electronic toll systems, as this is the technology that would be used in many of the world's planned cordons. The difference in cash vs. non-cash responses to time-of-day toll systems suggests that factors that increase the salience of electronic tags (e.g, variable message signs displaying cordon costs) may lead to larger temporal adjustment. Figure A.5 compares bunching behavior at 6 months, 1 year, and 5 years after the beginning of peak hour pricing: the bunches become smaller over time, and the additional density is spread over a larger off-peak time zone at year 5 than it is at six months. Thus, while some drivers may be able to adjust their ideal arrival times in the long run, the parameters that drive bunching (schedule costs and ideal arrival times) appear stable for a large fraction of road users.

1.7.3. Comparisons to Parameter Estimates from the Literature

Several studies from the transportation economics literature provide valuable context for the logit and bunching estimator results presented in this subsection. A common heuristic for the value of travel time is 50% of the wage rate, which reflects seminal work by Lave (1969), as well as research collated by Small (2012). According to the 2010 - 2012 California Household Transportation Survey, the median Bay Area household earned roughly $66,000 per worker, equivalent to $31.74 per hour. The 50% heuristic therefore implies a median value of travel time of just under $16. Recent empirical estimates suggest slightly higher travel time: using a field experiment among Lyft riders, Goldszmidt et al. (2020) recover estimates of the value of travel time in San Francisco equal to roughly $20, or roughly 75% of the 2017 after-tax wage rate ($17.79 in 2010 dollars).

Estimates of scheduling costs (γ_e and γ_l) are less common in the economics literature. In general, existing studies accord with the canonical analysis by Small (1982), which found that a) it is more costly for drivers to be late than early, b) on a per-hour basis, the cost of being early is lower than the value of travel time, and c) the cost of being late can be higher or lower than the value of travel time depending on the setting. Kreindler (2018), for example, estimates that for drivers in Bangalore, India, early-arrival schedule costs are roughly a quarter of the value of travel time, and late-arrival is more costly than early arrival. In a 2005 choice experiment, Tseng, Ubbels, and Verhoef (2005) find that for drivers in the Netherlands, the cost of early arrival (€4.9/hour) is roughly half of the value of travel time (€9.8/hour), but late arrivals are very costly (€19.7/hour).

The scheduling costs I recover using discrete choice and bunching estimators are qualitatively similar to previous findings: The bunching estimator suggest that drivers prefer being early to being late. Both estimation strategies suggest that (on a per-hour basis) early and late costs are lower than the value of travel time.

1.8. Second-Best Optimal Cordon Prices

In this section I use the discrete choice model estimated in Section 1.7 together with the tax framework from Section 1.2 to calculate optimal cordon prices. I first demonstrate this procedure using San Francisco's proposed cordon, and then consider cordon zones in Los Angeles and New York.

At a high level, calculating optimal cordon prices in any city takes four steps: First, use travel survey data (e.g., the National Household Transportation Survey) to identify a representative sample of trips that pass through a city's proposed cordon. Second, assign externalities to those trips using information about the vehicle driven in each trip, and the

traffic density along the trip (this process is similar to the process described in Section 1.5). Third, use the model estimated in Section 1.7 to calculate substitution elasticities between different trips available to drivers. And fourth, apply the optimal tax formula outlined in Section 1.2 to the ingredients from steps 1-3.

1.8.1. San Francsico's Proposed Cordon Zone

The San Francisco County Transportation Authority (SFCTA) intends to pilot a downtown congestion pricing program in the next 3-5 years, with the goal of implementing cordon pricing by the end of the decade (San Francisco County Traffic Authority, 2021). Figure 1.8 shows a map of the proposed cordon zone, and Table 1.4 the proposed tolling schedule.

In the main results presented in this section, I treat as fixed the shape of the cordon and the time periods where prices will be charged. Doing so accords with the setup of the second-best tax problem described in Section 1.2, where the set of taxable goods is an exogenous constraint. Here the set of taxable goods, J, includes only two goods: morning and evening peak-hour trips that pass through the cordon zone. I present results from expanding the set of taxable goods in Section 1.8.8.

For simplicity, I also assume that all passenger vehicles will be charged the same price for using the cordon zone. This assumption abstracts from the low-income cordon price exemptions being considered by planning organizations in many cities. In Appendix A.11, I show that because the majority of commuters would not qualify for this exemption, the changes in welfare, congestion, and pollution that would result from exempting low-income drivers in the San Francisco Bay Area are second-order. As acknowledged in Section 1.2, the setup of this problem also assumes that policymakers do not weigh marginal utility across income groups. For a characterization of optimal corrective taxation under preferences for redistribution, see Allcott, Lockwood, and Taubinsky (2019).

Table 1.4 — San Francisco's Proposed Congestion Pricing Scheme

	Income Group			
Time Period	High	Middle	Low	Very Low
Peak Hours (6-10 a.m., 3-7 p.m.)	6.50	4.33	2.17	Free
Off-Peak Hours	Free	Free	Free	Free

Table 1.4: San Francisco's proposed cordon pricing scheme as of September 1, 2021. Trips that enter the cordon (see Figure 1.8) would be charged during peak hours according to the income of the registered vehicle. An individual's *Income Group* depends both on income and family size. For single individuals, the annual income thresholds for high, middle, low, and very low income are $150,000, $116,000, $66,000 and $46,000, respectively. For a household of four, the thresholds are $65,000, $95,000, $142,000, $166,000.

Figure 1.8 — San Francisco's Proposed Congestion Pricing Zone

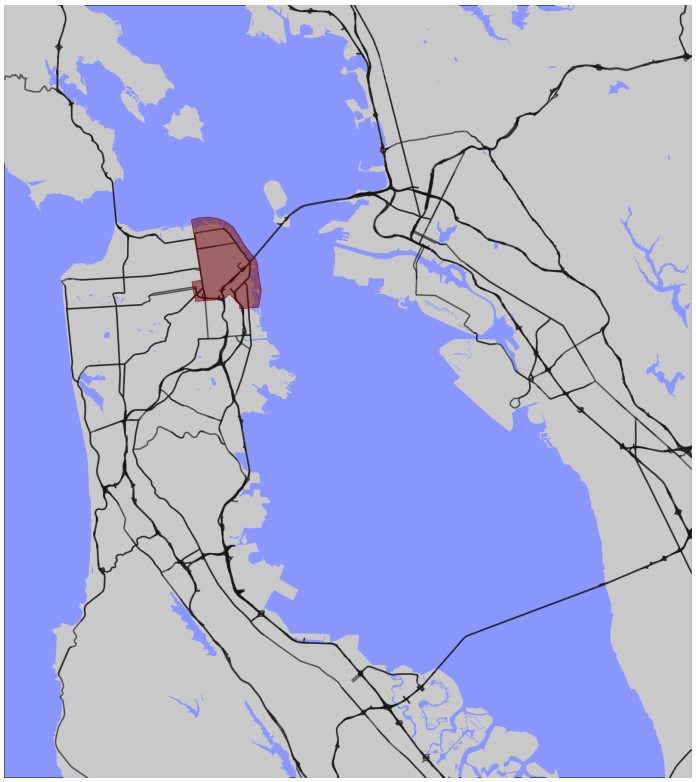


Figure 1.8: San Francisco's proposed cordon pricing scheme as of September 1, 2021. Trips that enter the cordon would be charged during peak hours according to the pricing scheme outlined in Table 1.4. Tolls will be levied using electronic tag readers mounted on gantries that span roadways on the border of the cordon region.

1.8.2. Personal Vehicle Trips in the San Francisco Area

The National Household Transportation Survey (NHTS) is a survey of US individual travel habits administered by the Federal Highway Administration. Participants in this survey are recruited via mail; survey responses are incentivized by small ($5 to $20) rewards, and can be completed through mail-back forms or online. The 2017 NHTS garnered responses from 381,975 individuals, each of whom filled out "Travel Diaries" that detailed their travel habits during one randomly selected 24-hour period. In addition to information about the attributes of the trip taken, the NHTS also collects demographic information about surveyed persons and their households.

I use the 2017 NHTS California Add-On[10] to build a representative dataset of trips that cross San Francisco's proposed cordon zone.[11] Each *trip* in this dataset consists of a start location (zip code or Census Block), an end location (zip code or Census Block), information about the vehicle that took the trip (make, vintage, fuel type), and the time of day that the trip was taken. I determine whether or not a trip passes through the cordon using the HERE Technology's *Routes* API. The resulting dataset contains 1,891 routes that cross the cordon zone during weekdays between the hours of 4 a.m. and 10 p.m., which I plot in the left pane of Figure 1.9.

To predict substitution in time and space under San Francisco's cordon, I construct a set of alternatives for each trip. For every cordon trip in the NHTS, I construct alternative departure times at 12-minute intervals throughout the day. Using HERE Technology's *Routes* API, I can assign travel times to each of these alternative trips by varying the departure time. For trips with termini that lie outside of the cordon zone (i.e., trips that only pass through the cordon zone en route to their destination), I identify the most direct detour that circumvents the cordon zone. I then calculate travel times for this non-cordon route for each 12-minute interval throughout an average traffic day. These detour routes are plotted in the right pane of Figure 1.9.

The result of this data collation is a set of trip endpoints for the San Francisco area, where drivers can choose over *route* $\in$ {cordon, non-cordon} and *time of day* $\in$ {4.0, 4.2, ... , 22.0}, as well as a generic outside option. This choice set allows me to predict how drivers would choose between options based on the attributes (travel time, time early, time late, and toll price) specified by the discrete choice model estimated in Section 1.3.

[10]The NHTS Add-On program allows States and Cities to pay the Federal Highway Administration to administer additional surveys in their region for the purposes of developing a more comprehensive sample.

[11]Note that the FasTrak microdata used in Section 1.7 are ill-suited for this task because many of the trips that cross San Francisco's proposed cordon do not use any bridge (e.g., trips with two termini within the city of San Francisco).

[This space is left intentionally blank]

FIGURE 1.9 — CORDON AND NON-CORDON ROUTES FOR BAY-AREA TRIPS

Cordon Routes *Non-Cordon Routes*

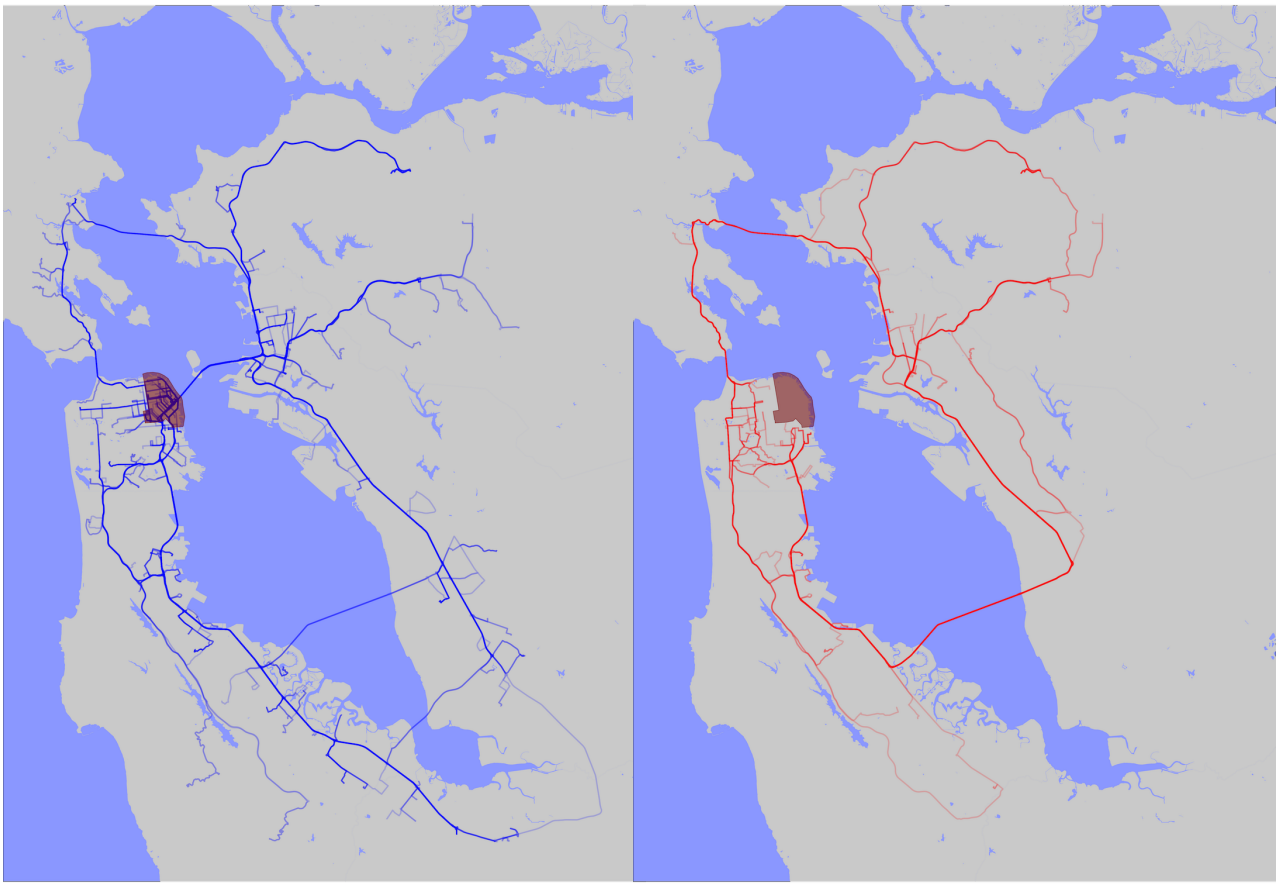


Figure 1.9: This figure plots cordon and non-cordon routes constructed from the 2017 National Household Transportation Survey (NHTS) California Add-On. The left pane plots 1,891 trips that cross San Francisco's proposed congestion zone, according to suggested routes generated with the HERE Technology's *Routes* API. The right pane plots detour routes for the subset of these trips where it is possible to circumvent the congestion zone (i.e., trips with both start and end points that are outside of the cordon). Each driver's choice set consists of a cordon route (the left pane) for every 12-minute time of day interval, as well as a non-cordon route (the right pane), if such a detour exists, for every 12-minute time of day interval. The choice sets of all drivers also include a generic outside option.

1.8.3. Trip-Level Externalities

For each trip described above (trips in the NHTS with suggested routes that pass through the cordon, as well as alternative trips in space and time), I assign traffic and pollution

externalities in a manner similar to the process described in Section 1.5. The detail of the NHTS survey data, however, allows for more precise estimation of both congestion and pollution externalities relative to trips observed in the FasTrak tolling data.

As shown in Figure 1.5, emissions vary by vehicle attributes as well as travel speed. The NHTS includes information about the vehicle used on each trip, including the vehicle vintage, make, and fuel type (gasoline, diesel, EV, or hybrid). Using the travel time and distance information for each trip returned by the HERE Routes API, I assign an average speed to each trip. I then merge emissions factors onto each trip based on vehicle vintage, fuel type, and travel speed, using data from California's EMFAC database. I plot the emissions externalities for the 1,891 NHTS trips that cross the proposed cordon in Figure 1.10.

To assign congestion externalities to trips, I use estimates from Yang, Purevjav, and Li (2020), who show that the marginal external (travel time) cost of traffic is convex in traffic density. Following the procedure used to assign externalities to FasTrak trips, I rely on a comprehensive network of traffic sensors on roadways in the Bay Area to estimate the traffic density along each route at different times of day. Concretely, this requires first identifying sensors along the trip's route, then assign a marginal external congestion cost (in dollars per mile) to this point based on the average traffic density at that sensor at the time of day associated with the trip. A trip's total congestion externality is then the average of the external congestion costs (in dollars per mile) along the route, multiplied by the length of the trip. I plot the trip-level externalities for the 1,891 NHTS trips that cross the proposed cordon in Figure 1.10.

Figure 1.10 — External Costs for Trips Crossing San Francisco's Proposed Cordon

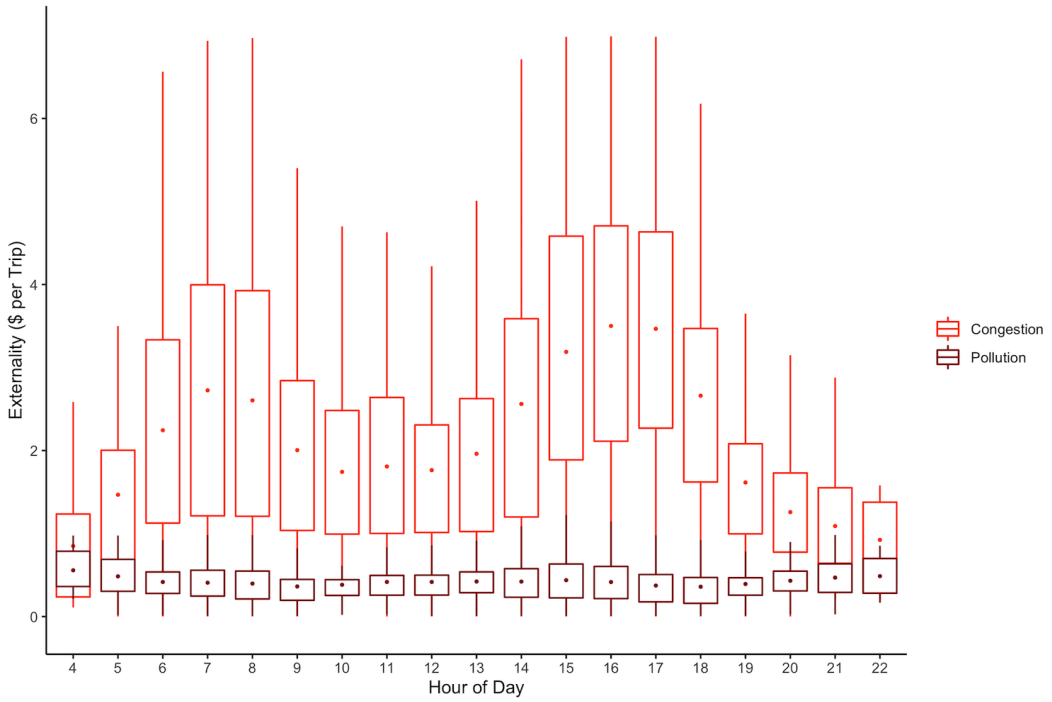


Figure 1.10: This Figure plots pollution (brown) and congestion (orange) externalities by hour for trips in the 2017 National Household Transportation Survey (NHTS) with suggested routes that pass through San Francisco's proposed cordon zone. The mean externality within any given hour is represented by a dot; the box spans the 25^{th} to 75^{th} externality percentile, and the bars span the 5^{th} to 95^{th} externality percentile. Trip routes reflect the suggested directions from HERE Technology's *Routes* API. Congestion costs were calculated by identifying traffic sensors along a given route and assigning per-mile congestion costs to each sensor using estimates of the density-congestion relationship from Yang, Purevjav, and Li (2020) and an average value of travel time of $20, as per Goldszmidt et al. (2020). Pollution emissions were calculated by merging emissions factors from California Air Resources Board's EMFAC database to trips based on vehicle fuel type, vehicle age, and average trip travel speed. I convert emissions to externalities using EPA social costs for global pollutants and EAISUR costs for local pollutant emissions in San Francisco. All values are in 2020 dollars.

1.8.4. Substitution Between Trips

The last set of parameters necessary for calculating optimal cordon prices are the parameters that govern how substitutable trips are in time and space. Specifically, calculating optimal prices using equation 1.10 requires *leakage shares* between trips: $\frac{dh_k}{dp_j} / \frac{dh_j}{dp_j}$. Recall that if j and k are trips (defined as a specific route $\in$ {cordon, non-cordon} at a specific hour of day $\in$ {4.0, 4.2, ..., 22.0}) the leakage share between trip k and trip j represents the share of the reduction in usage of trip k that shifts to trip j as a result of the increase of the price of taking trip j. For a concrete example, imagine that a one dollar increase in the price of driving through a cordon zone between the hours of 8 a.m. and 9 a.m. reduces trips by 10%, with 6% of all trips shifting one hour earlier (call these trips y) and 4% of trips shifting to routes that circumvent the cordon (call these trips z). The leakage shares are $\frac{dh_y}{dp_x} / \frac{dh_x}{dp_x} = 0.6$ and $\frac{dh_z}{dp_x} / \frac{dh_x}{dp_x} = 0.4$, respectively.

The leakage shares are implied directly from parameters of the mixed logit regression estimated in Section 1.7.[12] Formally, for any two trips $\{h^l, r^m\}$ and $\{h^j, r^k\}$, where h is the hour of day for a given trip and r is an indicator for whether or not the trip crosses San Francisco's cordon, the leakage share between these two trips for individual i is:

$$\frac{\frac{\partial \{h^j, r^k\}}{\partial p\{h^l, r^m\}}}{\frac{\partial \{h^l, r^m\}}{\partial p\{h^l, r^m\}}} = \int_\theta \beta_i s_{\{h^j, r^k\}}(\theta) s_{\{h^l, r^m\}}(\theta) f(\theta) d\theta \tag{1.24}$$

where θ is the joint distribution of random coefficients in the mixed logit model, β_i is the logit parameter governing price responsiveness for individual i, and $s_{\{h^j, r^k\}}(\theta)$ is the of share predicted trips that take route k at time j under the random coefficient vector θ.

1.8.5. Optimal Prices

Figure 1.11 plots three lines relevant for understanding optimal cordon prices. The blue (solid) line plots the average externalities for trips that pass through San Francisco's cordon zone by hour of day, estimated using the process detailed above. The green (dotted) line shows these externalities re-weighted as per Diamond (1973) to account for the correlation between the price-responsiveness of trips and idiosyncratic trip-level externalities, as reported in Table 1.2. Finally, the red line plots the second-best optimal prices for San Francisco's proposed cordon when tolling is restricted to morning and evening peak hours (6-10 a.m. and 3-7 p.m., respectively). The second-best optimal scheme charges $2.20 during morning peak hours, and $2.85 during evening peak hours. These second-best optimal prices are calculated

[12] Note the distinction between this formula, which recovers a cross-price *derivative*, and the canonical formula for a mixed logit cross-price *elasticity* (e.g., page 144 of Train (2009))

using equation 1.10, and take into account both the correlation between externalities and elasticities, as well as the substitution to unpriced alternatives in time and space.

FIGURE 1.11 — SECOND-BEST OPTIMAL ROAD PRICES

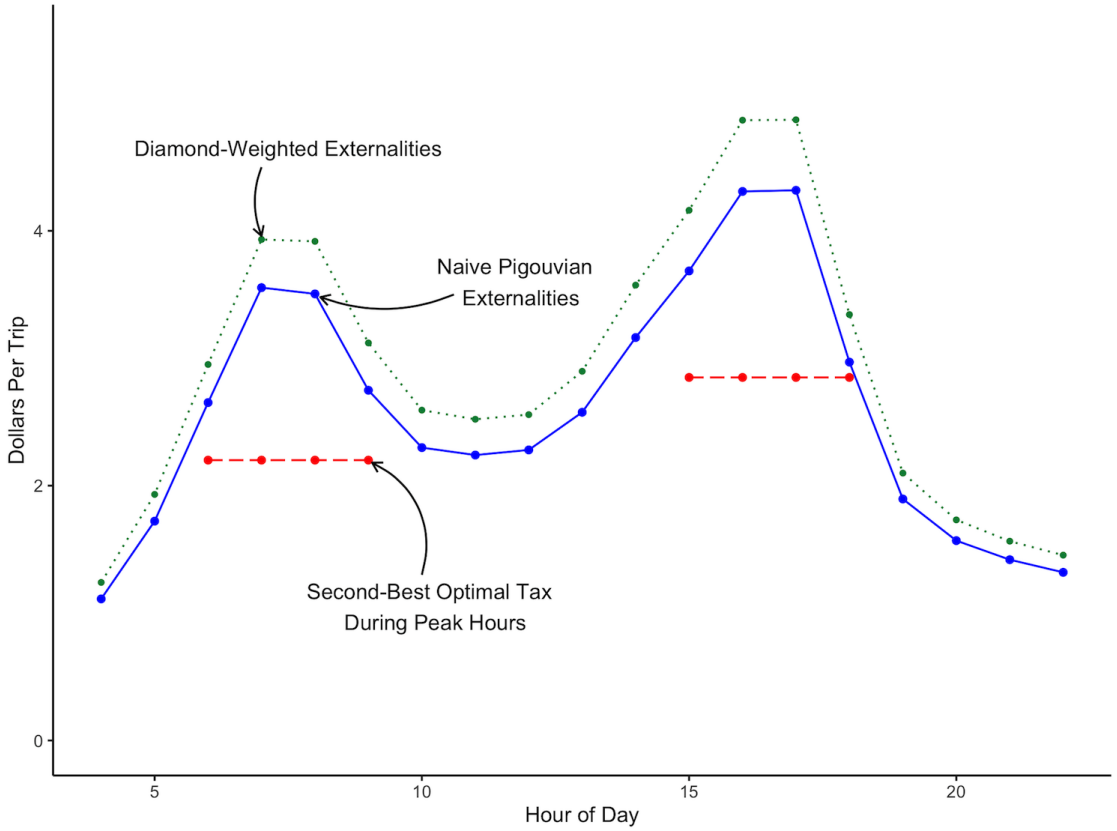


Figure 1.11: This figure plots three prices relevant for understanding optimal second-best cordon tolls. The blue (solid) line plots the average externality (pollution and congestion) for trips that cross San Francisco's cordon by hour of day, estimated using data from the 2017 NHTS (see Section 1.8.3). The green (dotted) line plots externalities re-weighted to account for the correlation between trip-level externalities and trip-level elasticities, as per Diamond (1973). The red line plots the second-best optimal price for San Francisco's proposed cordon when tolling is restricted to morning and evening peak hours (6-10 a.m. and 3-7 p.m., respectively). These second-best optimal prices are calculated using equation 1.10, and takes into account both the correlation between externalities and elasticities ("Diamond weights"), as well as the substitution (leakage) to unpriced alternative trips in time or space.

The results plotted in Figure 1.11 reflect social damages calculated using driving conditions that exist in the current, untaxed equilibrium. Consistent with the literature on externality taxation, the second-best tax formula presented in Section 1.8 phrases optimal taxes as a function of externalities at the optimum. As shown in figures 1.4 and A.15, the marginal damages associated with driving are non-constant in traffic density/speed, meaning that in general, damages at the taxed equilibrium will be different (lower) than those observed in the

untaxed equilibrium. Whether or not the difference between marginal damages calculated at versus away from the optimum is a first-order concern depends on the slope of the marginal damages function and the responsiveness of drivers to taxation. In Appendix A.6, I use simulations where I iteratively calculate taxes and traffic density to bound the second-best optimal cordon prices in San Francisco. The fixed point from this exercise constitutes a lower bound because it ignores "induced demand", which will tend to attenuate the difference in traffic conditions between taxed and untaxed equilibria (Duranton and Turner, 2011). I recover lower bounds of $1.60 and $1.80 for the morning and evening peak-hours, respectively.

1.8.6. The Impact of Pricing on Congestion, Emissions, and Welfare

Figure 1.11 shows that because tolls would incentivize drivers to substitute to routes that avoid the cordon zone (where they would still cause congestion and pollution), the optimal peak-hour cordon prices are below the marginal social damages associated with the average vehicle trip using the cordon zone. In this subsection, I estimate counterfactual driving behavior under a number of tax scenarios to understand the extent to which the imperfections in cordon pricing policies undermine the congestion, pollution, and welfare gains engendered by road pricing policies. These three scenarios are:

1. **No congestion pricing**. This is the status quo; the only charges that trips may face are the existing Bay Area bridge tolls, set to 2020 levels.

2. **First-best (Pigouvian) pricing**. Every trip a driver could choose would be priced according to its social damages, which include both congestion and pollution externalities.

3. **Second-best optimal peak-hour cordon prices**. These prices are calculated using equation 1.10. Trips that pass through the cordon area are charged $2.20 during morning peak hours, and $2.85 during evening peak hours (see Figure 1.11).

I plot outcomes from these simulations in Figures A.7 through A.9, and summarize the results in Table 1.5. Two themes emerge: first, on all three outcome measures — trips, congestion externalities, and pollution externalities — second-best optimal peak-hour pricing more closely resembles the status quo than the first-best policy. Second, there are distinct bunches in total trips, congestion, and pollution just outside peak-hour pricing periods.

Table 1.5 — Congestion, Pollution, and Welfare Effects of San Francisco's Cordon Zone

Outcome	Performance Relative to First-Best (%)
Reduction in Total Externalities	30.276
Reduction in Congestion	31.043
Reduction in Pollution	23.699
Welfare Gain	28.840

Table 1.5: This table compares the second-best optimal cordon pricing scheme in San Francisco to a first-best policy where all vehicle trips (all times of day; cordon and non-cordon) are charged according to marginal social damages. The four outcomes of interest are total externalities (pollution and congestion), congestion alone, pollution alone, and total welfare (the utility of all drivers, in dollars, less total externalities). The results in this table reflect 600,000 simulated choices by drivers in the NHTS dataset constructed above — 600,000 is roughly the number of weekday trips that pass through San Francisco's proposed cordon, according to the San Francisco County Transportation Authority. The choice probabilities for different alternatives (cordon vs. non-cordon trips at different times of day, and a generic outside option) were generated by applying the mixed logit model shown in Table 1.2 to the NHTS driver choice sets constructed in Section 1.8.

1.8.7. Cordon Pricing in New York and Los Angeles

In addition to San Francisco, city governments in New York and Los Angeles are also considering implementing cordon pricing zones (mapped in Figure 1.12). In this section, I calculate optimal peak-hour cordon prices for each of these cities, and evaluate the performance of the second-best optimal cordon pricing scheme relative to a policy that prices every trip at social marginal damages.

FIGURE 1.12 — PROPOSED CORDONS IN NEW YORK AND LOS ANGELES

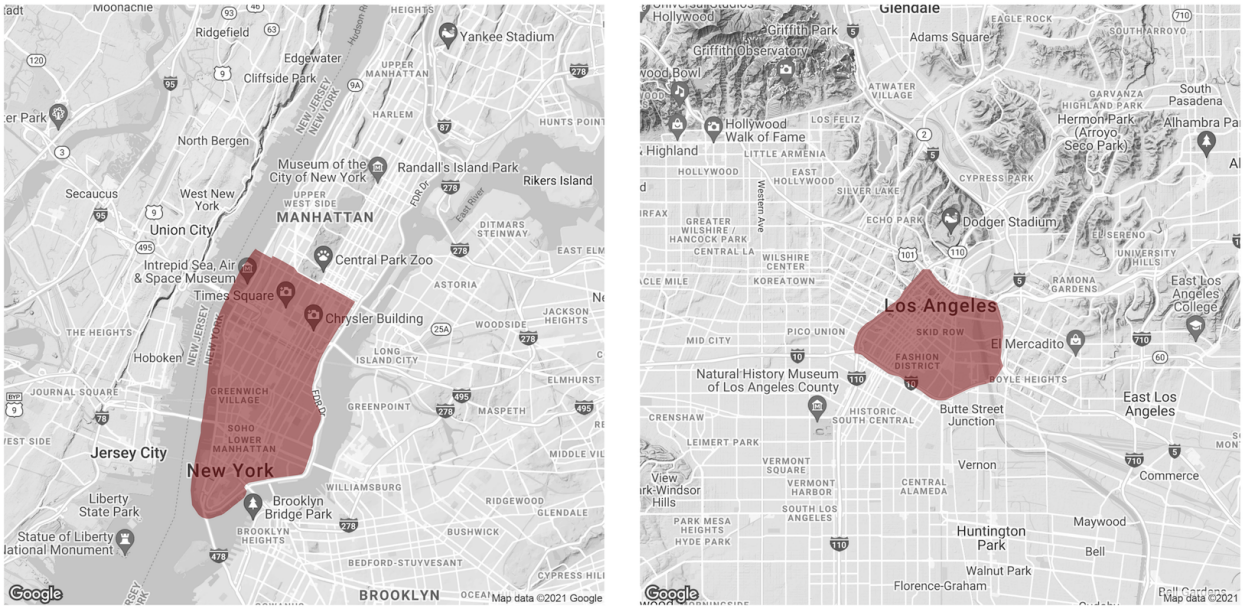


Figure 1.12: Proposed cordon pricing schemes in New York and Los Angeles. All proposals are as of August, 2021. The New York congestion map is courtesy of the *Regional Plan Association*; The Los Angeles map is courtesy of the *LA Metro*.

As outlined in Section 1.2, calculating the second-best optimal cordon prices requires information about the marginal damages of trips that cross through a cordon zone, as well as information about the elasticity and substitutability of these trips. For each of the above cities, I follow the same general template as in San Francisco (see Sections 8.2 through 8.4): First, I use survey data[13] and Here Technology's *Routes* API to identify trips where the fastest route passes through the city's proposed cordon. Second, I use vehicle attributes and travel speed to assign pollution externalities, and use traffic density data[14] from city roads to assign congestion externalities to those trips. Third, I calculate substitution parameters between those trips.

Ideally, there would be a natural experiment in each city that would allow for the estimation of city-specific driving demand primitives (price responsiveness, β, scheduling costs, γ_e and γ_l, and the value of travel time, α) that are used to calculate substitution parameters, as well as city-specific correlations between externalities and price responsiveness (Diamond weights). Absent such experiments, I calculate optimal cordon prices and welfare outcomes

[13]The NHTS does not report detailed trip start and end locations for states that are not part of the NHTS Add-On program. The trip-level data for New York come from the 2018 NY Citywide Mobility Survey.

[14]Traffic density data for Los Angeles is publicly available through PeMS. Traffic density for NY is courtesy of the NYSDOT Traffic Monitoring Section.

in New York and Los Angeles using the driving demand primitives and Diamond weights estimated in San Francisco (see Table 1.2). These results are reported in Tables 1.6 and 1.7.

In Appendix A.9, I use questions from the 2017 NHTS to examine the external validity of the model estimated in San Francisco. Specifically, the NHTS asks respondents to report their schedule flexibility (Yes/No) as well as their responsiveness to gasoline demand (Scale of 1 to 5). These proxies for demand primitives are broadly similar across the three cities I examine in this paper. In Appendix A.10, I document substitution to public transit in response to the increase in Bay Area bridge tolls, and discuss how optimal cordon prices may differ based on the availability of public transportation options. While estimates from a regression discontinuity performed on data from the Bay Area Rapid Transit (BART) system suggest that transit ridership increased after July 2010 (see Table A.7), the implied magnitude of mode shifting is small: These estimates suggest that only 6% of drivers who chose not to drive in response to the higher toll prices substituted those trips with BART. In Table A.8, I test whether access to public transit impacts the price responsiveness of Bay Area drivers. Point estimates suggest that drivers who live in zip codes near transit stations may be modestly more price responsive than those who live far away from transit stations, but this difference is not statistically significant. Broadly, the public transit ridership patterns in the Bay Area imply that while some drivers do shift to public transit when the price of their commuting trips increase, these shifts are relatively small, even in transit-rich areas.

Table 1.6 — Comparing Second-Best Cordon Prices to Social Damages

	Value ($)		
Period	San Francisco	Los Angeles	New York
Second-Best Price, AM Peak (6-10)	2.201	3.298	7.294
Second-Best Price, PM Peak (3-7)	2.850	4.533	7.919
Average Social Damages, AM Peak (6-10)	3.115	4.877	8.186
Average Social Damages, PM Peak (3-7)	3.821	5.724	12.635

Table 1.6: This table compares second-best optimal peak hour prices for the proposed cordons in San Francisco, Los Angeles, and New York to the average social damages associated with trips that pass through the cordon zones during this period. "Social damages" include both congestion and pollution damages. The second-best optimal cordon prices were calculated using Equation 1.10 — they reflect both heterogeneity in trip-level externalities, and leakage in time and space.

Table 1.7 — Congestion, Pollution, and Welfare Effects of Peak-Hour Cordon Pricing

	Performance Relative to the First-Best (%)		
Outcome	San Francisco	Los Angeles	New York
Reduction in Total Externalities	29.846	20.908	39.105
Reduction in Congestion	30.741	21.361	39.489
Reduction in Pollution	22.031	16.320	34.938
Welfare Gain	28.370	15.060	42.717

Table 1.7: This table compares the second-best optimal peak-hour cordon pricing scheme in 3 US cities to a first-best policy where all vehicle trips (all times of day; cordon and non-cordon) are charged based on the social damages they generate. "Peak hours" are defined as 6-10 a.m. and 3-7 p.m.; second-best cordon prices are constrained to be uniform during these hours. The four outcomes of interest are total externalities (pollution and congestion), congestion alone, pollution alone, and total welfare (the utility of drivers, in dollars, less total externalities). The results in this table reflect the simulated choices of 600,000 (SF and LA) to 1 million (NY) drivers. The choice probabilities for different alternatives (cordon vs. non-cordon trips at different times of day, and a generic outside option) were generated by applying the mixed logit model shown in Table 1.2 to the driver choice sets constructed using transportation survey data (see Section 1.8).

1.8.8. Hourly Cordon Pricing

Tables 1.5, 1.6, and 1.7 describe results where the policymaker is restricted to only price cordon trips during peak hours, as is proposed by the San Francisco County Traffic Authority. In this section I relax this constraint, allowing the policymaker to set a fixed hourly toll schedule during normal commuting times. In the notation of the second-best tax model outlines in Section 1.2, The set J now includes 14 taxable "goods," where each good covers all cordon trips for a given hour of day $\in \{6, 7, ..., 19\}$.

Tables 1.8 and 1.9 display estimates of welfare outcomes under second-best tax with hourly cordon pricing versus a first-best policy where every trip is charged according to the social damages associated with that trip. Relaxing this constraint leads to significant welfare improvements in three cities relative to a peak-hour scheme — the welfare gains are roughly twice as large in San Francisco and Los Angeles, and 16% higher in New York City. In each of the three cities, however, this more flexible policy fails to achieve more than half of the welfare gains that would be realized under a first-best policy, due to a combination of externality leakage, and uniform prices charged to heterogeneous users of the cordon zone

within each hour.

Table 1.8 — Congestion, Pollution, and Welfare Effects of Hourly Cordon Pricing

	Performance Relative to the First-Best (%)		
Outcome	San Francisco	Los Angeles	New York
Reduction in Total Externalities	44.459	36.596	49.358
Reduction in Congestion	45.097	37.044	50.086
Reduction in Pollution	38.975	32.072	41.457
Welfare Gain	51.104	29.392	49.706

Table 1.8: This table summarizes the performance of second-best optimal cordon pricing schemes when policymakers are allowed to set a fixed schedule of tolls between 6 a.m. and 7 p.m., relative to a first-best policy where every trip is charged according to its social damages. The four outcomes of interest are total externalities (pollution and congestion), congestion alone, pollution alone, and total welfare (the utility of drivers, in dollars, less total externalities). The results in this table reflect simulated choices using the mixed logit model shown in Table 1.2. The number of simulated trips reflects the number of daily trips in the cordon regions proposed in each city (600,000 in San Francisco and Los Angeles, and 1,040,000 in New York), as estimated by the planning authorities responsible for designing the cordon in each respective city. The "first-best" is a policy where all trips (regardless of the time of day or whether they pass through the cordon) are charged according to the marginal damages associated with that trip.

Table 1.9 — Back of the Envelope Welfare Gains From Cordon Pricing

	Welfare Gain Relative to the Status Quo ($ Million)		
Policy	San Francisco	Los Angeles	New York
First-Best	852	1,100	1,477
Second-Best (Peak Only)	246	166	630
Second-Best (Fixed Hourly)	412	323	715

Table 1.9: This table displays back of the envelope calculations for the annual welfare gains under three road pricing policies: 1) The first-best policy where all trips (including those that re-route to avoid a city's cordon) are priced according to marginal congestion and pollution damages; 2) second-best peak hour (6-10 a.m. and 3-7 p.m.) prices (see Table 1.6); and 3) second-best-optimal time-of-day prices, which are allowed to vary by hour according to a fixed schedule between 6 a.m. and 7 p.m. The cordon prices in rows (2) and (3) are calculated using Equation 1.10 — they reflect both heterogeneity in trip-level externalities, and leakage in time and space. The figures in this table reflect simulated choices using the mixed logit model shown in Table 1.2. The number of simulated trips reflect the number of daily trips in the cordon regions proposed in each city (600,000 in San Francisco and Los Angeles, and 1,040,000 in New York), as estimated by the planning authorities responsible for designing the cordon in each respective city.

1.9. Discussion

Cordon prices differ from a first-best driving tax in two important ways: incomplete coverage allows for leakage, and uniform prices cannot reflect the heterogeneity in trip-level damages. I find that these two imperfections are in tension as they apply to optimal road prices. Absent leakage, the correlation between price-responsiveness and trip-level externalities (Table 1.2) implies second-best prices that are *above* marginal damages. Absent heterogeneity, leakage to unpriced roads or times of day implies second-best prices that are *below* marginal damages.

The results from Figure 1.11 show that the leakage effect strongly dominates in the case of San Francisco's cordon zone: Optimal prices are $2.20 for the morning peak period and $2.85 for the evening peak period — roughly half of the average social cost for trips that pass through the cordon at those times.

My findings suggest that if the primitive determinants of driver decisions (price responsiveness, value of travel time, schedule flexibility) are similar across cities, then optimal cordon prices are also below the average of social damages generated by downtown trips in New York and Los Angeles. Table 1.6 shows that in New York, for example, the second-best optimal cordon prices are about $7 for both the morning and evening peaks, which is below

the average social damages associated with cordon trips in each of those periods ($8.18 and $12.64, respectively). In Los Angeles, the optimal morning and evening peak prices are $3.30 and $4.53, compared to average social damages of $4.87 and $5.72.

Cordon zones charging the second-best prices described in Table 1.6 would generate significant welfare gains for commuters and city residents in all three cities. The benefits from optimal peak-hour cordon prices range from $246 million annually in San Francisco to $426 million annually in New York (Table 1.9). To put these figures in perspective, the 2021 annual budget of the City of San Francisco is $13.7 billion, and the 2021 annual budget of New York is $88.2 billion. These annual welfare gains are therefore on the order of 0.5 to 2% of city budgets.

Despite these gains, the results in Section 1.8 suggest that the blunt nature of cordon pricing limits their effectiveness relative to an ideal policy. Optimal peak-hour cordons achieve between 15% (Los Angeles) and 41% (New York) of the welfare gains that would be realized under a first-best policy. Notably, peak-hour pricing policies are less effective at internalizing pollution externalities than they are at internalizing congestion externalities. This reflects the fact that a) congestion externalities represent the majority of the social damages from an average cordon trip and are therefore implicitly more heavily weighted in the optimal tax formula, and b) trip-level pollution externalities are not highly temporally correlated with congestion externalities, as shown in Figure 1.10.

What adjustments could improve the performance of the proposed cordon zones in the United States? Relative to a peak-only tolling scheme, allowing policymakers to set a fixed schedule of prices that vary by time of day (Table 1.8) provides sizeable welfare gains: $166 million in San Francisco, $151 million in Los Angeles, and $85 million in New York. In each city, however, this flexible pricing strategy fails to achieve half of the welfare gains relative to the first-best.

This paper takes as given the spatial layout of cordon zones. The central role of leakage in determining optimal prices, however, highlights the importance of a city's choice of cordon boundaries. The theory provided in Section 1.2 suggests that policymakers may want to set boundaries to preempt spatial leakage. Depending on the idiosyncratic geography of a city, an optimal cordon zone may include outlying or relatively uncongested routes that provide close substitutes for congested central roads. Expanding cordon zones, however, comes at a cost; regardless of the design of the tolling system at the boundaries, trips that remain entirely inside the cordon are not priced. Expanding the cordon too far may therefore undermine the policy's overall coverage. A full characterization of this tradeoff is beyond the scope of this paper, but may prove an interesting question for future research.

1.10. Conclusion

This paper makes three contributions: First, this paper generates the first estimates of optimal cordon prices that account for both pollution and congestion externalities. While optimal prices vary across proposed cordon zones in the US, several themes emerge: Congestion externalities constitute the bulk of marginal damages that determine optimal cordon prices, generally outweighing pollution externalities five- to ten-fold. This finding accords with work by Parry and Small (2005), who suggest that congestion (rather than pollution) is the largest component of an optimal gasoline tax. Additionally, optimal cordon prices tend to be *below* the average social damages associated with trips that cross through a cordon because of externality leakage in time and space. This leakage effect dominates the heterogeneity effect (the correlation between trip-level externalities and trip-level elasticities), which, all else equal, pushes second-best optimal prices above average social damages (see Diamond, 1973).

Second, this paper presents the first estimates of the welfare losses that result from imperfections in real-world cordon policies. Back of the envelope calculations suggest that while a second-best peak hour cordon price in San Francisco would produce roughly $200 million dollars worth of welfare gains, this policy would fall short of the first-best policy by $600 million annually. This foregone welfare is significant: $600 million is roughly 4% of the City of San Francisco's 2020-2021 Budget ($13.7 billion). The predicted performance of proposed cordons in New York and Los Angeles are qualitatively similar. Notably, among these imperfect policies, the peak-hour cordon zone in New York performs the best (capturing 42% of possible welfare gains), and the peak-hour cordon in Los Angeles performs the worst (capturing just 15% of possible welfare gains). This reflects the fact that it is much more difficult to find substitute routes in New York than it is in Los Angeles due to idiosyncratic geography.

Lastly, this paper contributes to the literature in public and environmental economics by extending existing models of second best-taxation to simultaneously account for leakage and heterogeneity in externalities. Accounting for these policy imperfections implies subtly different policy prescriptions than the canonical "Principle of Targeting" Sandmo (1975). When externality leakage and externality heterogeneity are present, the policy instrument that generates the largest welfare improvements may not be the tax that best targets the naive average of externalities. Instead, for each good, the optimal instrument balances the magnitude of externality reduction with the damages that would result from leakage. The results in this paper highlight a case where, due to policy imperfections, the optimal policy differs significantly from a tax that best targets the average of consumption externalities. While applying the second-best tax framework outlined in this paper requires detailed infor-

mation about externalities and consumer demand, the increasing availability of microdata continues to lower the costs for credible estimation of demand systems. This trend, together with the ubiquity of imperfections in externality taxation, suggest that this framework will be useful for future research in settings outside of optimal road pricing.

Chapter 2

What Drives Support for Inefficient Environmental Policies? Evidence from a Nevada Ballot Initiative

2.1. Introduction

Negative externalities are often regulated with performance standards or quantity thresholds where economic theory suggests that price-based mechanisms offer a more cost-effective alternative. Examples include the US Corporate Average Fuel Economy standards, low-carbon fuel standards, and the US Clean Air Act. While certain inefficient policies can be tied to regulatory capture or legislative lobbying, the ubiquity of performance regulation also reflects the preferences of US voters. Standard-based policies for reducing electricity emissions, for example, enjoy bipartisan voter support in many states, where state-level carbon tax ballot initiatives have repeatedly failed in the US. Among the many idiosyncratic attempts to explain voters' aversion to price-based regulation, a 2018 meta-analysis highlighted cost salience, perceived (in)effectiveness, and fairness concerns as common themes (Carattini, Carvalho, and Fankhauser, 2018). A robust voting literature demonstrates that these three considerations also matter in other settings: In general, voters tend to prefer policies that they perceive as cheaper, fairer, and more effective (Healy, Persson, and Snowberg (2017), Huber, Wicki, and Bernauer (2020), Kuziemko, Norton, Saez, and Stantcheva (2015)).

Under these preferences, it is puzzling why voters routinely support performance standards over price-based policies. Given the cost-effectiveness of Pigouvian taxation and the ability of governments to pair these policies with redistribution, it should be possible to construct a price-based regulation that is superior to a performance-based regulation on at least one of the three dimensions of efficacy, fairness, or cost, holding fixed the others. One expla-

nation for the tension between voters' stated preferences and the attributes of policies they support is that voters are misinformed. A growing literature has documented voter misperception of policy features (Sapienza and Zingales, 2013), lending credibility to early models of voter inattention proposed by Downs et al. (1957), Sims (2003), and others. Positively attributing qualities of voter behavior to misperceptions, however, is a difficult empirical task. While demonstrating misperception is relatively straightforward, drawing causal conclusions drawn from stated preference data may be confounded by omitted variables, or suffer from incentive-incompatible survey formats.

In this paper, I solve these identification problems using an information provision experiment conducted around a vote on Nevada's Renewable Portfolio Standard (RPS). The variation induced by this experiment allows me to study how voters *perceive* and *respond to* policy attributes. The information provision experiment had three stages: First, I surveyed a pool of Nevadans on their support for both a 50% RPS (which was on the ballot in Nevada in 2020) and for a hypothetical alternative price-based policy (a $25 dollar carbon tax). I also recorded their initial perceptions of the cost, effectiveness, and regressivity of these policies in an incentive-compatible manner. Second, I provided respondents with source-randomized information about these policies. And third, in a follow-up survey I recorded voting behavior and posterior beliefs about both of these policies.

Leveraging the variation in beliefs about policy attributes induced by this information provision experiment, I estimate logit and linear models of voter support for corrective policies. These models allow me to answer three research questions: First, how do voter perceptions of policy attributes (cost, effectiveness, and regressivity) influence voter behavior? Second, do misperceptions of policy attributes explain voter preferences for non-tax corrective policies? And third, given the answers to the aforementioned questions, can policy design or information provision bolster support for price-based corrective policies?

Results from the initial survey confirm that respondents prefer the performance-based policy (RPS) to the price-based policy (carbon tax), and suggest significant inaccuracy and bias in beliefs about the attributes of these policies. For example, on average, respondents believed that Nevada's 50% RPS would generate emissions reductions roughly five times larger than estimates from academic research. Using variation in beliefs induced by the information provision experiment, I recover elasticities of policy support with respect to perceived policy attributes. I find that respondents are relatively unresponsive to perceived policy cost and perceived regressivity: point estimates on the *cost* coefficient suggest that decreasing the average voter's perception of a given policy's cost by $1000 annually (roughly four times the mean cost across policies) would increase the probability that the voter supports said policy by just 1.3 percent. Conversely, I estimate that policy support is relatively elastic with respect to perceived policy effectiveness.

Armed with models of voter behavior, I then investigate the extent to which (mis)perceptions of policy attributes explain the gap in policy support. While misperceptions of policy attributes are significant, Oaxaca-Blinder decompositions suggest that they do not explain a large portion of the gap in support between the two policy types. According to my estimated models of voter behavior, holding fixed perceptions of all policy attributes, Nevadan voters are still 13.8% more likely to support Nevada's RPS than they are the carbon tax alternative. To the extent that differences in perceptions of policy attributes do explain the premium that respondents placed on RPS policies, my estimates suggest that this is largely a result of their overly optimistic views of the effectiveness of these policies. Finally, I use these estimated models of voter behavior to investigate several counterfactual scenarios, which generally demonstrate the difficulty in achieving majority support for carbon taxation either through policy design, or information provision.

2.2. Related Literature

My research is connected to the existing economic literature in three ways.

First, studying how and whether voters respond to policy attributes is related to the existing literature on the rationality of voters. A substantial portion of this literature is devoted to empirically testing for behavioral 'types.' For example, it has been demonstrated that voters are sociotropic (Hansford and Gomez, 2015), retrospective Bischoff and Siemers (2013), and time-inconsistent (Banzhaf and Oates, 2012; Dell'Anno and Mourao, 2012).

It has also been previously noted that voter irrationality or misperception may lead to inefficient policy. Downs et al. (1957), for example, argues that rational ignorance (resulting from the low probability of any individual vote changing an electoral outcome together with the costs of acquiring information) could lead to inefficient policy. Alternatively, Caplan (2001) presents a model of rational inattention where voters actively resist updating their priors because they have preferences over *beliefs*. That is, religious and social identities lead people to prefer holding certain beliefs over others. He argues that inefficiencies arise from the externalities borne of this inattention: The private cost of inattentive activity is near zero, but in aggregate these actions lead to suboptimal policy.

My research will tangentially touch on two specific voter types: *inattentive* and *altruistic*. The choice to directly test for voter altruism reflects voter model advances by Jankowski (2007) and Edlin, Gelman, and Kaplan (2007), who added 'social preferences' to the egoist model of voter decisions as a way of solving the paradox of voting. Models that allow for altruism have also been substantiated in laboratory experiments (Fowler (2006); Dawes, Loewen, and Fowler (2011)). Recent work has also demonstrated significant misperceptions of the costs and benefits of public policies (Blaufus, Chirvi, Huber, Maiterth, and Sureth-

Sloane, 2020; Stantcheva, 2020; Sausgruber and Tyran, 2011), which impairs the ability of voters to choose optimally between options regardless of altruism. Taken together, this body of work provides a strong case for why voters may not follow the 'pocketbook' model implicitly assumed in many early models of voter behavior.

Second, this paper relates to a significant body of work investigating the prevalence of inefficient policies. Although none of my research questions explicitly require the policy in question to be efficient, the answer to these questions will allow me to contribute to a more general conversation on the political economy of efficient policies.

The political Coase theorem (PCT) is a central idea in the study of the (in)efficiency of institutions and policies. It stipulates that political actors should agree on policies that maximize efficiency, regardless of the original distribution of political power (Vira, 1997). This principle fails in many settings (Acemoglu, 2003). Prominent examples include unpriced road congestion, limits on free trade, and reliance on inefficient standards to regulate air pollution and vehicle emissions.

Broadly speaking, commitment and rent-seeking have been proposed as explanations for the failure of states to enact policies that could make everyone better off: North and Weingast (1989) first outlined why commitment issues may undermine the political Coase theorem: The ability to enforce contracts is crucial to a functioning Coase theorem. Because one of the parties entering into a political contract (the state, or a politician) is granted enforcement power, they cannot commit to not using this power to later alter the contract. With no guarantee that gains from an efficient policy will be distributed to citizens, voters are reluctant to pledge their support to policies or platforms that promise to improve efficiency. This idea was formalized in a game-theoretical model by Acemoglu (2003), and has been demonstrated empirically to suppress support for efficient policies Galiani, Torrens, and Yanguas (2014).

Special interest groups also contribute to the failure of the political Coase theorem. Rent-seeking was first introduced by Tullock (1967) and Krueger (1974), who both describe how this behavior leads to inefficient *outcomes*. Several models describe how rent-seeking could similarly lead to inefficient *policies*: Acemoglu and Robinson (2000), for example, models a monopolist who has the political capital to oppose the introduction of a new efficient technology in order to preserve rents, Becker (1983) models policymaking as competition for political pressure between taxpayers and special interests, and Grossman and Helpman (1994) model political contributions as bids that determine subsequent policy. Empirical studies have demonstrated lobbying's influence on efficient policies in several settings, including climate policy (Meng and Rode, 2019) and free trade Goldberg and Maggi (1999).

This paper will contribute to understanding the completeness of the current criticisms of the political Coase theorem. In a setting with an independent judicial system, ballot

initiatives effectively solve the commitment problem that plagues the political Coase theorem in other settings. The results from this paper will speak to whether misperception of policy attributes erodes the ability political actors to correctly identify and support efficient policies.

Lastly, the methods I use to investigate voter responsiveness are synthesized from a suite of papers that leverage misperception and information provision to understand how beliefs map to actions. Broadly, these papers identify the causal effect of changing beliefs on actions by first eliciting priors (in an incentive-compatible way) about the cost or benefit associated with some outcome (e.g., a tax, good, or insurance policy), and then randomizing information treatments. By matching actions (e.g., support, purchase, or uptake) with posterior beliefs, the econometrician can produce causal estimates for how beliefs change the variable of interest. These methods have been used to understand behavior in a wide range of settings, including preferences for income redistribution (Kuziemko et al. (2015)), preferences over relative income (Bottan and Perez-Truglia, 2017), participation in political protests (Cantoni, Yang, Yuchtman, and Zhang, 2019), perceived incidence (Rees-Jones and Taubinsky, 2016), female labor force participation (Bursztyn, González, and Yanagizawa-Drott, 2018), and support for carbon taxation (Douenne and Fabre, 2020).

2.3. Background: State-level emissions policies in the US

In this section, I provide a brief overview of two varieties of state-level carbon emissions policies in the United States, and touch on the political and efficiency considerations of each type of regulation.

The US does not comprehensively regulate emissions of carbon dioxide at the federal level.[1,2] Instead, to the extent that US carbon emissions are regulated, this is largely accomplished through a patchwork of state-level policies. In this study, I contrast two large-scale emissions policies: *Renewable portfolio standards* and *carbon pricing schemes*.[3]

Renewable portfolio standards are policies designed to reduce emissions from state-level electricity grids by mandating a shift toward renewable generation. While the details of RPSs vary across states, these policies generally require that a specified fraction of all electricity sold by utilities be generated from renewable sources.[4] Additionally, RPSs allow utilities to come into compliance by purchasing renewable electricity credits (RECs) from other utilities that exceed the RPS requirement. These policies are the most prevalent state-level carbon emissions policies: as of 2020, 30 states have instituted RPSs, and seven states have adopted similar non-binding renewable energy goals.

RPSs have been criticized for failing to incentivize decarbonization in the electricity sector along all possible margins (Reguant, 2018). RPSs do not, for example, provide incentives for utilities to supply electricity from relatively clean fossil sources (natural gas) over relatively dirty fossil sources (coal). Similarly, as with all output-based performance standards, RPSs introduce a second inefficiency via an implicit subsidy to overall production (Goulder, Long, Lu, and Morgenstern, 2019). Consistent with these inefficiencies, estimates of the price of CO_2 emissions avoided by RPSs tend to exceed most estimates of the social cost of carbon (Greenstone and Nath, 2019).

A smaller number of states (see Figure 2.1) have instituted price-based schemes as a means for reducing state-level emissions. These are the Regional Greenhouse Gas Initiative (RGGI), covering 10 Northeastern States, and California's Cap-and-Trade system, established in 2010

[1] Exceptions include Corporate Average Fuel Economy Standards (CAFE) and federal tax incentives for renewable production

[2] Several bills have failed in Congress, including the Climate Stewardship Act (in 2003 and 2005), the Global Warming Pollution Reduction Act (2007), and the American Clean Energy and Security Act (2009). The Clean Power Plan, first proposed by the EPA in 2014, was repealed by President Trump in 2017.

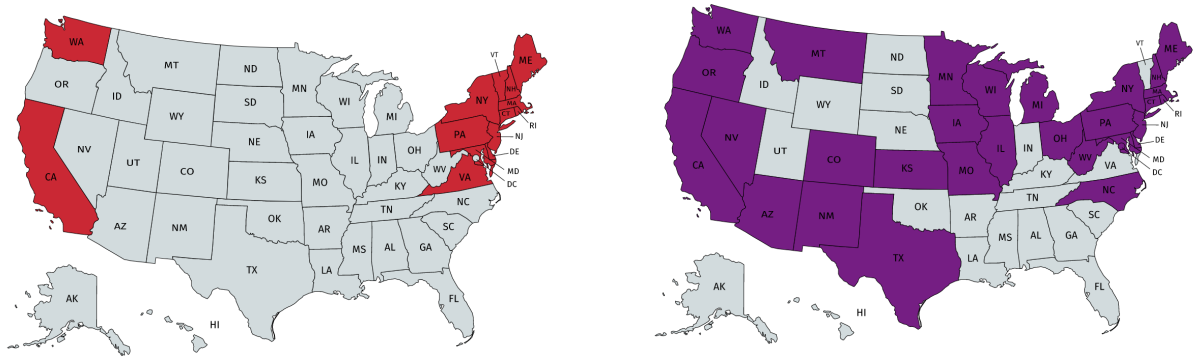
[3] Two other large-scale state-level electricity emissions policies are worth mentioning: feed-in tariffs and production subsidies. For a comprehensive analysis of the efficiency qualities of each of these policies, see Reguant (2018).

[4] "Renewable" sources generally include wind, solar, geothermal, and biomass. In some cases, hydroelectric and/or nuclear power are also included as "renewable" or "clean."

under AB32, and Washington state's Clean Air Rule, established in 2016. These policies allocate permits for carbon emissions and allow emitters to trade these permits. The overall number of permits available decreases over time in accordance with the idiosyncratic climate goals of the state or region. By tying compliance to total emissions rather than renewable generation, a price-based policy will equate the cost of emissions abatement across all possible margins, thereby achieving any given emissions reduction at the lowest cost (Boyce, 2018; Wilson and Staffell, 2018).

Despite the efficiency advantages of price-based emissions policies, they remain rare in the US. The paucity of carbon pricing or trading schemes reflects failures to pass policies both legislatively and electorally. Carbon tax ballot initiatives failed in Washington State in 2016 and 2018, and propositions failed to meet the signature requirements necessary to make the ballot in 2020 in both Oregon and Utah. Since 2018, bills instituting carbon pricing have failed in the state legislatures of Connecticut, Hawaii, Maryland, Massachusetts, Minnesota, Montana, New Hampshire, New Mexico, New York, Rhode Island, Texas, Utah, and Vermont.[5] As an anecdote illustrative of the political advantage enjoyed by renewable portfolio standards over carbon pricing schemes, four of these states (Connecticut, Massachusetts, New Mexico, Maryland) increased their RPS requirements within two years of rejecting a carbon tax legislative bill (EIA (2018, 2019)).

In the remainder of this paper, I leverage an RPS level increase in Nevada to understand how perceptions of the attributes of these policies informs support for both price- and performance-based emissions policies.



(a) Carbon Pricing Schemes

(b) Renewable Portfolio Standards

Figure 2.1: State-level emissions policies in the US

[5]See the legislative tracker maintained by *Price on Carbon*.

2.4. Setting: Nevada Question 6

Nevada's Question 6 was an instituted constitutional amendment that was approved by Nevada voters during the 2020 election. This amendment increased Nevada's Renewable Portfolio Standard from requiring that 25% of the State's electricity come from renewable sources by 2030 to requiring 50% by 2030. The ballot language for Question 6 is as follows:

> ***Question 6***
>
> *Shall Article 4 of the Nevada Constitution be amended to require, beginning in calendar year 2022, that all providers of electric utility services who sell electricity to retail customers for consumption in Nevada generate or acquire incrementally larger percentages of electricity from renewable energy resources so that by calendar year 2030 not less than 50 percent of the total amount of electricity sold by each provider to its retail customers in Nevada comes from renewable energy resources?*

This initiative passed with 57.94% support. There are two peculiarities of this initiative worth noting: First, this initiative was also on the ballot in 2018, as initiated constitutional amendments in Nevada require passage in two consecutive even-year elections. The 2018 initiative passed with 59.28% support. Second, the Nevada State Legislature had already adopted a 50% RPS target via SB 358. Because subsequent state legislatures could easily change this target, this 2020 initiative was advertised as bill to prevent backsliding of the RPS target. Nevada does not have a carbon pricing system, nor has it voted on a carbon pricing system through either a ballot initiative of a senate bill.

2.5. Survey Design

2.5.1. Participants

I recruited participants through three online platforms: Prolific, Amazon's CloudResearch, and UC Berkeley's Xlab. The survey was made available only to users on each platform who had registered as residents of Nevada, and who were 18 years old or older. To verify that respondents did not take the survey on multiple platforms, the survey presented to users on the *XLab* and *Prolific* platforms screened participants based on whether they had accounts with the other platforms.

Obtaining a sufficiently large sample was a significant challenge due to the relatively low number of Nevadans on these online platforms. I received 359 responses to the initial

survey and 316 responses to the follow-up survey. Of the 316 participants who returned for the follow-up, I was able to verify (by zip code or IP address) that 275 of them resided in Nevada. Table 2.1 displays the characteristics of the participant pool relative to the demographics of the Nevada electorate. Respondents tend to be younger, are more likely to identify as Democrats, and more likely to support Nevada Ballot Question 6 than the average Nevadan voter.

To account for this demographic and ideological bias engendered by selection into these platforms, I re-weight my sample using R's `anesrake` package, following (Battaglia, Izrael, Hoaglin, and Frankel, 2009). This package implements iterative proportional fitting (or "raking"), which aims to generate a set of sample weights that best match population proportions subject to user-specified objective functions and constraints on the magnitude of weights.

Descriptive Statistics

Respondents from Prolific, mTurk, and Xlab

Variable	Sample	Sd	Electorate
Female	0.524	0.500	0.510
White	0.652	0.477	0.730
Democrat	0.502	0.501	0.500
Under 44	0.791	0.407	0.480
Over 64	0.044	0.205	0.180
Support RPS (prior)	0.751	0.433	0.570
N=275			

Table 2.1: Descriptive statistics for the 275 Nevadans who participated in both the prior and the posterior surveys. The first column ("Sample") displays the fraction of participants with each row's trait. The third column ("Electorate") displays the same figures for the Nevada electorate, as per the US Census.

Figure 2.2 – Respondent Locations by Zip Code

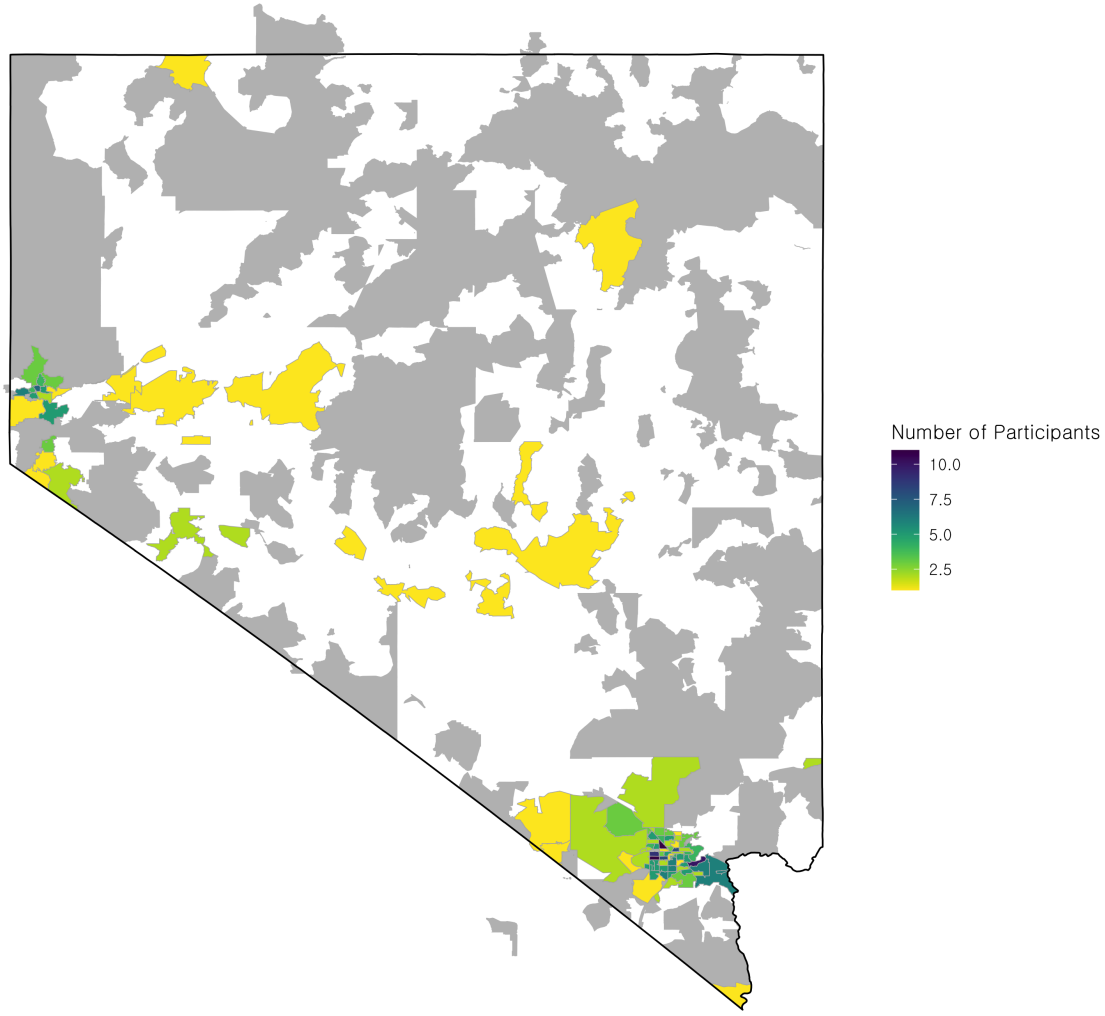


Figure 2.2: Locations of 275 individuals who participated in both prior and posterior surveys by zip code. The black outline is the Nevada State border. Zip codes colored gray did not have any respondents. White areas are uninhabited.

2.5.2. Information Provision Experiment

To understand how voter perceptions about policy attributes influence policy support, I solicited respondents' views on Nevada Ballot Question 6 as well as on a hypothetical price-based alternative policy, and tracked how their *beliefs* and *preferences* changed in response to information about these policies. Figure 2.3 shows this information with abbreviated survey questions. See Appendix B.1 for more details.

Prior Survey (Starting October 8th)

Beginning on October 8th, 2020, I opened the prior survey to Nevada residents. The rationale for this start date was to ask voters about their preferences as close to the election date as possible, while minimizing the probability that respondents had already voted by mail (Nevada began mailing ballots on October 9th, 2020). This initial survey had three parts:

Part 1: Elicit priors. In this section, participants were shown the official text from Nevada Ballot Question 6. They were asked whether or not they planned on voting for the bill, and asked to share their beliefs on three attributes of the bill: *cost, effectiveness, and regressivity*.

To elicit beliefs about perceived (private) policy costs, participants were asked whether they believed the RPS would financially cost or financially benefit their household. Based on the response to this question, they were then asked to report how much they expected the policy to cost/benefit them per year, in dollars. To understand perceptions of policy *regressivity*, participants were asked whether they believed the RPS would financially cost or financially benefit the average low-income household (a household making $27,000 annually), and then (as in the cost belief solicitation) asked to report how much they expected the policy to cost/benefit a low-income household per annum, in dollars. Finally, participants were asked to report whether they believed that the policy would reduce CO_2 emissions in Nevada. Participants who believed that the policy would reduce emissions were asked to report how much they expected the policy to reduce total state-level emissions by 2030, in percent.

Following the information provision literature, these attribute questions were be performed in an incentive-compatible manner: Before soliciting these beliefs, respondents were told that the 5% of respondents who answered these questions most accurately would be awarded a $10 bonus.

After responding to these questions about Question 6, respondents were shown a hypothetical alternative ballot initiative ("Question 7") that would impose a $25 carbon tax in Nevada and cut the state sales tax by 1.5%. The language of this initiative is modeled after Washington State's 2016 carbon tax ballot initiative (Initiative 732); the full text of this hypothetical initiative can be found in Appendix B.1. Respondents were asked how they would vote on this policy if it were on the ballot instead of Question 6, as well as analogous questions about cost, effectiveness, and regressivity of this hypothetical alternative initiative.

Part 2: Economic and Demographic Information In this section, respondents provided information about their age, income, energy expenditure, and employment. This information was used to tailor the information they receive about the private incidence of these policies.

Part 3: Information Provision. In this section, all participants[6] received information about the cost, effectiveness, and regressivity of both of the policies (the RPS and the carbon tax). Within each attribute-policy pair (e.g. the *cost* of the *carbon tax*), the respondents were randomly shown one of two possible academic information sources.

Follow-up Survey (Starting November 3rd)

Beginning after polls closed on November 3rd (National Election Day in 2020), I opened the follow-up survey. The survey was only displayed to individuals who completed the prior survey, and had two parts:

Part 1: Record Posteriors and Voting Behavior. I recorded respondents' (self-reported) voting behavior on Nevada Question 6, and posterior support for the hypothetical carbon tax alternative ballot initiative,"Question 7." Additionally, I collected posterior beliefs on cost, effectiveness, and regressivity for both initiatives using the same questions outlined above.

Part 2: Record Additional Voter Information. The final stage of the survey involved collecting information that may have 'primed' voters toward certain responses

[6]All respondents received information treatment because the goal of this experiment was to induce randomization in beliefs, not to identify the impact of information provision per se.

had it been collected prior to eliciting beliefs. This information includes political affiliation, voting method, and exposure to advertising.

Figure 2.3 – Survey Design

Figure 2.3: Information provision flowchart. The questions displayed in the RHS of this figure have been abbreviated for ease of exposition. For the full survey, see Appendix A.

2.6. Results

In this section, I outline the results of the information provision experiment. Section 2.6.1 and 2.6.2 cover initial support and initial perceptions of policy attributes, respectively. Sections 2.6.3 and 2.6.4 cover posterior support and posterior beliefs. I present results from my regression models in section 2.6.5 and present counterfactual results in section 2.6.6.

2.6.1. Prior Support

Figures B.2 presents prior support for Question 6 (a 50% RPS) and the hypothetical alternative price-based policy (a $25 carbon tax) after re-weighting the sample to match the demographic characteristics of Nevada's electorate. For the unweighted fractions, see Appendix B.2.

Figure 2.4: Self-reported policy support for Nevada's Ballot Question 6 (a 50% RPS), and a price-based alternative policy (a $25 carbon tax with a 1.5% sales tax cut) among 275 Nevadans. The proportions in this figure reflect re-weighting to account for demographic and ideological differences between the sample and the Nevada electorate.

2.6.2. Prior Beliefs

Figure 2.5 presents respondent beliefs for each attribute (*cost, effectiveness, and regressivity*) of each of the two policies. Reported perceptions about private costs (panel 1) are similar for the two policies, with respondents on average viewing the carbon tax alternative as slightly

more costly. Similarly, respondents reported similar beliefs about incidence on low-income households (panel 3), with carbon taxes viewed as slightly more costly. Respondents did, however, report significant differences in initial views about policy effectiveness: On average, participants expected Question 6 to reduce emissions by 24.2% by 2030, as compared to an expected 13.2% reduction under the carbon tax alternative.

These results constitute significant misperceptions. Table 2.2 presents average beliefs alongside academic estimates. For example, measures of mean absolute error suggest that on average, respondents misperceive annual RPS and carbon tax costs by $250 and $1000, respectively. These initial beliefs also suggest *biased* perceptions of certain attributes. The third column of Table 2.2 suggests that while the misperceptions in certain policy attributes (e.g., the effectiveness of carbon taxes) are relatively symmetric about the 'truth', the misperceptions of other attributes (e.g., the effectiveness of renewable portfolio standards) are asymmetric.

Figure 2.5 – Initial Beliefs about Policy Attributes

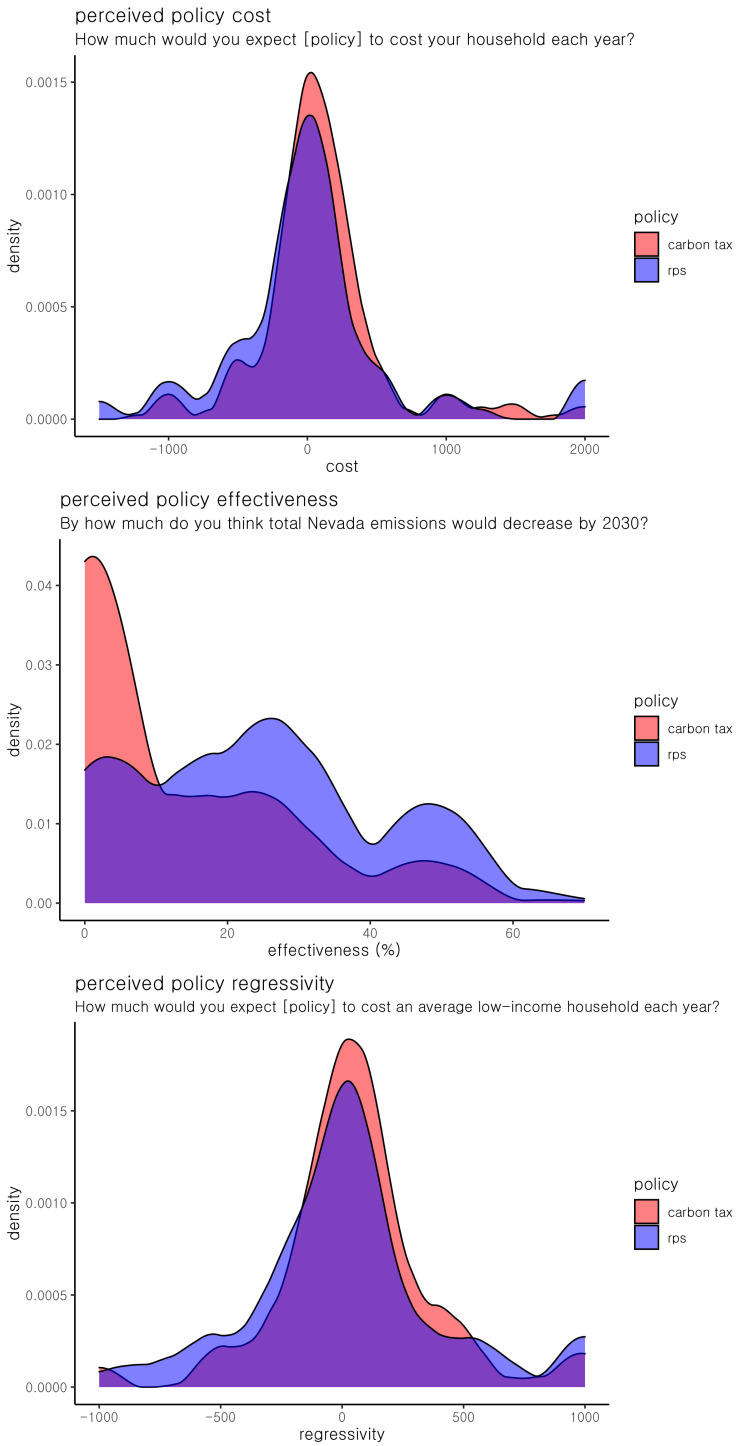


Figure 2.5: This figure displays (unweighted) the initial beliefs among surveyed Nevadans over three policy attributes (private cost, effectiveness at reducing emissions, and regressivity) for each policy option (RPS and carbon tax).

Initial Perception of Policy Attributes				
attribute	mean belief (prior)	academic research estimate	mean error (bias)	mean absolute error
cost, rps ($)	-117.6	126.2	-243.7	951.5
cost, carbon tax ($)	-6.9	999.5	-1006.4	1504.2
effectiveness, rps (%)	24.2	4.5	19.7	21.2
effectiveness, carbon tax (%)	13.2	14.5	-1.3	14.4
regressivity, rps ($)	21.9	334.5	-656.7	1256.7
regressivity, carbon tax ($)	350.3	1007.0	-312.6	806.3

Table 2.2: This table displays summary statistics for initial perceptions of carbon tax and RPS attributes for 275 voters in Nevada. The first column is the mean belief of a given attribute among survey participants. The second column displays the mean estimate of the policy attribute among the two academic sources used in the information provision experiment. Column three displays the mean error (i.e., the bias) among survey participants, and column 4 displays the mean *absolute* error (i.e., the level of misperception).

2.6.3. Posterior Support

Figure 2.6 presents posterior support for Question 6 (a 50% RPS) and the hypothetical alternative price-based policy (a $25 carbon tax), after re-weighting the sample to match the demographic characteristics of Nevada's electorate. See appendix B.2 for unweighted results.

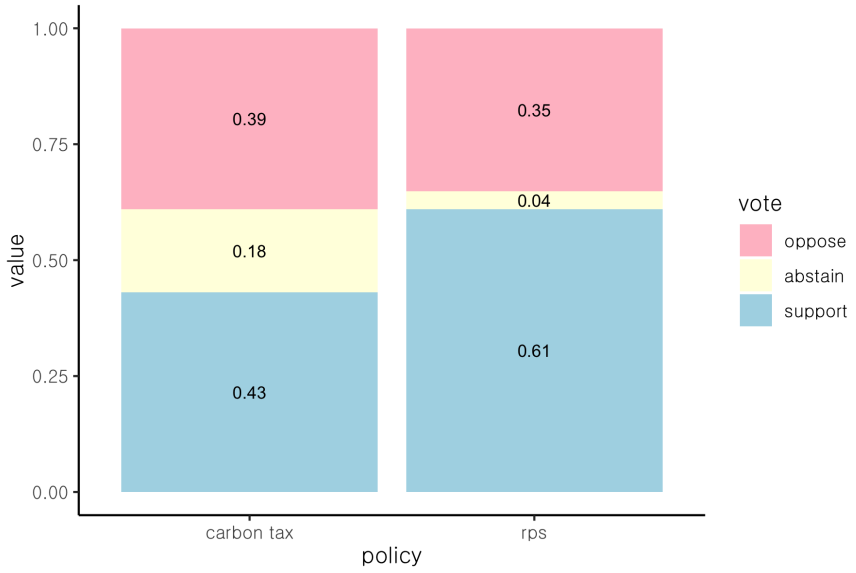


Figure 2.6: Self-reported policy support for Nevada's Ballot Question 6 (a 50% RPS), and a price-based alternative policy (a $25 carbon tax with a 1.5% sales tax cut) among 275 Nevadans, after receiving information treatment about policy attributes. The proportions in this figure reflect re-weighting to account for demographic and ideological differences between the sample and the Nevada electorate.

2.6.4. Posterior Beliefs

Figure 2.7 presents respondent beliefs about each policy recorded during the follow-up survey. As in the prior survey, reported perceptions about both private costs (panel 1) and beliefs about incidence on low-income households (panel 3) are similar across the policies. In contrast to initial beliefs, however, respondents now believe, on average, that Question 6 is *more* expensive than the carbon tax alternative, and that Question 6 places a *higher* burden on low-income households than does the carbon tax alternative. Similarly, responses in the follow-up survey indicate that participants revised their initial beliefs about policy effectiveness to view carbon taxes as more effective, and RPSs as less effective.

Figure 2.7 – Posterior Beliefs about Policy Attributes

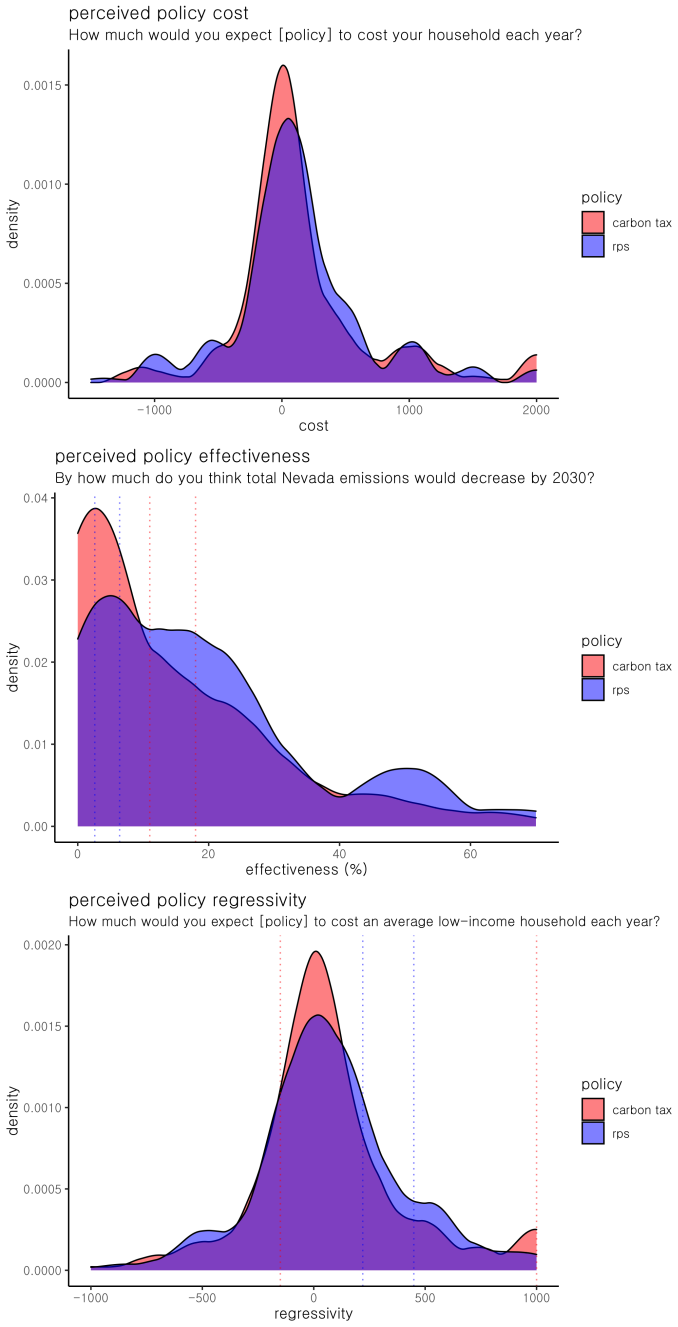


Figure 2.7: This figure displays (unweighted) the beliefs of 275 surveyed Nevadans over three policy attributes (private cost, effectiveness at reducing emissions, and regressivity) for each policy option (RPS and carbon tax), *after* they received information about each of these policy attributes. The red and blue dotted lines represent information provision provided to respondents about carbon taxes and RPSs, respectively. Panel 1 does not display information provision because it was tailored by income level.

2.6.5. Regressions

I now use the variation in beliefs induced by the information provision experiment to estimate models of voter support for corrective policies. Table 2.3 presents results from an ordered logit regression and Table 2.4 presents results from a linear probability model. In Table 2.5, I present an Oaxaca-Blinder Decomposition of the difference in support between carbon tax and RPS policies.

Ordered Logit Model

First, I use an ordered logit model to study voter behavior. I model voter i's support for policy p at time t, $Vote_{i,p,t} \in \{support, abstain, oppose\}$, as a function of the latent utility that would be realized under the passage of a given policy. This latent utility, $u_{i,p,t}$, is in turn a function of voter i's beliefs about the private costs of the policy $c_{i,p,t}$, i's beliefs about the regressivity of the policy, $r_{i,p,t}$, and i's beliefs about the effectiveness of the policy $e_{i,p,t}$. Γ_p, θ_t, and η_i are fixed effects for policy, period, and individual, respectively. By employing an ordered rather than binary logit, I capture the information encoded in abstaining votes. This regression therefore relies on the structural assumption that increasing the latent utility associated with a given policy would lead voters to be more likely so support the policy over abstaining from voting, and more likely abstain from voting over opposing the policy.

$$u_{i,p,t} = \alpha + \beta_1 c_{i,p,t} + \beta_2 r_{i,p,t} + \beta_3 e_{i,p,t} + \Gamma_p + \theta_t + \eta_i + \epsilon_{i,p,t} \tag{2.1}$$

$$Vote_{i,p,t} = \begin{cases} oppose, & \text{if } u < \mu_1 \\ abstain, & \text{if } \mu_1 < u < \mu_2 \\ support, & \mu_2 < u \end{cases}$$

Ordered Logit Model				
variable	coef	se	t	p
cost	−0.022	0.024	−0.918	0.358
effectiveness	0.032	0.006	4.981	0.000
regressivity	−0.011	0.052	−0.211	0.833
periodpre	−0.237	0.150	−1.579	0.114
policyrps	0.591	0.158	3.738	0.000

Table 2.3: Results from a two-period ordered logit regression modeling voter behavior as a function of voter perceptions (Equation 2.1). The dependent variable in this regression is a ternary variable where *support*>*abstain*>*oppose*. *Cost* is the perceived private cost of a given policy, *effectiveness* is the perceived effectiveness of the policy (expected percent reduction in state-level emissions), and *regressivity* is the expected incidence on a low-income household. Fixed effects in this regression include policy (rps vs. carbon tax), period (pre vs. post), and individual. Both perceived *cost* and *regressivity* are measured in thousands of dollars.

Linear Probability Model

I now present results of a linear probability model, where I regress an indicator for voter i's support for a policy ($I(vote\ \ yes_{i,p,t})$) on beliefs about the costs of the policy $c_{i,p,t}$, beliefs about the regressivity of the policy, $r_{i,p,t}$, and beliefs about the effectiveness of the policy $e_{i,p,t}$. As above, Γ_p, θ_t, and η_i are fixed effects for policy, period, and individual, respectively.

$$I(vote\ \ yes_{i,p,t}) = \alpha + \beta_1 c_{i,p,t} + \beta_2 r_{i,p,t} + \beta_3 e_{i,p,t} + \Gamma_p + \theta_t + \eta_i + \epsilon_{i,p,t} \tag{2.2}$$

Linear Probability Model			
Variable	Beta	95% CI	P
cost	-0.0013	-0.0050, 0.0024	0.5
effectiveness	0.0047	0.0028, 0.0067	<0.001
regressivity	-0.0010	-0.0121, 0.0101	0.9
factor(period)			
post	—	—	
pre	-0.0461	-0.0956, 0.0033	0.067
factor(policy)			
carbon tax	—	—	
rps	0.1367	0.0856, 0.1878	<0.001

Table 2.4: Results from a two-period linear probability regression (Equation 2.2). The dependent variable is coded 1 if a voter reports that they support a given policy during a given period, and 0 if they either abstain or oppose. *Cost* is the perceived private cost of a given policy, *effectiveness* is the perceived effectiveness of the policy (expected percent reduction in state-level emissions), and *regressivity* is the expected incidence on a low-income household. Fixed effects in this regression include policy (rps vs. carbon tax), period (pre vs. post), and individual. Both perceived *cost* and *regressivity* are measured in thousands of dollars.

Oaxaca-Blinder Decomposition

Finally, following Oaxaca (1973) and Blinder (1973), I partition the gap in support between RPS policies and carbon taxes into a portion that is explained by differences in beliefs about the attributes of these policies (X, below), and an 'unexplained' portion that results from the differential responsiveness of voters to the same attributes of the two policy types (β, below).

If individual i's support for policy $p \in \{rps, tax\}$ at time $t \in \{pre, post\}$ is a function of perceptions about policy attributes X:

$$I(support_{i,p,t}) = X_{i,p,t}\beta_p + \epsilon_{i,p,t}$$

Then for a given time period, the difference in mean between the two policy types can be written as follows. The results of this exercise (in both levels and percent) are displayed in Table 2.5.

$$\Delta\bar{I}(support) = \underbrace{(\bar{X}_{tax} - \bar{X}_{rps})\beta_{reference}}_{\text{“explained”}} + \underbrace{\bar{X}_{tax}(\beta_{tax} - \beta_{reference}) + \bar{X}_{rps}(\beta_{reference} - \beta_{rps})}_{\text{“unexplained”}} \tag{2.3}$$

Oaxaca-Blinder Decmposition	
difference	value
Levels, explained	-0.068
Levels, unexplained	-0.228
Levels, Total	-0.297
Percent, explained	23.072
Percent, unexplained	76.928
Percent, Total	100.000

Table 2.5: Results from a twofold Oaxaca-Blinder decomposition of the gap in prior support for carbon taxes vs. renewable portfolio standards. The model used for this decomposition is a two-period linear probability model with individual and time fixed effects. *Explained* components represent the share of the gap in support between the two policy types that can be attributed to differences in “endowments” (differences in initial beliefs about policy attributes), conditional on the estimated model of voter behavior. *Unexplained* components represent the portion of the gap that cannot be attributed to differences in perceptions of attributes. The unexplained share results from differences in responses to perceived attributes between policy types (differences in “slope”), and the interaction between the endowment and slope effects.

2.6.6. Counterfactuals

Armed with the models of voter behavior estimated in section 2.6.5, I estimate voter support under three counterfactual scenarios: *perfectly informed voters*, *targeted redistribution of carbon tax revenues*, and *targeted information provision*. For each counterfactual, I predict vote shares using an ordered logit model. Broadly speaking, the goal of these counterfactual exercises is to speak to the degree of mutability in policy support.

Counterfactual 1: Perfectly Informed Voters. In this scenario, I replace voters prior beliefs of each of the policy attributes with the 'true' policy attributes,[7] and estimate the vote shares (*support, abstain, oppose*) for each of the two policies. The results from this exercise are displayed in Figure 2.8.

Figure 2.8 — Perfectly Informed Voters

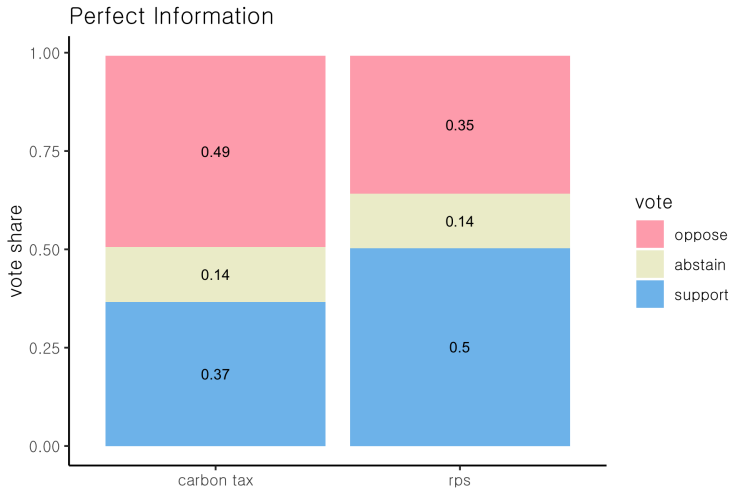


Figure 2.8: Counterfactual vote shares under perfectly informed voters. Shares reflect vote probabilities generated by applying an ordered logit model (model X) to data where initial voter beliefs have been replaced with beliefs that accord with (average) conclusions academic research about the cost, effectiveness, and regressivity of each policy.

Counterfactual 2: Targeted redistribution of carbon tax revenues. In this scenario, I model support for a carbon tax policy where rather than returning revenue to voters through a sales tax, revenue is used to minimize the incidence (maximize the transfers) to marginal ("swing") voters. Note that the transfers used in this section are likely infeasible:

[7] It is worth noting that participants may have private information about the expected costs of these policies, specifically carbon pricing. Given the average magnitude of cost misperceptions, however, the bias engendered by private information is likely second-order.

Sallee (2019) details the difficulty in targeting compensation to ameliorate tax burdens from corrective policies given the information set and policy levers available to regulators. As such, the results from exercise can be viewed as bounds on the extent to which compensation to marginal voters can be expected to impact support for carbon taxation.

Figure 2.9 — Rebates Targeted at Swing Voters

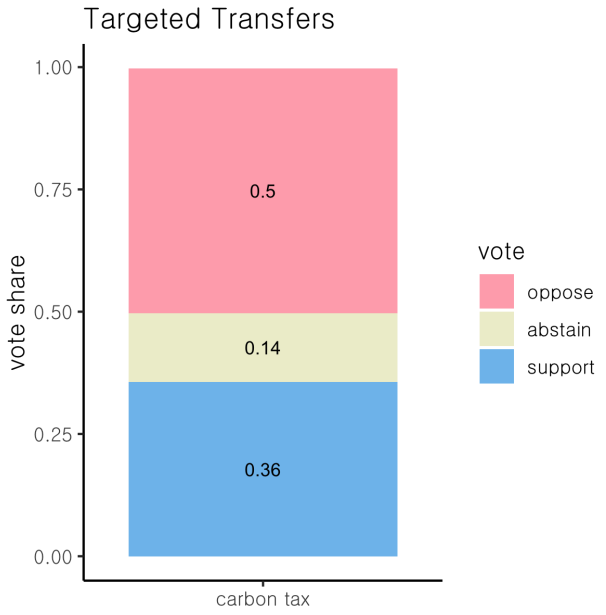


Figure 2.9: Counterfactual vote shares under targeted rebates.

Counterfactual 3: Targeted information provision. In this counterfactual, I imagine providing information on policy *effectiveness* alone to respondents who either abstained from voting or opposed carbon taxes in either period. This counterfactual takes advantage of the fact that respondents have biased perceptions about carbon tax cost and regressivity that bolster support.

Figure 2.10 — Targeted Information Provision

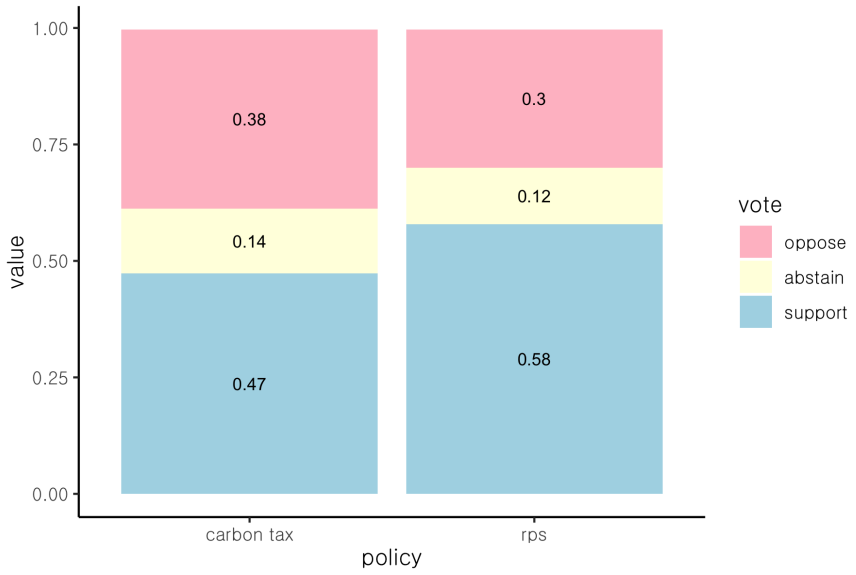


Figure 2.10: Counterfactual vote shares under targeted information.

2.7. Discussion

Results from the initial survey confirm that respondents prefer the performance-based policy (RPS) to the price-based policy (carbon tax), and suggest significant misperceptions about the attributes of these policies. On average, respondents misperceive the private costs by 211% in absolute terms, costs to low-income households by 286%, and the effectiveness of each policy by 183%. Unweighted survey results imply a 29 percentage point advantage for RPSs before information provision; weighted results imply an 18 percentage point advantage.

After providing source-randomized information provision, respondents updated their priors to view carbon taxes more favorably, and RPSs less favorably. That is, relative to reported beliefs in the initial survey, respondents in the follow-up survey reported believing RPS (carbon taxes) to be more (less) costly, more (less) regressive, and less (more) effective.

Panel fixed-effects regression models using the variation in beliefs induced by the information provision suggests that respondents who updated their priors to view a given policy as more effective, less costly, or less regressive were more likely to support said policy. Notably, the estimated coefficients on *cost* and *regressivity* are small, with 95% confidence interval generally including zero. The largest elasticities admitted by these confidence intervals are quite modest. For example, point estimates from Equation 2.2 imply that decreasing the average voter's perceived cost of a corrective policy by $1000 annually (four times the mean cost across policies) would increase the probability that the average voter supports said pol-

icy by just 1.3 percent. Conversely, I recover large elasticities of policy support with respect to perceived policy effectiveness. The coefficient in Table 2.4, for example, suggests that a voter who updates her prior from believing that a given policy would reduce emissions by 15% in a decade to believing that a given policy would reduce emissions by 25% in a decade would be 5% more likely support said policy. Importantly, this magnitude of change in perception is not uncommon in my sample. Pooling across the two polices, the average change in perceived effectiveness was roughly 9 percentage points.

The decomposition analysis presented in Table 2.5 suggests that a meaningful portion (23%) of the unweighted difference in support between the carbon tax and RPS policies can be attributed to differences in perceptions of policy attributes. On its face, this result would suggest that it may be possible to generate majority support for carbon taxes by correcting misperceptions, especially if respondents tended to hold initial beliefs about carbon taxes that were biased toward pessimism. Counterfactual exercises, however, suggest the opposite type of bias: to the extent that misperceptions of policy attributes explain the gap in support for these two policies, the misperceptions are *asymmetric* and *optimistic* toward Nevada's RPS. For example, counterfactual estimates in Figure 2.8 imply that replacing voter perceptions with 'true' policy attributes would reduce support RPSs by roughly 10%, while increasing support for carbon taxes by just 4%.

The relatively low elasticities of support with respect to perceived policy cost and perceived regressivity also imply that targeted transfers will have limited effectiveness in bolstering carbon tax support. Figure 2.10 illustrates that even a policy implausibly-well designed to minimize the incidence to swing voters may fail to achieve majority support.

Approaching the Oaxaca-Blinder decomposition results from a glass-half-empty is also informative for the understanding of voter beliefs about corrective policies. Conditional on observable characteristics, voters are 13.8% (22.8% in raw terms) more likely to vote for RPSs than they are taxes — I call this unexplained residual 'tax aversion,' as it captures the dispreference for a tax-based policy relative to a non-tax alternative, all else equal.[8]

[8]It should be noted that 'all else equal' assumes that the attributes that I include in this paper cover the attributes relevant to Nevadan voters.

2.8. Conclusion

I use an information provision experiment conducted around a vote on Nevada's renewable portfolio standard (RPS) to study voter preferences for public policies. Before summarizing the findings of this paper, I stress that these conclusions are drawn from a relatively small sample of Nevada voters. Some of the findings presented in this paper are statistically imprecise, and the precisely-estimated results may not hold in different political or economic environments.

With these caveats in mind, I present four conclusions. First, using incentive-compatible elicitation of beliefs, I demonstrate significant inaccuracy and bias in the perceptions of the cost, effectiveness, and regressivity of carbon taxes and renewable portfolio standards. For example, on average, respondents overestimated the emissions reductions resulting from RPS policies by a factor of five. Second, among those surveyed, I recover small elasticities of policy support with respect to perceived costs and perceived regressivity, and relatively large elasticities of policy support with respect to perceived policy effectiveness. Third, I find that differences in perceptions of policy attributes explain roughly a quarter of the gap in support for carbon taxes and RPS policies, implying tax aversion—the dispreference for taxes holding other policy attributes fixed—may play a significant role in the political barriers facing price-based policies in this setting. Returning to the research questions outlined in the introduction, this decomposition trivially suggests that misperceptions alone cannot explain the entire policy support gap. To the extent that misperceptions of policy attributes *can* explain differences in support for these two policies, the explained gap results mainly from respondents' optimistic views of RPS policies rather than pessimistic views about carbon tax attributes. Fourth, and finally, I predict counterfactual estimates of vote shares to shed light on the practical mutability of the observed shares of policy support. These exercises suggest that even implausibly well-targeted transfer schemes using carbon tax revenue are unlikely to significantly increase support for these policies.

Regarding the design of policies, the results from this experiment suggest that sophisticated policy design is unlikely to generate significant political gains for carbon taxation. Instead, my results highlight the relative importance of increasing public confidence in market-based policies. While the efficacy of price instruments as methods to address externalities is consensus among economists and energy policy sphere, this notion is not reflected in public perceptions. The results in this paper suggest that this lack of communication may significantly hinder the political success of price-based policy instruments.

Chapter 3

The Congestion Costs of Uber and Lyft

3.1. Introduction

Transportation network companies (TNC) like Uber and Lyft have grown rapidly over the past decade to become integral parts of urban transportation systems. A small but growing literature has attributed to these companies benefits that include billions in annual consumer surplus (Cohen, Hahn, Hall, Levitt, and Metcalfe, 2016), reductions in drunk driving (Greenwood and Wattal, 2015), and flexible work (Cramer and Krueger, 2016; Angrist, Caldwell, and Hall, 2017).

The costs of TNC expansion, however, have yet to receive commensurate treatment in the economics literature. Most notably, TNCs have been accused of contributing to traffic congestion (San Francisco Transit Authority, 2018; Schaller Consulting, 2018), but existing studies of the impact of TNCs on congestion are few, arrive at varied conclusions, and do not quantify the implied congestion costs (Li, Hong, and Zhang, 2019; Erhardt, Roy, Cooper, Sana, Chen, and Castiglione, 2019). Back of the envelope calculations suggest these costs could be substantial. A 2017 Inrix report, for example, placed the annual cost of congestion to US drivers at $305 billion (Inrix, 2017)—roughly two orders of magnitude larger than estimates of national consumer surplus provided by Uber (Cohen et al., 2016). This suggests that if TNCs have even a modest impact on traffic congestion, the negative externalities associated with lengthening travel times could offset consumer surplus benefits. Understanding how and whether TNCs impact traffic congestion therefore plays a crucial role in determining appropriate policy response to the continued growth of these companies.

Two identification problems, however, make causal inference difficult when studying the relationship between TNC activity and traffic congestion. First, Uber and Lyft likely select

entry locations based on trends in city-level characteristics unobservable to the econometrician. Comparisons that leverage differences in TNC entry dates across locations may therefore suffer from reverse causality. Second, within-city time series regressions may be biased by omitted variables (e.g., gentrification) which are serially correlated with TNC activity and also impact congestion.

In this paper I leverage a natural experiment in Austin, TX to circumvent these identification challenges: On May 9th, 2016, both Uber and Lyft unexpectedly exited Austin following a vote that upheld a city ordinance requiring driver background checks. I combine this variation in TNC activity with novel and granular Bluetooth traffic speed data, and setting-specific estimates of the value of travel time to answer two research questions. First, do transportation network companies impact traffic congestion? And if so, what are the travel-time related costs or benefits of TNC operation?

This setting informs two empirical strategies: a difference in differences comparing pre- versus post-May 9th traffic speeds in 2015 (where both companies operated year round) to 2016 (where both companies exit on May 9th), and a regression discontinuity in time. Across specifications, I find evidence of modest increases in traffic speeds following the exit of Uber and Lyft. Difference in differences results suggest that across all hours, traffic speeds increased roughly 1% following the exit of Uber and Lyft. 7 am. to 7 p.m. traffic speeds increased by 2.3%, with the largest TNC-related slowdowns occurring during the middle of the day (11 a.m. to 2 p.m.). Using setting-specific estimates of value of the travel time, I calculate that Austinites would be willing to pay roughly $33 to $52 million annually to avoid these slowdowns. Back of the envelope calculations suggest that these figures are a small fraction (4-6%) of total Austin-area congestion costs, and are roughly the size of estimates of the consumer surplus associated with TNC operation in Austin.

These findings improve on the existing literature in three ways. First, this is to my knowledge the only paper to use the exit of Uber and Lyft to study the impacts of TNCs on congestion. This translates to weaker identifying assumptions than those imposed in analyses leveraging the staggered expansion of these companies. Second, I extend existing analyses by mapping changes in travel speeds to changes in travel time costs, providing the first estimates of the congestion costs associated with TNC activity. And third, the spatial and temporal granularity in the Bluetooth data allows me to perform analyses that contribute to a more complete picture of the heterogeneous impacts of TNC activity on traffic congestion.

These findings also provide several important takeaways for policymakers. First, TNC activity can be viewed roughly as a transfer, as the consumer surplus enjoyed by TNC passengers is of similar size to the time loss incident on incumbent drivers. Second, it is difficult to rationalize TNC quantity restrictions purely on welfare grounds, as the lost consumer surplus may outweigh travel time gains. In other words, even if TNC regulation is

more politically achievable than are price-based congestion controls, TNC regulation appears (at least in the Austin case) to be a poor tool to address congestion-related externalities. Relatedly, the relatively modest impacts of TNCs on traffic congestion in Austin suggest that congestion taxes targeted specifically at ridesharing companies are unlikely to result in large traffic-related welfare gains. Lastly, the fact that speeds slow in response to TNC activity suggests TNCs add vehicles miles traveled (VMT) to the transit system. In other words, the VMT avoided by sharing rides are outweighed by additional trips induced by the availability of TNCs.

The rest of this paper is organized as follows. Section 3.2 describes related literature and background. Section 3.3 details the events that precipitated the departure of Uber and Lyft from Austin. Section 3.4 outlines the data sources. I describe my empirical strategy and threats to identification in Section 3.5, and present results in Section 3.6. Section 3.7 concludes.

3.2. Background and Related Literature

Traffic congestion is a significant urban disamenity. It is costly (Inrix, 2018), it is associated with lower self-reported happiness (Anderson, Lu, Zhang, Yang, and Qin, 2016), and it comes with considerable co-costs in terms of noise and pollution (Currie and Walker, 2011b). Although a tax is the canonical policy prescription for congestion (Vickrey, 1969), both theory and empirics suggest that because targeting individual contributions to congestion is difficult, realistic congestion pricing instruments (e.g., cordon charges) may fall well short of the welfare gains achievable by a hypothetical first best policy (Knittel and Sandler, 2018; Prud'Homme and Bocarejo, 2005). This, coupled with the potential political advantage of TNC regulation over comprehensive congestion taxation suggests that understanding the sign and magnitude of TNC related time costs or savings will be important for informing city-level policy. Indeed, several cities have already moved to regulate TNCs in the name of congestion. New York City, for example, cited congestion as a motivation for its 2018 ridesharing cap (New York Times, 2018). As of 2020, San Francisco, New York, and Chicago have all imposed "congestion fees," levied on TNC trips in the city center (New York Times, 2019). Outside of the US, cities like London and Vancouver have weighed congestion impacts as they deliberate over TNC policy (Reuters, 2019; Vancouver Sun, 2019).

As a number of other observers have noted, however, the impact of TNCs on traffic speeds is theoretically ambiguous. While survey data from Rayle, Shaheen, Chan, Dai, and Cervero (2014) and Clewlow and Mishra (2017) suggest TNCs induce trips, and Mangrum and Molnar (2018) demonstrate that taxis—the closest analog to TNCs—increase congestion on the margin, Cramer and Krueger (2016) show that in five of six US cities, Uber drivers

spend a significantly higher fraction of their time with a passenger in their vehicle than do taxi drivers. This ride-sharing effect could attenuate or outweigh the effect of induced trips. There may also be complementarities between TNCs and public transit: Hall, Palsson, and Price (2018) use a difference in differences design on measures from the National Transit Database to conclude that Uber is indeed a complement to public transportation. It is unclear, though, whether complementarity between TNCs and public transit will result in more or fewer vehicle trips.

To date there exists little econometric work on whether TNCs cause traffic congestion, and existing results arrive at varied conclusions. Li, Hong, and Zhang (2019), for example, use city-level congestion measures and differences in Uber's entry date to estimate the company's impact on congestion, concluding that Uber improves city-level congestion measures. Erhardt et al. (2019), on the other hand, use 2010 and 2016 Inrix traffic data and scraped measures of Uber activity to calibrate a traffic engineering model of San Francisco. They conclude that ridesharing companies were responsible for significant (30%) increases in vehicle hours traveled. In addition to the fact that these studies reach contradicting conclusions, the identification concerns outlined in the introduction suggest value in reassessing this question using a natural experiment.

3.3. Natural Experiment

Austin, TX, is the 11th largest incorporated place in the United States and suffers from considerable congestion: According to Inrix, Austin ranked 14th nationally and 72nd globally in the number of average hours lost to congestion per driver. Cities with similar levels of per-driver congestion costs include San Diego, Berlin, and Manchester. Both Uber and Lyft began operating in Austin in 2014.

In December 2015, the Austin City Council passed Ordinance No. 20151217-075, which imposed a series of regulations on TNCs, including data requirements, restrictions on idling locations, and most controversially, fingerprinting requirements to facilitate driver background checks (The City Council of Austin, 2015). Proposition 1, sponsored by Uber and Lyft, attempted to overturn this ordinance. On May 7th, 2016, the Proposition was defeated in a citywide vote, with 56% of voters casting against (The Texas Tribune, 2016). In protest, Uber and Lyft exited the Austin market on May 9th (New York Times, 2016). 13 months later, Uber and Lyft re-entered Austin as Governor Greg Abbott signed into law HB 100, which overturned Austin's local ordinance (The 85th Texas Legislature, 2017). This variation in TNC activity provides the basis for my empirical identification.

During the yearlong absence of Uber and Lyft, Austin was not without ridesharing. A number of smaller TNCs entered the market or expanded their Austin presence following

the defeat of Proposition 1. In date of their arrival in Austin, these companies are: GetMe (December 2015), Fare (Mid-May 2016), Fasten (June 1st, 2016), Tride Technologies (June 15th, 2015), and RideAustin (June 16th, 2016). Wingz, which provides rides to and from the airport, also started operating in Austin in May of 2016. A survey of Austin commuters conducted in November 2016 by Hampshire, Simek, Fabusuyi, Di, and Chen (2017) offers a view of take up of these alternative rideshare companies. RideAustin held the largest market share (47.4%), followed by Fasten (34.5%), Fare (12.9%), GetMe (2.8%), Wingz (1.6%), and Tride (0.4%). Informed by the Hampshire et al. (2018) survey and the universe of RideAustin's 2016 trip-level data, I am able to infer the level of total TNC activity in Austin following the exit of Uber and Lyft. I can therefore identify a window following the Proposition 1 vote where alternative TNC activity is negligible (see Section 3.5.1).

3.4. Data

I use data collected from an array of Bluetooth sensors along major roadways (both highway and surface-level) operated by the Austin Department of Transportation. Located inside traffic signal cabinets, these sensors detect unpaired Bluetooth devices (e.g., smartphones, car systems) and estimate traffic speeds based on the movement of single devices (which are given unique anonymous identifiers) through the network of sensors.

I use an aggregated version of this dataset prepared by Post Oak Traffic Systems, which isolates device movements through specific road segments (henceforth *segments*), which are short sections along just one road. This company pre-processes the data in several ways. Data are aggregated at 15 minute bins and represent the average speed across the segment for devices that appear at the origin reader first, and then the destination reader, and do not appear at any other sensors in the interim. These data are also filtered for outliers: only observations that fall within 75% of the IQR of the previous 15 observations are used in calculating speeds. This type of filtering is applied to combat bias from the movement of non-vehicle Bluetooth devices (like those carried by pedestrians) through the sensor network.

In addition to the data cleaning performed by Post Oak Traffic Systems, I further restrict my sample to consistently reporting sensors. Of the 430 total segments, I drop segments that report in fewer than 70% of days during each year (2015 and 2016) of the study period, leaving me with a panel of 79 segments. For robustness I also report results using a) all segments that report in more than 30% of study period days and b) only segments that report during 100% of study period days.

The 79 segments I use in my preferred specification are plotted in Figure 3.1 and summarized in Table 3.1. The mean segment length is 0.72 miles, with minimum and maximum lengths of 0.06 and 3.8 miles, respectively. As shown in Figure 1, my sample covers a range

of road types. The smallest roads in my sample are two-lane roads, the largest are 7-lane roads, and the median segment is a 5-lane road. I observe 966,301 15-minute speed reports during my study period. On average, a segment sees 4.77 devices move from origin to destination during each 15 minute period, meaning that my data summarize roughly 4.6 million segment traverses. The average travel speed is 2.99 minutes per mile, which corresponds to 20.06 miles per hour. This figure is consistent with periods of significant congestion.

My variable of interest is minutes per mile, which has two advantages over miles per hour. First, a change of one mile per hour does not represent a constant damage over the domain of this variable: In terms of time lost, changing from 5 to 4 miles per hour is roughly 20 times as costly as changing from 20 to 19 miles per hour. Second, multiplying outcomes in minutes per mile by estimates of the value of time is a straightforward way to arrive at cost calcuations from changes in traffic delays.

While novel and granular, the Bluetooth data bring challenges for estimation. First, in the raw data available on the Austin Open Data Portal, 61 of the 79 segments used in my analysis show the segment length changing over the course of the study period. While most of these adjustments are minor, and personal correspondence with Austin Transportation Department employees suggests that these adjustments likely reflect updated length measurements and not relocation of Bluetooth sensors, I nonetheless investigate the possibility that these segment length changes constitute a threat to identification in Appendix C.5. I use the updated length measurements for all speed calculations in all time periods. A second challenge is the possibility of Bluetooth sensors measuring the movement of pedestrians. If filtering does not eliminate all measurement error originating from Bluetooth devices used by Austinites walking or biking, and the use of these modes of transit is correlated with the period where Uber and Lyft exited Austin, the empirical strategies I describe below will arrive at biased estimates. I further investigate this in Section 3.5.4.

I compile several other datasets to augment my analysis. To control for weather-related shocks, I use precipitation and temperature data accessed through the National Oceanographic and Atmospheric Administration's National Centers for Environmental Information. To isolate a period of time where the impact of other TNCs is minimal, I use RideAustin's trip-level data. These data range from June 2nd, 2016 to April 13th, 2017, and are publicly available online (RideAustin, 2017). Lastly, I use two datasets to arrive at setting-specific value of time estimates. The first is the National Household Travel Survey (NHTS), which contains information on income and commuting habits. The second is a toll price and travel time dataset from the MoPac variable price freeway in Austin. These data were provided courtesy of the Central Texas Regional Mobility Authority and are further detailed in Appendix C.1.

3.5. Empirical Strategy

3.5.1. Timeframe

I use Bluetooth traffic data from 2015 and 2016 to study the relationship between TNCs and congestion. I truncate this window to isolate periods where the variation in traffic speeds can be credibly attributed to the failure of Proposition 1. As described in Section 3.3, a number of TNCs entered the market following the exit of Uber and Lyft. Estimations using the entire yearlong suspension period as a comparison would therefore underestimate any changes relative to a TNC-free counterfactual. Informed by the universe of trips from RideAustin—the TNC with the largest market share during Uber and Lyft's absence—I truncate my estimation period on August 1st, 2016. Similarly, Austin hosts the South by Southwest Music Festival (SXSW) each March. I restrict my analysis to exclude the 2015 and 2016 festivals. This leaves me with data from March 20th to August 1st for both 2015 and 2016. The 2016 study period is plotted with TNC data in Figure 3.2. Note that although the Austin Bluetooth data extend through 2019, significant portions of the spring are missing data from years 2017, 2018, and 2019, including Uber and Lyft's re-entry in May of 2017. While this rules out difference in differences specifications using later years, I am able to use data from different parts of 2017-2019 to perform placebo regression discontinuity estimates (see Appendix C.4).

3.5.2. Difference in Differences

To study the effect of the exit of Uber and Lyft on travel times, I compare traffic speeds pre and post May 9th in 2016 (where Uber and Lyft exited) to 2015 (where both companies operated year-round). To capture heterogeneity in the congestion impacts across time of day, I perform this comparison within each hour of day, h (or equivalently, interacting each right-hand side term below with an hour of day dummy):

$$s_{i,y,t} = \alpha + \beta_h \delta_y \eta_t + \gamma_1 \delta_y + \gamma_2 \eta_t + \boldsymbol{\gamma_3} \delta_y \boldsymbol{\theta_i} \cdot t + \boldsymbol{\gamma_4} \boldsymbol{\theta_i} + \boldsymbol{\Gamma} \boldsymbol{X}_{y,t} + \epsilon_{i,y,t} \tag{3.1}$$

Where $s_{i,y,t}$ is the speed (in minutes per mile) measured over segment i on day t of year y. δ_y is a dummy that equals one for the year 2016, and η_t is a dummy that equals one for days (in any year) after May 9th. $\boldsymbol{\theta_i}$ is a set of dummies for each road segment, and t is the signed number of days between a given date and May 9th of that year. $\boldsymbol{X_t}$ is a vector of controls that includes day of week fixed effects, holiday fixed effects, 10 10-degree daily temperature bins, and 10 daily precipitation level bins. The interacton betwen $\delta_y \eta_t$ is the treatment indicator, as it takes a value of 1 for observations after May 9th, 2016, and zero otherwise. $\delta_y \boldsymbol{\theta_i} \cdot t$ are segment-year specific linear time trends.

The identifying assumption in the estimation of β_h—the effect of Uber and Lyft operation on travel speeds during a given hour of day h—is that conditional on seasonality and weather, the difference in travel speeds between 2016 and 2015 at hour h does not change after May 9$^{\text{th}}$ for reasons other than the operation of Uber and Lyft.

I calculate hour-specific congestion impacts with the goal of producing more accurate cost estimates. As I show in Appendix C.1, variable-toll data suggest that the value of travel time in Austin varies significantly from hour to hour. Similarly, the number of vehicles on the road peaks during rush hours. Together, this information suggests that the same change in traffic speeds could produce different aggregate congestion costs at different times of day. By matching hour-specific estimates of the impact of TNCs to hour-specific vehicle miles traveled (VMT) and hour-specific estimates of the value of travel time, my cost calculations account for temporal heterogeneity that pooled estimates may not reflect. To determine whether the convolution between hourly congestion impacts and hourly VOT is a first-order consideration, I also estimate a model pooling across hours of day. This estimator is Equation 3.1, but run without interacting hour of day fixed effects with the right hand side variables. The rationale for this regression is to simulate what estimation and inference might look like using temporally aggregated data.

To investigate spatial heterogeneity, I estimate a model pooling over hours of day and allowing an idosyncratic treatment effect for each road segment. This model is equivalent to Equation 3.1, but interacts the set of segment dummies with the treatment indicator, $\delta_y\eta_t$. $\boldsymbol{\beta}$ is now a 1x79 vector of segment-specific treatment effect estimates. Note that in this pooled Equation hour of day fixed effects are included in $\boldsymbol{X}_{y,t}$.

$$s_{i,y,t} = \alpha + \boldsymbol{\beta}\delta_y\eta_t\boldsymbol{\theta}_i + \gamma_1\delta_y + \gamma_2\eta_t + \boldsymbol{\gamma}_3\delta_y\boldsymbol{\theta}_i \cdot t + \boldsymbol{\gamma}_4\boldsymbol{\theta}_i + \boldsymbol{\Gamma}\boldsymbol{X}_{y,t} + \epsilon_{i,y,t} \tag{3.2}$$

3.5.3. Regression Discontinuity

Lastly, I estimate a regression discontinuity model, again estimating hour-specific treatment effects (β_h) by interacting each term in the regression Equation with a set of hour of day fixed effects.

$$s_{i,t} = \alpha + \beta_h\eta_t + \boldsymbol{\gamma_1}\boldsymbol{\theta}_i + \boldsymbol{\gamma_2}\boldsymbol{\theta}_i \cdot t + \gamma_2\boldsymbol{\theta_i} \cdot t^2 + \boldsymbol{\gamma_4}\boldsymbol{\theta_i} + \boldsymbol{\Gamma}\boldsymbol{X_t} + \epsilon_{i,t} \tag{3.3}$$

The identifying assumption for β_h is that conditional on weather, potential outcomes (traffic speeds) in hour of day h are continuous about May 9$^{\text{th}}$, 2016. While the identifying assumption for the RD is arguably weaker than that of the difference in differences estimator, the RD will produce estimates of the short-term response to the exit of Uber and Lyft. As such, I rely on the difference in difference estimator to produce my preferred annual congestion

cost figures.

3.5.4. Threats to Identification

Threat 1: Contemporaneous Shocks. The identifying assumptions in both the RD and DID estimates rely on the absence of *year*post*-specific shocks to Austin area travel speeds. The end of the University of Texas, Austin (UT) school year, for example, presents a potential threat to identification if university-related traffic activity differed substantially between 2015 and 2016. In Appendix C.4, I use placebo exit dates to determine whether or not shocks that create regression discontinuity estimates on the order of my reported coefficients are empirically common. Figure C.2 displays coefficient estimates using the actual exit date in relation to the distribution of coefficients from 134 regression discontinuities using placebo exit dates, 13 of which were chosen to line up with the beginning/end of a UT semester. 5 of the 134 placebo coefficients (4%) are more negative than the estimates using the actual TNC exit date, zero of which correspond to the start/end of a UT semester. This placebo test therefore suggests that shocks that produce RD estimates on the order of my estimates are empirically uncommon, and that my results are not likely a result of changing traffic patterns related to activity at UT.

Threat 2: Other Modes of Transportation. If the exit of Uber and Lyft led Austinites to substitute toward walking or biking and these trips were not dropped as outliers during data processing, β will not be identified. In other words, for other modes of transportation to bias my estimates, traffic speed must be mismeasured, and that mismeasurement must be correlated with the treatment.

Data on mode shares and mode speeds suggest that this type of bias cannot alone account for my results. Hampshire et al. (2017) suggest 1.8% of TNC users switched to bikes following Uber's exit. If TNCs made up 10% of Austin trips, and bikes constituted 1.53% (United States Census Bureau, 2015), this mode shifting represents an 11.8% increase in total bike trip volume. The average car in my sample took 2.99 minutes to traverse a mile—3.01 minutes per mile fewer than the 6 minutes per mile (10 mph) assumed by Google biking directions. These figures imply that for changes in bike shares to alone account for a change of 0.1 minutes per mile (roughly the average treatment effect across daytime hours), bikes would need to constitute roughly 28% of observed Bluetooth samples *after* dropping extreme travel time outliers. This figure is inconsistent with the travel speeds implied by the movement of Bluetooth devices, which greatly exceed 10 miles per hour on average.

Nonetheless, I draw on a second traffic speed dataset to empirically examine this concern. In addtion to Bluetooth sensors, the city of Austin also maintains pneumatic sensors that take

periodic measurements of traffic speeds. While these measurements are not frequent enough to act as a replacement dependent variable, they do allow me to study the relationship between Bluetooth speed measurements and true traffic speeds by matching segments to pneumatic sensors.

While we should not expect pneumatic sensors to match segment speeds exactly (segments often include intersections), if there is significant switching to non-vehicular modes of transport that biases the Bluetooth speed measurements, this would be reflected in a change in the relationship between the two measurements. For example, say we have a segment-sensor pair, and prior to May 9th, 2016, when the pneumatic sensor reports a speed of 25 mph, the Bluetooth segment on average reports a speed of 20 mph. If there is bias from mode-switching, we would expect this relationship to change in the post period. Now, when the pneumatic sensor again registers 25 mph, the incresed number of non-filtered pedestrian datapoints biases the segment measurement downward, to, say, 18 mph.

To operationalize this anecdote, I match segments to pneumatic sensors, and run a regression of segment speeds on sensor speeds, allowing for a differential slope term interacted with a post May 9th 2016 dummy. If I find a statistically (and economically) significant difference in slopes, I treat this as evidence of mode choice related bias. This exercise is detailed in Appendix C.2. I match 39 Bluetooth segments to pneumatic road sensors. In a simple regression with month of year and road segment fixed effects, I find little evidence to support pedestrian-induced bias in my estimates. As shown in Table C.1, the coefficient on the interaction between the post dummy and the pneumatic segment speed is not statistically different from zero, nor is it of meaningful magnitude.

Threat 3: TNC Driving Speeds. If TNC vehicles drive significantly slower or faster than the average non-TNC vehicle in a way that remains after filtering, the above estimates of β_h will be biased. During congested conditions it is unlikely that this should occur: if congestion slows all drivers, then travel time measurements from any subset of vehicles should be representative of average speeds. At free-flow traffic speeds, however, it is possible that TNCs drive faster (due to profit motive) or slower (idling to find riders) than non-TNC vehicles.

To test these concerns, I use public trip-level data from the startup RideAustin, which entered the market following the departure of Uber and Lyft. Following Mangrum and Molnar (2018), who construct "taxi races" to test whether different types of taxi travel at different speeds, I match RideAustin trips to Bluetooth segments, allowing me to test the null hypothesis that TNC vehicles drive at the same speeds as the average mix of vehicles.

This exercise is detailed in Appendix C.3. Over 221 trip-segment matches, I find that on average RideAustin vehicles traveled 0.03 minutes per mile slower while traversing a given

segment than did the average device during the same time period. This difference is not statistically significant, nor should it meaningfully bias my results. Assuming TNCs account for 10% of vehicle trips, for example, this difference in speeds implies a bias on the order of 0.003 minutes per mile—one to two orders of magnitude smaller than my estimates of the impact of TNCs on traffic speeds. To the extent that speed differences do generate bias, they will lead me to overstate improvements in traffic speeds resulting from a TNC ban.

3.6. Results and Discussion

3.6.1. Traffic Speeds

Across multiple specifications, I find evidence of modest increases in traffic speeds following the exit of Uber and Lyft. Results from my preferred specification (Equation 3.1) are displayed in Table 3.2 and Figure 3.3. Point estimates of changes in minutes per mile are largely negative, suggesting reduced congestion after the exit of Uber and Lyft. While the 95% confidence intervals for hour-specific estimates of changes in travel times generally include zero, an F-test rejects the null hypothesis of $\beta_h = 0 \ \forall h$ ($p < 0.0001$). Although TNCs appear to negatively impact morning rush hour conditions, I estimate little change in evening rush hour speeds. The largest improvements in travel times following TNC exit come, surprisingly, between 11 a.m. and 2 p.m. Point estimates for off-peak hours (8 p.m. to 6 a.m.) are small and straddle zero. This pattern could be a result of TNCs comprising a higher share of vehicles during the middle of the day than during peak hours. Additionally, evening rush hour effects could be muted if TNC users are more likely to share cars during the evening than they are during the morning and early afternoon.

Figure 3.6 and Table 3.3 display results from Equation 3.3, a regression discontinuity by hour of day. These figures are qualitatively similar to, but larger in magnitude than the difference in differences results, suggesting that the short-term impacts of TNC exit may be more pronounced than the medium-term impacts. Figure 3.7 plots residuals from a pooled regression discontinuity performed on daytime traffic speeds in 2016 (when Uber and Lyft exited) and 2015 (where both companies operated year-round). The 2015 regresion discontinuity estimates a null effect, offering evidence that the 2016 regression discontinuity results are not driven by seasonal changes in traffic patterns.

Table 3.4 shows the results from running versions of Equations 3.1 and 3.3, pooling across hours. The pooled difference in differences results suggest that on average, speeds increase by 0.026 minutes per mile ($p = 0.15$), or roughly 0.9% following TNC exit. Consistent with the hour-specific estimates, restricting the pooled DID analysis to daytime hours (7 a.m. to 7 p.m.) generates larger estimates of speed increases following TNC exit ($\beta = -0.068$,

$p = 0.2$). This coefficient translates to a 2.3% increase in daytime traffic speeds. Figure 3.4 displays the raw speed data for daytime traffic by week of year for my study window, and provides evidence of the absence of pre-trends. Figure 3.5 plots an event study version of Equation 3.1, where separate treatment effects are estimated for each week. Consistent with the growth of RideAustin and other Austin-area TNC alternatives through the second half of 2016, the event study shows the treatment effect decaying over time: 2016 traffic speeds are significantly lower than those in 2015 for 10 weeks following the exit of Uber and Lyft, but by week 17, point estimates suggest that traffic speeds had returned to the baseline 2015-2016 difference.

As in the hour-specific estimates, the pooled RD estimates are larger in magnitude than are the DID results: Travel times decreased by 0.102 minutes per mile ($p = 0.01$) across all hours and by 0.134 ($p = 0.003$) minutes per mile for daytime hours. These coefficients correspond to travel time reductions of 3.4% and 4.5%, respectively.

As noted in Section 3.3, a number of ridesharing firms entered the market after the exit of Uber and Lyft. If alternative TNC activity was substantial during the study period, my results will be attenuated relative to the counterfactual of a TNC-free Austin. Although RideAustin's data provide some insight into the level of alternative TNC activity in Austin, it is unclear whether the growth of RideAustin in 2016 is representative of the growth of all alternative TNCs, or whether RideAustin grew by cutting into the market share of firms like Fasten and Fare, which arrived earlier. In Table 3.5, I report estimates of the impact of TNCs on traffic speeds in Austin under each of these two possible trajectories of TNC activity in 2016: In rows 1 and 3 (*RideAustin Data*), I assume 10,200 TNC trips per day during the pre-period (see Uber (2015)), and use RideAustin's time-series data—inflated by the reciprocal of its market share—to produce a time-varying measure of TNC activity following the failure of Proposition 1. This time series is plotted in red in Figure 3.2. Under this assumption, neither the fuzzy DID nor the fuzzy RD specification differs significantly from the results in Table 3.4. In rows 2 and 4 (*Hampshire Data*), I again use 10,200 TNC trips per day for the pre-exit figure, but then assume that a constant 4,180 (41% of 10,200) TNC trips per day during are completed during the treatment period. This figure reflects results from a November 2016 survey conducted by Hampshire et al. (2017), where 41% of survey respondents reported that they completed an Uber or Lyft reference trip using another TNC following the failure of Proposition 1. This trajectory assumes that RideAsutin's growth is not representative of the alternative TNC market, and instead entirely reflects RideAustin winning over customers from other already-established Uber and Lyft alternatives.

Intuitively, the results using the *Hampshire Data* assumption are roughly 1.7 times larger than the estimates from Table 3.4. This offers a useful bound for this exercise investigating attenuation. If the RideAustin data is even partially representative of the growth of alterna-

tive ridesharing companies in Austin, then TNC activity in May-August of 2016 was lower than the 41% replacement reported by Hampshire et al. (2018) in November. I therefore view row 4—which implies a 7.6% reduction in travel speeds when moving from zero TNC activity to full TNC activity—as an upper bound for the congestion impacts of TNCs in Austin.

My estimates of changes in traffic speeds together with data on the number of total TNC and non-TNC vehicle trips in Austin allow me to estimate the implied elasticity of congestion with respect to TNC volumes. According to the 2017 NHTS, Austin-area households take an average of 3.6 vehicle trips a day. Austin's 370,00 households, then, generate roughly 1.35 million vehicle trips per day. The available data on Uber and Lyft suggest that the two services together completed roughly 10,200 trips per day prior to their 2016 exit from Austin (see Figure 3.2). Multiplying this figure by a factor of two to reflect the capacity factor estimated by Cramer and Krueger (2016) suggests that Uber and Lyft together accounted for 1.5% of Austin-area vehicle trips prior to the failure of Proposition 1. My pooled estimates of the impact of TNCs on Austin-area congestion therefore imply congestion elasticities with respect to TNC volume of between 0.6 and 2.3. My preferred specification, which suggests a 2.3% increase in daytime traffic speeds following the exit of Uber and Lyft, corresponds to an elasticity 1.5. These estimates lie on the lower end of the range of congestion elasticities from existing studies. Anderson et al. (2016), for example, estimate a congestion elasticity of 2.7 in Beijing; findings from Leape (2006) imply an elasticity of 2.5 in London, and results from Eliasson (2009) imply an elasticity of 1.5 in Stockholm. The relatively small congestion elasticity implied by my estimates may reflect the lower levels of congestion in Austin relative to cities like London and Beijing, or the offsetting effect of trips saved by the 'ridesharing effect' of TNCs.

In Appendix C.4, I investigate the robustness of the results presented in this section. Figure C.1 plots both the difference in differences and regression discontinuity results for bandwidths ranging from 20 to 70 days around May 9th. The conclusion that daytime traffic speeds increase following the exit of Uber and Lyft holds across bandwidth choices. To test the likelihood that the regression discontinuity estimates presented above are the result of a contemporaneous shock to Austin-area traffic speeds, I compare my estimates to coefficients from 134 regression discontinuities using placebo exit dates. The results of this exercise are shown in Figure C.2. 5 of the 134 placebo coefficients (4%) fall below the estimate using the true exit date, suggesting that it is empirically unlikely that my RD estimates are the result of an unobserved Austin-area transit shock.

3.6.2. Heterogeneity and External Validity

Results from Equation 3.2, which allows for segment-specific congestion responses, are plotted in Figures 3.9 and 3.10. Two themes emerge. First, there is no clear spatial pattern in congestion impacts: I estimate negative and positive travel time impacts both for segments in the city center and for outlying roads. Second, Figure 3.9 shows that segments that experienced exceptionally large changes in traffic speed were characterized by exceptionally high levels of pre-period traffic congestion, suggesting construction or other segment-specific shocks may explain these estimates. Absent these outliers, the segment-specific effects exhibit relatively low variance. An important avenue for future work would be to investigate whether these outliers indeed represent extreme congestion reductions from TNC operation in select locations, or whether they can be explained by data absent from this setting.

The external validity of the results presented in this paper hinges on whether Austin is representative of other metropolitan areas in terms of commuter preferences and the substitutability of transit options. To determine which cities have transit systems that resemble Austin's, Figure 3.11 depicts how public transit use and personal vehicle travel vary across the 20 largest metro areas in the US. Commuters in the majority of large American cities (especially those located in the "Sun Belt") exhibit mode choices similar to those in Austin, where commuters heavily favor solo commutes in personal vehicles. In cities with extensive public transit systems (e.g. New York, Washington, San Francisco), however, commuter choices are quite different than they are in Austin. This suggests caution when applying the results described in this paper to address policy questions in these metro areas.

3.6.3. Equilibrium response

It is worth discussing the extent to which the brief disruption of TNC activity in Austin provides insights into the equilibrium differences in congestion levels between a city with and a city without TNCs. In addition to the first-order changes in traffic flow caused by the absence of TNCs, the full equilibrium response to the exit of TNCs would reflect a combination of short and long-run adjustments made by road users. More specifically, road users will spatially re-optimize in response to differential speed changes, and city residents may change long run by vehicle purchase or sorting decisions.

Because spatial re-optimization over route choices likely occurs in the short-run, and I use a large sample of segments covering different types of roadways, my estimates likely reflect this spatial substitution. My results do not, however, reflect long-term adjustments: The above analysis compares traffic speeds in Austin with and without TNCs, holding fixed decisions on car purchasing behavior and locational sorting. Some of the reduction in congestion that I measure likely comes from individuals who, prior to 2016, chose not to purchase a

vehicle because they had access to ridesharing. In the long run, the actions of these marginal car owners would erode the traffic improvements that resulted from the exit of Uber and Lyft.

3.6.4. Congestion Costs

Armed with estimates of hour-specific changes in travel times, I calculate the external congestion cost associated with TNC operation as follows:

$$\Delta\text{congestion cost} = \sum_h \Delta \text{ minutes per mile}_h * \text{miles driven}_h * \text{value of time}_h \tag{3.4}$$

Δ*minutes per mile* are the coefficients, by hour of day, h, estimated above. To estimate *miles driven*$_h$, I use periodic traffic counts to estimate the share of VMT by hour of day in Austin, and multiply these shares by estimates of daily VMT provided by the Texas Department of Transportation. Note that this operation assumes that my estimates represent an average effect for all VMT within Austin City limits. Table 3.6 provides evidence in support of this assumption: According to data maintaned by the Texas Department of Transportation, roads included in my preferred specification resemble those not included in terms of congestion, VMT, and speed. Finally, I calculate Austin-specific hourly value of travel time (VOT) estimates (*value of time*$_h$, above) using data from Austin's MoPac freeway (see Appendix C.1). I also present results using a VOT heuristic from Small and Verhoef (2007): 50% of the wage rate. I estimate the wages of Austin-area drivers by calculating the per-worker income for car-commuting Austin households in the 2017 National Household Travel Survey (NHTS). Standard errors for all congestion cost estimates follow formulas developed by Goodman (1960).

I report the results of this exercise in Table 3.7. Both rows reflect changes in travel times estimated in my preferred specification (Equation 3.1), a DID across years. Using time-varying (MoPac) and uniform (NHTS) VOT estimates, I calculate the daily congestion costs associated with Uber and Lyft activity at \$92,071 and \$127,983, respectively. These estimates correspond to annual costs of \$33 million (p=0.181) and \$46 million (p=0.049). Disaggregating this sum by weekend and weekday effects produces slightly larger figures annual cost figures: \$39 million (p=0.041) and \$52 million (p=0.003). In Appendix C.4, I report results from this exercise using a regression where observations are weighted by the number of Bluetooth devices recorded in each 15-minute window. The rationale for this specification is to investigate whether the above results are biased when segments with differing traffic flows are implicitly given the same weight in determining regression coefficients. The cost estimates from this weighted regression are similar to the non-weighted results and imply an annual congestion cost associated with Austin-area TNC activity of \$54 million. Note

that each of these aggregate cost measures relies on the assumption that VOT is uniform across the city. While misattributing VOT estimates by location may bias these estimates, there are two reasons why this bias is likely small: First, in a recent investigation of the heterogeneity of urban VOT, Buchholz, Doval, Kastl, Matějka, and Salz (2020) find that the majority of the variation in VOT is across individuals rather than across locations within a city. Second, the lack of a spatial gradient in the segment-level congestion estimates (see Figure 3.10) means that a cost calculation using a modest VOT gradient between the city and the suburbs (as in Figure 8 of Buchholz et al. (2020)) would produce similar aggregate welfare measures as those reported above in Tables 3.4 and 3.7, so long as the VOT estimates I use reflect the rough spatial average of the Austin metro area.

Several outside studies provide valuable context when interpreting these congestion cost estimates. First, according to the Inrix Global Scorecard, the aggregate 2017 travel time cost in Austin was $2.8 billion, $810 million of which was attributed to traffic slowdowns (Inrix, 2017). Back of the envelope calculations using my estimates therefore suggest that Uber and Lyft together accounted for 1-2% of all travel time costs in Austin, and 4-6% of congestion costs. Second, estimates of consumer surplus associated with TNCs offer a useful benchmark for policymakers weighing the benefits of TNC operation against the costs. The results from Cohen et al. (2016) allow me to produce two back of the envelope estimates of consumer surplus in Austin. First, Cohen et al. (2016) conclude that in the four cities they examine, $1.57 of consumer surplus are generated for every dollar spent on TNCs. Uber reported that in 2015, its drivers grossed $27 million in the Austin area (Uber, 2015). According to Uber's S1 filing, 83% of the payments to the Uber app went to drivers in 2016, suggesting that Austin consumers pay roughly $32.5 million to Uber annually. Inflating this figure by the multiplier from Cohen et al. (2016), and assuming an equal expenditure-consumer surplus ratio between Lyft and Uber implies a total TNC-related consumer surplus for the city of Austin of $72.9 million annually. As a second way of estimating consumer surplus in Austin, I rescale the Cohen et al. (2016) estimate of national consumer surplus from Uber ($6.8 billion). Multiplying this figure by Austin's share of the US urban population (0.49%) yields an estimate of the 2015 Austin-area consumer surplus of $33 million annually. Again inflating this figure to account for Lyft's market share suggests a TNC-related consumer surplus for the city of Austin of $47 million. These exercises therefore suggest that the congestion costs resulting from Uber and Lyft activity in Austin are similar in magnitude to the consumer surplus benefits provided by these companies.

3.7. Conclusion

Using a natural experiment in Austin, TX, I study whether transportation network companies impact traffic speeds. I estimate that TNCs are responsible for a 2.3% increase in Austin-area travel times between 7 a.m. and 7 p.m. This figure masks important heterogeneity, with the largest TNC-related slowdowns occurring between 11 a.m. and 2 p.m. By matching setting-specific changes in traffic speeds to hour-specific estimates of the value of travel time, I find that Austinites would be willing to pay $33 to $52 million annually to avoid the slowdowns induced by TNC activity. Back of the envelope calculations using estimates of TNC consumer surplus from Cohen et al. (2016) suggest that the cost of TNC-related congestion in Austin is similar in magnitude to the consumer surplus generated by these companies.

These results have important implications for urban transportation policy. While a comprehensive congestion tax is the most efficient response to congestion externalities, charging all road users the social marginal cost of their actions is both technologically and politically challenging (King, Manville, and Shoup, 2007). Given these difficulties, a natural question is whether policies that target a related good (TNCs) would improve welfare, and if so, how such a policy would perform relative to the first-best.

My results suggest that quantity restrictions—like those imposed by New York in 2016—are unlikely to produce substantial welfare benefits, as the congestion benefits from restricting TNC activity are roughly offset by lost consumer surplus. Estimates of the congestion costs associated with TNC activity also speak to the efficacy of TNC taxation as a means of addressing traffic externalities. Models of imperfect Pigouvian taxation show that the welfare gains of a second-best tax relative to a first-best tax are a function of how well second best taxes target heterogeneous externalities, as well as the correlation between demand elasticities and idiosyncratic externalities (Knittel and Sandler (2018) and Diamond (1973)). Taxing ridesharing, then, will perform well as a second-best congestion pricing scheme if a) ridesharing activity is well-correlated with congestion, and b) if demand for ridesharing is elastic relative to the demand for other congesting trips. The results from this natural experiment suggest that TNC activity is not highly correlated with urban congestion externalities: back of the envelope calculations suggest TNCs are responsible for a small fraction of the congestion costs in Austin, and TNC-related externalities often occur at times where the value of travel time is relatively low. Taken together with results from Cohen et al. (2016), who characterize TNC trips as inelastic, uniform TNC taxation is unlikely to produce welfare gains that approach those realized under a first-best congestion price. For policymakers politically constrained to target congestion taxes at TNCs, the significant temporal and spatial variation in congestion impacts estimated in this paper suggests that there may be

substantial improvements in policy efficiency from including spatial and temporal variation in pricing. TNC congestion fees levied in New York and San Francisco, for example, may be made more efficient by adding variation to the tax rate by time of day or origin/destination zone.

These results also pose several questions that may inform future research. First, identifying the drivers of the spatial and temporal heterogeneity in the congestion impacts of TNCs would be useful to city policymakers attempting to better target TNC-based congestion policies. Second, replicating this type of analysis in other settings with similar identification opportunities would provide a valuable test of external validity, especially in cities with markedly different commuting and public transit landscapes. Finally, the fact that speeds slow in response to TNC activity suggests TNCs add vehicles to the road. In other words, my results suggest that the ride-induction effect dominates the ride-sharing effect. This conclusion will be important to test in other cities, as the impact of TNCs on VMT is an important uncertainty in the prediction of transportation sector emissions.

3.8. Figures

FIGURE 3.1—BLUETOOTH SEGMENT LOCATIONS IN AUSTIN, TX

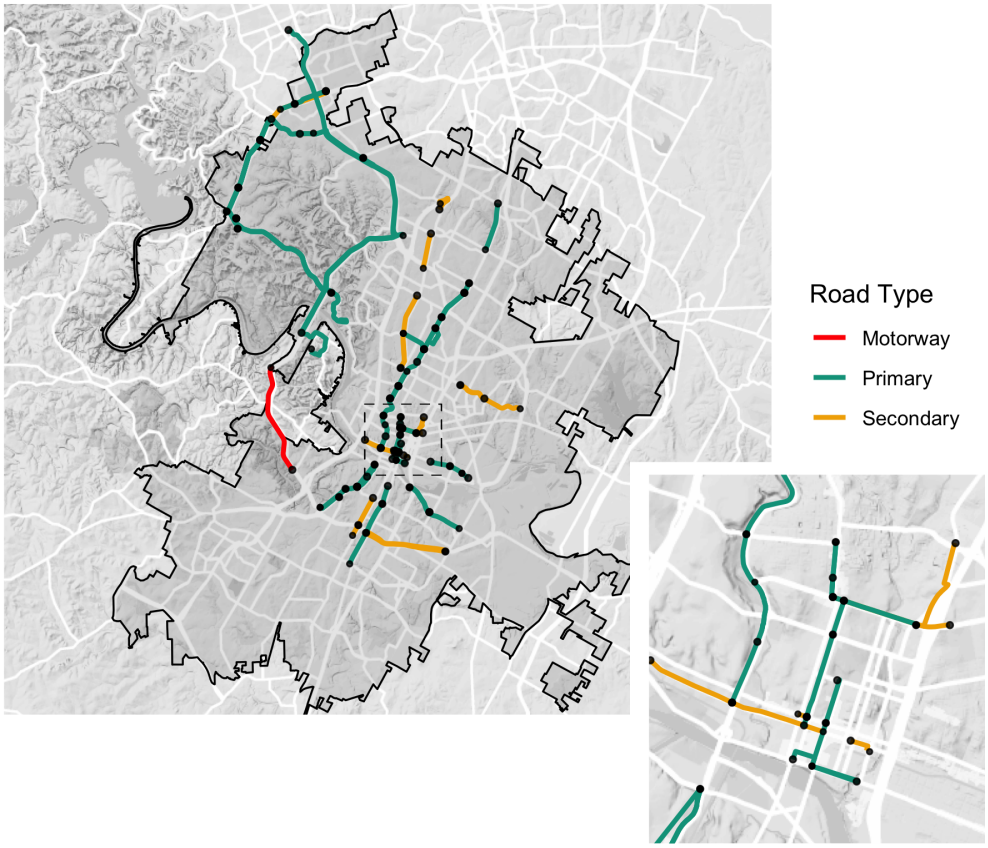


Figure 3.1: *Notes:* Nodes represent terminal Bluetooth sensor locations for each of the 79 segments used in my analysis. Note that some sensors act as both origin and destination readers for different segments. Paths represent Google Maps recommended driving directions between endpoints of a given segment, colored by Open Street Map road type. Motorways are major divided highways, primary roads are large multi-lane roads that may or may not be divided. Secondary roads are typically two- to four-lane surface streets. The black line is the Austin city limit.

Figure 3.2—Variation in TNC Activity

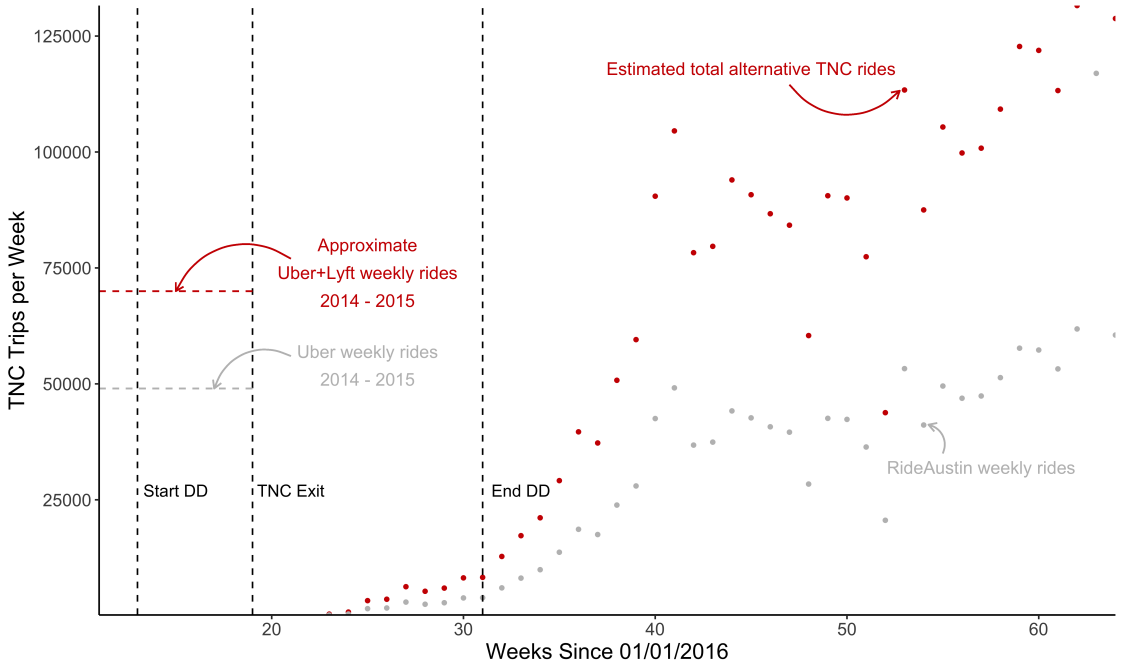


Figure 3.2: *Notes:* This figure displays the variation in ridesharing activity I use to identify the impact of TNCs on congestion. From left to right, the vertical lines represent the start of the 2016 difference in differences period (March 20th), the failure of Proposition 1 (May 9th), and the end of the 2016 difference in difference period (August 1st). The grey dotted line is the average number of Uber trips per week (as per an Uber Report on 2014-2015 operations). The red dotted line represents an estimate of total Uber and Lyft pre-exit activity, assuming a 30% Lyft market share. Note that because both Uber and Lyft entered Austin in 2014, the actual number of Uber trips in early May 2015 was likely much larger than 70,000 per week. The grey dots plot weekly RideAustin activity for the first 9 months of the company's operation. The red dots inflate the RideAustin data by the reciprocal of its November 2016 market share (47%) to provide an estimate for the total level of post Proposition 1 alternative TNC activity.

FIGURE 3.3—DIFFERENCE IN DIFFERENCES RESULTS

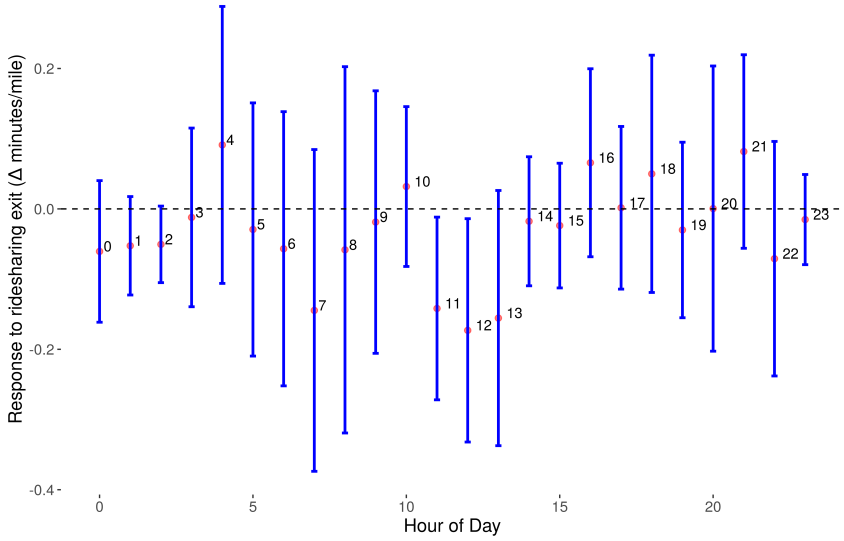


Figure 3.3: *Notes:* Results from Equation 3.1, a difference in differences comparing pre vs. post May 9th traffic speeds in 2015 (where both Uber and Lyft operated in Austin) to pre vs. post May 9th traffic speeds in 2016 (where both TNCs exited Austin). Points represent the estimated effect of TNC departure on traffic speeds (in minutes per mile) by hour of day. Controls include day of week, holiday, and segment fixed effects, segment-specific linear trends in days since May 9th, and flexible controls for temperature and precipitation. Bars reflect 95% confidence intervals from two-way standard errors clustered by segment-week. Traffic speed data were accessed through the City of Austin's Open Data Portal.

FIGURE 3.4—PARALLEL TRENDS

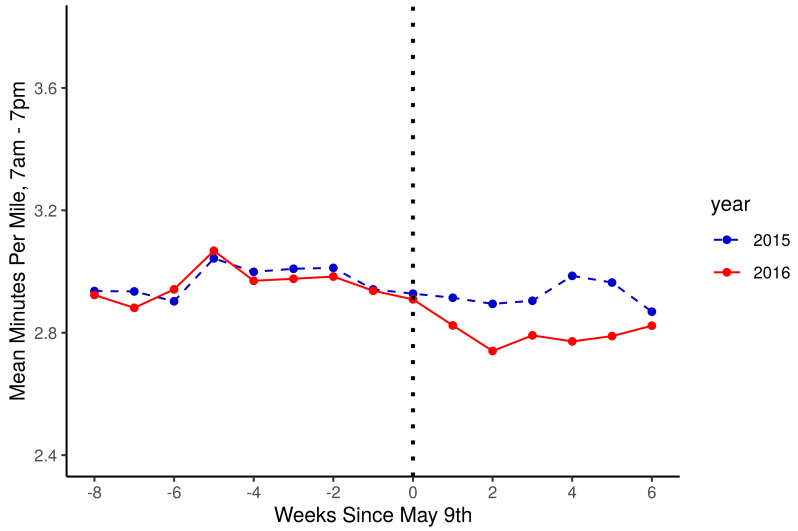


Figure 3.4: *Notes:* This figure shows raw average speed (minutes per mile) between 7 a.m. and 7 p.m. over 79 road segments in Austin, TX, plotted by week of year for 2015 and 2016. Data were accessed through the City of Austin's Open Data Portal. The dotted line represents the week of May 9$^{\text{th}}$, where Uber and Lyft ceased operation in Austin in 2016. Note that week zero is partially treated, as May 9$^{\text{th}}$, 2016 was a Monday.

FIGURE 3.5—EVENT STUDY

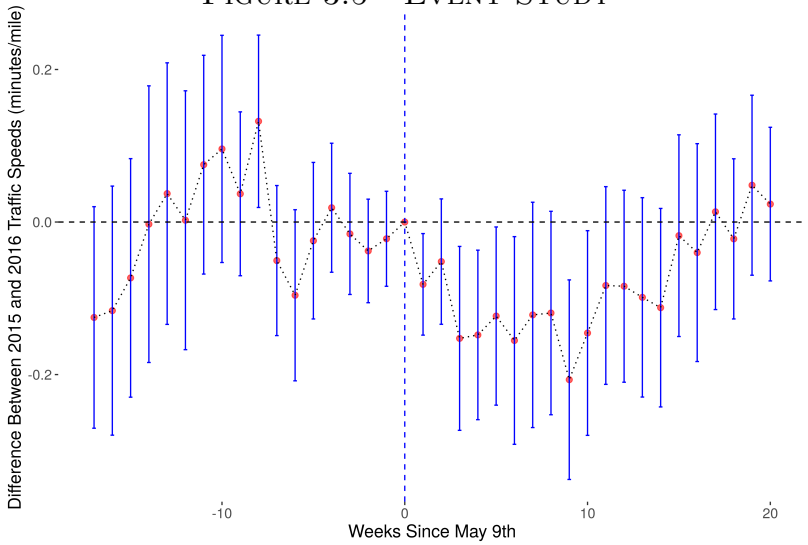


Figure 3.5: *Notes:* This figure plots results from a difference in differences regression with separate coefficients for 17 pre and 20 post-exit weeks. Points represent the estimated difference between 2015 and 2016 traffic speeds (in minutes per mile) relative to the difference in the week leading up to May 9$^{\text{th}}$. Controls include day of week, holiday, SXSW, and segment fixed effects, as well as flexible controls for temperature and precipitation. Bars reflect 95% confidence intervals from standard errors clustered by segment.

FIGURE 3.6—REGRESSION DISCONTINUITY RESULTS

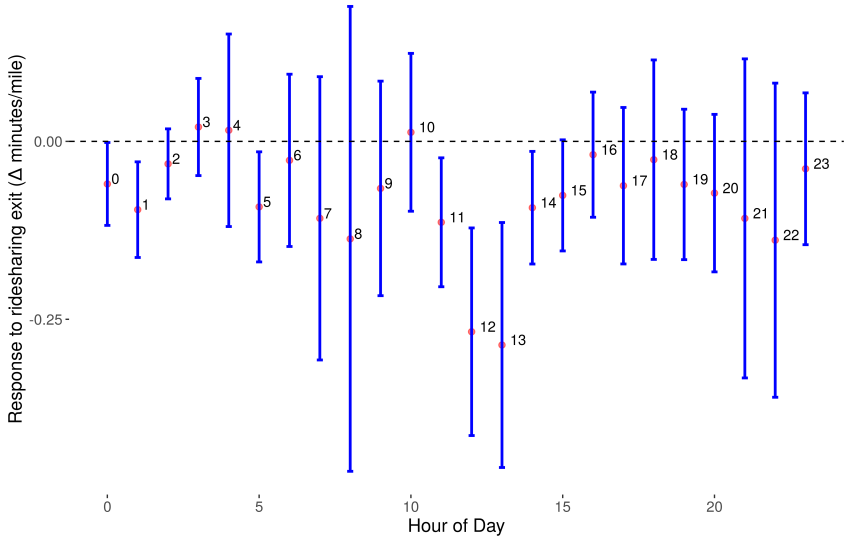


Figure 3.6: *Notes:* Results from Equation 3.3, a regression discontinuity performed on traffic speeds across 79 road segments in Austin, TX. The bandwidth is March 20th - August 1st of 2016, which (asymmetrically) spans the May 9th departure of Uber and Lyft. Points represent the estimated effect of TNC exit on traffic speeds by hour of day. A negative point indicates an estimated increase in traffic speed. Controls include day of week, holiday, and segment fixed effects, segment-specific second degree polynomials in days since May 9th, and flexible controls for temperature and precipitation. Bars reflect 95% confidence intervals from two-way standard errors clustered by segment-week. Traffic speed data were accessed through the City of Austin's Open Data Portal.

FIGURE 3.7—REGRESSION DISCONTINUITY RESIDUAL PLOTS

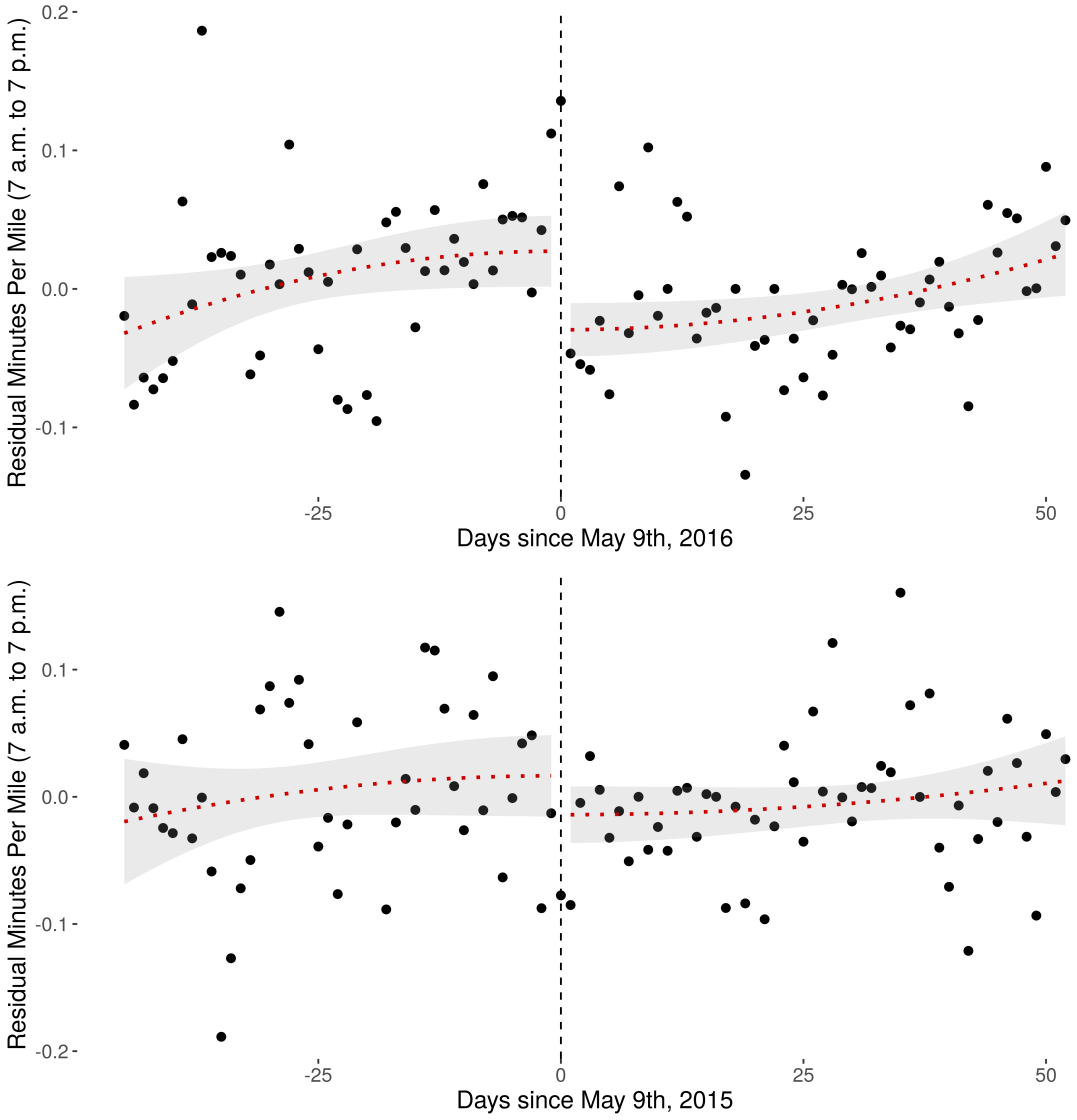


Figure 3.7: *Notes:* This figure plots the daily mean residuals from a pooled version of Equation 3.3, omitting the $\delta_y \eta_t$ (year*post) indicator. The dependent variable is minutes per mile, measured over 79 road segments in Austin, TX. The bandwidth is March 20$^{\text{th}}$ - August 1$^{\text{st}}$ of 2016 (or 2015), which (asymmetrically) spans the May 9$^{\text{th}}$ departure of Uber and Lyft. Controls include hour of day, day of week, holiday, and segment fixed effects, segment-specific second degree polynomials in days since May 9$^{\text{th}}$, and flexible controls for precipitation and temperature. The dotted line represents a second degree polynomial in days since May 9$^{\text{th}}$; the shaded region is the 95% confidence interval. Traffic speed data were accessed through the City of Austin's Open Data Portal.

Figure 3.8—Weekday vs. Weekend Effects

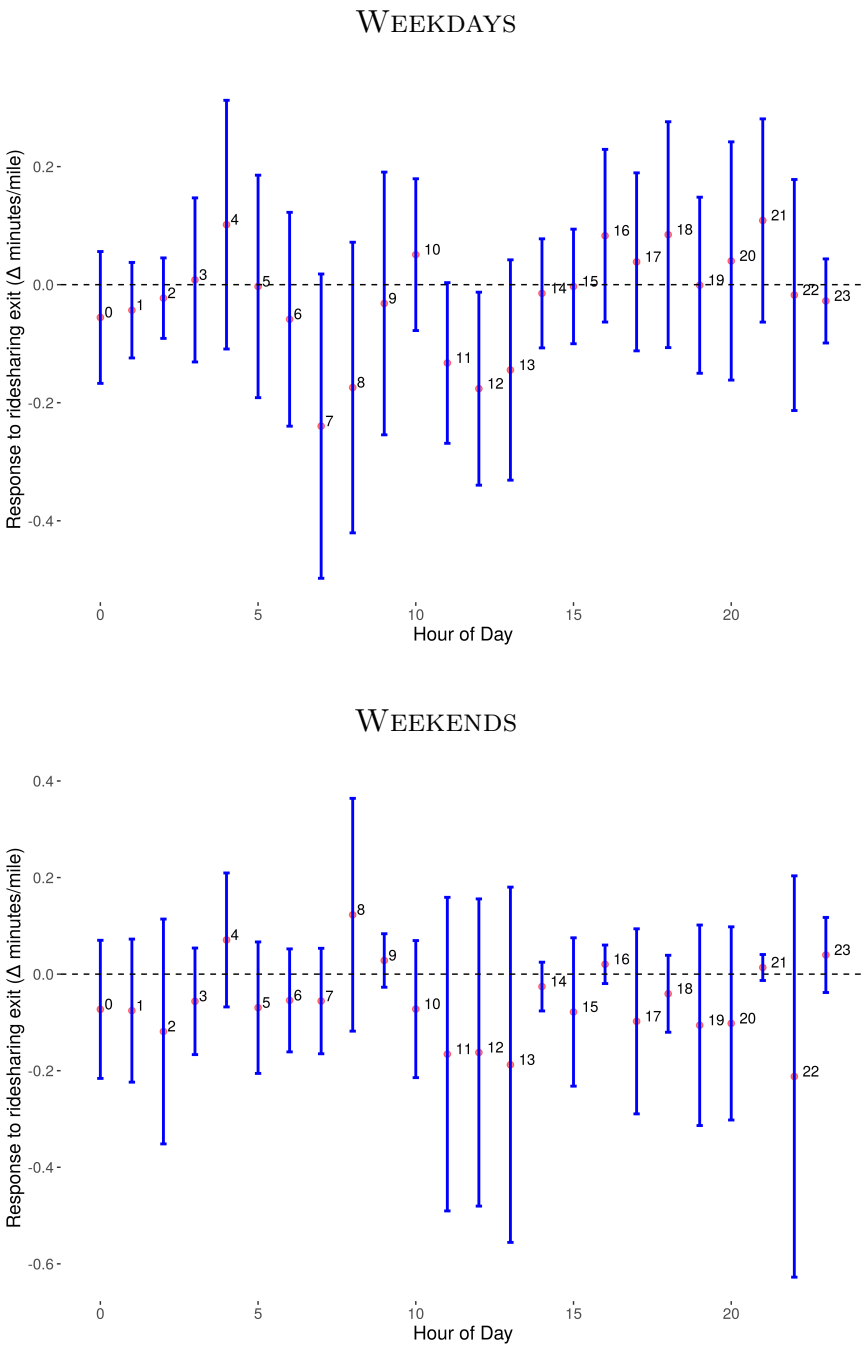


Figure 3.8: *Notes:* This figure plots difference in differences estimates (Equation 3.1) of the impact of Uber and Lyft's exit on travel speeds separately for weekdays and weekends. Points represent the estimated effect of TNC departure on traffic speeds (in minutes per mile) by hour of day. Controls include day of week, holiday, and segment fixed effects, segment-specific linear trends in days since May 9^{th}, and flexible controls for temperature and precipitation. Bars reflect 95% confidence intervals from two-way standard errors clustered by segment-week. Traffic speed data were accessed through the City of Austin's Open Data Portal.

Figure 3.9—Distribution of Segment-Specific Responses

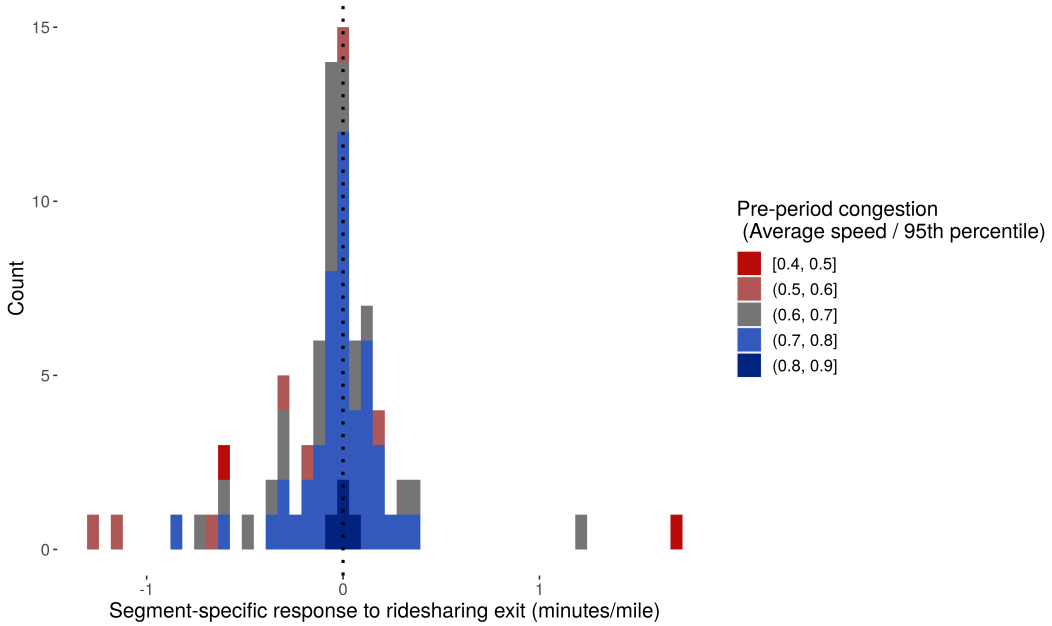


Figure 3.9: *Notes:* Results from Equation 3.2, a difference in differences comparing pre vs. post May 9th traffic speeds in 2015 (where both Uber and Lyft operated in Austin) to pre vs. post May 9th traffic speeds in 2016 (where both Uber and Lyft exited Austin), allowing for segment-specific congestion responses. Bars represent the number of segments with idiosyncratic changes in traffic speeds falling withing a given bin. Cells are colored by the pre May 9th 2016 congestion level, as measured by the ratio of average speed to the 95th percentile of speed. Traffic speed data were accessed through the City of Austin's Open Data Portal.

Figure 3.10—Segment-Specific Responses

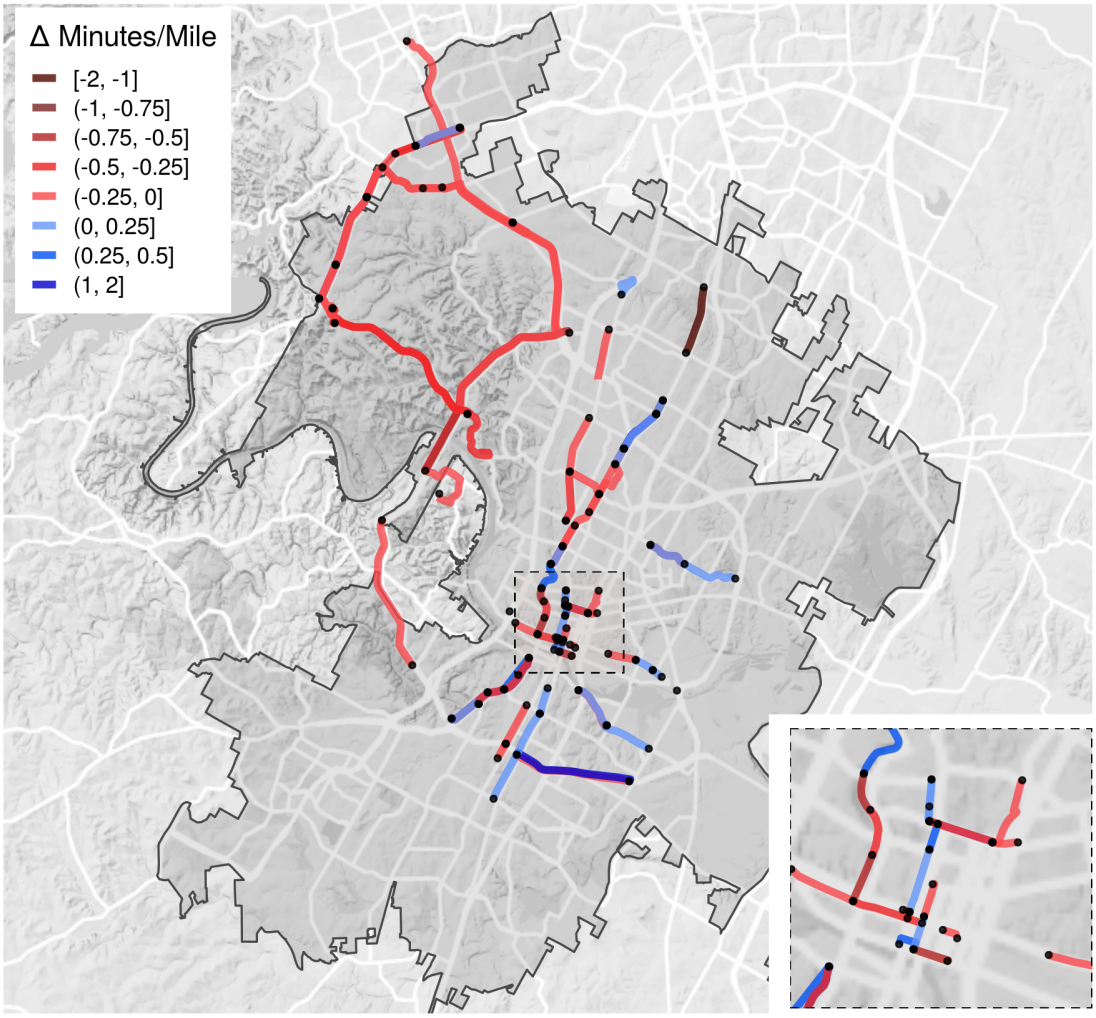


Figure 3.10: *Notes:* Results from Equation 3.2, a difference in differences comparing pre vs. post May 9th traffic speeds in 2015 (where both Uber and Lyft operated in Austin) to pre vs. post May 9th traffic speeds in 2016 (where both Uber and Lyft exited Austin), allowing for segment-specific congestion responses. Paths represent Google Maps recommended driving directions between endpoints of a given segment, colored by the sign and magnitude of the estimated segment-specific change in traffic speed. The black line is the Austin city limit. Traffic speed data were accessed through the City of Austin's OpenData Portal.

Figure 3.11—External Validity

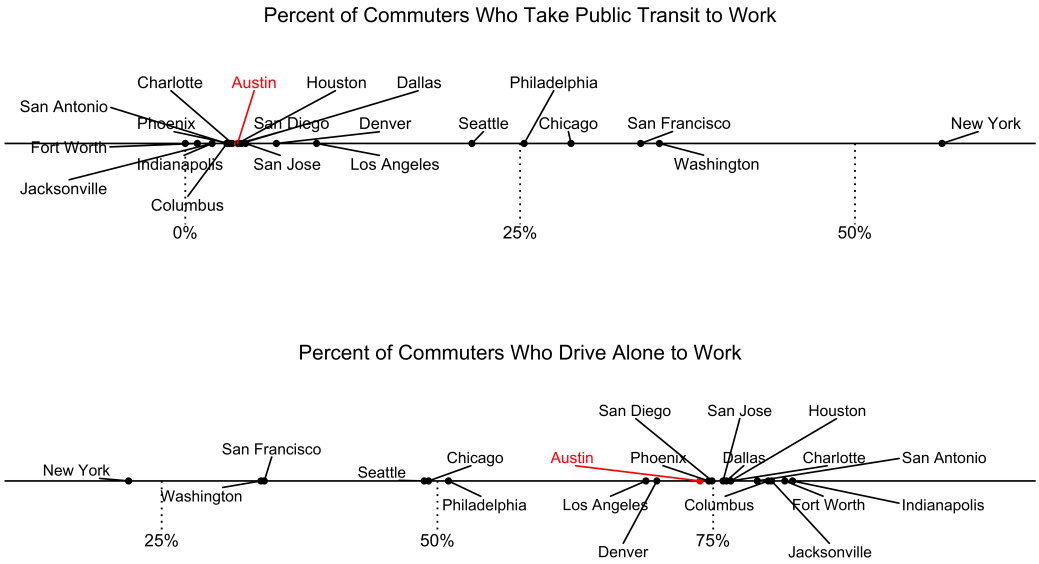


Figure 3.11: *Notes:* This figure depicts vehicle and public transit use across the 20 largest US cities for the year 2017. Data for both subfigures come from the U.S. Census Bureau's 2017 American Community Survey. Similarities in commuting behavior between Austin and other "Sun Belt" cities suggests that the findings in this paper may be most applicable to this group of metros.

3.9. Tables

Table 3.1—Road Segment Summary Statistics

	mean	sd	min	max
Average Speed (mph)	24.74	9.61	2.17	95.04
Minutes Per Mile	2.99	1.79	0.63	27.69
Segment Length	0.72	0.57	0.06	3.80
Number samples	4.77	3.77	1.00	45.00
Number of Lanes	4.70	0.91	2.00	7.00

Table 3.1: Summary statistics for traffic data along 79 road segments in Austin, TX. Speed data reflect the average travel time for Bluetooth devices that move from origin sensor to destination sensor during a given 15-minute interval. As described in Section 3.4, data are also filtered for outliers. Traffic speed data were accessed through the City of Austin's OpenData Portal.

Table 3.2—Difference in Differences Results

Hour of Day	β_h	se	p
0	-0.0606	0.0515	0.2588
1	-0.0525	0.0358	0.1640
2	**-0.0505**	0.0278	0.0908
3	-0.0121	0.0649	0.8547
4	0.0911	0.1006	0.3805
5	-0.0293	0.0920	0.7547
6	-0.0568	0.0996	0.5778
7	-0.1446	0.1169	0.2365
8	-0.0583	0.1331	0.6680
9	-0.0188	0.0954	0.8469
10	0.0318	0.0581	0.5921
11	**-0.1418**	0.0664	0.0508
12	**-0.1730**	0.0812	0.0512
13	-0.1555	0.0927	0.1157
14	-0.0176	0.0469	0.7125
15	-0.0238	0.0453	0.6083
16	0.0657	0.0683	0.3524
17	0.0016	0.0591	0.9790
18	0.0500	0.0862	0.5715
19	-0.0300	0.0638	0.6451
20	0.0004	0.1036	0.9966
21	0.0817	0.0703	0.2646
22	-0.0709	0.0852	0.4191
23	-0.0152	0.0328	0.6497
F-test			0.0000
N			966,301

Table 3.2: *Notes*: Results from Equation 3.1, a difference in differences comparing pre vs. post May 9$^{\text{th}}$ traffic speeds in 2015 to pre vs. post May 9$^{\text{th}}$ traffic speeds in 2016 (where both Uber and Lyft exited Austin). Controls include segment-specific linear in day trends, a precipitation dummy, day of week fixed effects, and year and post May 9$^{\text{th}}$ dummies. Standard errors are clustered by segment-week. β_h represent the estimated effect of TNC departure on traffic speeds (in minutes per mile) by hour of day. Bold coefficients are significant at the 10% level. The final row reports the p-value from a joint hypothesis test of $\beta_h = 0 \; \forall h$.

TABLE 3.3—REGRESSION DISCONTINUITY RESULTS

Hour of Day	β_h	se	p
0	**-0.0598**	0.0297	0.0637
1	**-0.0959**	0.0343	0.0142
2	-0.0316	0.0251	0.2287
3	0.0202	0.0348	0.5711
4	0.0156	0.0690	0.8245
5	**-0.0920**	0.0395	0.0352
6	-0.0267	0.0618	0.6727
7	-0.1081	0.1016	0.3050
8	-0.1370	0.1666	0.4248
9	-0.0661	0.0769	0.4047
10	0.0128	0.0566	0.8249
11	**-0.1137**	0.0462	0.0275
12	**-0.2675**	0.0744	0.0029
13	**-0.2861**	0.0878	0.0057
14	**-0.0932**	0.0404	0.0368
15	**-0.0759**	0.0398	0.0775
16	-0.0187	0.0449	0.6833
17	-0.0623	0.0561	0.2856
18	-0.0257	0.0715	0.7246
19	-0.0605	0.0539	0.2808
20	-0.0727	0.0565	0.2188
21	-0.1082	0.1144	0.3601
22	-0.1388	0.1126	0.2379
23	-0.0385	0.0544	0.4910
F-test			0.0000
N			501,010

Table 3.3: *Notes*: Results from Equation 3.3, a regression discontinuity performed on traffic speeds across 79 road segments in Austin, TX. The bandwidth is March 20$^{\text{th}}$ - August 1$^{\text{st}}$ of 2016, which (asymmetrically) spans the May 9$^{\text{th}}$ departure of Uber and Lyft. Controls include hour of day, day of week, holiday, and segment fixed effects, segment-specific second degree polynomials in days since May 9$^{\text{th}}$, and flexible controls for temperature and precipitation. Standard errors are clustered by segment-week. β_h represent the estimated effect of TNC departure on traffic speeds (in minutes per mile) by hour of day. Bold coefficients are significant at the 10% level. The final row reports the p-value from a joint hypothesis test of $\beta_h = 0 \; \forall h$.

Table 3.4—Pooled Estimates

	β (Δ minutes/mile)	*se*	*p*	Implied annual cost ($)
Difference in Differences (All hours)	-0.0261	0.0170	0.1479	-33,096,514
Difference in Differences (7 a.m. - 7 p.m.)	-0.0684	0.0529	0.2004	-63,985,000
Regression Discontinuity (All Hours)	-0.1015	0.0353	0.0052	-129,010,337
Regression Discontinuity (7 a.m. - 7 p.m.)	-0.1335	0.0433	0.0028	-124,930,327

Table 3.4: *Notes*: The first two rows display results from a variation of Equation 3.1, a difference in differences specification that estimates the pooled impact of TNC exit on traffic speeds across hours of day. β represent the estimated effect of TNC departure on traffic speeds, measured in minutes per mile. Controls include segment-specific linear in day trends, controls for precipitation, day of week fixed effects, hour of day fixed effects, and year and post-May 9^{th} dummies. Standard errors are clustered by segment-week. Row 1 shows the results of this regression using speed data on all hours, and column 2 shows results restricted to 7 a.m. to 7 p.m. The final column displays annual costs implied by multiplying β by annual Ausin-area VMT, and then by $15.40, which is 50% of the average per-worker wage rate for Austin households, according to the 2017 NHTS. Traffic data were accessed through the City of Austin's OpenData Portal. Rows 3 and 4 repeat this exercise for Equation 3.3.

Table 3.5—Fuzzy RD and Fuzzy DD

	β (Δ minutes/mile)	*se*	*p*
Fuzzy DID, RideAustin Data	0.0721	0.0573	0.2084
Fuzzy DID, Hampshire Data	0.0978	0.0237	0.0010
Fuzzy RD, RideAustin Data	0.1324	0.0133	0.0000
Fuzzy RD, Hampshire Data	0.2263	0.0227	0.0000

Table 3.5: *Notes*: The first two rows display results from a fuzzy difference in differences specification that estimates the impact of TNC activity on Austin traffic speeds between 7 a.m. and 7 p.m. The regression coefficient represents the change in travel times (in minutes per mile) resulting from a change from full TNC operation [1] to no TNC operation [0]. The first row (*RideAustin Data*) uses RideAustin's trip-level data together with estimates of RideAustin's market share to construct a measure of daily TNC activity. The second row assumes that the level of TNC activity during the period following the exit of Uber and Lyft was a constant 41% of pre-exit levels. This assumtion is based on a November 2016 survey conducted by Hampshire et al. (2018). Rows three and four report results from fuzzy regression discontinuity designs that estimate the impact of TNC activity on Austin traffic speeds between 7 a.m. and 7 p.m., using the same TNC activity assumptions.

Table 3.6—Austin Metro Validity

	Not Sampled	Sampled	p
Annual Delay per Mile (person-hours)	118,372.98	101,368.11	0.58
Texas Congestion Index	1.36	1.39	0.61
Peak Period Average Speed	33.22	29.52	0.16
Freeflow Speed	41.20	37.97	0.28
Average Daily VMT	204,488.84	166,805.50	0.46
Peak Period Annual Hours of Delay (person-hours)	377,230.86	241,909.86	0.34

Table 3.6: *Notes*: This table uses data maintained by the Texas Department of Transportation (TXDoT) to compare observable characteristics of Austin-area roads that do not appear in my sample (column 1) to those that do (column 2). Column (3) reports p-value of the corresponding t-test for a difference in means. The data cover 86 road segments in Austin, 35 of which overlap with the 79 Bluetooth segments used in the above analysis. Note that roads sections are coded as sampled even if the Bluetooth segment does not cover the entire corresponding TXDoT road segment.

Table 3.7—Congestion Cost Estimates

	daily cost ($)	*se*	*p*	annual cost ($)
Time-varying VOT	-92,071	97,547	0.1806	-33,605,827
Uniform $15.40 VOT	-127,983	72,156	0.0489	-46,713,725

Table 3.7: *Notes*: Estimates of the travel-time congestion costs of TNC operation in Austin, TX. The first row displays the result of the exercise described in Equation 3.4, which matches hour-specific changes in travel time to hour-specific willingess to pay estimates (detailed in Appendix C.1) and hour-specific traffic volume measurements. The second row uses a VOT of $15.40, which is 50% of the average per-worker wage rate for car-commuting Austin households, according to the 2017 NHTS. Standard errors for cost figures are calculated following Goodman (1960).

Table 3.8—Summary Statistics for Components of Cost Calculation

Hour of Day	β	*se*	VOT	sd	VMT	sd	Cost/hour (MoPac VOT)	Cost/hour (NHTS VOT)
0	-0.06	0.04	8.66	3.05	201,112	1,667	-1,759	-1,759
1	-0.05	0.04	8.4	2.96	125,170	825	-921	-921
2	-0.05	0.04	8.29	2.91	95,873	566	-669	-669
3	-0.01	0.05	8.3	2.92	64,993	265	-109	-109
4	0.09	0.05	8.32	3	45,746	73	578	578
5	-0.03	0.06	9.69	3.57	79,183	89	-375	-375
6	-0.06	0.07	12.67	8.28	216,472	1,181	-2,595	-2,595
7	-0.14	0.08	25.64	30.81	510,299	7,646	-31,539	-31,539
8	-0.06	0.12	28.78	37.99	680,662	11,764	-19,034	-19,034
9	-0.02	0.08	14.67	17.84	729,092	8,322	-3,344	-3,344
10	0.03	0.05	10.16	4.08	726,981	3,464	3,919	3,919
11	-0.14	0.04	10.1	3.75	786,248	4,234	-18,771	-18,771
12	-0.17	0.04	10.05	4.83	910,484	4,754	-26,387	-26,387
13	-0.16	0.04	10.41	6.06	943,347	5,733	-25,444	-25,444
14	-0.02	0.04	11.75	9.67	919,202	6,131	-3,176	-3,176
15	-0.02	0.04	20.79	19.93	937,669	5,146	-7,716	-7,716
16	0.07	0.04	29	26.84	950,548	4,552	30,189	30,189
17	0	0.04	24.5	22.83	953,432	4,493	618	618
18	0.05	0.04	23.27	24.29	940,423	4,676	18,220	18,220
19	-0.03	0.04	11.55	9.64	834,215	3,861	-4,822	-4,822
20	0	0.04	9.76	3.68	644,943	3,076	47	47
21	0.08	0.04	8.42	3.4	529,941	2,619	6,082	6,082
22	-0.07	0.04	8.67	3.36	429,036	2,491	-4,397	-4,397
23	-0.02	0.04	8.57	3.06	306,256	2,127	-665	-665
Daily Cost							**-92,071**	**-127,983**

Table 3.8: *Notes*: This table displays the components of the aggregate cost calculations in Table 3.7. The first two columns reproduce the difference in differences results from Table 3.2. Columns 3 and 4 show the mean and standard deviation of the value of travel time (VOT) estimates from the MoPac Freeway (see Appendix C.1). Columns 5 and 6 report the mean and standard deviation of hourly VMT. These estimates are generated by multiplying the hourly share of total Austin Traffic (as per Austin traffic count data) by estimates of aggregate Austin-area VMT (as per the Texas Department of Transportation). The penultimate column reports the hourly costs estimates using the VOT from column 3; the final column reports hourly cost estimates applying a constant $15.40 per hour.

Table 3.9—Congestion Costs Disaggregating Weekdays and Weekends

	annual cost ($)	*se*	*p*
Time-varying VOT	-39,989,479	21,348,072	0.0410
Uniform $15.40 VOT	-52,084,544	15,784,030	0.0026

Table 3.9: *Notes*: Estimates of the travel-time congestion costs of TNC operation in Austin, TX, using separate estimates for weekdays and weekends. The first row follows Equation 3.4, and uses separate hourly Value of Travel Time (VOT) schedules for weekdays and weekends derived from MoPac Data. The second row applies a uniform VOT time across all hours. Following the heuristic provided by Small and Verhoef (2007), this VOT is 50% of the per-worker wage rate for Austin-area car commuting households, as per the 2017 NHTS. Standard errors for cost figures are calculated as described in Section 3.6.

Bibliography

Acemoglu, Daron. 2003. "Why not a political Coase theorem? Social conflict, commitment, and politics." *Journal of comparative economics* 31 (4):620–652.

Acemoglu, Daron and James A Robinson. 2000. "Political losers as a barrier to economic development." *American Economic Review* 90 (2):126–130.

Allcott, Hunt, Benjamin B Lockwood, and Dmitry Taubinsky. 2019. "Regressive sin taxes, with an application to the optimal soda tax." *The Quarterly Journal of Economics* 134 (3):1557–1626.

Anderson, Michael L. 2020. "As the wind blows: The effects of long-term exposure to air pollution on mortality." *Journal of the European Economic Association* 18 (4):1886–1927.

Anderson, Michael L and Maximilian Auffhammer. 2014. "Pounds that kill: The external costs of vehicle weight." *Review of Economic Studies* 81 (2):535–571.

Anderson, Michael L, Fangwen Lu, Yiran Zhang, Jun Yang, and Ping Qin. 2016. "Superstitions, street traffic, and subjective well-being." *Journal of Public Economics* 142:1–10.

Angrist, Joshua, Sydnee Caldwell, and Jonathan Hall. 2017. "Uber vs. Taxi: A Driver's Eye View." *NBER Working Paper 23891* .

Arnott, Richard, Andre De Palma, and Robin Lindsey. 1990. "Economics of a bottleneck." *Journal of Urban Economics* 27 (1):111–130.

———. 1993. "A structural model of peak-period congestion: A traffic bottleneck with elastic demand." *The American Economic Review* :161–179.

Auffhammer, Maximilian and Ryan Kellogg. 2011. "Clearing the air? The effects of gasoline content regulation on air quality." *American Economic Review* 101 (6):2687–2722.

Banzhaf, H Spencer and Wallace E Oates. 2012. "On fiscal illusion and Ricardian equivalence in local public finance." Tech. rep., National Bureau of Economic Research.

Barron, Alexander R, Allen A Fawcett, Marc AC Hafstead, James R McFarland, and Adele C Morris. 2018. "Policy insights from the EMF 32 study on US carbon tax scenarios." *Climate Change Economics* 9 (01):1840003.

Battaglia, Michael P, David Izrael, David C Hoaglin, and Martin R Frankel. 2009. "Practical considerations in raking survey data." *Survey Practice* 2 (5):1–10.

Becker, Gary S. 1983. "A theory of competition among pressure groups for political influence." *The quarterly journal of economics* 98 (3):371–400.

Bischoff, Ivo and Lars-HR Siemers. 2013. "Biased beliefs and retrospective voting: why democracies choose mediocre policies." *Public Choice* 156 (1-2):163–180.

Blaufus, Kay, Malte Chirvi, Hans-Peter Huber, Ralf Maiterth, and Caren Sureth-Sloane. 2020. "Tax Misperception and Its Effects on Decision Making–a Literature Review." *TRR* 266.

Blinder, Alan S. 1973. "Wage discrimination: reduced form and structural estimates." *Journal of Human resources* :436–455.

Blomquist, Sören, Whitney K Newey, Anil Kumar, and Che-Yuan Liang. 2021. "On bunching and identification of the taxable income elasticity." *Journal of Political Economy* 129 (8):000–000.

Börjesson, Maria, Jonas Eliasson, Muriel B Hugosson, and Karin Brundell-Freij. 2012. "The Stockholm congestion charges—5 years on. Effects, acceptability and lessons learnt." *Transport Policy* 20:1–12.

Bottan, Nicolas L and Ricardo Perez-Truglia. 2017. "Choosing your pond: location choices and relative income." Tech. rep., National Bureau of Economic Research.

Boyce, James K. 2018. "Carbon pricing: effectiveness and equity." *Ecological economics* 150:52–61.

Buchholz, Nicholas, Laura Doval, Jakub Kastl, Filip Matějka, and Tobias Salz. 2020. "The value of time: Evidence from auctioned cab rides." Tech. rep., National Bureau of Economic Research.

Bursztyn, Leonardo, Alessandra L González, and David Yanagizawa-Drott. 2018. "Misperceived social norms: Female labor force participation in Saudi Arabia." Tech. rep., National Bureau of Economic Research.

Cantoni, Davide, David Y Yang, Noam Yuchtman, and Y Jane Zhang. 2019. "Protests as strategic games: experimental evidence from Hong Kong's antiauthoritarian movement." *The Quarterly Journal of Economics* 134 (2):1021–1077.

Carattini, Stefano, Maria Carvalho, and Sam Fankhauser. 2018. "Overcoming public resistance to carbon taxes." *Wiley Interdisciplinary Reviews: Climate Change* 9 (5):e531.

Chetty, Raj, John N Friedman, Tore Olsen, and Luigi Pistaferri. 2011. "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records." *The Quarterly Journal of Economics* 126 (2):749–804.

City of San Francisco. 2021. "City Performance Scorecards." .

Clewlow, Regina and Gouri Mishra. 2017. "Disruptive Transportation: The Adoption, Utilization, and Impacts of Ride-Hailing in the United States." *Institute of Transportation Studies, University of California, Davis Research Report UCD-ITS-RR-17-07* .

Cohen, Peter, Robert Hahn, Jonathan Hall, Steven Levitt, and Robert Metcalfe. 2016. "Using big data to estimate consumer surplus: The case of uber." Tech. rep., National Bureau of Economic Research.

Cramer, Judd and Alan B Krueger. 2016. "Disruptive change in the taxi business: The case of Uber." *American Economic Review* 106 (5):177–82.

Cronin, Julie Anne, Don Fullerton, and Steven Sexton. 2019. "Vertical and horizontal redistributions from a carbon tax and rebate." *Journal of the Association of Environmental and Resource Economists* 6 (S1):S169–S208.

Currie, Janet and Reed Walker. 2011a. "Traffic congestion and infant health: Evidence from E-ZPass." *American Economic Journal: Applied Economics* 3 (1):65–90.

———. 2011b. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Review* 3:65–90.

Davis, Lucas W. 2008. "The effect of driving restrictions on air quality in Mexico City." *Journal of Political Economy* 116 (1):38–81.

———. 2017. "Saturday driving restrictions fail to improve air quality in Mexico City." *Scientific Reports* 7 (1):1–9.

Davis, Lucas W and James M Sallee. 2020. "Should electric vehicle drivers pay a mileage tax?" *Environmental and Energy Policy and the Economy* 1 (1):65–94.

Dawes, Christopher T, Peter John Loewen, and James H Fowler. 2011. "Social preferences and political participation." *The Journal of Politics* 73 (3):845–856.

Dell'Anno, Roberto and Paulo Mourao. 2012. "Fiscal illusion around the world: an analysis using the structural equation approach." *Public finance review* 40 (2):270–299.

Deryugina, Tatyana, Garth Heutel, Nolan H Miller, David Molitor, and Julian Reif. 2019. "The mortality and medical costs of air pollution: Evidence from changes in wind direction." *American Economic Review* 109 (12):4178–4219.

Diamond, Peter A. 1973. "Consumption externalities and imperfect corrective pricing." *The Bell Journal of Economics and Management Science* :526–538.

Douenne, Thomas and Adrien Fabre. 2020. "French attitudes on climate change, carbon taxation and other climate policies." *Ecological Economics* 169:106496.

Downs, Anthony et al. 1957. "An economic theory of democracy." *New York: Harper* .

Duranton, Gilles and Matthew A Turner. 2011. "The fundamental law of road congestion: Evidence from US cities." *American Economic Review* 101 (6):2616–52.

Edlin, Aaron, Andrew Gelman, and Noah Kaplan. 2007. "Voting as a rational choice: Why and how people vote to improve the well-being of others." *Rationality and society* 19 (3):293–314.

EIA, U.S. 2018. *Updated renewable portfolio standards will lead to more renewable electricity generation*.

———. 2019. *Four states updated their renewable portfolio standards in the first half of 2019*.

Eliasson, Jonas. 2009. "A cost–benefit analysis of the Stockholm congestion charging system." *Transportation Research Part A: Policy and Practice* 43 (4):468–480.

Erhardt, Gregory D, Sneha Roy, Drew Cooper, Bhargava Sana, Mei Chen, and Joe Castiglione. 2019. "Do transportation network companies decrease or increase congestion?" *Science advances* 5 (5):eaau2670.

Finkelstein, Amy. 2009. "E-ztax: Tax salience and tax rates." *The Quarterly Journal of Economics* 124 (3):969–1010.

Foreman, Kate. 2016. "Crossing the bridge: The effects of time-varying tolls on curbing congestion." *Transportation Research Part A: Policy and Practice* 92:76–94.

Fowler, James H. 2006. "Habitual voting and behavioral turnout." *The Journal of Politics* 68 (2):335–344.

Galen, L Barbose. 2018. "US Renewables Portfolio Standards: 2018 Annual Status Report." Tech. rep., Tech. rep. Nov.

Galiani, Sebastian, Gustavo Torrens, and Maria Lucia Yanguas. 2014. "The political coase theorem: experimental evidence." *Journal of Economic Behavior & Organization* 103:17–38.

Gibson, Matthew. 2019. "Regulation-induced pollution substitution." *Review of Economics and Statistics* 101 (5):827–840.

Gibson, Matthew and Maria Carnovale. 2015. "The effects of road pricing on driver behavior and air pollution." *Journal of Urban Economics* 89:62–73.

Giuliano, Genevieve. 1992. "An assessment of the political acceptability of congestion pricing." *Transportation* 19 (4):335–358.

Goldberg, Pinelopi Koujianou and Giovanni Maggi. 1999. "Protection for sale: An empirical investigation." *American Economic Review* 89 (5):1135–1155.

Goldszmidt, Ariel, John A List, Robert D Metcalfe, Ian Muir, V Kerry Smith, and Jenny Wang. 2020. "The Value of Time in the United States: Estimates from Nationwide Natural Field Experiments." Tech. rep., National Bureau of Economic Research.

Goodman, Leo A. 1960. "On the exact variance of products." *Journal of the American statistical association* 55 (292):708–713.

Goulder, Lawrence H, Xianling Long, Jieyi Lu, and Richard D Morgenstern. 2019. "China's Unconventional Nationwide CO 2 Emissions Trading System: The Wide-Ranging Impacts of an Implicit Output Subsidy." Tech. rep., National Bureau of Economic Research.

Grainger, Corbett A and Charles D Kolstad. 2010. "Who pays a price on carbon?" *Environmental and Resource Economics* 46 (3):359–376.

Green, Colin P, John S Heywood, and Maria Navarro. 2016. "Traffic accidents and the London congestion charge." *Journal of Public Economics* 133:11–22.

Green, Colin P, John S Heywood, and Maria Navarro Paniagua. 2020. "Did the London congestion charge reduce pollution?" *Regional Science and Urban Economics* 84:103573.

Green, Jerry and Eytan Sheshinski. 1976. "Direct versus indirect remedies for externalities." *Journal of Political Economy* 84 (4, Part 1):797–808.

Greenstone, Michael and Ishan Nath. 2019. "Do Renewable Portfolio Standards Deliver?" *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2019-62).

Greenwood, Brad N and Sunil Wattal. 2015. "Show me the way to go home: an empirical investigation of ride sharing and alcohol related motor vehicle homicide." *Fox School of Business Research Paper* (15-054).

Grossman, Gene and Elhanan Helpman. 1994. "Foreign investment with endogenous protection." Tech. rep., National Bureau of Economic Research.

Hall, Jonathan D, Craig Palsson, and Joseph Price. 2018. "Is Uber a substitute or complement for public transit?" *Journal of Urban Economics* 108:36–50.

Hampshire, Robert, Chris Simek, Tayo Fabusuyi, Xuan Di, and Xi Chen. 2017. "Measuring the impact of an unanticipated disruption of Uber/Lyft in Austin, TX." *Lyft in Austin, TX (May 31, 2017)* .

Hanna, Rema, Gabriel Kreindler, and Benjamin A Olken. 2017. "Citywide effects of high-occupancy vehicle restrictions: Evidence from "three-in-one" in Jakarta." *Science* 357 (6346):89–93.

Hansford, Thomas G and Brad T Gomez. 2015. "Reevaluating the sociotropic economic voting hypothesis." *Electoral Studies* 39:15–25.

Healy, Andrew J, Mikael Persson, and Erik Snowberg. 2017. "Digging into the pocketbook: Evidence on economic voting from income registry data matched to a voter survey." *American Political Science Review* 111 (4):771–785.

Heo, Jinhyok, Peter J Adams, and H Oliver Gao. 2016. "Public health costs of primary PM2. 5 and inorganic PM2. 5 precursor emissions in the United States." *Environmental Science & Technology* 50 (11):6061–6070.

Hernandez-Cortes, Danae and Kyle C Meng. 2020. "Do environmental markets cause environmental injustice? Evidence from California's carbon market." Tech. rep., National Bureau of Economic Research.

Herrnstadt, Evan, Ian WH Parry, and Juha Siikamäki. 2015. "Do alcohol taxes in Europe and the US rightly correct for externalities?" *International Tax and Public Finance* 22 (1):73–101.

Holland, Stephen P. 2012. "Emissions taxes versus intensity standards: Second-best environmental policies with incomplete regulation." *Journal of Environmental Economics and Management* 63 (3):375–387.

Huber, Robert A, Michael L Wicki, and Thomas Bernauer. 2020. "Public support for environmental policy depends on beliefs concerning effectiveness, intrusiveness, and fairness." *Environmental Politics* 29 (4):649–673.

Inrix. 2017. *INRIX Global Traffic Scorecard*. URL http://inrix.com/scorecard/.

———. 2018. *INRIX Global Traffic Scorecard*. URL http://inrix.com/scorecard/.

Jacobs, Bas and Ruud A De Mooij. 2015. "Pigou meets Mirrlees: On the irrelevance of tax distortions for the second-best Pigouvian tax." *Journal of Environmental Economics and Management* 71:90–108.

Jankowski, Richard. 2007. "Altruism and the decision to vote: Explaining and testing high voter turnout." *Rationality and Society* 19 (1):5–34.

Johnson, M Bruce. 1964. "On the economics of road congestion." *Econometrica: Journal of the Econometric Society* :137–150.

King, David, Michael Manville, and Donald Shoup. 2007. "The political calculus of congestion pricing." *Transport Policy* 14 (2):111–123.

Kleven, Henrik J and Mazhar Waseem. 2013. "Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from Pakistan." *The Quarterly Journal of Economics* 128 (2):669–723.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics* 8:435–464.

Knittel, Christopher R, Douglas L Miller, and Nicholas J Sanders. 2016. "Caution, drivers! Children present: Traffic, pollution, and infant health." *Review of Economics and Statistics* 98 (2):350–366.

Knittel, Christopher R and Ryan Sandler. 2018. "The welfare impact of second-best uniform-Pigouvian taxation: evidence from transportation." *American Economic Journal: Economic Policy* 10 (4):211–42.

Kopczuk, Wojciech. 2003. "A note on optimal taxation in the presence of externalities." *Economics Letters* 80 (1):81–86.

Kreindler, Gabriel E. 2018. "The welfare effect of road congestion pricing: Experimental evidence and equilibrium implications." *Unpublished paper* .

Krueger, Anne O. 1974. "The political economy of the rent-seeking society." *The American economic review* 64 (3):291–303.

Kuziemko, Ilyana, Michael I Norton, Emmanuel Saez, and Stefanie Stantcheva. 2015. "How elastic are preferences for redistribution? Evidence from randomized survey experiments." *American Economic Review* 105 (4):1478–1508.

Lave, Charles A. 1969. "A behavioral approach to modal split forecasting." *Transportation Research UK* 3 (4).

Leape, Jonathan. 2006. "The London congestion charge." *Journal of Economic Perspectives* 20 (4):157–176.

Lehe, Lewis. 2019. "Downtown congestion pricing in practice." *Transportation Research Part C: Emerging Technologies* 100:200–223.

Li, Ziru, Yili Hong, and Zhongju Zhang. 2019. "Do Ride-sharing Services Affect Traffic Congestion? An Empirical Study of Uber Entry." *Working Paper* .

Mangrum, Daniel and Alejandro Molnar. 2018. "The marginal congestion of a taxi in New York City." *Working Paper* .

Marron, Donald B, Eric J Toder, and Lydia Austin. 2015. "Taxing Carbon: What, Why, and How." *Why, and How (June 25, 2015)* .

Meng, Kyle C and Ashwin Rode. 2019. "The social cost of lobbying over climate policy." *Nature Climate Change* 9 (6):472–476.

Muller, Nicholas Z and Robert Mendelsohn. 2007. "Measuring the damages of air pollution in the United States." *Journal of Environmental Economics and Management* 54 (1):1–14.

New York Times. 2016. *Uber and Lyft End Rides in Austin to Protest Fingerprint Background Checks*. URL https://www.nytimes.com/2016/05/10/technology/uber-and-lyft-stop-rides-in-austin-to-protest-fingerprint-background-checks.html.

———. 2018. *Uber Hit With Cap as New York City Takes Lead in Crackdown*. URL https://www.nytimes.com/2018/08/08/nyregion/uber-vote-city-council-cap.html.

———. 2019. *Your Taxi or Uber Ride in Manhattan Will Soon Cost More*. URL https://www.nytimes.com/2019/01/31/nyregion/uber-taxi-lyft-fee.html.

North, Douglass C and Barry R Weingast. 1989. "Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England." *The journal of economic history* 49 (4):803–832.

Oaxaca, Ronald. 1973. "Male-female wage differentials in urban labor markets." *International economic review* :693–709.

Parry, Ian WH. 2009. "Pricing urban congestion." *Annu. Rev. Resour. Econ.* 1 (1):461–484.

Parry, Ian WH and Kenneth A Small. 2005. "Does Britain or the United States have the right gasoline tax?" *American Economic Review* 95 (4):1276–1289.

Parry, Ian William Holmes. 2002. "Comparing the efficiency of alternative policies for reducing traffic congestion." *Journal of Public Economics* 85 (3):333–362.

Percoco, Marco. 2016. "The impact of road pricing on accidents: a note on Milan." *Letters in Spatial and Resource Sciences* 9 (3):343–352.

Prud'Homme, Remy and Juan Pablo Bocarejo. 2005. "The London congestion charge: a tentative economic appraisal." *Transport Policy* 12 (3):279–287.

Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *The Economic Journal* 37 (145):47–61.

Rausch, Sebastian and Valerie J Karplus. 2014. "Markets versus regulation: The efficiency and distributional impacts of US climate policy proposals." *The Energy Journal* 35 (Special Issue).

Rausch, Sebastian and Matthew Mowers. 2014. "Distributional and efficiency impacts of clean and renewable energy standards for electricity." *Resource and Energy Economics* 36 (2):556–585.

Rayle, Lisa, Susan Shaheen, Nelson Chan, Danielle Dai, and Robert Cervero. 2014. "App-Based, On-Demand Ride Services: Comparing Taxi and Ridesourcing Trips and User Characteristics in San Francisco." *University of California Transportation Center* .

Rees-Jones, Alex and Dmitry Taubinsky. 2016. *Heuristic perceptions of the income tax: Evidence and implications for debiasing*. National Bureau of Economic Research.

Regional Plan Association. 2021. "Congestion Pricing in New York City." Tech. rep., Regional Plan Association.

Reguant, Mar. 2018. "The efficiency and sectoral distributional implications of large-scale renewable policies." Tech. rep., National Bureau of Economic Research.

Reuters. 2019. *Uber and other taxi firms to pay London congestion charge*. URL https://www.reuters.com/article/us-britain-taxi/uber-and-other-taxi-firms-to-pay-london-congestion-charge-idUSKBN1OI14H.

RideAustin. 2017. *Comprehensive Ride Data*. URL https://data.world/ride-austin.

Saez, Emmanuel. 2010. "Do Taxpayers Bunch at Kink Points?" *American Economic Journal: Economic Policy* 2 (3):180–212.

Sallee, James M. 2019. "Pigou creates losers: On the implausibility of achieving pareto improvements from efficiency-enhancing policies." Tech. rep., National Bureau of Economic Research.

San Francisco County Traffic Authority. 2021. "Downtown Congestion Pricing Study: Winter 2021 Update." Tech. rep., San Francisco County Traffic Authority.

San Francisco Transit Authority. 2018. *TNCs and Congestion*. URL https://www.sfcta.org/projects/tncs-and-congestion.

Sandmo, Agnar. 1975. "Optimal taxation in the presence of externalities." *The Swedish Journal of Economics* :86–98.

———. 1978. "Direct versus indirect Pigovian taxation." *European Economic Review* 7 (4):337–349.

Sapienza, Paola and Luigi Zingales. 2013. "Economic experts versus average Americans." *American Economic Review* 103 (3):636–42.

Sausgruber, Rupert and Jean-Robert Tyran. 2011. "Are we taxing ourselves?: How deliberation and experience shape voting on taxes." *Journal of Public Economics* 95 (1-2):164–176.

Savidge, Nico. 2021. "Chart: Five ways COVID changed Bay Area traffic." .

Schaller Consulting. 2018. *The New Automobility: Lyft, Uber and the Future of American Cities*. URL http://www.schallerconsult.com/rideservices/automobility.pdf.

Sekar, Samantha and Brent Sohngen. 2014. "The effects of renewable portfolio standards on carbon intensity in the United States." *Resources for the Future Discussion Paper* (14-10).

Sims, Christopher A. 2003. "Implications of rational inattention." *Journal of monetary Economics* 50 (3):665–690.

Small, Kenneth and Erik Verhoef. 2007. "The Economics of Urban Transportation." *Routledge* .

Small, Kenneth A. 1982. "The scheduling of consumer activities: work trips." *The American Economic Review* 72 (3):467–479.

———. 2012. "Valuation of travel time." *Economics of Transportation* 1 (1-2):2–14.

Small, Kenneth A, Erik T Verhoef, and Robin Lindsey. 2007. *The economics of urban transportation*. Routledge.

Stantcheva, Stefanie. 2020. "Understanding tax policy: How do people reason?" Tech. rep., National Bureau of Economic Research.

The 85th Texas Legislature. 2017. *Texas House Bill 100*. URL https://legiscan.com/TX/bill/HB100/2017.

The City Council of Austin. 2015. *Ordinance 20151217-075*. URL http://www.austintexas.gov/content/december-17-2015-austin-city-council-regular-meeting.

The Congressional Budget Office. 2016. "Impose a Tax on Emissions of Greenhouse Gases." .

The Texas Tribune. 2016. *Austin's Proposition 1 Defeated*. URL https://www.texastribune.org/2016/05/07/early-voting-austin-proposition-against/.

Tonne, Cathryn, Sean Beevers, Ben Armstrong, Frank Kelly, and Paul Wilkinson. 2008. "Air pollution and mortality benefits of the London Congestion Charge: spatial and socioeconomic inequalities." *Occupational and Environmental Medicine* 65 (9):620–627.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Tseng, Yin Yen, Barry Ubbels, and Erik Verhoef. 2005. "Value of time, schedule delay and reliability." In *ERSA conference papers*.

Tullock, Gordon. 1967. "The welfare costs of tariffs, monopolies, and theft." *Economic Inquiry* 5 (3):224–232.

Uber. 2015. *Case Study Shows Our Impact in Austin*. URL https://www.uber.com/blog/austin/case-study-shows-our-impact-in-austin/.

United States Census Bureau. 2015. *American Commnity Survey*. URL https://www.census.gov/programs-surveys/acs.

Vancouver Sun. 2019. *Dan Fumano: Vancouver wants to charge Uber and Lyft users a congestion fee*. URL https://vancouversun.com/news/politics/dan-fumano-vancouver-wants-to-charge-uber-and-lyft-users-a-congestion-fee.

Verhoef, Erik, Peter Nijkamp, and Piet Rietveld. 1995. "Second-best regulation of road transport externalities." *Journal of Transport Economics and Policy* :147–167.

Vickrey, William S. 1963. "Pricing in urban and suburban transport." *The American Economic Review* 53 (2):452–465.

———. 1969. "Congestion theory and transport investment." *The American Economic Review* 59 (2):251–260.

Vira, Bhaskar. 1997. "The political Coase theorem: identifying differences between neoclassical and critical institutionalism." *Journal of Economic Issues* 31 (3):761–780.

Wilson, IA Grant and Iain Staffell. 2018. "Rapid fuel switching from coal to natural gas through effective carbon pricing." *Nature Energy* 3 (5):365–372.

Yang, Jun, Avralt-Od Purevjav, and Shanjun Li. 2020. "The marginal cost of traffic congestion and road pricing: Evidence from a natural experiment in Beijing." *American Economic Journal: Economic Policy* 12 (1):418–53.

Zhang, Wei, C-Y Cynthia Lin Lawell, and Victoria I Umanskaya. 2017. "The effects of license plate-based driving restrictions on air quality: Theory and empirical evidence." *Journal of Environmental Economics and Management* 82:181–220.

Zhong, Nan, Jing Cao, and Yuzhu Wang. 2017. "Traffic congestion, ambient air pollution, and health: Evidence from driving restrictions in Beijing." *Journal of the Association of Environmental and Resource Economists* 4 (3):821–856.

Appendix A

Appendix for *For Whom the Bridge Tolls: Congestion, Air Pollution, and Second-Best Road Pricing*

A.1. Theory Appendix

A.1.1. Substitution with Many Goods

Setup: A representative consumer chooses quantities of M goods, $(h_1, ..., h_M)$ and a numeraire, z. Each non-numeraire good has an associated externality, ϕ_m. A policymaker can choose tax levels for goods $j \in \{1, ..., J\}$ where $J < M$. I assume goods $k \notin \{1, ..., J\}$ are un- or under-taxed.

The consumer's problem: An agent maximizes utility over M goods $(h_1, ..., h_M)$ and a numeraire good z.

$$max\{U(h_1, ..., h_M) + z\} \quad s.t. \tag{A.1}$$

$$(p_1 + \tau_1)h_1 + (p_J + \tau_J)h_J + p_{J+1}h_{J+1} + ... + p_M h_M + z \leq I \tag{A.2}$$

The first-order conditions for an interior solution to the consumer's problem are:

$$U_j = \lambda(p_j + \tau_j) \quad \forall \quad j \in \{1, ..., J\} \tag{A.3}$$

$$U_k = \lambda(p_k) \quad \forall \quad k \notin \{1, ..., J\} \tag{A.4}$$

$$\lambda = 1 \tag{A.5}$$

The planner's problem: I assume that the planner seeks to maximize aggregate welfare, which is the utility of the representative consumer less the aggregate social cost of consumption, $\sum_1^M \phi_m h_m$. The planner's choice variables are tax levels $\tau_1 ... \tau_J$, which are applied to the taxable goods $j \in \{1, ..., J\}$.

$$\begin{aligned} max\{U(h_1, ..., h_M) + z - \sum_1^M \phi_m h_m\} \; st. \\ p_1 h_1 + ... + p_N h_N + z \leq I \end{aligned} \tag{A.6}$$

Assuming an internal solution, first-order condition wrt p_j (where $j \in \{1, ..., J\}$) is:

$$0 = \frac{\partial h_j}{\partial p_j}[U_j - \phi_j - p_j] + \sum_{k \neq j}^{M} \frac{\partial h_k}{\partial p_j}[U_k - \phi_k - p_k] \tag{A.7}$$

Plugging in the consumer's first order conditions and solving for τ_m...

$$0 = \frac{\partial h_j}{\partial p_j}[\tau_j - \phi_j] + \sum_{k \neq j}^{J} \frac{\partial h_k}{\partial p_j}[\tau_k - \phi_k] + \sum_{l=J+1}^{M} \frac{\partial h_l}{\partial p_j}[\phi_l] \tag{A.8}$$

$$\tau_j = \phi_j + \frac{1}{\frac{\partial h_j}{\partial p_j}}(\sum_{k \neq j}^{J} \frac{\partial h_k}{\partial p_j}[\phi_k - \tau_k] + \sum_{l=J+1}^{M} \frac{\partial h_l}{\partial p_j}\phi_l) \tag{A.9}$$

This intermediate results is intuitive. Holding fixed all taxes other than τ_j, the optimal value for this final tax is its externality, ϕ_m, minus a term that captures the extent to which consumers switch to other goods, and the level of unpriced externality of those goods.

Identifying the optimal tax level for *all* taxable goods requires solving J equations simultaneously:

$$\tau_j + \frac{1}{\frac{\partial h_j}{\partial p_j}}(\sum_{k \neq j}^{J} \frac{\partial h_k}{\partial p_j}\tau_k) = \phi_j + \frac{1}{\frac{\partial h_j}{\partial p_j}} \sum_{l=1}^{M} \frac{\partial h_l}{\partial p_j}\phi_l \tag{A.10}$$

This gives us J equations, each linear in the J tax levels:

$$a_1^j \tau_1 + ... + a_k^j \tau_k + ... + a_J^j \tau_J = b_j \qquad \forall j \in \{1, ..., J\} \tag{A.11}$$

Where a_k^j and b_m are defined as:

$$a_k^j = \frac{\frac{\partial h_k}{\partial p_j}}{\frac{\partial h_j}{\partial p_j}} \tag{A.12}$$

$$\beta_j = \phi_j + \sum_{l=1}^{M} \frac{\frac{\partial h_l}{\partial p_j}}{\frac{\partial h_j}{\partial p_j}} \phi_l \tag{A.13}$$

The a and β terms have an intuitive interpretation. a_k^j is the share of the reduction in overall consumption of good j that shifts to good m as a results of an increase in the price of good j. β_j is the overall reduction in externalities that results from the increase in the price of good j; this consists of a direct component, ϕ_j plus a (negative) leakage term, $\sum_{l=1}^{M} \frac{\partial h_l}{\partial p_j} / \frac{\partial h_j}{\partial p_j} \phi_l$.

This system can be written as:

$$\begin{bmatrix} a_1^1 & ... & a_J^1 \\ & ... & \\ a_1^j & ... & a_J^J \end{bmatrix} \begin{bmatrix} \tau_1 \\ ... \\ \tau_J \end{bmatrix} = \begin{bmatrix} b_1 \\ ... \\ b_J \end{bmatrix} \tag{A.14}$$

$$\boldsymbol{A\tau} = \boldsymbol{b} \tag{A.15}$$

$$\boldsymbol{\tau} = \boldsymbol{A}^{-1}\boldsymbol{b} \tag{A.16}$$

A.1.2. Heterogeneity and Leakage

Setup: N Heterogeneous consumers choose between M externality-generating goods and a numeraire, z. I denote individual i's consumption of good m as h_i^m. Each individual has an exogenous income μ_i. I assume that each consumer's utility is a function of their consumption of these M goods and a quasilinear numeraire, as well as other's consumption of these goods (which generate externalities and decrease i's utility): $U_i(h_1^1, ...h_1^M, ..., h_i^1, ...h_i^M, ...h_N^1, ...h_N^M) + z_i$.

As in section 2.3, a policymaker can choose tax levels for goods $j \in \{1, ..., J\}$ where $J < M$. I assume goods $k \notin \{1, ..., J\}$ are un- or under-taxed. I denote τ^j as the tax on good j.

The consumer's problem: Agent i maximizes utility over M goods $(h_i^1, ..., h_i^M)$ and their consumption of the numeraire good z_i.

$$\begin{aligned} max\{U_i(h_1^1, ...h_1^M, ..., h_i^1, ...h_i^M, ...h_N^1, ...h_N^M) + z_i\} \; st. \\ (p^1 + \tau^1)h_i^1 + (p^J + \tau^J)h^J + p^{J+1}h_i^{J+1} + ... + p^M h_i^M + z_i \leq \mu_i \end{aligned} \tag{A.17}$$

The first-order conditions for this problem are:

$$\begin{aligned} \frac{\partial U_i}{\partial h_i^j} &= \lambda(p^j + \tau^j) \qquad \forall \;\; j \in \{1, ..., J\} \\ \frac{\partial U_i}{\partial h_i^k} &= \lambda(p^k) \qquad \forall \;\; k \notin \{1, ..., J\} \\ \lambda &= 1 \end{aligned} \tag{A.18}$$

The planner's problem: I assume that the planner seeks to maximize aggregate welfare, $\sum_1^N (U_i + z_i)$. The planner's choice variables are tax levels $\tau^1 ... \tau^J$, which are applied to the taxable goods $j \in [1, J]$.

$$max\{\sum_i^N (U_i(h_1^1, ...h_1^M, ..., h_i^1, ...h_i^M, ...h_N^1, ...h_N^M) + z_i)$$

$$st. \qquad (p^1)\sum_i^N h_i^1 + ... + (p^J)\sum_i^N h_i^J + (p^{J+1})\sum_i^N h_i^{J+1} + ... + (p^M)\sum_i^N h_i^M + \sum_i^N z_i \leq \sum_i^N \mu_i \tag{A.19}$$

Assuming an internal solution, first-order condition wrt p^j (where $j \in [1, J]$) is:

$$0 = \sum_{i=1}^N \frac{\partial U_i}{\partial h_i^l}\frac{\partial h_i^l}{\partial p_j} + \sum_{i=1}^N \sum_{g \neq i}^N \frac{\partial U^i}{\partial h_g^1}\frac{\partial h_1^g}{\partial p_j} + ... + \frac{\partial U^i}{\partial h_g^M}\frac{\partial h_g^M}{\partial p_j} - p^1 \sum_i \frac{\partial h_i^1}{\partial p_j} - ... - p^M \sum_i \frac{\partial h_1^M}{\partial p_j} \tag{A.20}$$

Plugging in the consumer's first order conditions and solving for τ_j...

$$\tau_j = \frac{\sum_{i=1}^N \sum_g^N (\frac{\partial U^i}{\partial h_g^1}\frac{\partial h_g^1}{\partial p_j} + ... + \frac{\partial U^i}{\partial h_g^M}\frac{\partial h_g^M}{\partial p_j})}{\sum_{i=1}^N \frac{\partial h_i^j}{\partial p_j}} + \frac{\sum_{k \neq j}^J \frac{\partial h_i^k}{\partial p_j}\tau_k}{\sum_{i=1}^N \frac{\partial h_i^j}{\partial p_j}} \tag{A.21}$$

This expression for the optimal level of a given tax is equivalent to the equation for substitutes with homogeneous damages where each of the marginal damages have been replaced by a "Diamond" term which accounts for heterogeneity.

A.2. Additional Figures and Tables

FIGURE A.1 — TRAFFIC SENSORS IN THE BAY AREA

Figure A.1: This figure plots traffic sensors from the Caltrans Performance Measurement System (PeMS). Each sensor reports hourly vehicle count and speed data that are converted to traffic density (vehicles/lane/mile) using the fundamental equation of traffic flow. These traffic density readings are then used to assign congestion externalities to vehicle trips based on route and time of day, as described in sections 1.5 and 1.8.

[This space is left intentionally blank]

Figure A.2 — Mixed Logit Results

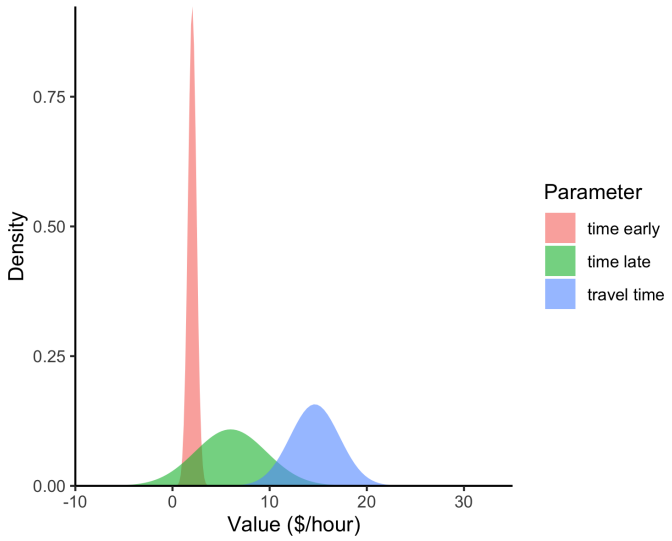


Figure A.2: This figure displays the mean and standard deviation from a random coefficients ("mixed") logit regression used to estimate the parameters in Equation 1.13. The data used in this regression are the FasTrak tolling microdata described in section 1.5: A panel of 32,104 individuals and 1,078,044 bridge crossings between June 15, 2010 and July 15, 2010, excluding weekends and holidays. The dependent variable is an indicator variable for whether an individual i elects to take a trip on route r at time of day h. *Travel time* is the travel time (in hours) that driver i would incur by traveling via route r at time h. *Time early* is the number of hours that that driver i would arrive before their ideal arrival time if they were to travel via route r at hour h; *time late* is defined analogously. *Price* is the toll that driver i would incur by traveling via route r at hour h. The mean and standard deviation of all time-related variables have been normalized relative to the coefficient on the *price* variable.

Table A.2 — Mixed Logit Results

Variable	**Mean**	**sd**
Time Early ($/hr)	2.033	0.175
Time Late ($/hr)	5.984	1.032
Travel Time ($/hr)	14.647	1.491
Price	1.000	0.135

Table A.2: This table displays the mean and standard deviation from a random coefficients ("mixed") logit regression used to estimate the parameters in Equation 1.13. The data used in this regression are the FasTrak tolling microdata described in section 1.5: A panel of 32,104 individuals and 1,078,044 bridge crossings between June 15, 2010 and July 15, 2010, excluding weekends and holidays. The dependent variable is an indicator variable for whether an individual i elects to take a trip on route r at time of day h. *Travel time* is the travel time (in hours) that driver i would incur by traveling via route r at time h. *Time early* is the number of hours that that driver i would arrive before their ideal arrival time if they were to travel via route r at hour h; *time late* is defined analogously. *Price* is the toll that driver i would incur by traveling via route r at hour h. The mean and standard deviation of all time-related variables have been normalized relative to the coefficient on the *price* variable. The mean and standard deviation of all time-related variables have been normalized relative to the coefficient on the *price* variable. All values are in 2010 dollars.

Figure A.3 — Bunching at Price Notches

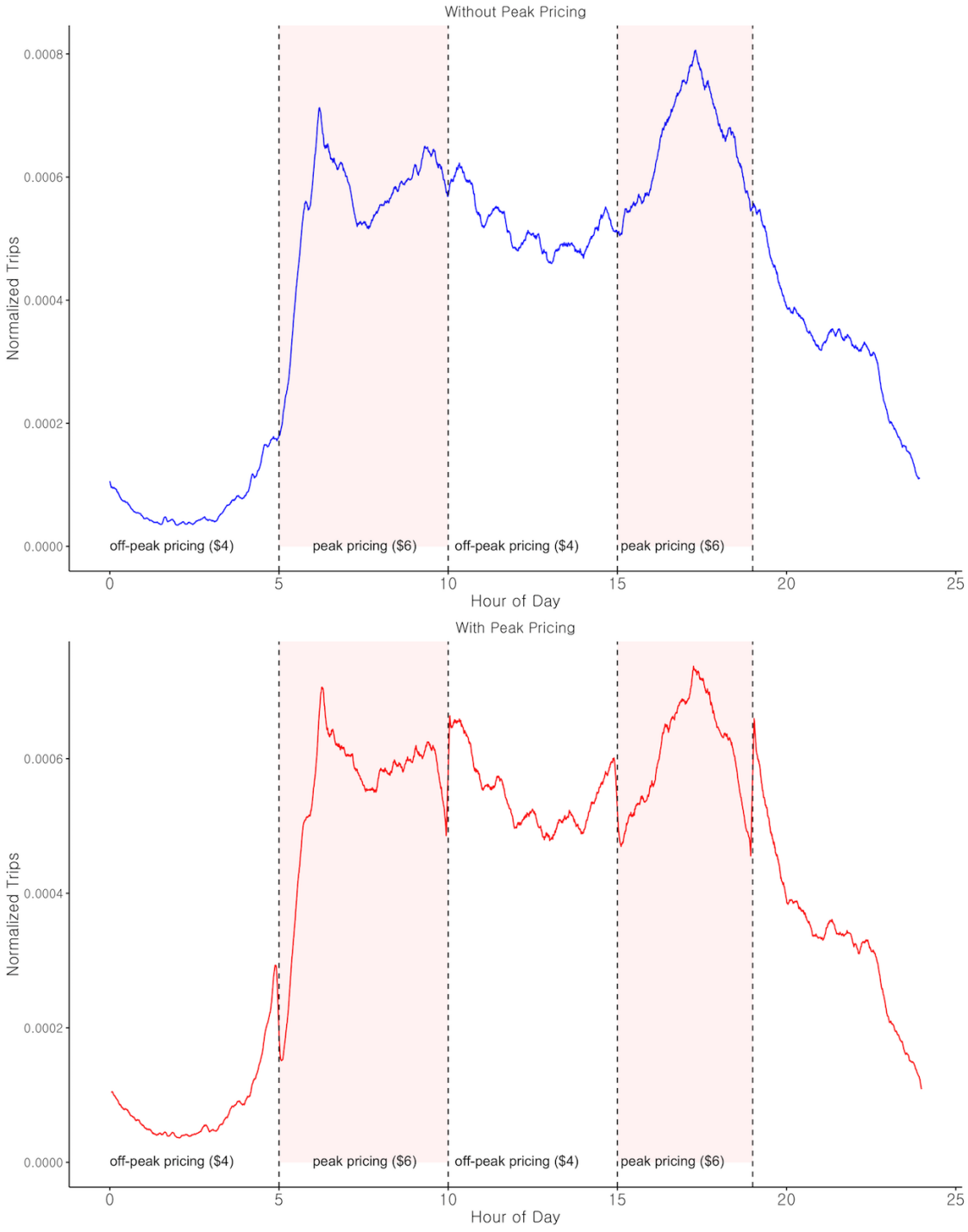


Figure A.3: This figure plots the density (the share of total daily crossings) of passenger vehicle trips crossing San Francisco's *Bay Bridge* in the 6 months before (blue) and 6 months after (red) the imposition of peak hour pricing on July 1, 2010. This plot excludes trips that use the carpool lane, as well as eligible electric vehicles, each of which faced a different pricing scheme. The red shaded regions demarcate times of day that were subject to peak-hour pricing after July 1, 2010.

Figure A.4 — Detail of Bunching at Price Notches

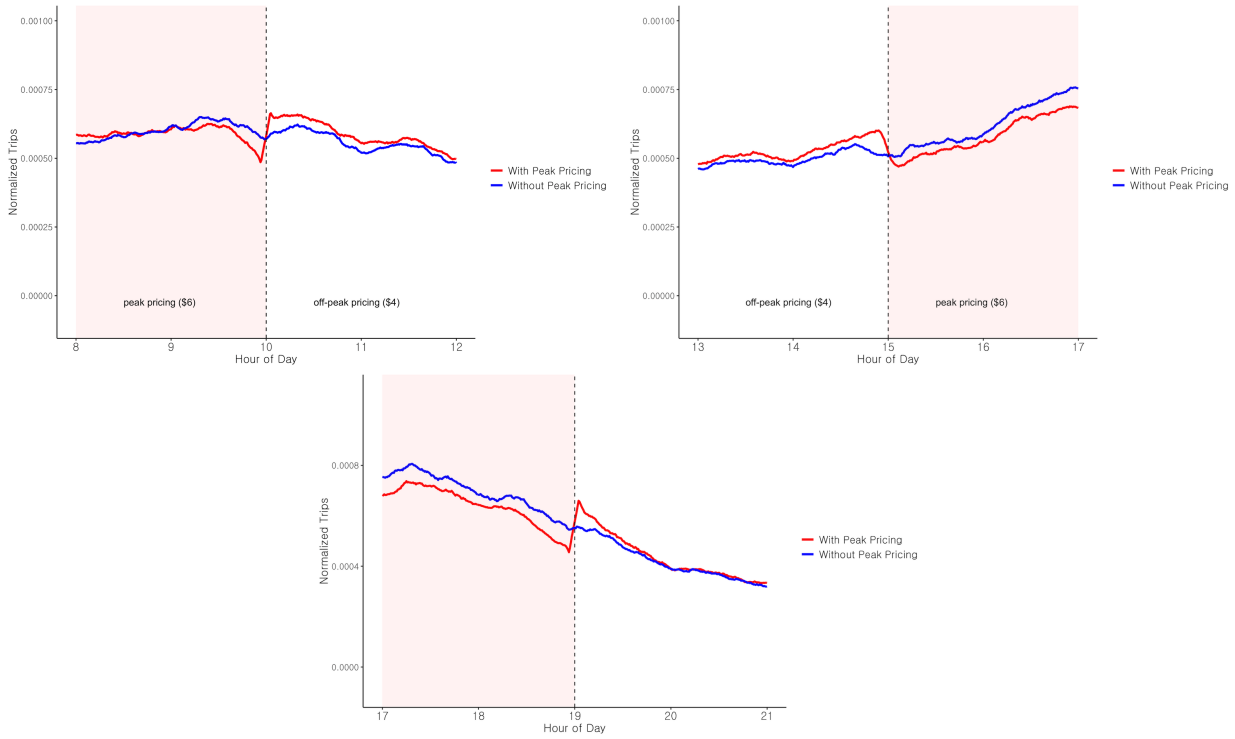


Figure A.4: San Francisco's *Bay Bridge* imposed peak hour pricing on July 1, 2010 (see Section 1.4). This figure plots the density of passenger vehicle trips crossing the Bay Bridge in the 6 months before (blue) and 6 months after (red) the imposition of peak hour pricing for the 10 a.m., 3 p.m., and 7 p.m. price notches. This plot excludes trips that use the carpool lane, as well as eligible electric vehicles, each of which faced a different pricing scheme. The red shaded regions demarcate times of day that were subject to peak-hour pricing after July 1, 2010.

Figure A.5 — Bunching in the Short and Long Run

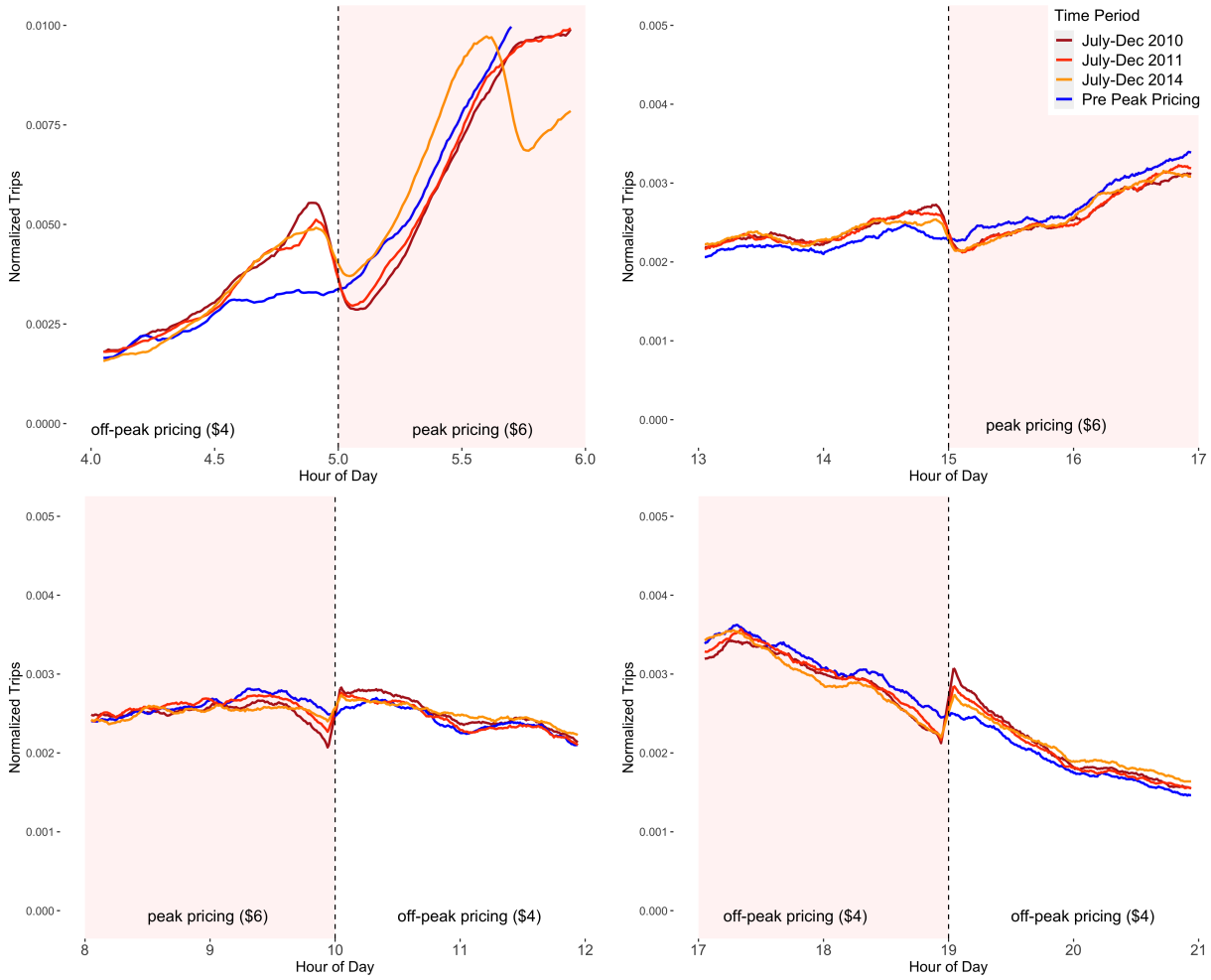


Figure A.5: San Francisco's *Bay Bridge* imposed peak hour pricing on July 1, 2010 (see Section 1.4). This figure plots the density of passenger vehicle trips crossing the Bay Bridge in the hours surrounding each of the price notches in January-July of 2010 (blue) against three other time periods: July-December of 2010 (brown), July-December of 2011 (red), and July-December of 2014 (orange). This plot excludes trips that use the carpool lane, as well as eligible electric vehicles, each of which faced a different pricing scheme. The red shaded regions demarcate times of day that were subject to peak-hour pricing after July 1, 2010.

Figure A.6 — Bunching in Electronic Tolls vs. Cash Tolls

Figure A.6: San Francisco's *Bay Bridge* imposed peak hour pricing on July 1, 2010 (see Section 1.4). This figure plots the density of passenger vehicle trips crossing the Bay Bridge in the hours surrounding each of the price notches in 2014 (records of cash payments are unreliable prior to 2014). The red line shows trips using an electronic FasTrak device; the grey line shows cash payments. This plot excludes trips that use the carpool lane, as well as eligible electric vehicles, each of which faced a different pricing scheme. The red shaded regions demarcate times of day that were subject to peak-hour pricing after July 1, 2010.

Figure A.7 — Simulated Travel Choices Under Peak-Hour Cordon Pricing in San Francisco

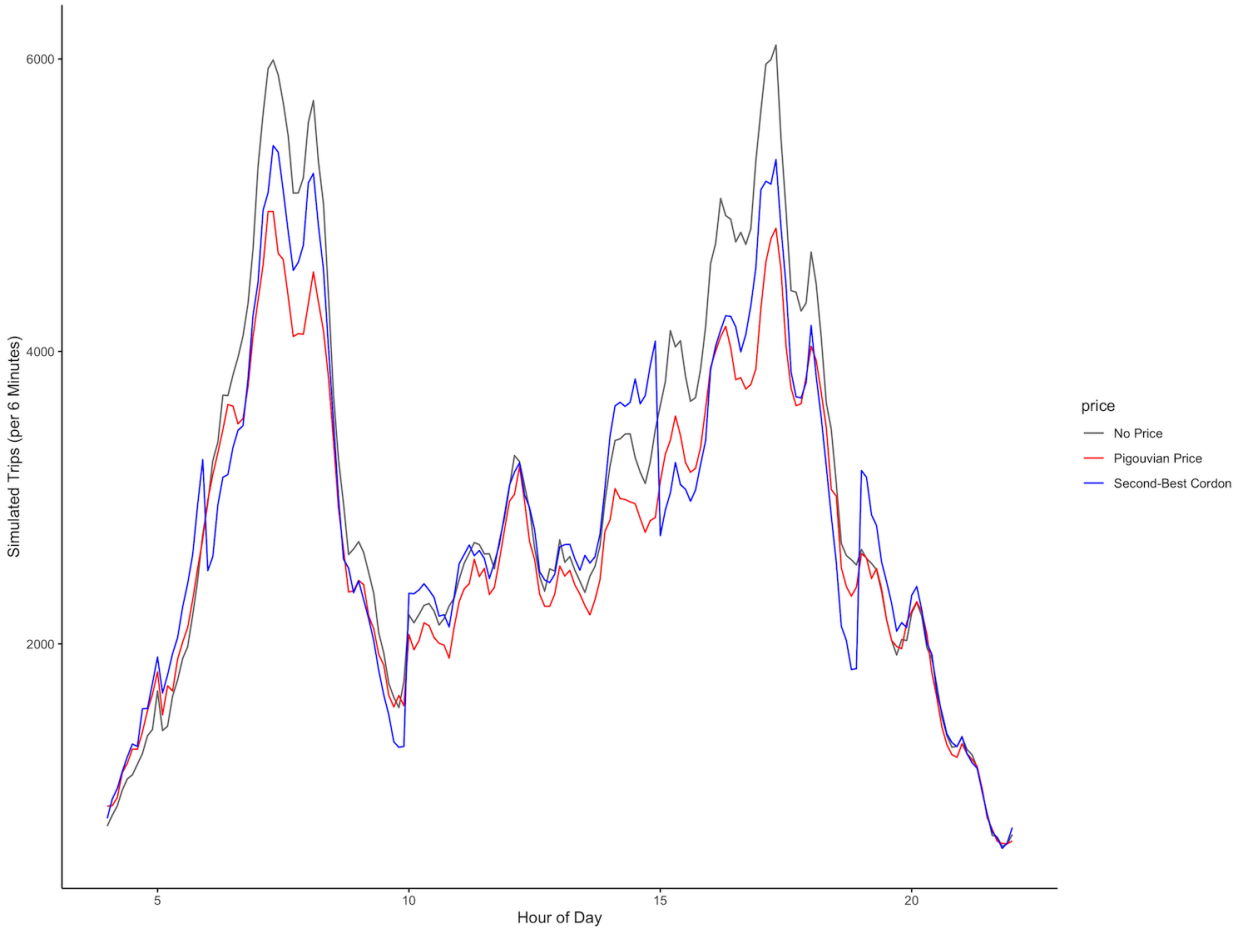


Figure A.7: In this figure I plot the number of trips that pass through or near the cordon under three simulations using the mixed logit model estimated in Table 1.2 of Section 1.7 together with the NHTS trip dataset described in Section 1.8. In each scenario, I predict 600,000 choices — roughly daily total of vehicle trips that pass through San Francisco's proposed cordon (San Francisco County Traffic Authority, 2021). The grey line plots predicted trips by time of day without any pricing (the status quo). The blue line plots trips under the first-best scheme where every trip a driver could choose (including non-cordon trips) would be priced according to it's marginal pollution and congestion externalities. The red line plots trips under the second-best optimal peak-hour cordon price from Figure 1.11. Note that all lines include both trips that cross through the cordon, and "detour" trips that circumvent the cordon.

Figure A.8 — Simulated Congestion Under Peak-Hour Cordon Pricing in San Francisco

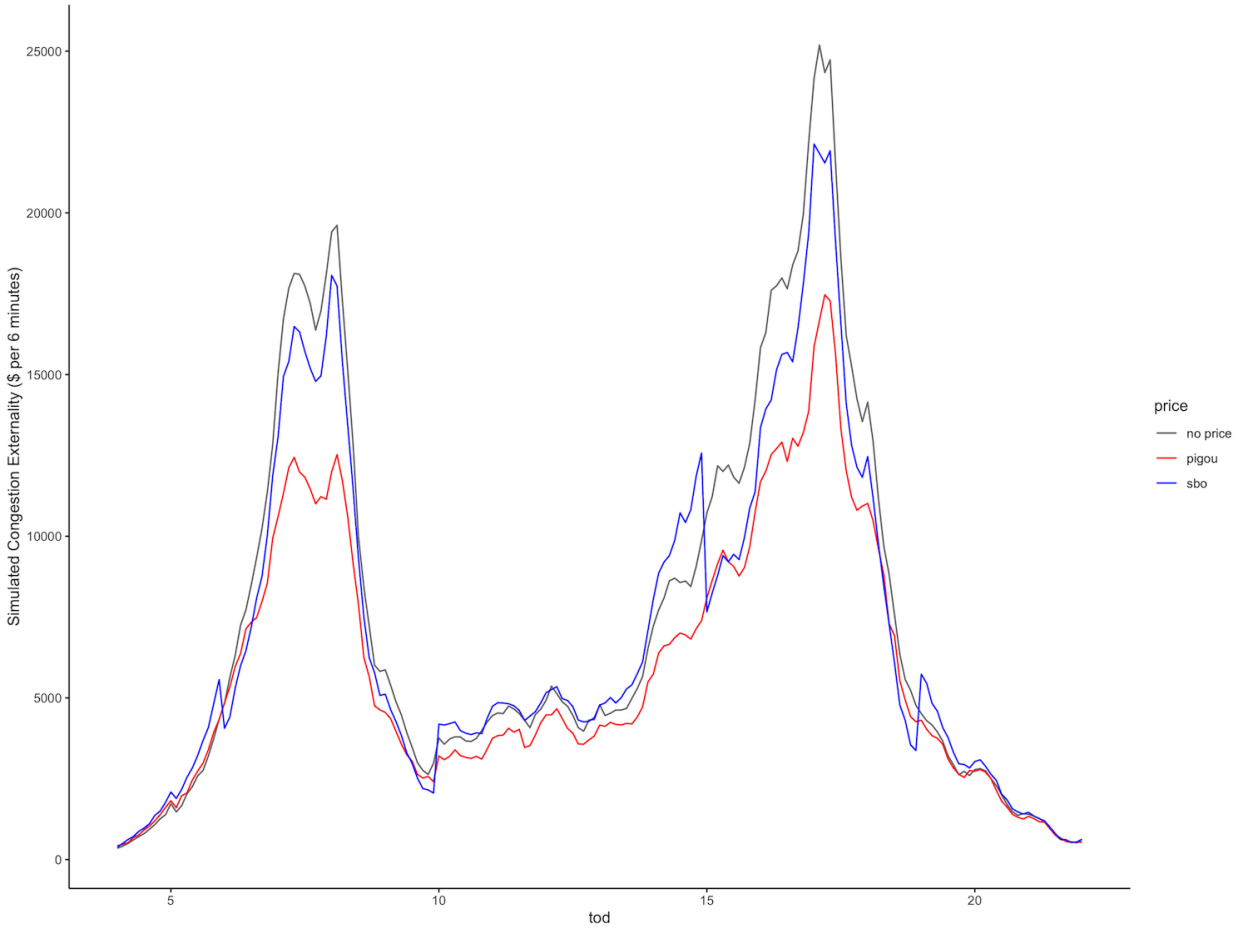


Figure A.8: In this figure I plot the total congestion externalities under three simulations using the mixed logit model estimated in Table 1.2 of Section 1.7 together with the NHTS trip dataset described in Section 8. In each scenario, I predict 600,000 choices — roughly daily total of vehicle trips that pass through San Francisco's proposed cordon (San Francisco County Traffic Authority, 2021). The grey line plots the sum of congestion externalities by time of day without any pricing (the status quo). The blue line plots congestion under the first-best policy where every trip a driver could choose (including non-cordon trips) would be priced according to it's marginal pollution and congestion externalities. The red line plots sum of congestion externalities under the second-best optimal peak-hour cordon price from Figure 1.11. Note that all lines include congestion from trips that cross through the cordon, as well as "detour" trips that circumvent the cordon.

Figure A.9 — Simulated Pollution Under Peak-Hour Cordon Pricing in San Francisco

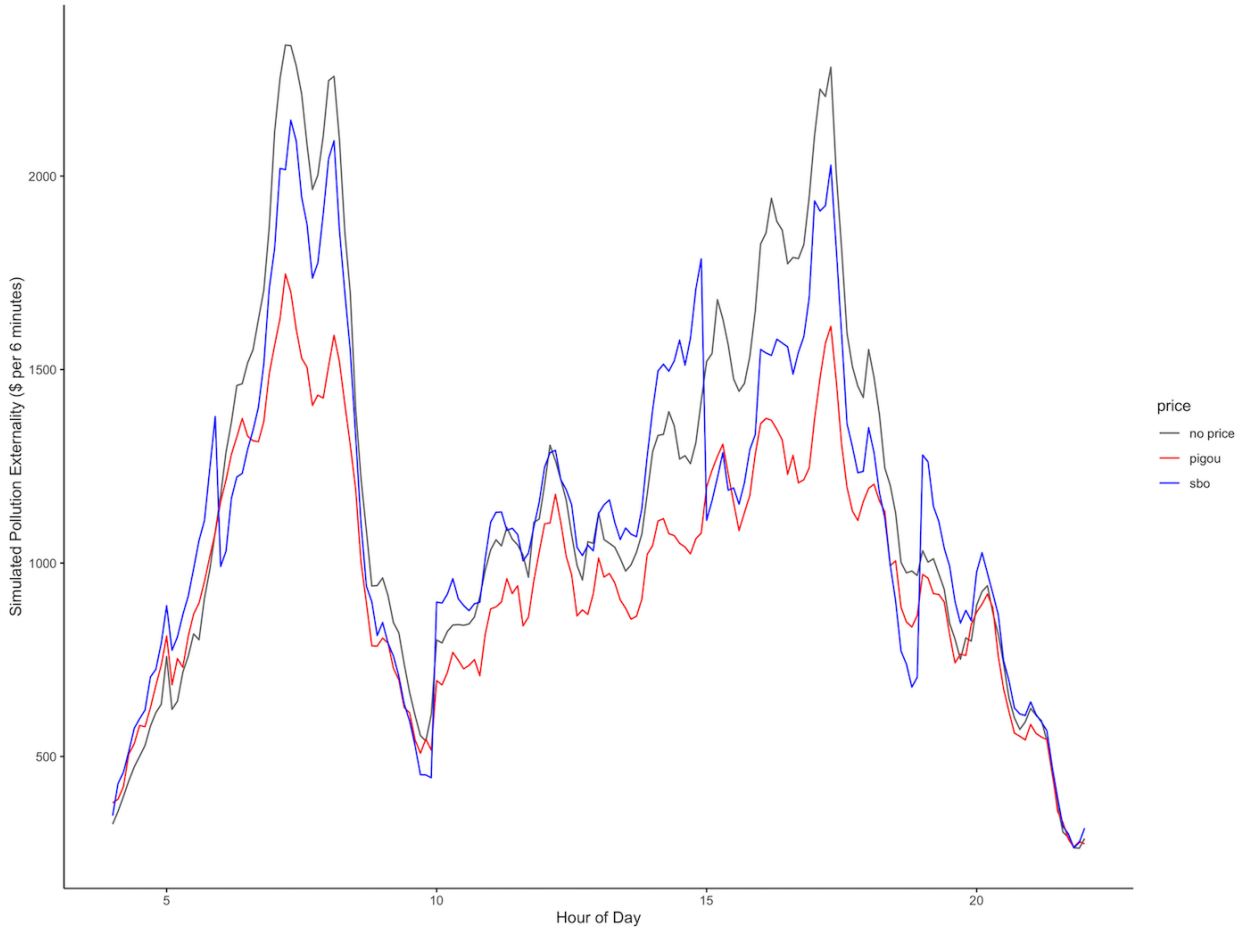


Figure A.9: In this figure I plot the total pollution externalities under simulations using the mixed logit model estimated in Table 1.2 of Section 1.7 together with the NHTS trip dataset described in Section 8. In each scenario, I predict 600,000 choices — roughly daily total of vehicle trips that pass through San Francisco's proposed cordon (San Francisco County Traffic Authority, 2021). The grey line plots the sum of pollution externalities by time of day without any pricing (the status quo). The blue line plots pollution externalities under the first-best policy where every trip a driver could choose (including non-cordon trips) would be priced according to it's marginal pollution and congestion externalities. The red line plots sum of pollution externalities under the second-best optimal peak-hour cordon price from Figure 1.11. Note that all lines include pollution from trips that cross through the cordon, as well as "detour" trips that circumvent the cordon.

A.3. Calculating Emissions Externalities

This section details the process of estimating emissions externalities for each trip in the FasTrak dataset.

The California Emissions Factor (EMFAC) fleet database reports average vehicle emissions rates (measured in grams per mile) by county. These data are stratified by vehicle fuel type, vehicle vintage, and vehicle travel speed. The EMFAC database reports the following pollutant species: particulate matter (PM2.5, or PM), nitrogen oxides (NO_x), nitrous oxide (N_2O), reactive organic compounds (ROC), ammonia (NH_3), carbon dioxide (CO_2), sulfur oxides (SO_2), and methane (CH_4). The data underlying EMFAC aggregates reflect state vehicle registrations and data from the California Bureau of Automotive Repair's (BAR) Smog Check database. For each FasTrak trip, I assign emission factors for each pollutant based on the average travel speed for that trip (see Appendix A.5) and the county where the FasTrak device is registered. The total emissions of any pollutant is the estimated emissions *rate* for that trip multiplied by the trip *length*.

To convert trip-level emissions to costs, I use social cost estimates from two sources. For local pollutants, I use damages predicted by the EAISUR model (Heo, Adams, and Gao, 2016), which combines a state-of-the-art chemical transport model together with estimates from the economics and epidemiology literatures to predict the cost of emitting pollution in different areas of the United States. For global pollutants, I use social damages from the US EPA. These pollutant values are listed in Table A.5.

Table A.5 — Social Costs of Vehicle Pollution in San Francisco

Pollutant	**Damage ($/Ton)**
$PM_{2.5}$	772,000
SO_2	65,800
NO_x	24,200
NH_3	1,24,000
CO_2	51
CH_4	1,500
N_2O	18,000
ROC	2,392

Table A.3: This table display the social costs of emitting 1 ton of various pollutants in San Francisco. Estimates of local pollutants (PM2.5, nitrogen oxides (NO_x), nitrous oxide (N_2O), reactive organic compounds (ROC), ammonia (NH_3), sulfur oxides (SO_2)) reflect annual averages from the EAISUR model (Heo, Adams, and Gao, 2016). Global pollutants (carbon dioxide (CO_2) and methane (CH_4)) are values used by the US EPA.

Table A.5 — Social Costs of Vehicle Pollution in Los Angeles

Pollutant	**Damage ($/Ton)**
$PM_{2.5}$	1,270,000
SO_2	44,750
NO_x	52,750
NH_3	825,750
CO_2	51
CH_4	1,500
N_2O	18,000
ROC	2,392

Table A.4: This table display the social costs of emitting 1 ton of various pollutants in Los Angeles. Estimates of local pollutants (PM2.5, nitrogen oxides (NO_x), nitrous oxide (N_2O), reactive organic compounds (ROC), ammonia (NH_3), sulfur oxides (SO_2)) reflect annual averages from the EAISUR model (Heo, Adams, and Gao, 2016). Global pollutants (carbon dioxide (CO_2) and methane (CH_4)) are values used by the US EPA.

Table A.5 — Social Costs of Vehicle Pollution in New York City

Pollutant	**Damage ($/Ton)**
$PM_{2.5}$	$1,270,000$
SO_2	$44,750$
NO_x	$52,750$
NH_3	$825,750$
CO_2	51
CH_4	$1,500$
N_2O	$18,000$
ROC	$2,392$

Table A.5: This table display the social costs of emitting 1 ton of various pollutants in New York City. Estimates of local pollutants (PM2.5, nitrogen oxides (NO_x), nitrous oxide (N_2O), reactive organic compounds (ROC), ammonia (NH_3), sulfur oxides (SO_2)) reflect annual averages from the EAISUR model (Heo, Adams, and Gao, 2016). Global pollutants (carbon dioxide (CO_2) and methane (CH_4)) are values used by the US EPA.

A.4. Bunching Estimator

This appendix contains details of the bunching estimators used as a second empirical approach to recovering scheduling elasticities (see Section 1.6). The following two equations are bunching estimators that do, and do not account for changes in travel times for bunches, respectively:

$$\gamma_e = \frac{\beta \Delta p + \alpha \Delta T}{B/((1-a) f_0(h^*))}$$

$$\gamma_e = \frac{\beta \Delta p}{B/((1-a) f_0(h^*))}$$

Table A.6 shows estimates of each of the component parts of these estimators for the 5 a.m. price notch. The change in price (Δp) is the same ($2) for all notches. The excess mass (B) is the integral of the difference in densities in the period (half an hour) prior to the imposition of peak hour pricing. Following Kleven and Waseem (2013), I use the comparison of the pre and post July 2010 density within the 5 minutes after the beginning of peak-hour pricing to identify the fraction of unresponsive individuals (a, 76%). As an approximation for the change in travel time (ΔT), I use TomTom's Historic Traffic Stats to compute the difference in average travel travel times between 5:00 a.m. and 6:00 a.m. (see Figure A.10 below) for FasTrak drivers using the Bay Bridge. The components of bunching estimators for the other three notches (10 a.m., 3 p.m., and 7 p.m.) follow this same procedure.

Table A.6 — Bunching Estimator for Scheduling Costs (Shifting Earlier, 5 a.m.)

Parameter	**Estimate**
Fraction Unresponsive (a)	0.76058
Excess Mass at Notch (B)	0.00208
Baseline Density at Notch	0.00019
Mean Schedule Cost without Friction ($/hour)	18.65659
Mean Schedule Cost accounting for Frictions ($/hour)	4.46673
Mean Schedule Cost accounting for Frictions and Travel Time ($/hour)	6.19461

Table A.6: Rows 1-3 of this table show estimates of parameters used to infer scheduling costs from the additional density of trips just after the end of peak-hour pricing on San Francisco's Bay Bridge (equation 1.21). Rows 4-6 show estimates of scheduling costs. In Row 4, I calculate the naive average scheduling cost under the assumption that there are no optimization frictions. In row 5, I use the estimated fraction of non-responsive individuals from row 1 to account for optimization frictions. In row 6, I also account for the difference in travel times for drivers who reschedule their trips to avoid peak-hour pricing.

Figure A.10 — Travel Times in the Vicinity Price Notches

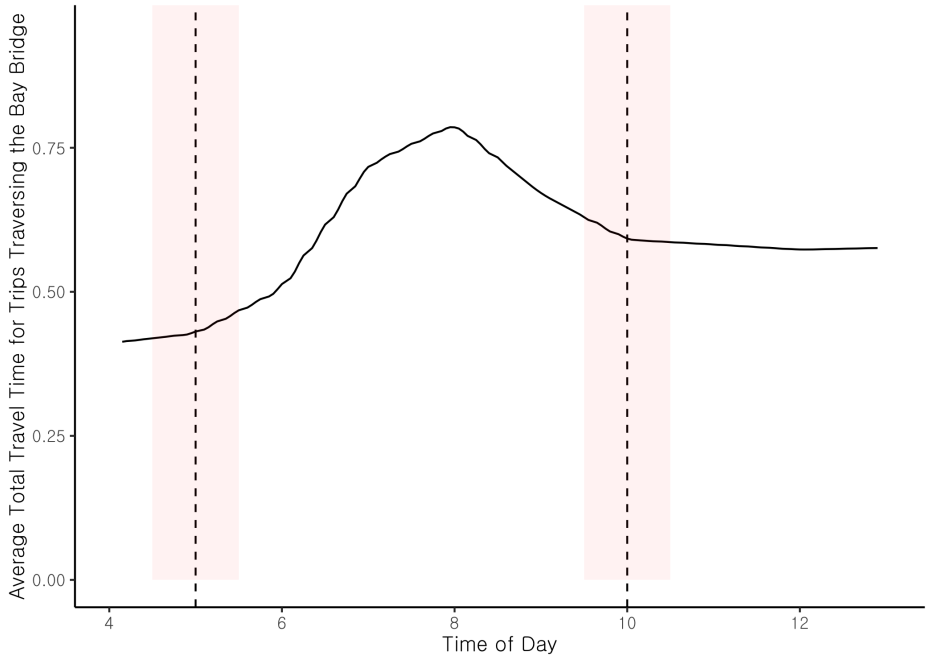


Figure A.10: This figure plots average travel times for trips traversing the Bay Bridge during the morning hours. The average travel times in this figure were calculated by 1) identifying all drivers that primarily use the Bay Bridge and b) using TomTom Historic Traffic Stats to calculate travel times for each individual, for each hour of day as described in Section 1.5. The red shaded area represent the approximate range where individuals adjust in response to the imposition of peak-hour pricing, according to FasTrak toll data. The relatively flat profile of travel times in the price notch neighborhood suggests that the first-order decision facing drivers who travel at this time of day is between price and scheduling costs, as opposed to changes in total travel time. As shown in Table A.6, estimates that account for differences in travel times in the bunching estimators are roughly 30% larger that estimates that ignore differences in travel times.

[This space is left intentionally blank]

A.5. Imputing Travel Times

Travel times, T_i, are not directly observed for FasTrak trips, and therefore must be imputed. In this appendix, I describe the process for inferring travel times, $T(h_i, r_i)$, for each trip in each individual's choice set.

The choice set of any individual consists of all *bridges* $\in$ {*Dumbarton Bridge, San Mateo Bridge, Bay Bridge, Richmond Bridge*} at all *times of day* $\in$ {4.0, 4.2, ..., 22}. A *trip* in this choice set constitutes a *bridge-time* pair, (h_i, r_i). I estimate travel times for each trip in each individual's choice set in three steps:

***Step 1*: Infer the distribution of endpoints**. The FasTrak tolling data include information about the bridges used, as well as the home zip code associated with each FasTrak device. Before calculating travel times using historic traffic data, I must make inferences about the missing endpoints for each driver. To do so, I use survey data from the 2010-2012 California Household Travel Survey (CHTS). This survey constitutes a representative sample of Bay Area commuters, an contains detailed information on the driving habits of respondents. To generate a probability distribution of "work" endpoints for each individual, I subset the CHTS survey data to trips that match based on home city and bridge used. The Bay Area is relatively unique in that it is a large metropolitan area that consists of many small cities. The 29 "cities" that serve as termini for travel time estimation are plotted in Figure A.11.

***Step 2*: Calculate travel times**. I use TomTom's Historic Traffic Stats to calculate the travel times. This traffic database contains detailed historic traffic data collected from TomTom devices, as well as data that TomTom purchases from other GPS providers. For each FasTrak device in the sample, I calculate the travel time between the device's home city and each of the end cities assigned positive probability for that device in *Step 1*. Importantly, I estimate travel times for both trips that were taken, as well as counterfactual trips that used a different bridge or were taken at a different hour of day.

***Step 3*: Aggregate travel times by bridge and time of day**. Lastly, I collapse the distribution of possible travel times within each *bridge-time* pair by the probability weights from *Step 1*. That is, the result of *Step 2* contain travel times for each choice (a *bridge-time* pair), for each device based on possible "work" locations associated with that device. *Step 3* then assigns a single travel time to each bridge-time choice for each device by taking the probability-weighted sum of the travel times associated these possible work locations, where

those probability weights are based on the CHTS survey data (*Step 1*).

The result of Steps 1-3 is a data set that contains estimated travel times for each trip taken by each device, as well as the travel times that a driver would have faced for each trip had they taken it at a different hour of day or using a different bridge.

Figure A.11 — TomTom Traffic Segments

Figure A.11: This figure plots the coverage of the historic travel time data purchased from TomTom (in red) together with the 29 most populous cities in the Bay Area. These road segments were selected using Google Maps suggested driving points between the origin and destination cities. These traffic data report the average weekday travel times for passenger vehicles traveling along each segment of road, by hour of day, for the year prior (July 1, 2009 - July 1, 2010) and the year following (July 1, 2010 - July 1, 2011) the 2010 adjustment to Bay-Area bridge tolls.

A.6. Equilibrium Considerations

The second-best cordon price results presented in Section 1.8 reflect social damages calculated using traffic conditions in untaxed equilibrium. Consistent with the literature on externality taxation, the second-best tax formula in Section 1.2 phrases optimal taxes as a function of externalities *at the optimum*. As shown in figures 1.4 and A.15, the marginal damages associated with driving are not constant in traffic density/speed, meaning that in general, damages at the taxed equilibrium will be different (lower) than those observed in the untaxed equilibrium. Whether the difference between marginal damages calculated at versus away from the optimum is a first-order concern depends on the slope of the marginal damages function and the responsiveness of drivers to taxation.

In this appendix, I simulate changes in traffic density under taxation to estimate a lower bound for the second-best optimal cordon prices in San Francisco. Specifically, I iteratively calculate traffic density, driver choices, and taxes until I reach a fixed point where driver's decisions under a given tax vector, τ^*, imply traffic densities (and associated externalities) such that applying equation 1.10 to these conditions again yields τ^*.

This algorithm is as follows:

Let $\boldsymbol{\tau_0}$ be the cordon taxes calculated using equation 1.10 (the second-best tax formula from the theory section) under current traffic conditions, and let $\boldsymbol{\phi_0}$ be the externalities under current conditions, as described in Section 1.8.

Repeat the following steps until the optimal cordon taxes calculated in any two subsequent iterations ($\boldsymbol{\tau_n}$ and $\boldsymbol{\tau_{n+1}}$) meet some arbitrary element-wise threshold for convergence: $|\tau_n^h - \tau_{n+1}^h| < \epsilon$, where τ_n^h is element h of the tax vector calculated in step n.

In any iteration, n:

Step 1. Use the NHTS dataset described in Section 1.8 to simulate 600,000[1] driver choices under $\boldsymbol{\tau_{n-1}}$.
In the first iteration, use $\boldsymbol{\tau_0}$, defined above.

[1] As per the San Francisco County Traffic Authority, roughly 600,000 vehicle trips cross San Francisco's proposed cordon daily.

Step 2. Re-scale the hourly sensor-level road densities by comparing the simulated number of trips that would pass over a given sensor in a given hour under the status quo to the number of trips that would pass over a given sensor in a given hour under the simulation from *Step 1.*

Step 3. Re-estimate the social damages associated with each trip according to the updated hourly traffic densities from *Step 2.* Call these updated damages $\boldsymbol{\phi_n}$. This details of assigning congestion externalities to routes are covered in Section 1.8.3.

Step 4. Apply equation 1.10 (the second-best tax formula from the theory section), using the updated damages, $\boldsymbol{\phi_n}$. Define this tax vector as $\boldsymbol{\tau_n}$.

Figure A.12 plots the results of applying this algorithm to cordon pricing in San Francisco using a convergence threshold of \$0.01. The initial points (iteration one) are the taxes calculated with trip-level damages that reflect current traffic conditions, and are therefore equivalent to the results shown in Section 1.8 (see row 1 of Table 1.6). After 9 iterations of recalculating traffic density and taxes, the morning and evening converge to \$1.59 and \$1.80, respectively.

The fixed point from this exercise constitutes a lower bound because it ignores "induced demand," or "rebound," that is, marginal drivers who would have chosen not to take a trip in the absence of road pricing, but choose to take the trip under road pricing due to lower travel times. For any step $n > 1$ in the above algorithm, induced demand would imply traffic densities higher than those estimated by the discrete choice model (Duranton and Turner, 2011). Induced demand would therefore attenuate the difference in traffic conditions between taxed and untaxed equilibria. Optimal taxes that take into account endogenous externalities therefore lie between the results presented in Section 1.8 and the fixed point calculated in this appendix.

Figure A.12 — Bounding Equilibrium Effects

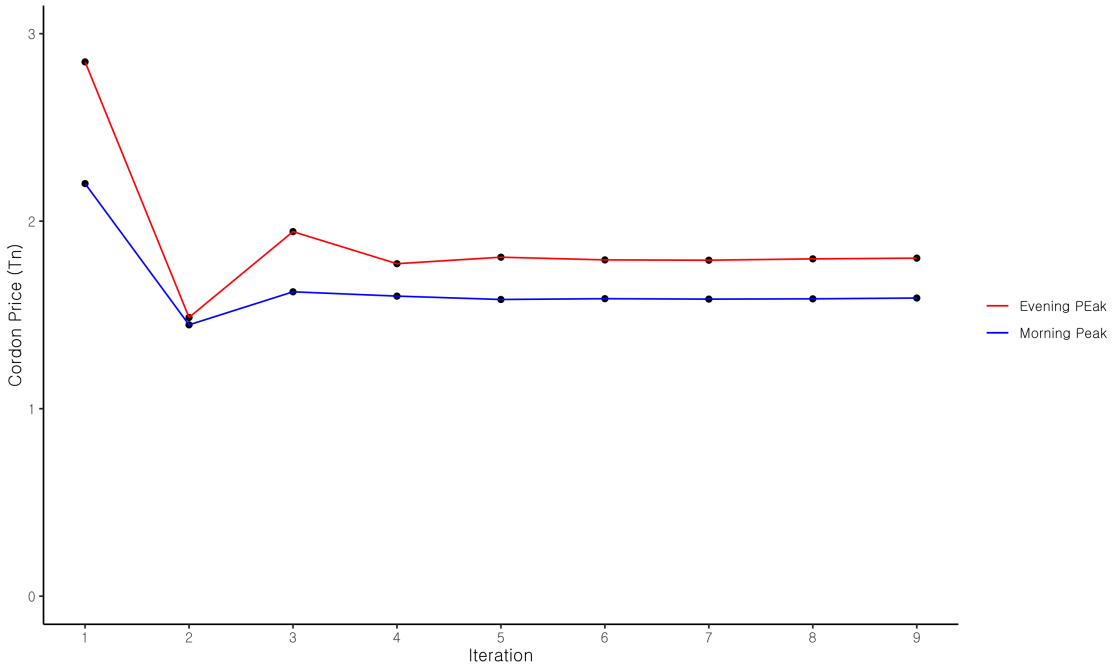


Figure A.12: Per-mile driving externalities are larger under denser traffic conditions (Yang, Purevjav, and Li, 2020). As a result, trip-level traffic externalities calculated using untaxed traffic conditions may overestimate optimal taxes. This figure displays the results of the simulation exercise where I iteratively calculate traffic density, driver choices, and taxes until reaching a fixed point where driver's decisions under a given tax vector, τ^*, imply traffic densities (and associated externalities) such that applying equation 1.10 yields τ^*. This optimal tax contains two elements: morning (blue) and evening (red) peak hour prices. The fixed point in this exercise is a lower bound for the second-best peak-hour cordon prices in San Francisco because it ignores "rebound," or "induced demand" —– drivers adding other trips or shifting from other modes in response to the improved traffic conditions under taxation.

A.7. Congestion Pricing and Accidents

In a manner similar to congestion and pollution externalities, the decision to drive imposes external accident risk on other drivers. Anderson and Auffhammer (2014) show that this externality relies crucially on vehicle weight, and exceeds congestion and pollution externalities for the average US driver.

Large accident externalities for the average US driver, however, may not translate to higher optimal cordon prices because of differences in the risks of accidents in urban vs. rural areas. Empirical studies of the impact of congestion charges on accidents suggest that the value of accident reductions are several orders or magnitude smaller than pollution and congestion externalities. Green, Heywood, and Navarro (2016), for example, find that the London cordon zone reduced overall accidents by 35%, and fatal accidents by 25 to 35%. Because of the relatively low number of fatal auto-related deaths in London, however, the authors value these safety improvements at just £28 million annually. For comparison, Leape (2006) estimates the congestion benefits from London's cordon zone were estimated at £230 million annually. Similarly, Percoco (2016) finds that while Milan's Cordon Zone reduced overall traffic accidents by 16 to 18%, there was no detectable impact on fatal accidents. Valuations of associated benefits are therefore dominated by the roughly $3 billion in reduced pollution and congestion externalities (Gibson and Carnovale, 2015).

The relatively small impact of congestion pricing on severe accidents may reflect the fact that many of the main risk factors severe traffic accidents — high traffic speeds, drinking and driving, and nighttime driving — are not well targeted by cordon zones. Relatedly, driving in cities in the US and Europe tends to be relatively safe overall, making it straightforward to put bounds on the accident-related benefits that may accrue from congestion pricing.

In San Francisco, for example, there are 20 to 30 fatal accidents (including pedestrian fatalities) each year (City of San Francisco, 2021). Under a $10 million value of a statistical life, reducing traffic fatalities in San Francisco by 30% would be worth roughly $90 million dollars —– an order of magnitude smaller than my estimated of the combined congestion and pollution benefits associated with cordon pricing in San Francisco. All indicators suggest that a cordon zone would fall well short of this mark. During 2020, for example, the number of traffic fatalities (31) did not fall amid the 30% pandemic-related decrease in Bay-Area traffic (City of San Francisco, 2021; Savidge, 2021).

Together, these pieces of evidence suggests that it is unlikely that accounting for accident externalities would substantively change the conclusions in this paper.

A.8. Interactions with Existing Taxes and Revenue Requirements

In this appendix, I cover the interaction between road pricing and existing environmental policies, as well as the literature on whether governmental revenue requirements impact the optimal Pigouvian tax.

A.8.1. Accounting for Existing Environmental Taxes

Broadly speaking, in the presence of existing Pigouvian taxes the optimal level for an *additional* tax covers the difference between the marginal damages associated with consumption and the existing corrective tax. It is therefore important to account for existing environmental policies that act as a tax on driving when calculating optimal Pigouvian road prices.

There are a number of State and Federal policies that regulate vehicle-related local pollution emissions in California. These policies largely fall into two categories: tailpipe emissions regulations (e.g., catalytic converter requirements) and fuel content regulations (e.g., volatile organic compound regulations). Below, I use a simple model to demonstrate that these two types of policies have different implications for designing an additional tax to internalize remaining externalities associated with driving. Regulations that impact vehicle costs should *not* be taken into account when calculating optimal road prices. The costs of fuel content regulations, however, should be subtracted from road prices to the extent that these regulations lead to higher per-mile driving prices.

Existing policies that impact vehicle cost:

Consider a representative household with exogenous income I that consumes two goods, driving x and a quasilinear numeraire good z. Driving is associated with an externality, $\phi(a)$. The per-mile magnitude of this externality can be abated (a) on the assembly line at cost $c(a)$. I assume that ϕ_a and c_a are each differentiable, with $c'(a) > 0$ and $\phi`(a) < 0$. The planner's problem is to choose an abatement level, a and a diving level x to maximize total welfare:

$$W = u(x) + z - \phi(a) \cdot x - c(a) \quad \text{s.t.} \quad I \geq z - p \cdot x$$

The Lagrangian associated with this maximization problem is:

$$\mathcal{L} = u(x) + z - \phi(a) \cdot x - c(a) + \lambda(I - z - p \cdot x)$$

The first-order conditions for an interior solution to this problem are:

$$\lambda = 1$$
$$u'(x) = \phi(a) + p$$
$$\phi'(a)x = c'(a)$$

These conditions imply that the planner equates marginal abatement costs and marginal abatement benefits, and (separately) equates marginal driving costs and marginal driving benefits. The fact that abatement costs do not enter directly into the first order condition for x implies that if a is set at some exogenous level, the policymaker would ignore the abatement cost when choosing the optimal level of driving, only weighing the utility of driving against the externalities that remain after abatement. I therefore ignore the costs of environmental policies that impact vehicle prices (e.g., requirements for catalytic converters) when calculating the level of "unpriced" externalities for drivers.

Existing policies that impact fuel cost:

Now consider the same consumer model, but the per-mile magnitude of this externality can be abated by altering fuel content at cost $c(a) \cdot x$. That is, the total abatement cost now depends on the amount of driving, x.

Again consider a policymaker who maximizes total social welfare, W:

$$W = u(x) + z - (\phi(a) - c(a)) \cdot x; \quad \text{s.t.} \quad I \geq z - p \cdot x$$

The Lagrangian associated with this maximization problem is:

$$\mathcal{L} = u(x) + z - (\phi(a) - c(a)) \cdot x + \lambda(I \geq z - p \cdot x)$$

The first-order conditions with respect to x and a are:

$$\lambda = 1$$
$$u'(x) = \phi(a) + c(a) + p$$
$$\phi'(a) = c'(a)$$

As above, these first-order conditions imply that the planner equates marginal abatement costs and marginal abatement benefits, and equates marginal driving costs and marginal driving benefits. The crucial difference in this case is that the marginal cost of driving now includes abatement costs. As a result, the social planner will still weight these costs when setting optimal road prices.

The results in the body of this paper are not adjusted for existing environmental policies that impact the variable cost of driving, namely fuel content regulation. Auffhammer and Kellogg (2011) estimate that fuel content regulations in California cost roughly 12 cents (in 2020 dollars) per gallon. If an average trip crossing San Francisco's cordon boundary travels roughly 10 miles per hour and has a fuel efficiency of 20 miles per gallon, the second-best optimal prices in this paper adjusted for pre-existing fuel regulation would be roughly $0.06 lower than the results shown in Section 1.8.

A.8.2. Accounting for Government Revenue Requirements

The stylized models above raise the question of whether *any* policy that increases the per-mile cost of driving about the competitive equilibrium should be accounted for when calculating optimal road prices. Work by Kopczuk (2003) and Jacobs and De Mooij (2015) suggests that optimal taxation and Pigouvian taxation are separable problems: The calculation of optimal road prices should not take into account taxes that exist as a result of governments balancing the distortions of various revenue sources.

As noted by Jacobs and De Mooij (2015), however, this argument relies on the fact that the marginal cost of public funds is one in an optimal tax system. If the marginal cost of public funds is *not* one, then the optimal second-best Pigouvian tax could be higher or lower than a tax set equal to marginal social damages. Absent strong evidence that the marginal cost of public funds is above or below one, I assume that the marginal cost of public funds is one in the San Francisco Bay Area, and therefore do not adjust optimal road prices to reflect their interactions with the tax system. As anecdotal evidence of this assumption, note that California state and local ballot initiatives frequently feature direct votes on taxation, bond issuance, and spending decisions. It is plausible that this low barrier to public finance reform allows California's tax code to reflect citizen's preferences for public goods and redistribution more accurately than do tax codes regions without ballot initiatives.

A.9. Assessing External Validity with the NHTS

The appropriateness of using of the driving demand model estimated using data from the San Francisco Bay Area (see Section 1.7) to cordon pricing in other cities depends on whether trips taken in other cities are similarly substitutable, and whether similar correlations between trip-level externalities and price responsiveness are present. In this appendix, is use data from the 2017 National Household Transportation Survey (NHTS) to investigate these relationships for two other US cities — New York and Los Angeles — that are currently considering implementing congestion pricing. I further investigate external validity in Appendix A.10, where I use public transit data from the Bay Area to examine whether the price-responsiveness of driving trips differs based on the availability of public transit.

Broadly, NHTS data suggest that the relevant relationships in each of these cities are similar to those in San Francisco. Drivers appear similarly able to shift trips temporally. Figure A.13, for example, shows that similar fractions of drivers report flexible work schedules in each of these cities. Figure A.14 shows that likelihood of a given trip being flexible varies in New York and Los Angeles in a manner similar to the within-day variation in San Francisco. Figures A.15 through A.17 provide suggestive evidence that the way that externalities generated by driving — congestion and pollution — vary with price responsiveness in New York and Los Angeles is similar to the way that these externalities vary with price responsiveness in San Francisco. In each city, drivers who "agreed" or "strongly agreed" that gasoline prices impacted their decision to drive were modestly more likely to drive an older, more polluting vehicle. Similarly, drivers that report being more responsive to gas prices report driving along more congested routes, measured as the difference in reported commute time with vs. without traffic.

Figure A.13 — Schedule Flexibility by Metro Area

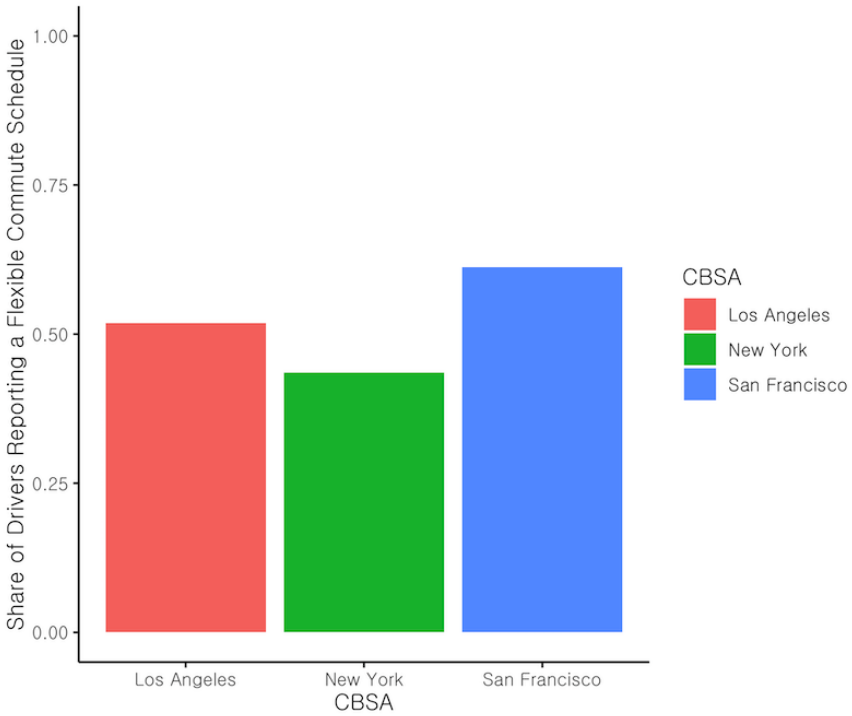


Figure A.13: This figure plots the share of drivers who report having a flexible work schedule by metro area, according to the 2017 National Household Transportation Survey.

Figure A.14 — Schedule Flexibility by Time of Day

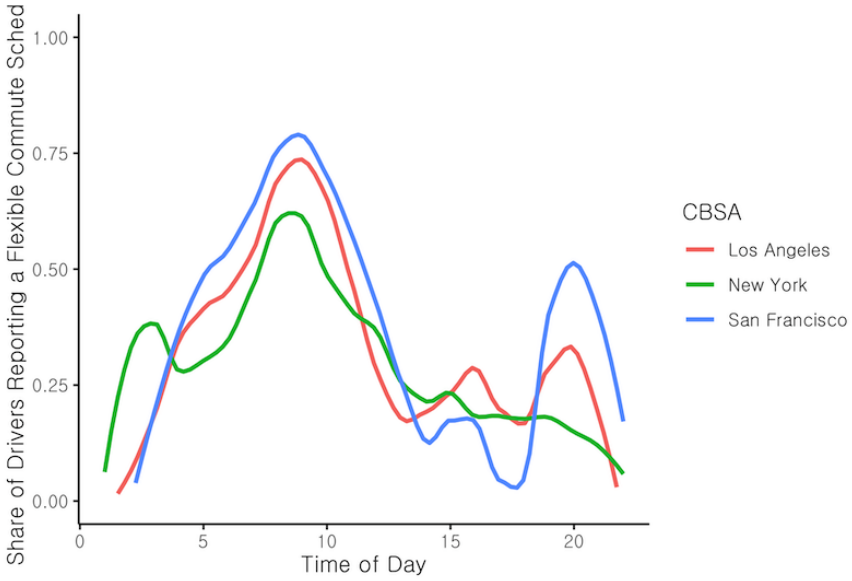


Figure A.14: This figure plots the share of drivers who report having a flexible work schedule by time of day and metro area, according to the 2017 National Household Transportation Survey.

Figure A.15 — Emissions Factors vs. Gas Price Responsiveness

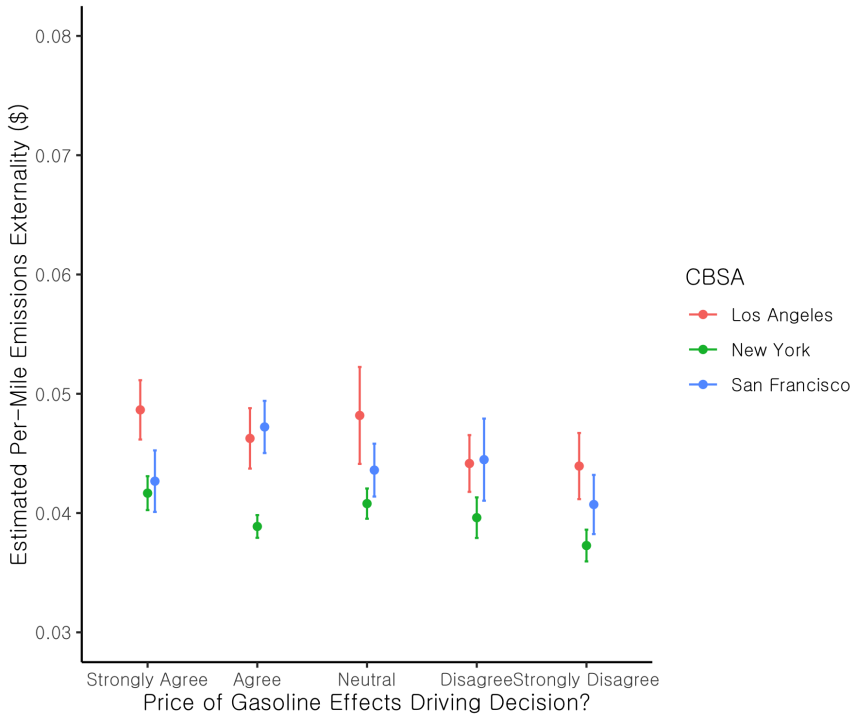


Figure A.15: This figure plots estimates emissions factors of vehicles in the 2017 National Household Transportation Survey against vehicle owners' self-reported responsiveness of driving demand with respect to gasoline prices. Emissions factors reflect vehicle age and fuel type.

Figure A.16 — Vehicle Age vs. Gas Price Responsiveness

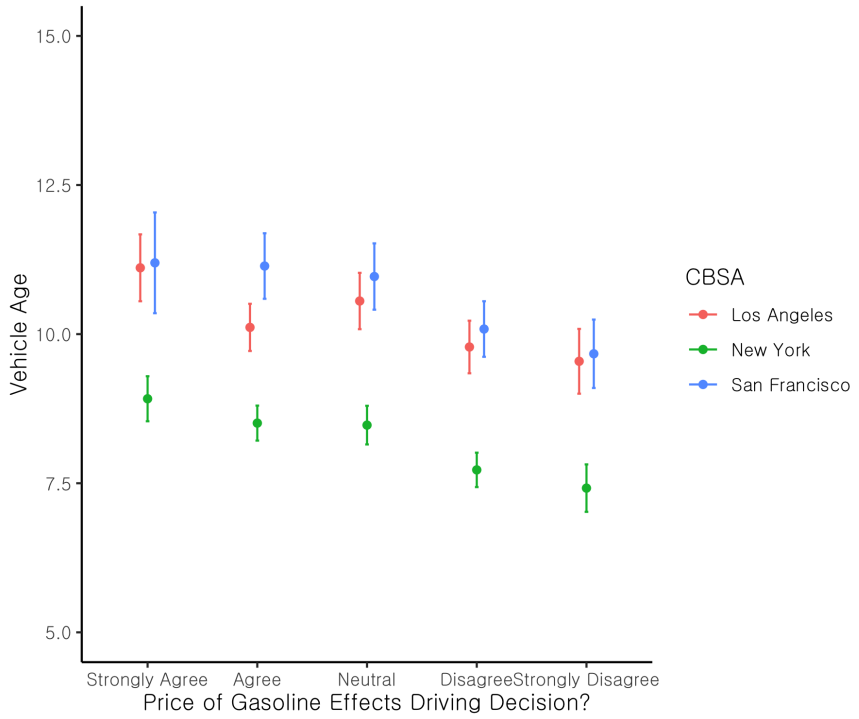


Figure A.16: This figure plots vehicle age against vehicle owners' self-reported responsiveness of driving demand with respect to gasoline prices using data from the 2017 National Household Transportation Survey.

Figure A.17 — Congestion vs. Gas Price Responsiveness

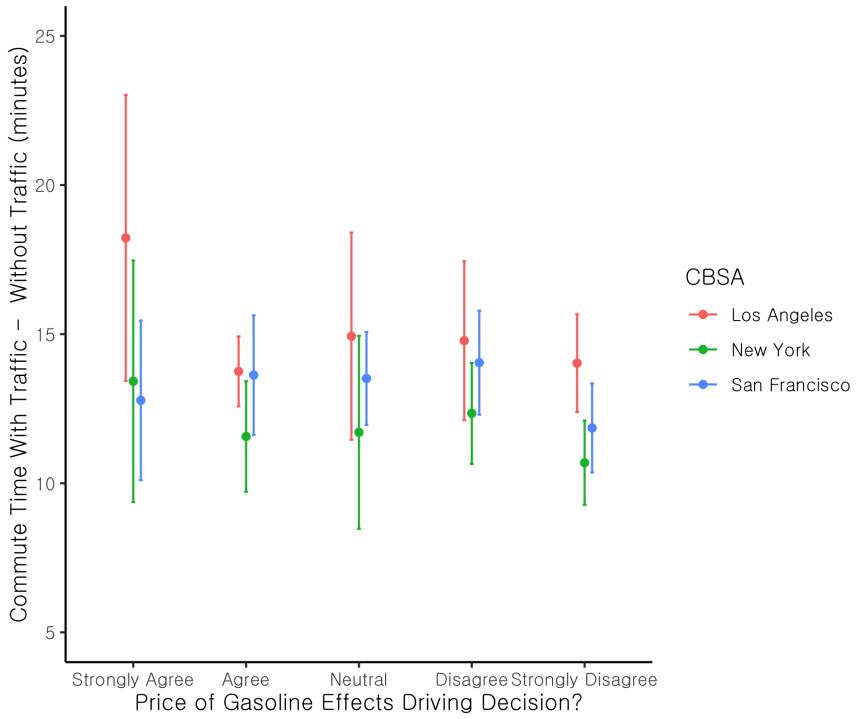


Figure A.17: This figure plots self-reported gasoline price responsiveness against the amount of time a driver reports loosing to traffic during their commute for drivers in the 2017 National Household Transportation Survey.

A.10. Public Transit

As outlined in Section 1.2, optimal cordon prices are determined in part by the unpriced social cost of substitutes to peak-hour cordon trips. Understanding how substitution to the outside option (any non-driving activity, including not traveling, or public transit) differs with access to public transportation is therefore crucial for applying the discrete choice model estimated in the Bay Area to other cities.

In this appendix, I first use data from the Bay Area Rapid Transit system to estimate the magnitude of substitution to public transportation in response to the 2010 change in toll prices on Bay Area Bridges. I then estimate an alternative specification of the logit model presented in Section 1.7 to test whether drivers with FasTrak devices registered in zip codes with easy walking access to public transit are more price responsive than are drivers who live in areas without access to public transit.

A.10.1. Public Transit in the San Francisco Bay Area

The Bay Area Rapid Transit (BART) system is a light rail network that connects the eastern Bay Area to the San Francisco Peninsula. BART is the most commonly-used public transportation system in the Bay Area, and the only rail system with trans-bay lines. The 46 stations that comprise the BART system are plotted in Figure A.19. Riders are charged based on the length of their trip; in 2010 the minimum price for a BART trip was $1.75, and most trans-bay trips cost between $3 and $6. Prices for the BART system did not change between July 2009 and July 2012.

BART publishes monthly ridership at the station level. Table A.7 and Figure A.18 show the change in BART ridership estimated using a regression discontinuity design around the July 1, 2010 change in bridge prices. In my preferred specification (column 3), I estimate that the increase in toll prices on the Bay Area bridges increased BART ridership by an average of 105 weekday rides per station per month. This point estimate corresponds to a 1.3% increase relative to baseline ridership levels. Multiplying this estimate by the number of BART stations (46) implies an estimate of 4,830 additional weekday BART trips per month, or 230 additional BART trips per weekday following the increase in bridge toll prices.

These point estimates suggest that while some drivers switched to public transit, the drivers who switched to public transit represent a small fraction of the total number of drivers who substituted away from driving. For reference, Foreman (2016) finds that the average change in hourly trips following the July 2010 price increases on the Bay, San Mateo, and Dumbarton Bridges were -87, -14, and -48, respectively, implying a total decrease 3,576 driving trips *per day* on these three bridges. Taking both of these estimates at face value implies that only 6.4% of the decrease in trans-bay trips were replaced by BART trips.

A.10.2. Price Responsiveness and Public Transit Access

In Table A.8, I re-estimate equation 1.15, allowing price-responsiveness to vary with access to public transportation. Specifically, I interact the *price* variable with an indicator variable for whether there is a BART station within 20 minutes walking distance of a given driver's zip code.[2] Point estimates suggest that drivers living in zip codes with transit stops nearby are slightly more price responsive than are those without transit stops nearby, but this difference is not statistically significant.

[2]According to the 2017 NHTS, roughly 90% of respondents who report taking public transit to work walk 20 minutes or fewer to the transit station.

In summary, although some Bay Area drivers responded to the increase in bridge tolls by shifting to public transit, the overall share of drivers who switched modes is quite low, and price sensitivity does not vary significantly based on public transit access. A possible explanation for the similarity in price responsiveness across drivers with different access to transit is that unobservable characteristics may determine selection into driving. Put differently, the people who live in transit-rich neighborhoods *but nonetheless still choose to drive* may have idiosyncratic preferences or pressures that lead them to be reluctant to switch modes, even though they happen to live near transit stations.

While these findings generally support the application of the discrete choice model estimated in San Francisco to areas with different public transit systems, several caveats bear mentioning. First, the BART system is a relatively expensive public transportation system. The magnitude price-induced substitution away from roads and toward transit undoubtedly depends on the price differential between modes. Conditional on the other attributes of transit trips, cities that have cheaper public transit (e.g., New York) may experience higher cross-price elasticities between vehicle transportation and public transit. Second, the results in this appendix (and this paper more generally) rely on short-term elasticity estimates, i.e., estimates of substitution elasticities holding fixed housing and work locations, as well as vehicle purchases. As firms and individuals sort in response to cordon prices, public transit access may lead to different long-run elasticities across regions where short-run elasticities look similar. While a full hedonic sorting model is beyond the scope of this paper; one would expect that cities with more connected and cheaper public transit systems would experience more drivers shifting to these modes. All else equal, this would (a) reduce leakage, and (b) increase second-best cordon prices relative to a city with poor public transit options.

Table A.7 — Changes Public Transit Ridership

	Specification		
Variable	(1)	(2)	(3)
Post	512.756	148.333	105.675
	(36.58)	(74.211)	(79.071)
Station FE	Yes	Yes	Yes
Month of Year FE	Yes	Yes	Yes
Bridge Closure Dummy	Yes	Yes	Yes
Linear Trend in Months	No	Yes	Yes
Second-Degree Trend in Months	No	No	Yes

Table A.7: This table displays the results of three regression discontinuity designs estimating the change in public transit ridership in the Bay Area following the July 2010 increase in driving tolls on trans-bay bridges. The *Post* variable is the reported change in monthly rides BART rides at the Station level; there are 46 Stations in the BART system. Standard errors clustered at the month level are shown in parenthesis. The data run from September of 2009 to June of 2012, and contain 1,462 station-month observations.

Figure A.18 — Public Transit Ridership Regression Discontinuity Plots

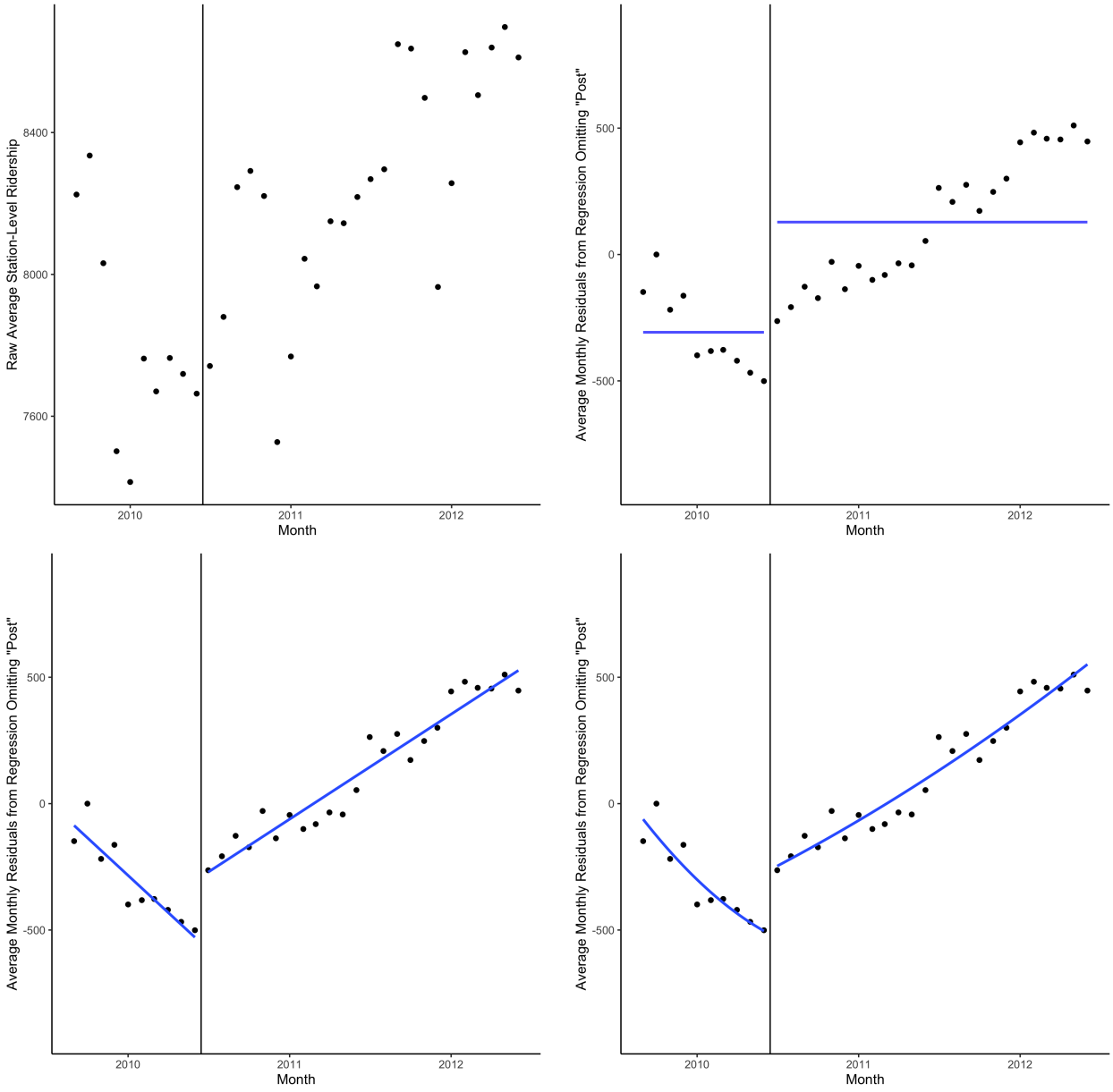


Figure A.18: The first pane in this figure plots the monthly station-level turnstyle exits averaged across the 46 stations on the Bay Area Rapid Transit (BART) system. The data run from September of 2009 to June of 2012, and contain 1,462 station-month observations. Panes 2 through 4 plot average monthly residuals from a regression of station-level exist on a set of station fixed effects, month-of-year fixed effects, and a dummy for months where there was a closure on the Bay Area's trans-bay bridges. Pane 2 fits a simple average to the pre vs. post residuals; pane 3 plots the pre and post residuals with a linear fit; pane 4 fits a second-degree polynomial to the pre and post residuals. The discontinuity between the fitted lines in these plots correspond to the treatment effects in specifications (1) through (3) in Table A.7, respectively.

FIGURE A.19 — ACCESS TO PUBLIC TRANSIT IN THE SAN FRANCISCO BAY AREA

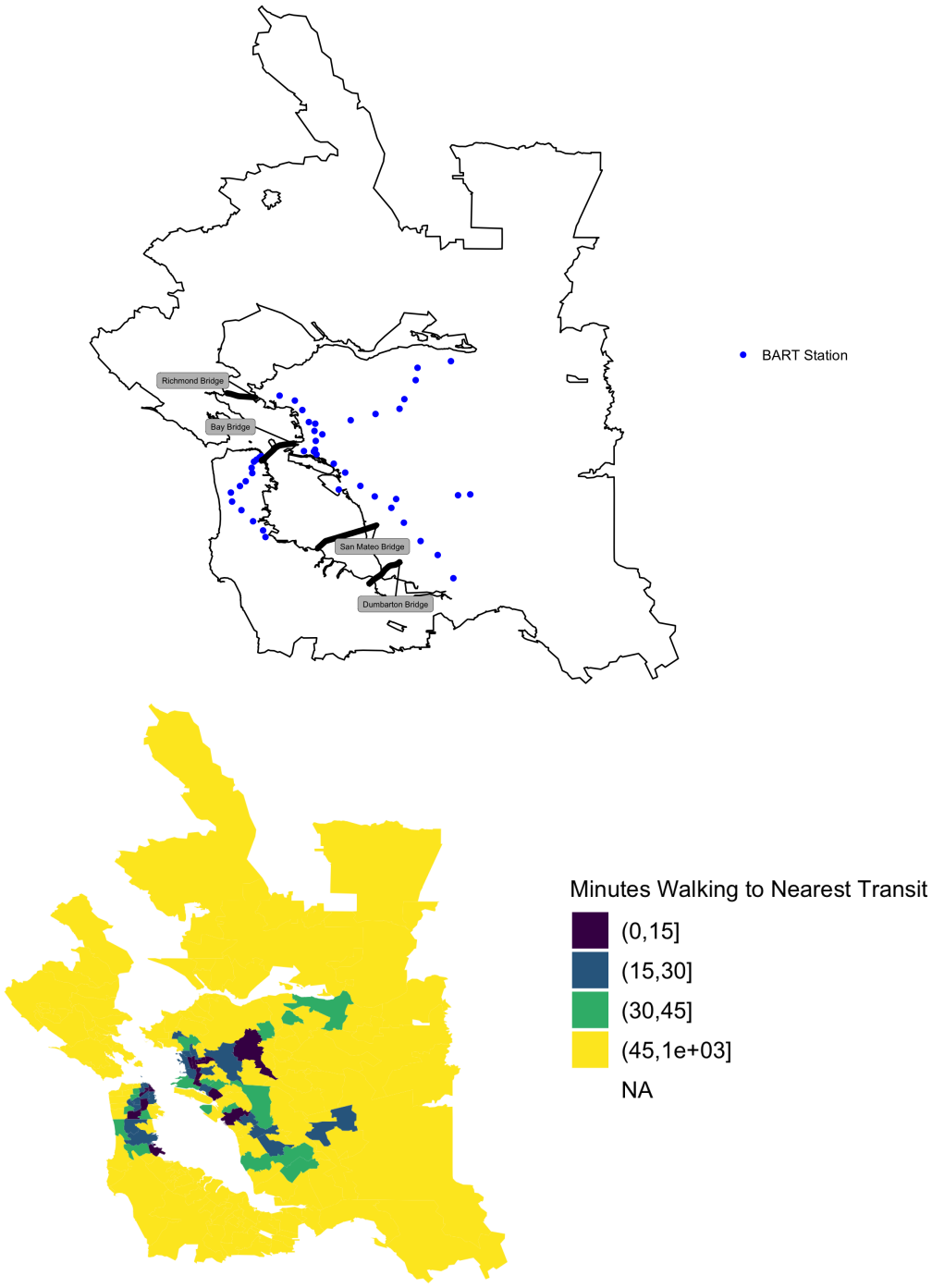


Figure A.19: The top panel in this figure plots Bay Area Rapid Transit (BART) light rail stations. The bottom panel plots the estimated walking time (as per google maps) from the google-registered address associated with a given Bay Area zip code (roughly the zip code centroid) to the nearest BART station.

Table A.8 — Logit Regressions with Price Responsiveness by Transit Access

	Specification	
Variable	(1)	(2)
Travel Time ($/hr)	-4.435	-4.434
	(0.151)	(0.149)
Time Early ($/hr)	-1.04	-1.04
	(0.052)	(0.052)
Time Late ($/hr)	-1.073	-1.073
	(0.045)	(0.045)
Price (Pooled)	-0.499	
	(0.027)	
Price (Transit Zipcodes)		-0.514
		(0.033)
Price (Non-Transit Zipcodes)		-0.496
		(0.033)
Day of Week FE	Yes	Yes
Bridge FE	Yes	Yes

Table A.8: Results from a variation of Equation 1.15, a discrete choice model where drivers choose over routes and times of day, estimated with a simple logit model. This model is estimated using FasTrak tolling microdata from the San Francisco Bay Area, as described in Section 1.5. The dependent variable is whether an individual i elects to take a trip on route r at time of day h. *Travel time* is the travel time (in hours) that driver i would incur by traveling via route r at time h. *Time early* is the number of hours that that driver i would arrive before their ideal arrival time if they were to travel via route r at hour h. *Time late* is analogously defined. *Price* is the toll that driver i would incur by traveling via router at hour h. In Column (2), I interact price with *transit access*, an indicator for whether a BART train station is within 20 minutes walking distance of a given driver's zip code. Two-way standard errors are clustered at the individual and zip code levels.

A.11. Low-Income Exemptions

While road pricing can increase economic efficiency, concerns about regressivity have prompted planners in many cities to consider road pricing schemes specifically designed to reduce the incidence on low-income road users. In San Francisco, for example, a majority of the congestion pricing proposals under consideration include some level of income-based exemption (San Francisco County Traffic Authority, 2021). Similarly, as of 2021, New York plans to refund congestion tolls for drivers who make under $60,000 per year (Regional Plan Association, 2021).

Table A.9 compares predicted reductions in pollution, congestion, and deadweight loss in San Francisco under (a) the second-best peak hour cordon prices estimated using equation 1.10, and (b) the same policy where drivers from low-income households (those with self-reported household income below $75,000 in the NHTS) are exempt from cordon fees. This exercise suggests that the efficiency costs of these exemptions are modest: exempting low-income drivers from cordon pricing in San Francisco would generate reductions in pollution and congestion externalities that are 1-3 percentage points smaller than under an optimal no-exemption policy. These efficiency costs are substantially smaller than the efficiency costs that reflect other imperfections in cordon pricing. For example, I estimate that restricting pricing to peak hours generates reductions congestion, pollution and deadweight loss that are 10 to 20 percentage points smaller than a pricing scheme where a policymaker can set a fixed hourly schedule of prices between 6 am and 10 pm (see Table 1.8).

The relatively small efficiency cost of exemptions reflects the low proportion of low-income drivers in trips that use the cordon. In the sample of 1,891 trips from the California NHTS with fastest routes that pass through California's cordon zone, just 9% are taken by drivers from houses with a total annual income of less than $75,000.

Table A.9 — Second-Best Cordon Pricing with Low-Income Exemptions

	Performance Relative to First-Best (%)	
Outcome	Second-Best Peak Hour	Low Income Exemption
Reduction in Total Externalities	30.576	28.173
Reduction in Congestion	31.367	28.782
Reduction in Pollution	23.779	22.939
Welfare Gain	32.304	32.601

Table A.9: Column (1) of this table reproduces the results from Table 1.5, which compares outcomes under second-best optimal peak hour cordon pricing to outcomes under Pigouvian pricing. The second column in this table compares the first-best policy to cordon pricing scheme that is identical to the scheme column (1), except that households making less than $75,000 per year are exempt from cordon fees. Income data reflect self-reported household income from the 2017 National Household Transportation Survey.

Appendix B

Appendix for *What drives support for inefficient environmental policies? Evidence from a Nevada ballot initiative*

B.1. Survey Details

In this appendix, I provide information on the language used in the survey to describe each of the policy alternatives, and provide references for the information provision section of the survey. PDF versions of the entire prior and posterior surveys are available here and here, respectively.

The following is how Nevada Question 6 was presented to respondents:

> *The following initiative will be on the 2020 ballot in Nevada:*
>
> *Question 6*
>
> *Shall Article 4 of the Nevada Constitution be amended to require, beginning in calendar year 2022, that all providers of electric utility services who sell electricity to retail customers for consumption in Nevada generate or acquire incrementally larger percentages of electricity from renewable energy resources so that by calendar year 2030 not less than 50 percent of the total amount of electricity sold by each provider to its retail customers in Nevada comes from renewable energy resources?*

How do you plan on voting for this initiative?

The following is how the carbon tax alternative was is presented to respondents:

Question 6 is a policy that addresses state-level carbon emissions. As an alternative to requiring that a certain percent of energy be produced by renewable sources, some states and countries put a price on carbon emissions. Consider the following hypothetical alternative to Question 6:

Question 7

Shall Article 4 of the Nevada Constitution be amended to levy, beginning in calendar year 2022, a carbon emissions fee of $25 per metric ton of carbon on the sale or use of certain fossil fuels and fossil-fuel-generated electricity, and reduce the sales tax by 1.5 percentage points, while freezing Nevada's renewable energy standard at its current level?

If this initiative were on the ballot instead of Question 6, how would you vote on this initiative?

Information Provision sources

The following are the sources for 12 possible information treatments that are randomized to survey participants."Question 6" is Nevada's RPS; "Question 7" is the hypothetical carbon tax ballot initiative.

Question 6 cost high

Greenstone and Nath (2019)

Question 6 cost low

Galen (2018)

Question 6 regressivity high

Rausch and Karplus (2014)

Question 6 regressivity low

Rausch and Mowers (2014)

Question 6 effectiveness high

Greenstone and Nath (2019)

Question 6 effectiveness low

Sekar and Sohngen (2014)

Question 7 cost high

Cronin, Fullerton, and Sexton (2019)

Question 7 cost low

Marron, Toder, and Austin (2015)

Question 7 regressivity high

Grainger and Kolstad (2010)

Question 7 regressivity low

Marron, Toder, and Austin (2015)

Question 7 effectiveness high

Barron, Fawcett, Hafstead, McFarland, and Morris (2018)

Question 7 effectiveness low

The Congressional Budget Office (2016)

B.2. Supplementary Figures

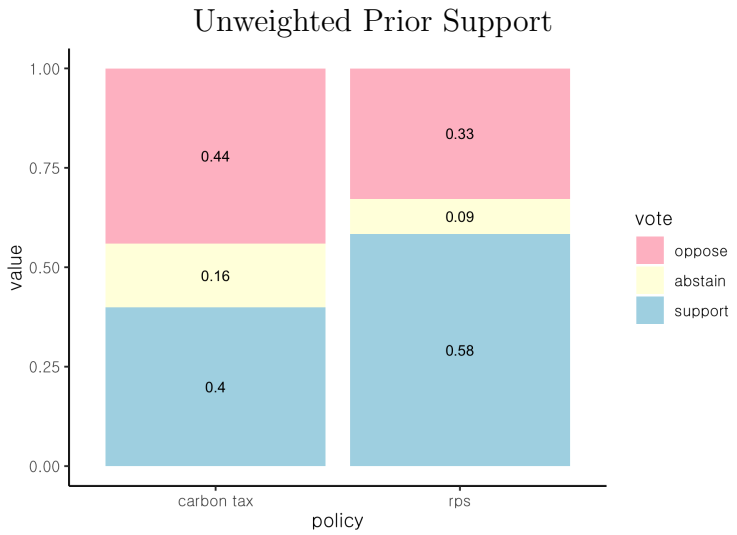


Figure B.1: Self-reported policy support for Nevada's Ballot Question 6 (a 50% RPS), and a price-based alternative policy (a $25 carbon tax with a 1.5% sales tax cut) among 275 Nevadans.

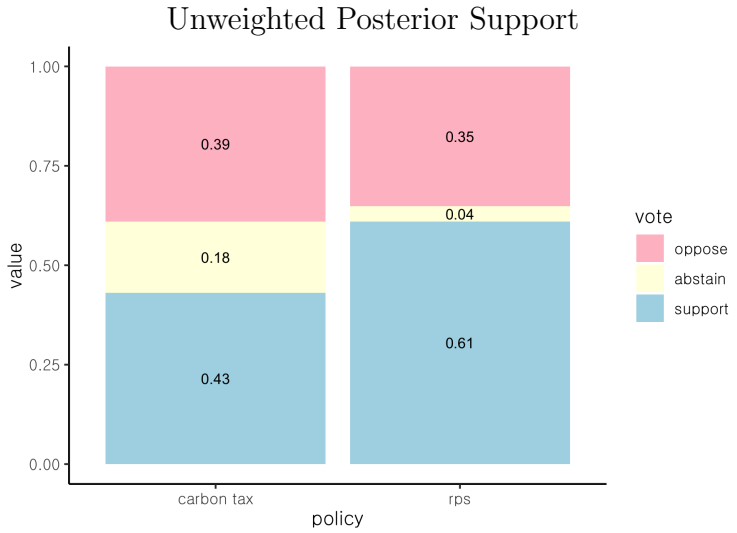
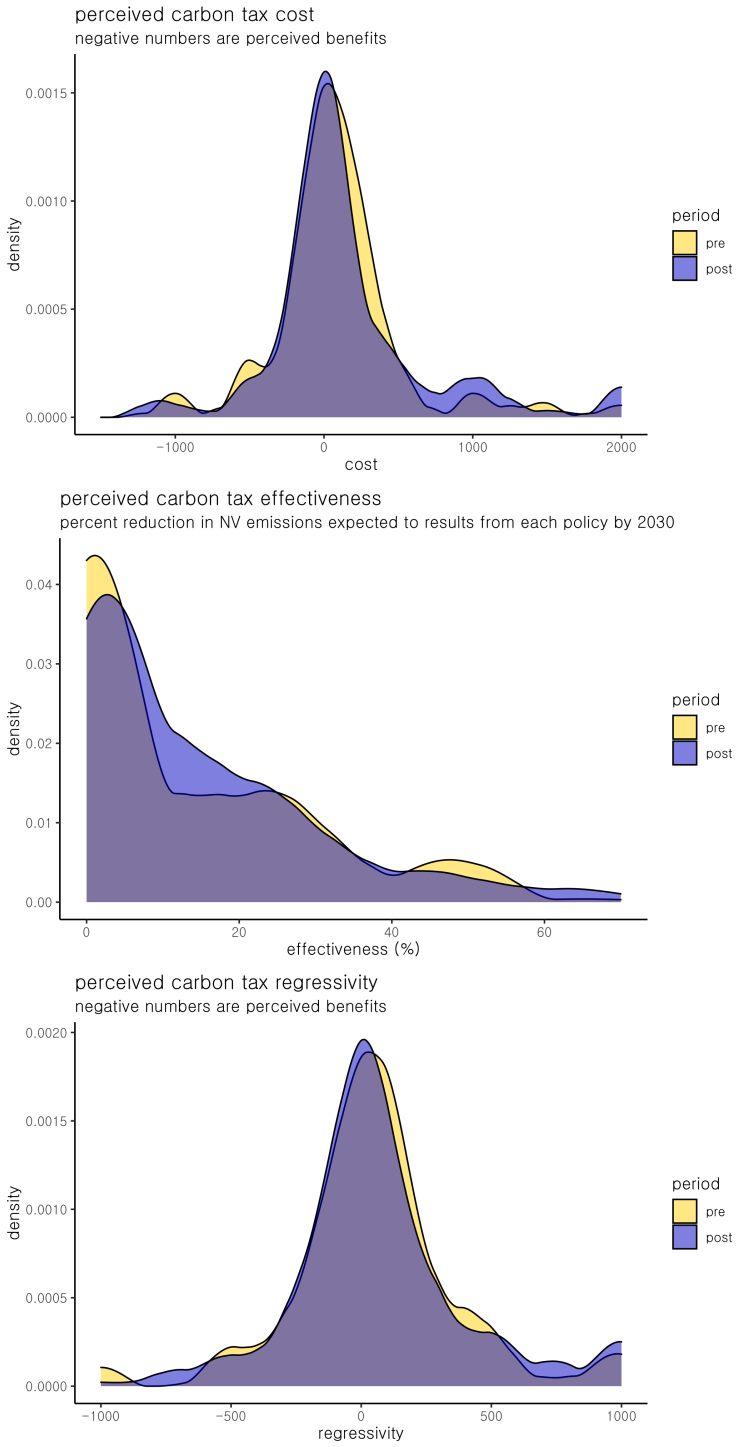
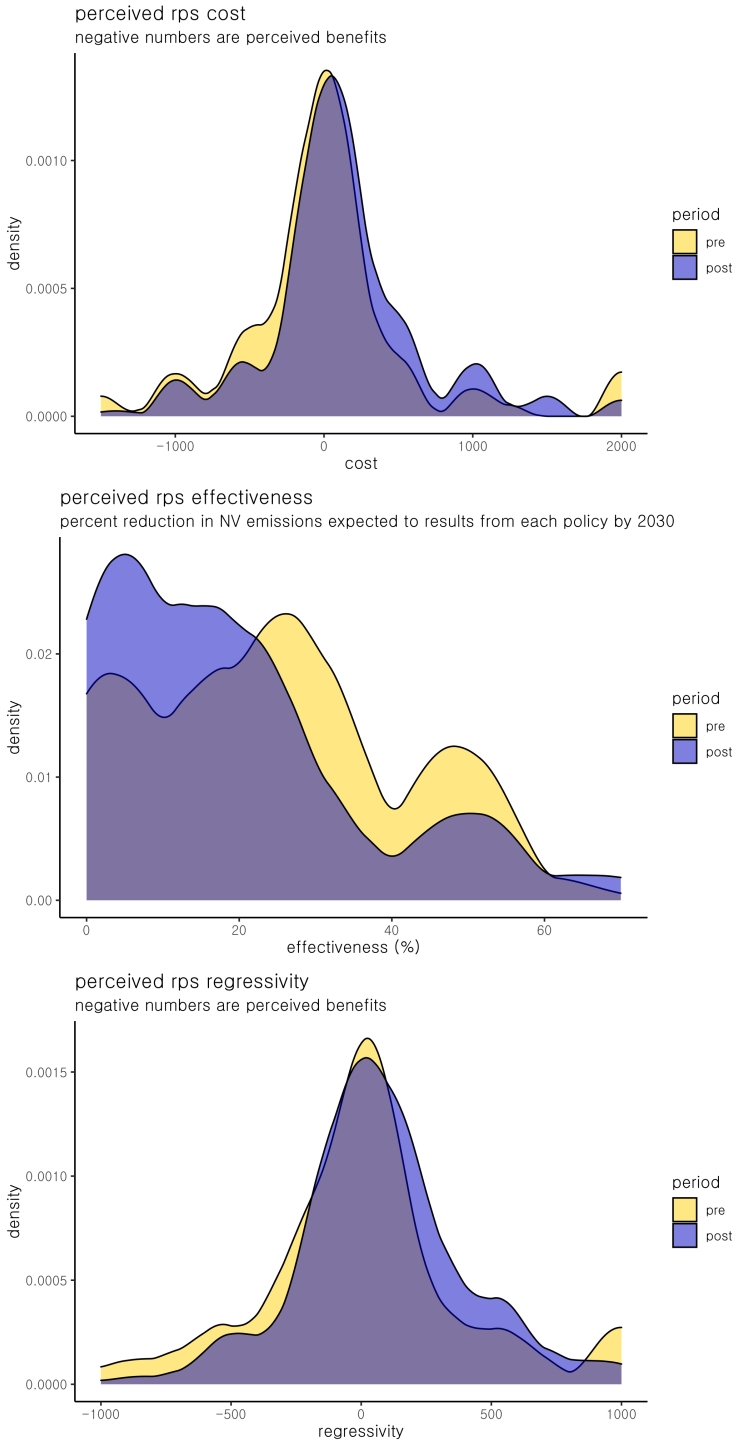


Figure B.2: Self-reported policy support for Nevada's Ballot Question 6 (a 50% RPS), and a price-based alternative policy (a $25 carbon tax with a 1.5% sales tax cut) among 275 Nevadans.

B.2.1. Prior vs. Posterior Beliefs about Carbon Taxes



B.2.2. Prior vs. Posterior Beliefs about Renewable Portfolio Standards



Appendix C

Appendix for *The Congestion Costs of Uber and Lyft*

C.1. Revealed Preference VOT Estimates

The MoPac (Texas State Highway 1) is a north-south route in Austin. Starting in November 2017, the Central Texas Regional Mobility Authority opened an 11-mile variable-price express lane on the MoPac (see Figure C.1). The price of using this lane adjusts in order to keep the express lane moving at free-flow speeds: tolls increase when the express lane is busy and decrease when it is underused. Toll rates are posted at the northbound and southbound entrances.

Using 30-minute resolution data provided by the Central Texas Regional Mobility Authority on MoPac prices and average travel times on the tolled and non-tolled lanes, I recover time-varying estimates of the value of travel time. Commuters entering the MoPac see the toll price, but not the difference in travel times between the tolled and non-tolled lanes. I therefore produce value of travel time estimates by dividing the observed toll price on a given date and time by the *expected* travel time savings for that time of day. To calculate the expected time savings, I take the average time difference between tolled and non-tolled lanes by half hour of day. For example, if the average difference between tolled and non-tolled lanes is 4 minutes between 9:00 a.m. and 9:30 p.m., and I observe an average price between 9:00 and 9:30 on a given day of \$1, then the implied value of travel time for that half-hour block on that day is \$15 per hour. I then aggregate these observations across days to recover an average value of travel time for each half hour of day. Note that because the tolled lane is always faster than the non-tolled lane *in expectation*, this method rationalizes driver's use of the toll lane when ex-post travel times are equal between tolled and non-tolled lanes. [1]

[1] Note that the MoPac has a price floor of \$0.25; 55% of the observations (largely in off-peak hours) are at the price floor. VOT estimates excluding these observations produce similar results.

These estimates are summarized in Figure C.1, and are broadly consistent with value of travel time estimates from related settings (Small and Verhoef, 2007). Importantly, however, VOT peaks during morning and evening rush hour periods, possibly reflecting different commuters, or heterogeneity in the value of time based on trip purpose. Note that the motivation for using MoPac data is to study the convolution between the value of travel time and estimated congestion impacts, not to produce novel estimates of the value of travel time.

FIGURE C.1—REVEALED PREFERENCE VALUE OF TRAVEL TIME ESTIMATES

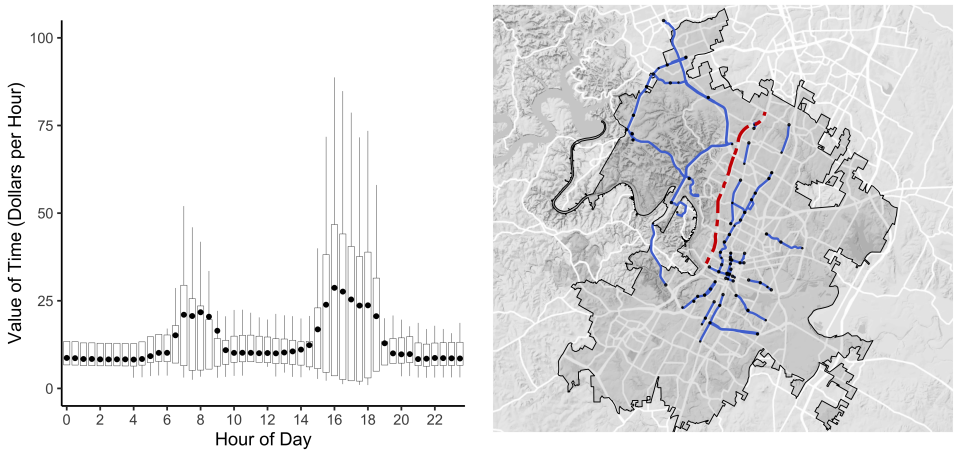


Figure C.1: *Left Pane:* Willingness to pay for travel time reductions in Austin, TX. Constructed using 2017 data from the MoPac variable-toll lane (Highway Loop 1), provided by the Central Texas Regional Mobility Authority. Dots represent means of observed equilibrium prices divided by expected time savings by hour of day.

Right Pane: MoPac expressway (red dashed) plotted with the 79 road segments used in estimation of the changes in travel speeds (blue). Nodes represent Bluetooth sensors, and the black line is the Austin City limit.

C.2. Threats to identification from other modes of transportation

In addition to Bluetooth sensors, the city of Austin maintains pneumatic sensors which take periodic measurements of traffic speeds. Pneumatic sensors are stretched across traffic lanes, and therefore will not be influenced by pedestrian activities. Additionally, pneumatic sensors classify observations by axel length, meaning activity from bicycles will not be reflected in vehicle speed measurements. I identify 39 instances where Bluetooth road segments overlie pneumatic sensors (see Figure C.1), and test whether the relationship between the two speed measures changes significantly after the Proposition 1 vote.

For other modes of transportation to bias my estimates, speed must be mismeasured on Bluetooth segments, and that mismeasurement must be correlated with the treatment. To test for this bias, I perform the following regression:

$$y_{i,t} = \alpha + \beta_1 s_{i,t} + \beta_2 D_t \cdot s_{i,t} + \gamma_1 \phi_t + \gamma_2 \boldsymbol{\theta}_i + \epsilon_{i,t} \tag{C.1}$$

Where $y_{i,t}$ is the Bluetooth speed measurement on segment i and time t, and $s_{i,t}$ is the pneumatic road segment speed measurement on segment i and time t. D_t is a treatment dummy, which equals one for days after May 9^{th}, 2016. ϕ_t is a set of month of year fixed effects, and θ_i is a set of segment fixed effects. I test the null hypothesis $\beta_2 = 0$, and interpret a significant result as evidence of bias in the Bluetooth speed masurements. Results from Equation C.1 are displayed in Table C.1.

Table C.1—Tests for Bias

	point estimate	*se*	*p*
β_1	0.0366	0.0221	0.0979
β_2	-0.0094	0.0222	0.6727

Table C.1: *Notes*: Results from Equation C.1, which tests whether the relationship between pneumatic speed measurements ($s_{i,t}$) and Bluetooth speed measurements ($y_{i,t}$) changes after the exit of TNCs in Austin. In addition to the variables listed, this regression includes segment and month of year fixed effects.

Figure C.1—Segment-sensor matching

Figure C.1: *Notes:* An example of a Bluetooth segment (blue path) matched to a pneumatic sensor (black dot). Segments were matched to sensors by both location and direction of travel

C.3. Threats to identification from TNC driving speeds

If TNC vehicles drive significantly slower or faster than the average non-TNC vehicle in a way that remains after filtering, the above empirical strategy will be biased. To test for this potential bias, I match RideAustin trips to segments, allowing me to test the null hypothesis that TNC vehicles drive at the same speeds as the average mix of vehicles.

I match TNC trips to segments based on the following criteria: for a given segment, the sum of the distance between segment and TNC trip termini must be less than 500 meters, and the distance traveled by the TNC vehicle must be within 10% of the segment length. I identify 1901 such matches, with several segments recording multiple TNC trips that fit this criteria. I then replicate the type of data filtering applied by Post Oak Traffic Systems. Recall that in the data I use for my analysis, only observations that fall within 75% of the interquartile range (IQR) of the 15 most recent observations are used to calculate average speeds. While I do not have access to the IQR data, I do have the standard deviation of speed measurements for any given 15-minute interval. I use this to estimate the IQR, and drop RideAustin trips that fall outside of IQR estimate for the corresponding segment, time, and date, as these trips likely be dropped from the Bluetooth speed measurements.

After applying these filters, I am left with 221 trip matches, which are summarized in Table C.1. The regression coefficient reflects a difference in means between TNC trip speeds and the segment speeds recorded at corresponding times. On average, TNC vehicles traversed segments 0.03 minutes per mile slower than did the average recorded vehicle. This difference is not statistically significant, and under reasonable assumptions about TNCs as a share of total vehicles (5-15%), should not generate meaningful bias in the results reported above.

A shortcoming of this test is that I only observe drivers while they have a passenger in the car. This test will not identify differences caused by passengerless TNCs driving systematically above or below the flow of traffic. Note that the most extreme cases of this activity (e.g., idling while waiting for passengers) will be dropped by the IQR filter.

Table C.1—TNC vehicle speeds

Δ Minutes per Mile	*se*	*p*
-0.0324	0.1643	0.8438

Table C.1: *Notes*: A comparison of means between Bluetooth-recorded travel speeds and TNC vehicle speeds. The coefficient Δ *Minutes per Mile* represents the difference in means between travel speeds (in minuts per mile) for these TNC trips and the average travel times recorded by the corresponding segment over the 15-minute period where the TNC trip occurred.

C.4. Robustness

In this section, I investigate the robustness of my results to alternative specifications and alternate road segment groups. Table C.1 shows the results of difference in difference regressions using alternate linear trend and weighting specifications. Table C.2 shows the results of running Equation 3.1 on different sub and supersets of the Bluetooth data used in the main analysis. My results are stable over these specifications: The F-test rejects the null that the hour-specific effects are jointly zero. Additionally, the estimated annual congestion cost of TNC activity is of a similar magnitude to my preferred specification in each of these alternative specifications.

Figure C.1 plots estimated regression discontinuity and difference in differences coefficients pooling over hours of day for symmetric bandwidths ranging from 20 to 70 days about May 9^{th}. The regression discontinuity results are stable over this range. The difference in differences results show positive, but not statistically significant point estimates for a small minority of bandwidths. Across all bandwidths, for both RD and DID specifications, point estimates of congestion impacts between 7 a.m. and 7 p.m. are negative. Taken together, the results presented in Figure C.1 are consistent with the conclusions in the body of this paper: daytime travel speeds in Austin likely increased following the exit of Uber and Lyft.

Figure C.2 displays regression discontinuity estimates of the impact of TNC departure on traffic speeds using the actual TNC exit date in relation to the distribution of coefficients from 134 regression discontinuities using placebo exit dates. The results of this exercise suggest that a) it is unlikely for shocks to the Austin-area traffic system to create changes in travel speeds on the order of the estimates reported in Table 3.4, and b) it is unlikely that the RD results reflect contemporaneous changes in traffic resulting from the end of the University of Texas, Austin semester.

TABLE C.1—ALTERNATE SPECIFICATIONS

	Model 1		Model 2		Model 3		Model 4	
Hour of Day	β_h	p	β_h	p	β_h	p	β_h	p
0	-0.0606	0.26	-0.0632	0.24	-0.0789	0.21	-0.1059	0.14
1	-0.0525	0.16	-0.0550	0.13	-0.0850	0.13	**-0.1405**	0.09
2	**-0.0505**	0.09	**-0.0508**	0.09	-0.1007	0.12	-0.0195	0.74
3	-0.0121	0.85	-0.0158	0.81	-0.1051	0.14	0.1506	0.27
4	0.0911	0.38	0.0841	0.41	-0.0795	0.38	0.1073	0.57
5	-0.0293	0.75	-0.0365	0.71	**-0.1850**	0.10	0.0886	0.66
6	-0.0568	0.58	-0.0765	0.43	**-0.2831**	0.03	-0.0599	0.75
7	-0.1446	0.24	-0.1647	0.16	**-0.3566**	0.07	0.0425	0.85
8	-0.0583	0.67	-0.1072	0.52	-0.1727	0.26	-0.3543	0.23
9	-0.0188	0.85	-0.0589	0.58	-0.1561	0.25	-0.0924	0.55
10	0.0318	0.59	0.0133	0.84	-0.1190	0.25	0.0001	1.00
11	**-0.1418**	0.05	**-0.1445**	0.04	-0.0621	0.29	-0.1327	0.14
12	**-0.1730**	0.05	**-0.1771**	0.04	**-0.1745**	0.05	**-0.1885**	0.10
13	-0.1555	0.12	**-0.1615**	0.09	-0.1082	0.12	**-0.2408**	0.08
14	-0.0176	0.71	-0.0224	0.63	-0.0379	0.42	-0.0103	0.84
15	-0.0238	0.61	-0.0304	0.49	-0.0257	0.60	-0.0026	0.97
16	0.0657	0.35	0.0578	0.37	0.0515	0.33	0.0715	0.32
17	0.0016	0.98	-0.0028	0.96	0.0248	0.60	0.0360	0.49
18	0.0500	0.57	0.0459	0.59	-0.0102	0.85	0.0615	0.33
19	-0.0300	0.65	-0.0285	0.65	-0.0254	0.63	-0.0327	0.57
20	0.0004	1.00	-0.0065	0.95	-0.0192	0.76	0.0324	0.73
21	0.0817	0.26	0.0775	0.27	0.0090	0.90	**0.1261**	0.08
22	-0.0709	0.42	-0.0763	0.37	-0.0491	0.51	-0.0502	0.57
23	-0.0152	0.65	-0.0235	0.53	**-0.1450**	0.07	**-0.1187**	0.04
F-test		0.00		0.00		0.00		0.00
Annual Cost		-33,605,827		-49,764,126		-91,753,234		-54,047,223

Table C.1: *Notes*: Results from Equation 3.1. Model 1 reproduces the results from my preferred specification, with year-segment specific linear trends. Model 2 includes only year-specific linear trends (i.e, pools across segments). Model 3 includes only segment-specific linear trends (i.e., pools across years). Model 4 uses the same specification as column 1, but weights observations (which are 15-minute average travel times) by the number of vehicles observed. Bold coefficients are significant at the 10% level. The penultimate row reports the p-value from a joint hypothesis test of $\beta_h = 0\ \forall h$; the final row reports annual costs using hourly VOT estimates from Austin's MoPac freeway (see Appendix C.1).

Table C.2—Alternate segment groups

	Group 1		Group 2		Group 3	
Hour of Day	β_h	p	β_h	p	β_h	p
0	-0.0606	0.26	-0.0379	0.44	-0.0399	0.78
1	-0.0525	0.16	**-0.0482**	0.08	-0.1551	0.12
2	**-0.0505**	0.09	-0.0412	0.19	-0.0854	0.42
3	-0.0121	0.85	-0.0125	0.74	0.0498	0.74
4	0.0911	0.38	0.0547	0.39	0.0785	0.64
5	-0.0293	0.75	-0.0124	0.87	-0.0177	0.88
6	-0.0568	0.58	-0.0147	0.84	0.2108	0.36
7	-0.1446	0.24	-0.0899	0.35	-0.1519	0.51
8	-0.0583	0.67	-0.1080	0.35	-0.1177	0.82
9	-0.0188	0.85	-0.0448	0.51	-0.2941	0.27
10	0.0318	0.59	-0.0162	0.73	-0.1159	0.42
11	**-0.1418**	0.05	**-0.1416**	0.05	**-0.2689**	0.07
12	**-0.1730**	0.05	-0.1246	0.19	**-0.3233**	0.03
13	-0.1555	0.12	-0.0956	0.28	-0.3944	0.13
14	-0.0176	0.71	0.0029	0.95	-0.1260	0.43
15	-0.0238	0.61	-0.0217	0.56	-0.0168	0.91
16	0.0657	0.35	0.0435	0.47	0.0059	0.96
17	0.0016	0.98	-0.0307	0.62	0.0587	0.49
18	0.0500	0.57	-0.0222	0.76	-0.0479	0.77
19	-0.0300	0.65	-0.0520	0.30	-0.1182	0.41
20	0.0004	1.00	0.0036	0.97	0.0546	0.70
21	0.0817	0.26	0.1056	0.10	0.0967	0.58
22	-0.0709	0.42	-0.0468	0.60	-0.1515	0.36
23	-0.0152	0.65	-0.0243	0.48	-0.1332	0.46
F-test		0.00		0.00		0.00

Table C.2: *Notes*: Results from Equation 3.1, applied different groups of road segments. Group 1 is my preferred specification, which uses all traffic segments which report in more than 70 percent of days in each year. Group 2 relaxes this level to segments that report in 30 percent of days. Group 3 uses only segments that report in every day of the study period. Bold coefficients are significant at the 10% level. The final row reports the p-value from a joint hypothesis test of $\beta_h = 0 \ \forall h$.

Figure C.1—Bandwidth Sensitivity

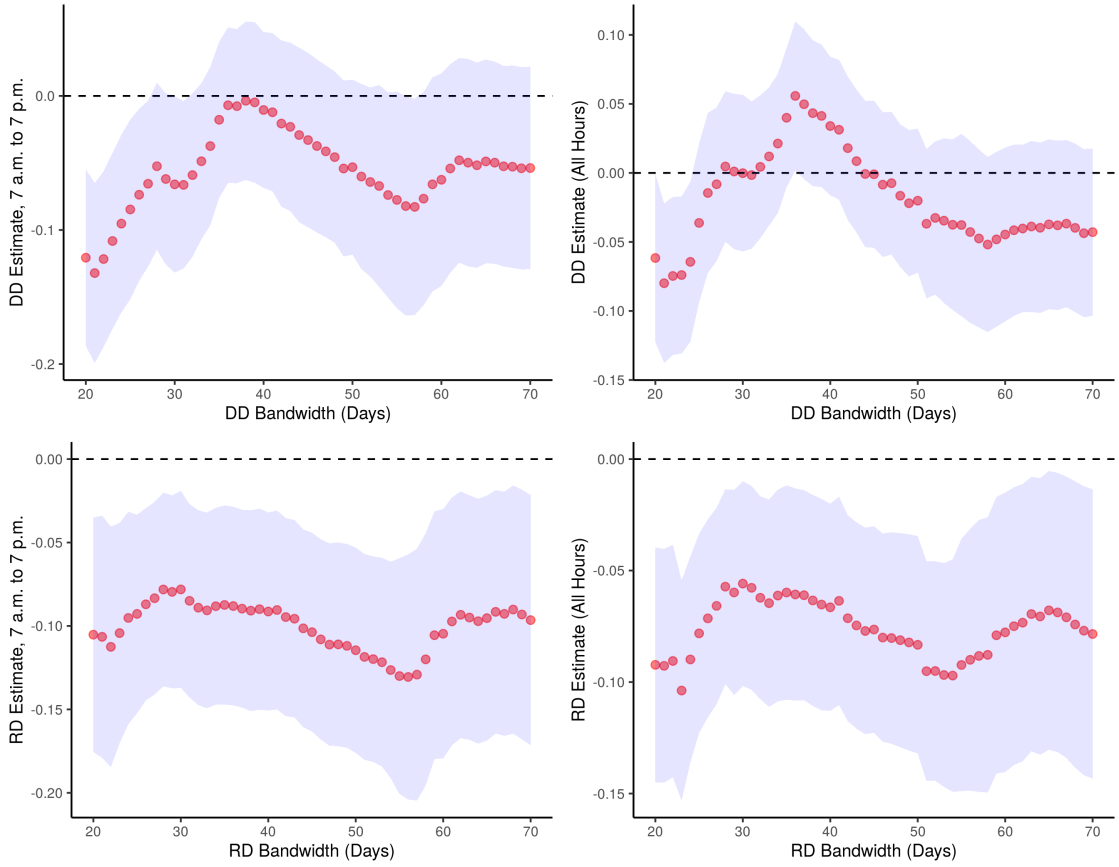


Figure C.1: *Notes:* This figure plots estimated difference in differences and regression discontinuity coefficients pooling over hours of day (a variation of Equation 3.3) for symmetric bandwidths ranging from 20 to 70 days about May 9$^{\text{th}}$, 2016. Bars represent the 95% confidence interval using standard errors clustered by segment-week. Specifications with a bandwidth of over 45 days include a dummy for the SXSW festival. For reference, my preferred specification uses an asymmetric bandwidth of 45 pre-period days and 52 post-period days.

FIGURE C.2—REGRESSION DISCONTINUITY PLACEBO TESTS

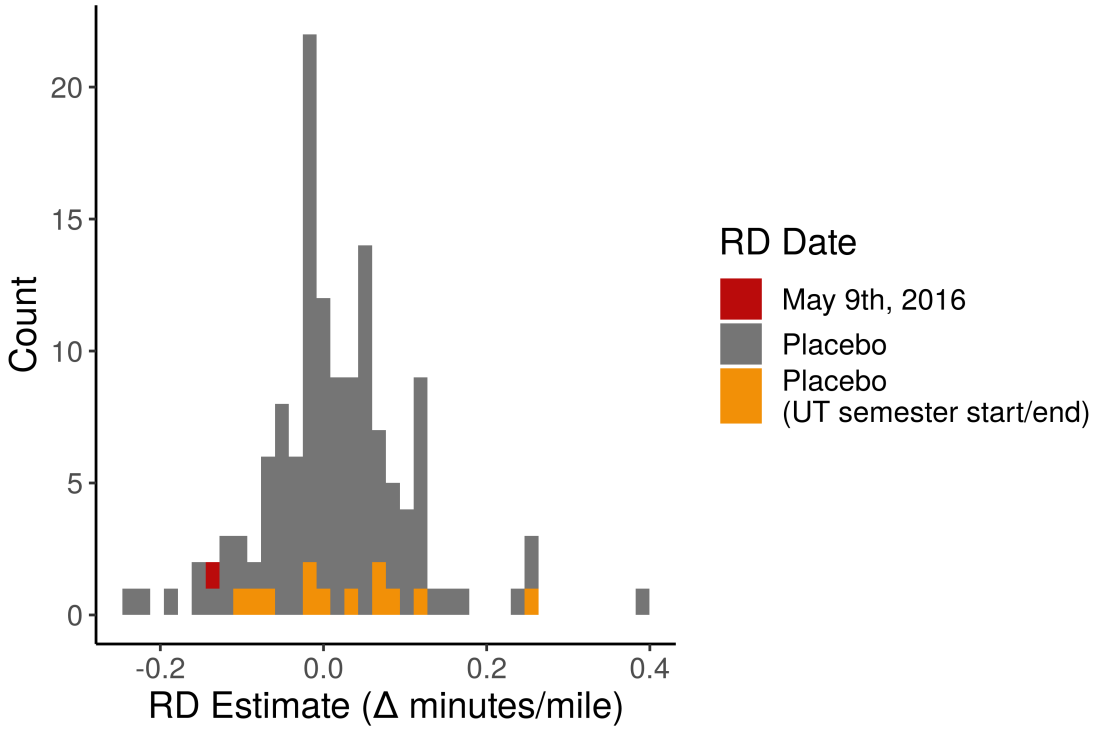


Figure C.2: *Notes:* This figure displays regression discontinuity estimates of the impact of TNC departure on traffic speeds in Austin, TX, using the actual TNC exit date (in red) in relation to the distribution of coefficients from 134 regression discontinuities using placebo exit dates (in grey and yellow). The grey cells are results from 121 placebo regression discontinuity dates that run every 20 days from 2015 to 2019 with a 30 day bandwidth. The yellow cells are results from using the start/end of the UT semester as placebo regression discontinuity dates. I omit placebo regressions with significant amounts of missing data (missing more than 75% of days within the RD bandwidth). Controls in each regression are day of week, holiday, and segment fixed effects, segment-specific second-degree polynomials in days since May 9th, and flexible controls for temperature and precipitation. Specifications with March dates include a dummy for the SXSW festival.

C.5. Segment Length Revisions

The raw Bluetooth traffic data available on the Austin Open Data Portal show the lengths of traffic segments varying over time. Of the 79 segments that I use in my analysis, 61 segment lengths changed during the study period, with the majority of these adjustments occurring on March 24th, 2016. On average, these adjustments were small: only four segment lengths were adjusted by more than 4%. According to the data providers, these adjustments most likely reflect updated distance measurements, not the repositioning Bluetooth readers. As such, I use updated segment lengths to calculate average speeds in all time periods.

To verify that this data quality issue does not constitute a threat to identification, I run a regression discontinuity about each segment length change. Of the 61 segments where lengths were adjusted, 47 had adequate data to run a regression discontinuity about the date where the segment length was changed. 8 of these 47 regression discontinuities registered statistically significant changes in traffic speeds. These results of these RDs are plotted in Figure C.1.

Table C.1 shows that omitting these 8 segments from my study pool does not substantively change the results presented in the body of this paper. Columns 1 and 2 reproduce estimates from Table 3.4, and columns 3 and 4 re-run these specifications excluding the 8 segments that registered a statistically significant change about their readjustment date. Across all four specifications, analyses using a restricted pool of segments suggest a modest increase in traffic speeds following the exit of Uber and Lyft from Austin. Note that the magnitude of the regression discontinuity results are sensitive to the exclusion of these 8 segments.

Table C.1—Sensitivity to Segments with Anomalous Length Revision Data

	β (Table 4)	*se*	β (Restricted sample)	*se*
Difference in Differences (All hours)	-0.02605	0.01701	-0.02972	0.02462
Difference in Differences (7 a.m. - 7 p.m.)	-0.06838	0.05295	-0.06137	0.03211
Regression Discontinuity (All Hours)	-0.10155	0.03535	-0.06498	0.01899
Regression Discontinuity (7 a.m. - 7 p.m.)	-0.13351	0.04331	-0.08567	0.02529

Table C.1: *Notes*: The first two columns reproduce the results from Table 3.4, which reports difference in differences and regression discontinuity results pooled across the hours of the day (Equations 3.1 and 3.3, respectively). Columns 3 and 4 show results from the same regressions run on a sample that excludes the 8 segments that registered a statistically significant change in traffic speeds about the adjustment date.

FIGURE C.1—SPEED AROUND SEGMENT LENGTH REVISIONS

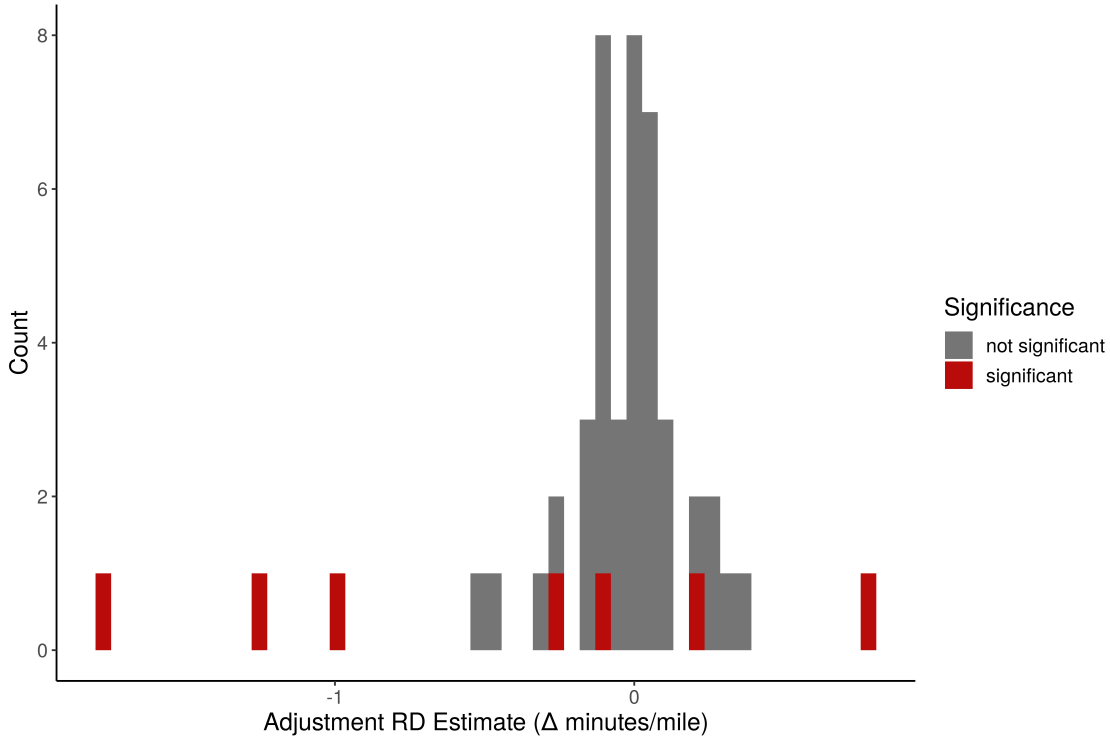


Figure C.1: *Notes:* This figure plots coefficients from 47 regression discontinuity specifications estimated where the raw Bluetooth data showed a change in a segment length. Each cell corresponds to the estimated change in traffic speed (in minutes per mile) estimated about the date of a single length revision. Controls in each regression include day of week, holidays, SXSW, and hour fixed effects, segment-specific linear trends in days since adjustment, and flexible temperature and precipitation functions. Significant results are potted in red. Standard errors in all regressions are Newey-West.