Can Usage-Based Pricing Reduce Traffic Congestion?*
[PRELIMINARY; DO NOT CIRCULATE]

Itai Ater[†] Adi Shany[‡] Brad Ross[§] Eray Turkel[¶] Shoshana Vasserman[‖]

January 2024

Abstract

This paper analyzes the effects of the largest field experiment to date that incentivizes drivers to limit driving during peak hours and congested areas via usage-based congestion pricing. The experiment monitored the driving behavior of 10,000 Israeli drivers who were recruited over the course of 2020. During the first six months of a driver's participation in the experiment, their driving behavior is monitored and recorded; afterwards, drivers receive an annual budget and are charged for each kilometer driven during historically congested times in congested areas. Whatever remains in each driver's budget is paid to them a year later. We use comprehensive data on driving behavior of participating drivers over the course of 2020 and 2021 to evaluate how usage-based congestion pricing affects driving behavior. The staggering of driver recruitment facilitates identifying treatment effects via a difference-in-differences approach. We find that drivers decrease their congested driving behavior by approximately 10% across a myriad of outcomes designed to detect both extensive margin (i.e. whether to take a trip) and intensive margin (i.e. when to take a trip and via which route) responses. We also find that there is significant treatment effect heterogeneity across drivers that can be predicted by pre-treatment driving behavior. The most affected drivers tend to be those who contributed more to congestion and who appear to have more flexibility in their driving choices and easier access to public transit, but they are not disproportionately socioeconomically advantaged or disadvantaged. Finally, we estimate how traffic density affects traffic speed on the largest highway in Israel using one year of high-frequency traffic sensor data and use this structural relationship to translate our treatment effects on driving behavior into increases in traffic speeds of up to 30% during the most congested times of day.

*First and foremost, we thank Noa Bar-Nir, Fabian Barrett, Hadar Fuks, Shai Gilat, Andrew Meyers, Yotam Nir, Doron Zamir, and Michelle Zheng for their exemplary research assistance. We are also grateful to our colleagues and numerous seminar and conference participants for their helpful comments.

[†]Tel Aviv University, Coller School of Business and CEPR. Email: ater@post.tau.ac.il.

[‡]Tel Aviv University, Coller School of Business. Email: adishany@tauex.tau.ac.il.

[§]Stanford Graduate School of Business. Email: bradross@stanford.edu

[¶]Google. Email: eturkel@stanford.edu

[‖]Stanford Graduate School of Business and the NBER. Email: svass@stanford.edu.

1 Introduction

Traffic congestion is a notorious source of inefficiency, frustration and environmental damage. On a social level, traffic delays are associated with billions of dollars of waste (Chandler-Wilde, 2023) and substantial greenhouse gas emissions (Barth and Boriboonsomsin, 2008). On an individual level, traffic congestion is associated with serious physical health problems (Currie and Walker, 2011), and time spent in congestion is reported to be among the worst moments of the day (Kahneman et al., 2006). Given these costs, policymakers have sought effective means to reduce congestion, increasingly through policies that charge fees for driving in highly demanded areas. Zone-based entry fees ("Cordon pricing") have been in place in London since 2003, Stockholm since 2006, and Milan since 2008, and proposals for similar fee structures in Seattle, San Francisco, and New York City are currently progressing to the final stages.

However, there is little causal evidence that these programs work to meaningfully reduce congestion. A direct inspection of London congestion shows that road speeds initially increased following the implementation of the Cordon price policy, but have since rebounded. Existing experimental studies—which took place prior to the COVID-19 pandemic, and were relatively small in size and scope—have found that drivers make almost no significant changes in their rush-hour driving patterns in response to small monetary incentives (Kreindler, 2022; Martin and Thornton, 2017). These studies suggest that the price levels that transportation officials are willing to impose on congested roads are not large enough to meaningfully change the driving choices of heavy commuters, who already incur substantial costs from driving in the form of lost time and parking expenses. Still, it is unclear whether these economically insignificant effect estimates were due to short treatment durations, ineffective pricing, smaller sample sizes, and/or pre-pandemic norms around schedule flexibility and working from home.

This paper describes the results of what is, to our knowledge, the largest and longest-running usage-based congestion pricing experiment that has ever been implemented: a government-sponsored experiment called "Derech Erech"[1] that tracked the driving behavior of 10,000 drivers in Israel over the span of two years. As in many other countries, congestion is a first-order policy problem in Israel. The experiment recruited drivers from across commuting zones throughout Israel in a staggered fashion between 2020 and 2021. Each participating driver installed a monitoring device in their vehicle that could access their GPS coordinates over the course of their trips. During the first six months of participation, participants' driving behavior was monitored, but they did not receive any communication or monetary incentives. After the initial six-month period, drivers began to incur a time-and-location-specific fee for each kilometer driven out of a virtual budget, the remainder of which was paid out in cash upon drivers' annual vehicle registration renewals.

Our analysis is composed of three parts. First, we consider the direct effects of the congestion pricing experiment: to what extent did treated individuals reduce their driving in congested times and places, and how? To examine these questions, we apply an event study approach in the spirit of Borusyak et al. (2021) to estimate average treatment effects across a number of driving behaviors over the course of 16 weeks after individuals are treated. Next, we ask whether and how treatment

[1] Derech Erech is a play on words in Hebrew, as "Derech" means both "route" and "direction/way", and "Erech" means "valuable" or "to value").

effects co-vary with observable driver characteristics. This line of questioning is informative about two important dimensions along which congestion pricing policies like the one we study should be evaluated. The first dimension is the extent to which a congestion pricing scheme is effective at reducing the driving behavior that contributes the most to congestion. The second dimension is the inefficiency or regressivity of a congestion pricing scheme: does it impose high costs on individuals who are highly sensitive to prices but do not have enough flexibility in their schedules to avoid being charged? To address these questions, we apply the Sorted Effects method developed in Chernozhukov et al. (2018) to compare the drivers with the smallest and largest estimated treatment effects across covariates measuring their pre-treatment driving behavior, socioeconomic status, and the accessibility of public transit alternatives to trips they take by car. Finally, we consider the welfare implications of this congestion pricing policy in equilibrium were it to be applied to all Israeli drivers. To do so, we estimate a structural model of the demand by drivers in our sample for different routes and travel times for their commuting trips, along with a nonparametric statistical model of how traffic speeds change with the number of cars on the main highway in Israel. We then simulate equilibrium flows along the highway if all drivers nationally were to be subjected to congestion pricing and evaluate the changes in road speeds, commute times and driver utility.

To identify the causal determinants of driving choices in our panel, we rely on two types of variation: variation in when drivers were recruited to the experiment, and variation in drivers' baseline commuting behavior that is nationally representative of drivers who would be impacted by proposed congestion pricing policies. Our dynamic treatment effect estimates are predicated on the staggered recruitment of individuals: an average of about 200 individuals were recruited per week over the course of 50 weeks. As such, on a typical week in which a driver switches from "monitoring" to "treated" status, there is a segment of other drivers are still in "monitoring" who can serve as a control group to separate individual-specific idiosyncrasies in driving behavior and time-varying aggregate conditions from the effects of our congestion pricing treatment. In keeping with Borusyak et al. (2021), our treatment effect estimator models each individual's baseline driving behavior using their "monitoring" period. It then infers the change in behavior due to the onset of pricing by comparing the observed "treated" behavior against a predicted baseline based on the same driver's previous ("monitoring") driving, and the concurrent driving of other drivers still in their "monitoring" phase.

Our dynamic treatment effect analysis suggests that drivers take some time to fully adjust to the congestion pricing scheme, as the treatment effects take several weeks to stabilize. However, after 10 weeks of operating under the pricing treatment, drivers substantially decrease their driving in congested areas and times. Aggregating across time, we find that drivers decrease a variety of measures of congested driving by an average of 10% or more relative to pre-treatment median outcomes, both on the extensive (e.g. number of trips) and the intensive (e.g. price per Km driven) margins.

Going beyond averages to understand the heterogeneity in effects, we focus on three broad dimensions: flexibility, socioeconomic status, and baseline demand for driving on congested roads. A well-designed congestion pricing intervention needs to take all of these factors into account. If a policy doesn't affect enough flexible drivers or flexible trips, it will raise revenue but fail to reduce congestion. Similarly, if prices are not high enough, drivers will not find it worthwhile to reduce their congestion-inducing driving behavior. On the flip side, if we mostly affect price-sensitive drivers who have lower incomes, we may reduce congestion but our intervention will have a regressive effect. Moreover, if we only affect the behavior of flexible and price-sensitive drivers but if they are not the main contributors to congestion, our policy will again be ineffective in reducing congestion.

Using the Sorted Effects method to test for differences along these dimensions between drivers who were most and least responsive to the congestion pricing treatment, we generally find support for the potential efficacy of a congestion pricing policy in our setting. Across different characterizations of flexibility and baseline demand, the most responsive drivers in our sample exhibited more flexibility in their arrival and departure times and took more expensive cross-city and metro-oriented trips during their pre-treatment "monitoring" period. Treatment effects also correlated with the availability of alternatives to driving. Considering an index characterizing the convenience of the best public transit alternative to each vehicle trip taken by car during each driver's "monitoring" period, we find that the most responsive drivers also tended to have better access to relevant public transit options. We did not find significant differences in treatment effects based on indicators of socioeconomic status based on individuals' home neighborhoods and vehicle characteristics. While this last exercise is limited by the availability of precise individual data due to privacy restrictions, our results are suggestive that the possibility of disproportionate regressive impacts may not be a major concern.

Having established that our usage-based congestion pricing scheme meaningfully decreases driving during the most congested times and in the most congested areas, we next assess whether our treatment effects on driving behavior translate into meaningful impacts on road speeds. To do so, we first estimate how reducing the number of drivers on each road affects the speed at which cars can travel in congested areas. Focusing on driving along the largest highway in Israel, the Ayalon, we use a year of high-frequency, contemporaneous measurements of traffic flows to estimate the structural relationship between traffic densities and speeds nonparametrically. Importantly, we instrument for traffic density using exogenous variation in sensor distances from geolocated car accidents in space and time since traffic density and speed are determined simultaneously in equilibrium due to drivers' endogenous route choices.

We use our estimate of the congestion function to assess how the treatment effects from our experiment would impact highway speeds if they were to be applied to the entire population of Israeli drivers, proportionally to their area of residence. Applying our average treatment effect estimates directly—under a ceteris paribus assumption in which drivers do not additionally react to the change in road speeds—we find that the 7% decrease in highway trips implied by our ATT estimates would lead to an increase in traffic speeds of up to 30% during the most congested times of day. In ongoing work, we are working to relax the ceteris paribus assumption to account for individuals' equilibrium adjustments to their driving choices. To do so, we estimate a model of each individual's demand for taking their most frequent trips at different departure times (or not at all) given the prices and travel durations that the individual would have expected from each option under their treatment status and the contemporaneous traffic conditions. We then simulate the counterfactual demand for each stretch of highway across time in equilibrium, where the shares of drivers expected to use each stretch at a given time due to their travel decisions takes into account the impact of the aggregate travel flows on the road speeds that they encounter.

Our paper offers several contributions to the literature: First, by studying a long-term experiment at a national scale, we are able to quantify the average impact of a congestion pricing mechanism that resembles the types of policies that are under consideration. The long duration of the experiment enables us to characterize drivers' responses over a longer period than previous studies. This point is particularly important since we find considerably larger effects on driving behavior when we extend the treatment time window, allowing time for participants to learn to adjust.

Second, observing driver-level incentives and driving responses allows us to identify the effect of congestion pricing on different subsamples of drivers. Understanding which populations are more sensitive to treatment and which are less could be instrumental for informing the design of maximally effective congestion pricing policies, as well as for assessing the scope for adverse distributional impacts. Third, we flexibly estimate the structural relationship between traffic density and speed while accounting for the endogeneity of these two quantities, which allows us to more accurately predict the counterfactual distribution of traffic speeds under our congestion pricing scheme.

In future work, we intend to extend our equilibrium framework to account not only for travel on major highways, but for travel throughout the Israeli road network writ large. To do this, we will combine data from Google Maps, public transit GPS records, and trips taken within our experiment to estimate a model of travel speeds along Israeli roads throughout our experiment. Using this panel of road speeds, we will scale up our travel demand model to account for all frequent trips (not just those that use the highway), estimate a non-highway congestion function, and simulate counterfactuals that target congestion in different areas. Finally, we will use our framework to examine the implications of a congestion pricing scheme not only in raw time savings but also in terms of shifts in consumer surplus across the population of drivers and to compare existing policy proposals against a hypothetical first best allocation of drivers to trips.

1.1 Existing Literature

Pigou (1912) and Knight (1924) first and subsequent theoretical contributions by Vickrey (1969) and Arnott, Palma and Lindsey (1993) show that congestion pricing policies can be efficiency-enhancing solutions to reduce traffic congestion. However, empirical research that examines to what extent drivers actually respond to monetary incentives to avoid congested roads, and how overall congestion and welfare may be impacted is limited. Due to data limitations, existing studies often rely on measuring aggregate effects of congestion charges on overall traffic using time variation for identification. These studies often struggle to identify the causal impact of congestion charges since factors other than the implementation of new charges typically also change simultaneously.

In a few cases, studies have used field experiments to examine the impact of congestion pricing. These experiments allow researchers to examine how drivers respond to congestion pricing, and which trips are more sensitive at the individual-driver level. The evidence from these experiments has thus far offered limited support for the effectiveness of congestion pricing. Martin and Thornton (2017) analyze the effects of a congestion pricing experiment in Melbourne, Australia. The experiment collected time and location data from 1,400 vehicles over nine months, during which participants were exposed to one of several congestion pricing treatments and a system of credit accounts similar to ours. A main finding by Martin and Thornton (2017) is that higher prices do not affect work commutes, the use of highly congested roads, or time spent driving at low speeds. They find that drivers adjust to higher prices primarily by taking fewer trips to malls and shops. Kreindler (2022) examines the effect of a field experiment in Bangalore, India. Using detailed travel data on 497 commuters and a usage-based pricing rule, Kreindler (2022) finds that commuters do not change the number of trips that they take, and make only minor changes to their routes otherwise. Extrapolating these results with an equilibrium model, Kreindler (2022) concludes that congestion pricing would not substantially reduce congestion in a city like Bangalore.

Other notable recent contributions include Barwick et al. (2022) who use data from household travel diaries and housing transaction data from Beijing. They estimate an equilibrium sorting model of housing location and commuting mode choice with endogenous traffic congestion to identify residents' home and work locations and evaluate the impact of alternative urban transportation policies. Durmeyer and Martinez (2022) use decennial survey data on commuting choices from Paris to estimate a structural model that considers individual transportation decisions, equilibrium road traffic levels, and speeds inside a city under different driving restrictions. Hahn, Metcalfe and Tam (2022) analyze a field experiment giving subsidies for off-peak public transit use in California and find evidence both that a policy of this type may increase overall welfare, and that a partial equilibrium assessment of welfare may overpredict the potential for welfare gains.

2 Data

2.1 The Experiment

The primary source of data for this paper is a large-scale field experiment structured toward assessing the impact of "usage-based" congestion pricing, in which drivers pay a per-kilometer fee (similar to a "Vehicle Miles Traveled" (VMT) fee) that varies by time and location. The experiment ran between November 2019 and November 2021 under the administration of Ayalon Highways Co. and was funded by the Israeli Ministries of Finance and Transportation.[2] Since late 2019, we have served as an independent research team collaborating with the Israeli government regarding the evaluation of the experiment. We joined the experiment after its design was determined.

Participating drivers were recruited in a staggered fashion across the months between January 2020 and October 2020. Approximately 10,000 drivers were recruited in total from across Israel to participate in the experiment, with a mean (median) of 648 (555) drivers recruited each month. Recruitment was done based on pre-defined geographic quotas, designed to capture the relative contribution to congestion from the driving behavior of each area. Figure 1 depicts the resulting spatial distribution of participating drivers. The majority of drivers reside in the broader Tel Aviv commuting zone, which is home to nearly half of the Israeli population.

To operate the 'driver side' of the experiment, two companies were contracted: (1) Pango, Israel's

[2] Ayalon Highways Co. is a metropolitan road company, also responsible for advanced transportation solutions and transportation innovation. It is a government entity similar to a national highway administration.

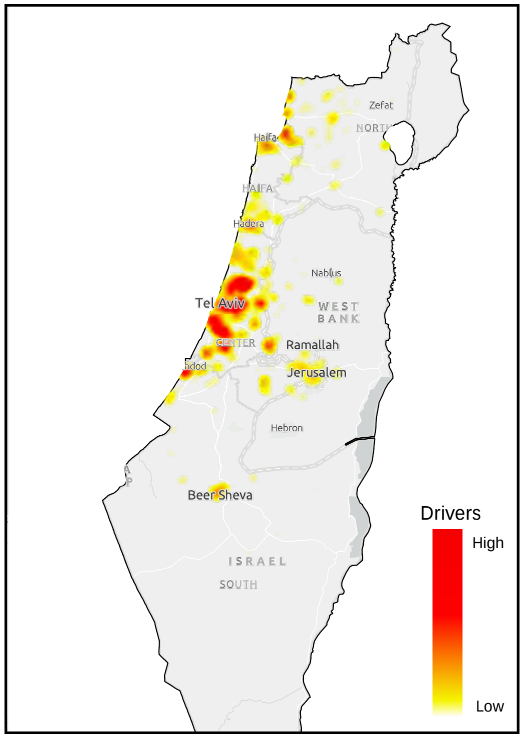
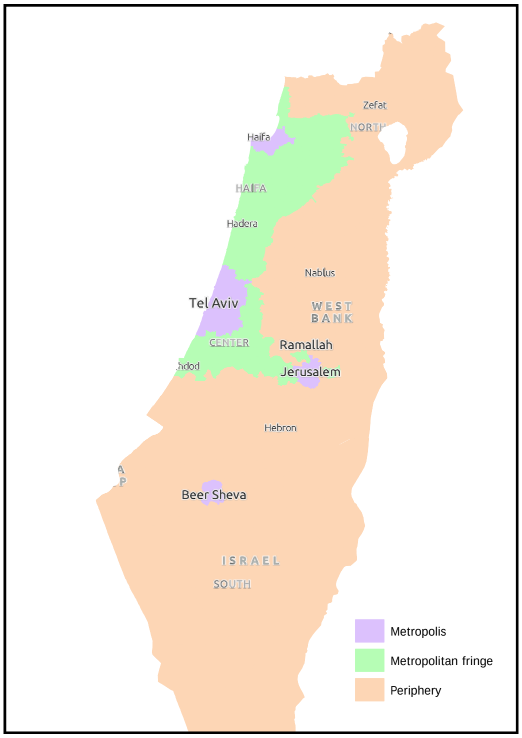


Figure 1: Areas of Recruitment, by percentage of drivers from each area.

leading parking payment company, and (2) Pointer, an Auto-Tech company that specializes in the protection and location of stolen vehicles. These companies recruited drivers directly through online marketing and telemarketing campaigns, as well as by approaching existing customers and presenting the experiment to them. Upon recruitment, drivers underwent a brief telephone interview, during which participants were asked to provide administrative information (like their bank account, car license, and car ownership documents) and were given a brief explanation of the experiment's rules. Once participants signed a participation agreement, they had to schedule the installation of a monitoring device, an inconspicuous, difficult-to-remove component that can track when a vehicle's engine is running or idle, as well as its GPS location.[3]

In the first six months after a driver joined the experiment, her driving behavior was monitored and recorded without monetary incentives for any type of driving (we refer to this as the driver's "monitoring" period). After six months in the monitoring period, drivers switched to an "active" status, and began to incur a time- and location-specific fee for each kilometer driven out of a virtual annual budget of 4,500 NIS (approximately $1,250). The virtual budget translated to real money: at each annual registration renewal, drivers were paid the pro-rated remainder of their budget up to 2,000 NIS by an automatic bank transfer. Subsequently, they continued in the experiment under the same conditions but with a renewed initial budget until the following year (The annual budget stayed the same, from one license renewal to another. In case the driver's initial charging period was less than 12 months, her initial budget was reduced proportionally).

[3]Most drivers already had a GPS tracking device installed as part of insurance requirements. In such cases, there was no need to install a second one.



(a) Pricing Zones.

	Hours	Metro	Metro Fringe	Periphery
	Weekdays			
Peak	6:45 - 9:30AM 3:30 - 6:30PM	1.5	0.3	0
Moderate	9:30AM - 3:30PM 6:30 - 8PM	0.1	0	0
Low	8PM - 6:45AM	0	0	0
	Weekends			
	All Hours	0	0	0

(b) price in NIS per kilometer traveled

Figure 2: Pricing schedule for the usage-based congestion pricing scheme.

The per-kilometer fee structure was designed to disincentivize driving on the most congested roads. Times were bucketed into "peak" (6:45–9:30am and 3:30–6:30pm), "moderate" (9:30–3:30am and 6:30–8:00pm) and "off-peak" (8:00pm–6:45am), while locations were bucketed into "metro" (the four major Israeli cities, Tel Aviv, Jerusalem, Haifa, and Beer Sheva), "sub-metro" (the commuting zones in which the majority of the Israeli population lives) and "periphery." See Figure 2b for the per-kilometer fees by time of day and geographic zones and Figure 2a for a map of the pricing zones.

Figure 3 summarizes the distribution of drivers' staggered recruitment and activation times. The y axis indexes drivers eventually recruited for our study, and the x axis indexes weeks relative to the first week of recruitment, which was January 5th, 2020. The coordinate corresponding to a given driver and week is colored red if that driver had not yet begun having their driving behavior monitored by that week, yellow if the driver was in their six-month monitoring period during that week, and blue if the driver was in the active, treated state during that week.

For each trip taken by a driver in the experiment, we observe the origin and destination at the level of a Traffic Analysis Zone (TAZ), the distance traveled, the trip duration, and the price that the driver paid (if she was in active status) or would have paid (if she was in monitoring status), based on the per-kilometer fee structure. Similar to US Census blocks, TAZs are geographic polygons that divide the land area of Israel into complete and non-overlapping parcels with relatively homogeneous sizes (in terms of population), but with the primary purpose of analyzing and modeling transit patterns. Due to privacy concerns, we cannot observe the specific location of a driver at any point in time. However, the TAZs can be quite small: metro area TAZs have an average (median) area of 0.7 (0.4)

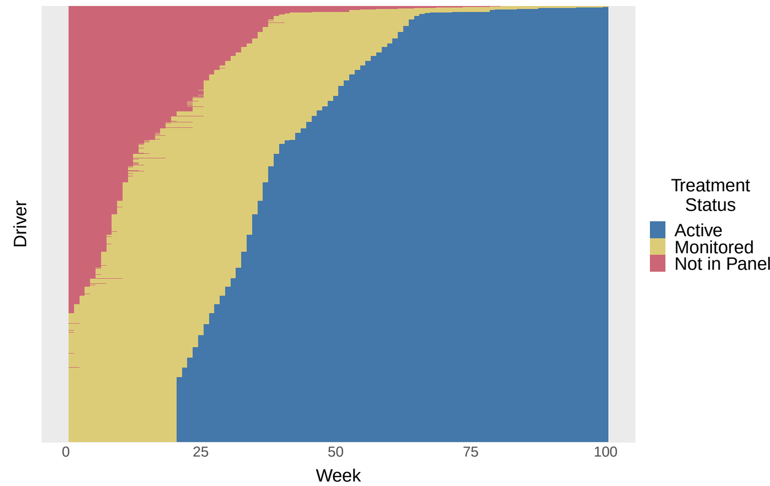


Figure 3: Participant status by week, starting on the week of January 5th, 2020. The y axis indexes drivers eventually recruited for our study, and the x axis indexes weeks relative to the first week of recruitment, which was January 5th, 2020. The coordinate corresponding to a given driver and week is colored red if that driver had not yet begun having their driving behavior monitored by that week, yellow if the driver was in their six month monitoring period during that week, and blue if the driver was in the active, treated state during that week.

square kilometers (naturally, the periphery area TAZs are bigger), and encode information regarding the types of establishments and land uses (residential, commercial, manufacturing, leisure, etc.) in each TAZ. In addition, while we cannot see the full route that drivers take, we do observe a breakdown of the trip based on the price formula: for each transition between time buckets or geographic zones, we observe the distance, duration, price, and start and end TAZs of that trip segment. Due to data quality issues, we leave incorporating this finer-grained trip information to future work.

In addition to trip characteristics for each driver, we observe each driver's participation status (monitoring or active) on each trip and the date of the status change; the time of payment of the prorated remainder budget and the amount paid; whether the dedicated mobile app of the experiment was downloaded after becoming active, as well as how much the driver used the app;[4] and the pollution level of the car and other car characteristics based on its license plate (we could not observe the license plate number itself, for privacy reasons).

2.2 Auxiliary Data

We also supplement our rich data on participants' driving behavior with the following data sources:

Highway sensor data The main highway on the Tel Aviv Metropolitan Area (which is the largest city in Israel) is called the Ayalon Highway and is operated and maintained by Netivey Ayalon Co. We received access to traffic volume and traffic speed information collected using Inductive Loop Detectors (ILD) sensors located at different locations along that highway. We use data on average

[4] A dedicated mobile app was designated the experiment to provide drivers information on the remaining balance of their budget and on estimated time and price for planned trips. drivers were offered to download the app once they switched to the "active" status.

traffic densities and speeds collected at 5 minute intervals over the course of one year to estimate the congestion function mapping traffic densities to speeds in Section 6.

Open Street Map Since we only observe the start and end TAZ for each trip in our data, we use open-source data on the Israeli road network from Open Street Map to identify which trips likely passed through highways like the Ayalon when assessing how our treatment effects on driving behavior would translate into road speed effects (OpenStreetMap contributors, 2017). To do so, we compute the fastest path through the road network between each trips start and end TAZ, taking road segment traversal speeds to be the speed limits reported in the Open Street Map data, and treat a trip as passing through a highway if any road segment in a fastest path was classified as a highway.

Public Transit Data We use information on historical public transit schedules and bus stop locations published and made publicly available by the Israeli Ministry of Transportation. We use this information to calculate possible public transit alternatives for the observed trips conducted by the drivers, as described in more detail in Section 3.

Demographics information We use demographic information obtained from the Israeli Central Bureau of Statistics (CBS) for each TAZ that includes information on population density and growth, the population age distribution, the percent of students and employees, the percent of minorities (Arab in particular), the number of Yeshivas (religious schools for orthodox Jews), and the TAZ socioeconomic ranking.[5]

3 Descriptive Statistics

Baseline Driving In our analysis, we include only trips that took place on weekdays (Sunday through Thursday in Israel) and exclude all trips that occur on holidays. Table 1 presents the descriptive statistics of our participants' driving weekly behaviour as observed during the "monitoring" (pre-treatment) period. The average (median) driver in our sample drives 14 (13) times each week (or 2.8 (2.6) trips each day) and spends 6 (5.3) hours behind the wheel each week (or 1.26 (1.1) hours per day). The average trip in our sample is 13 Km long. Drivers in the experiment take more trips than the average Israeli driver, who takes 1.9 trips per weekday (about 10 trips per work week). However, their typical trip tends to be shorter in distance and longer in duration, compared to the average Israeli driver who drives 21 Km per trip and spends less than one hour a day on the road.[6] Since the recruitment of the drivers for the experiment was determined based on riders' contribution to congestion, these differences are not surprising.

[5]The CBS formulated an index to characterize the local authorities in Israel according to the social level - economic of the population in them. The index is calculated based on several demographic and economic variables, including dependency ratio, average years of schooling in the adult population, percentage of academic degree holders, employment and income levels, etc. Lower values correspond to lower socio-economic status.

[6]Statistics on the general population is based on the most recent available cellular survey data.

				Quantiles					
Variable	Mean	SD	Min	10	25	50	75	90	Max
Weekly Trips	14.19	9.33	0	2.00	7.00	13.00	20.00	26.00	130.00
Weekly Km Driven	186.28	190.14	0	9.90	53.50	136.60	259.40	419.90	17309.90
Weekly Trips Dutration (Hours)	6.31	5.28	0	0.71	2.57	5.35	8.71	12.57	99.33
Weekly Price	45.11	51.66	0	0.00	4.43	27.01	67.90	119.82	646.00

Table 1: Summary Statistics of Main Outcome Variables, Pre-Treatment

Flexibility covariates To measure the relative flexibility of driving behavior and the relative costs of taking trips at other times, we construct multiple measures using monitoring period data. First, we focus on the consistency of departure times in the morning and in the evening. To capture whether a subject consistently drives earlier or later compared to highly congested peak hours, we measure the average end time of their last trip during morning peak hours, and the average departure time of their first trip during evening peak hours. To measure the variability in these trips, we use the entropy in the distribution of last morning trip and first evening trip departure times.

Personalized public transit indices To get a better understanding of the availability of travel options available to our subjects besides private vehicles, we construct a personalized public transit index. For each trip taken by our drivers during their monitoring period between an origin and a destination TAZ, we look at whether the trip could be made at that time of day with public transit, using our public transit data described in Section 2.2. If a potential public transit route exists between these two locations, we measure the predicted duration of this route at that time of day and take the ratio of the observed trip duration and the predicted duration of the equivalent public transit trip. We then take the average of these trip level duration ratios to compute the driver-level public transit index. If a public transit trip cannot be found for a given origin-destination pair, we set the ratio be equal to zero for that trip. Higher values of our index reflect the availability of public transit options between a driver's observed origin-destinations with relatively shorter travel times compared to the observed driving duration.

Socioeconomic covariates To learn about the demographic characteristics of the drivers in our sample, we match the driver home's TAZ with the demographic information from the CBS described in Section 2.2. To infer each driver's home TAZ, we calculate each driver's most frequent origin TAZ of their first trip of the day and their most frequent destination TAZ of their last trip of the day; whichever occurs more frequently is assumed to be the driver's home TAZ. The demographic information from the CBS provides the following information on each driver home TAZ: a socioeconomic index for which higher values reflect higher socio-economic status, log population density, population growth between 2012 to 2019, median population age, population share of students, population share of Arabs, and number of Yeshivas as a measure to the level of religiosity in the TAZ.

4 Average Effects of Usage-Based Pricing

Having detailed our setting and data, we turn to evaluating what average effects our congestion pricing scheme had on the driving behavior of participants in the experiment.

4.1 Identification and Estimation

To estimate the average effects of our congestion pricing scheme on driver behavior over time, we would ideally take advantage of random variation in the treatment statuses across the drivers in our sample in each week, e.g. random assignment of a group of drivers to a never-treated group or randomizing monitoring period lengths across drivers. However, recall from Section 2.1 that Ayalon Highway Co. designed the study so that drivers entered the study in a staggered fashion over time, and every driver was eventually treated after being monitored for the same number of weeks. As such, the only variation in treatment statuses across drivers in any given week of our study is driven by variation across drivers in recruitment times. Furthermore, the week that each driver entered the study may have depended on a litany of unobserved factors that could be correlated with potential driving behavior, such as the spatial recruitment priorities of Pango and Pointer over time and drivers' idiosyncratic eagernesses to begin participation in the study once recruited.

To circumvent these challenges, we rely on our repeated observations of drivers' pre-treatment behavior across weeks to account for persistent unobserved confounders across drivers and the presence of monitored drivers in our study each week to account for variation in driving behavior due to systematic time-varying shocks like COVID-19 prevalence. In particular, we apply the difference-in-differences framework for treatment effect identification and estimation described in Borusyak et al. (2021) and Callaway and Sant'Anna (2021). For each driver i in our study sampled from some superpopulation of Israeli drivers, we let M_i denote the length of their monitoring period and let C_i denote driver i's congestion pricing scheme activation week. Fixing some measure of weekly driving behavior as an outcome of interest, we let $Y_{it}(c)$ denote driver i's potential outcome in week t were the congestion pricing scheme to become active for i in period c, we let $Y_{it}(0)$ denote driver i's control potential outcome in week t if the congestion pricing scheme were not active for i by week t, and we let $Y_{it} := Y_{it}(C_i \mathbb{1}\{C_i < t\})$ denote driver i's observed outcome in week t. Our estimands of interest can be expressed as functions of *cohort average treatment effects on the treated* (CATTs) τ_{ct}, namely the average difference between treatment and control potential outcomes in week t amongst the *cohort* of drivers who were first treated in period $c \leq t$:

$$\tau_{ct} := \mathbb{E}[Y_{it}(c) - Y_{it}(0) \mid C_i = c].$$

To identify CATTs τ_{ct}, we impose a two-way fixed effects (TWFE) model of control potential outcomes as in Borusyak et al. (2021):

$$\mathbb{E}[Y_{it}(0) \mid \alpha_i] = \alpha_i + \gamma_t,$$

where α_i denote driver-specific fixed effects and γ_t denote week effects. The TWFE model implies the parallel trends assumption, namely that drivers' average outcomes would evolve in parallel across weeks in the absence of treatment. Assuming that the number of drivers treated in each week is non-negligible relative to the large sample size of drivers N,[7] our CATTs of interest are identified in a standard cohort-specific difference-in-differences fashion (Callaway and Sant'Anna, 2021; Borusyak et al., 2021).

To translate the difference-in-differences identification intuition into an estimation strategy that will also allow us to study treatment effect heterogeneity in Section 5, we follow the imputation approach of Borusyak et al. (2021). In particular, we can construct consistent estimates $\hat{\gamma}_t$ of the week effects γ_t and asymptotically unbiased estimates $\hat{\alpha}_i$ of the driver fixed effects α_i by running the following fixed effects regression on drivers' pre-treatment observed outcomes, where $T < \infty$ denotes the number of weeks in which we observe drivers' outcomes (Borusyak et al., 2021):

$$(\hat{\alpha}_i, \hat{\gamma}_t) := \underset{\alpha_i, \gamma_t}{\arg\min} \sum_{i=1}^{N} \sum_{t=1}^{T} \mathbb{1}\{C_i > t, C_i - M_i \geq t\} \left(Y_{it} - [\alpha_i + \gamma_t]\right)^2 .$$

We can then construct asymptotically unbiased imputed control potential outcomes $\hat{Y}_{it}(0) := \hat{\alpha}_i + \hat{\gamma}_i$ and individual treatment effect estimates $\hat{\tau}_{it} := Y_{it} - \hat{Y}_{it}(0)$ for each driver i and week $t \geq C_i$. We can also average these individual treatment effects across drivers in the same cohort to construct consistent estimates of CATTs τ_{ct}:

$$\hat{\tau}_{ct} := \sum_{i=1}^{N} \mathbb{1}\{C_i = c\}\hat{\tau}_{it} \Bigg/ \sum_{i=1}^{N} \mathbb{1}\{C_i = c\}.$$

4.2 Treatment Effect Dynamics

Our first set of results concern the effect dynamics of our usage-based congestion pricing treatment over time. Because our model of control potential outcomes (4.1) doesn't restrict treatment effect heterogeneity across drivers and weeks (Borusyak et al., 2021), we can aggregate our CATT estimates $\hat{\tau}_{ct}$ across cohorts to estimate average CATTs after e periods post-treatment τ_e:[8]

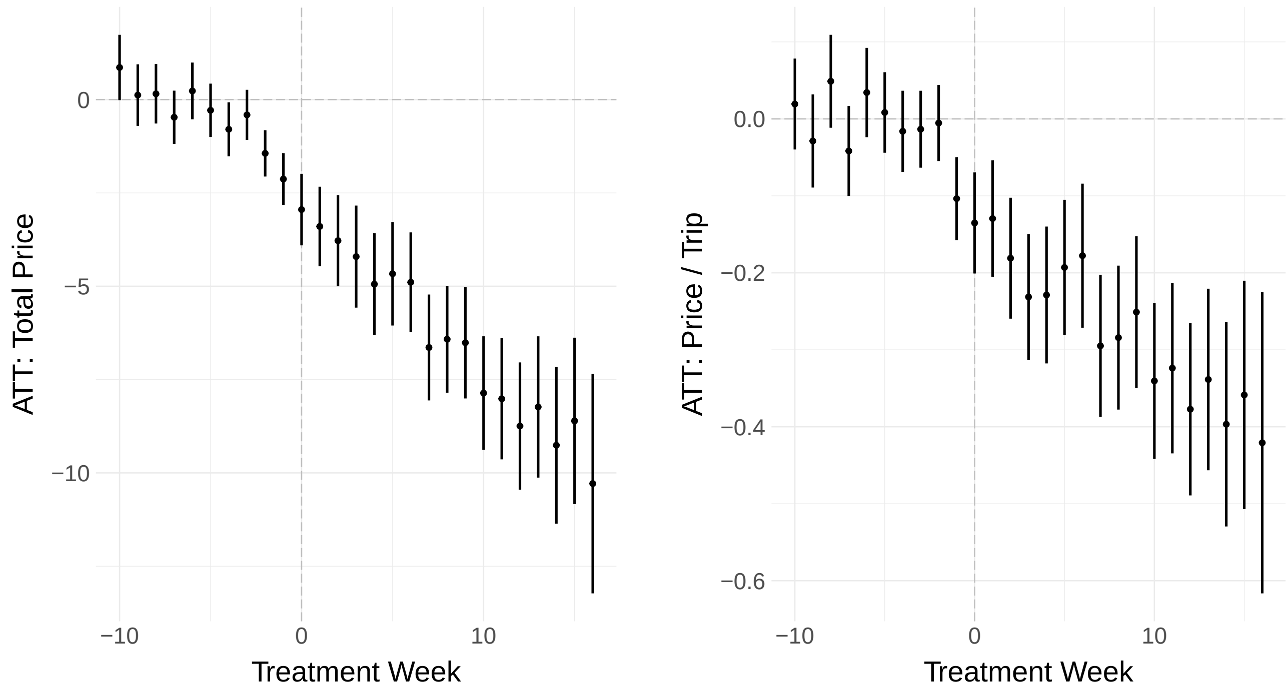
$$\tau_e := \sum_{c=1}^{T} \mathbb{P}(C_i = c)\tau_{c(c+e)}, \quad e = \underline{e}, \ldots, \overline{e}.$$

To be able to visualize potential pre-trends that would invalidate our control potential outcome model, we allow $\underline{e}$ to be negative. We conduct valid inference on dynamic treatment effect paths by constructing 95% simultaneous confidence bands via a weighted bootstrap procedure akin to the one proposed in Callaway and Sant'Anna (2021) with Dirichlet (i.e. Bayesian bootstrap) weights (Shao and Tu, 2012).

Figure 4 displays estimates of dynamic average treatment effects τ_e on several payment-related

[7] Formally, that $\mathbb{P}(C_i = c) > 0$ for all treatment weeks c.

[8] As discussed in Callaway and Sant'Anna (2021), for larger relative time horizons e, $\tau_{c(c+e)}$ is not estimable for some cohorts who are treated for fewer than $c+e$ weeks. Rather than change the set of cohorts we average over when estimating τ_e across different relative weeks e, we report dynamic effects free from contamination from these composition effects by estimating τ_e using only the cohorts of drivers for whom we observe outcomes for at least $\overline{e} = 16$ weeks post-treatment.



(a) Outcome: total weekly price paid

(b) Outcome: price per-trip

Figure 4: Event studies of dynamic average treatment effects across weeks relative to treatment week.

outcomes and associated 95% simultaneous confidence bands over the 10 weeks before and 16 weeks after each driver's treatment week on several outcomes of interest. In Figure 4a, we plot an event study treating the total weekly price deducted from each driver's budget as our outcome of interest. As prices are highest for driving in congested areas and times, a decrease in the total price most directly corresponds to a likely decrease in congestion-inducing behavior. In Figure 4b, we plot a similar event study, using the price per trip as an outcome. The price per trip averages out the total number of trips taken. As such, we interpret effects on it as intensive-margin responses: a lower price-per-trip indicates that treated drivers' trips were less likely to take place in high-priced areas and times on average. The effect on total weekly price seen in Figure 4a is thus driven by changes in drivers' trip times and routes, rather than just decreases in the total numbers of trips drivers take. The magnitudes on both the total weekly price effect Figure 4 are in NIS saved. To put these magnitudes in context, the pre-treatment medians of these variables are 35 NIS and 2 NIS/trip, respectively.

As can be seen in Figure 4, the treatment effects take several weeks to stabilize. This effect lag suggests that drivers may need several weeks to learn how best to adjust their behavior, and so studies with treatments that last a month or less like Kreindler (2022) may understate the full effect of the intervention. We also note that the anticipation effects depicted in the two weeks prior to treatment week in Figure 4 are likely due to noise in when drivers were actually alerted that they were going to begin getting treated, as well as minor inaccuracies in the recording of the exact treatment week. To assess the robustness of our treatment effect estimates to potential anticipation effects, we find similar results to those plotted in Figure 4 when we treat two weeks prior to each driver's measured treatment week as their effective treatment week (Callaway and Sant'Anna, 2021).

4.3 Longer-Term Effects Across Outcomes

We now turn to studying the average effects of our congestion pricing treatment on a wider variety of driving outcomes to better understand how drivers respond to our incentives. While the dynamic treatment effect paths reported in Section 4.2 provide rich information on how drivers' behavior changes over time, for parsimony in this section, we focus on treatment effect estimates $\tau := \frac{1}{12}\sum_{e=4}^{16}\tau_e$ that average the relative week-based average treatment effects τ_e over the fourth through 16th weeks relative to treatment. We use this aggregation to try to capture the longer-term effect of our intervention, since as discussed in Section 4.2, we see evidence that it takes time for drivers to re-optimize and adjust their behavior.

Throughout this section, to circumvent multiple testing problems when reporting results, we use thick lines to denote uncorrected confidence intervals computed via a Bayesian bootstrap, thin lines to denote Bayesian bootstrap-based confidence intervals that control the family-wise error rate (FWER) within each plot at 5% (Romano et al., 2010), and a non-gray color for a point estimate and its corresponding confidence intervals to indicate that the stepdown procedure developed in Romano and Wolf (2005) rejects the null hypothesis of zero effect size while controlling per-plot FWER at 5%.[9]

We expect drivers to respond differently to our incentives based on a range of observed and unobserved factors: the relative cost of being late or the reliability of outside options such as public transit, the relative flexibility or the convenience provided by driving, or the 'urgency' of each trip for the driver. For this reason, we broadly categorize trips to capture common driving patterns. We explain our trip categories and report results within each category below.

First, we categorize trips by departure time, looking at AM peak and PM peak separately, as well as the 'first' AM and PM trips taken by our drivers. Second, we categorize trips by their origin-destination characteristics: We use the label 'work trips' for trips that end in a TAZ that our location data set classifies as having a primary land use purpose related to a commercial activity [10]. We also label trips that end in a TAZ with a leisure related land use as 'leisure trips' [11]. We also split cross-city trips and within-city trips depending on the origin-destination pairs being in the same city or not, as well as separately categorizing trips ending in one of the four major Israeli metropolitan areas which we call 'Metro trips'. Generally, cross-city and metro trips are longer than within-city trips. Finally, we identify trips that represent the most frequently taken, habitual trips for our drivers - we compute the top 5 origin-destination pairs for each one of our drivers during their pre-activation period, and label trips between these as 'common trips.'

We plot estimated ATT's for total weekly price in Figure 5. Our estimated effect over all trip categories is a 7.5 NIS decrease in our drivers' weekly price paid, which is about a 0.15 standard deviation reduction. This overall decrease in prices paid is consistent with a meaningful reduction in

[9]The stepdown procedure developed in Romano and Wolf (2005) is able to be less conservative when rejecting null hypotheses than simultaneous confidence band-based testing procedures while maintaining the same FWER control; to our knowledge, no procedure exists for converting stepdown hypothesis tests into confidence intervals with simultaneous coverage guarantees but shorter widths.

[10]Specifically, TAZ's that have the primary land use types: Industrial, Engineering, Mining, Academic, Work, Education, Trade, Hospital, Government, Port, Recycling, Employment, Power, Commerce

[11]Specifically, TAZ's that have the primary land use types: Tourism, Stadium, Leisure, Cemetery, Sports, Nature, Heritage, Open Space

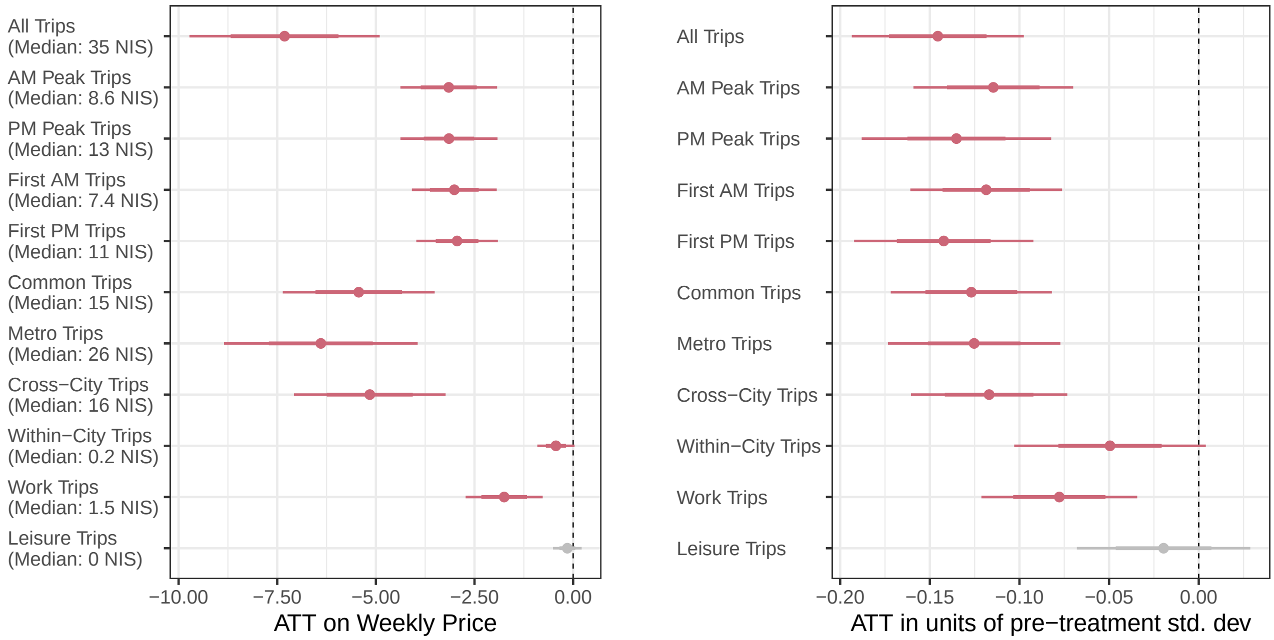


Figure 5: Average effects over week 4 to 16 post-treatment by trip category for total weekly price

congestion-inducing driving, because trips that are priced tend to be those that contribute the most to congestion.

The raw effects measured in NIS saved are not uniformly distributed over all trips. Metro trips, common trips and cross-city trips are where we observe the highest estimated effects around 5-6 NIS per week, whereas we see minor or insignificant effects for within-city and leisure trips (less than 1 NIS per week).

This pattern is explainable by the differences in the baseline prices: cross-city and metro trips are more likely to go through roads where congestion pricing is applied and tend to be longer. These trips cost significantly more than leisure or within-city trips, which tend to be shorter and may not pass through priced roads at all. The effects on AM and PM peak trips are similar in terms of magnitude (about 3 NIS per week), and lower than the effects for cross-city and metro trips since these time-based categories aggregate over shorter and longer distance trips.

Looking at effects measured in units of pre-treatment standard deviations, the difference between trip types is minor. We see estimated effects of about 0.13 standard deviations across all categories, except for within-city trips, work trips and leisure trips, where the baseline median prices are very low compared to other categories.

Next, we turn to estimated effects on total driving distance, price per trip, and price/kilometer to understand behavior change on the extensive versus intensive margins. Are our drivers still driving similar distances but taking cheaper routes or leaving at cheaper times? If they reduce overall distance driven and forgo some trips, are we still successful in disincentivizing them from driving during congested hours on busy roads for the trips that they still choose to take?

Driving distance is a relevant metric for our extensive margin impact, and also relevant for in-

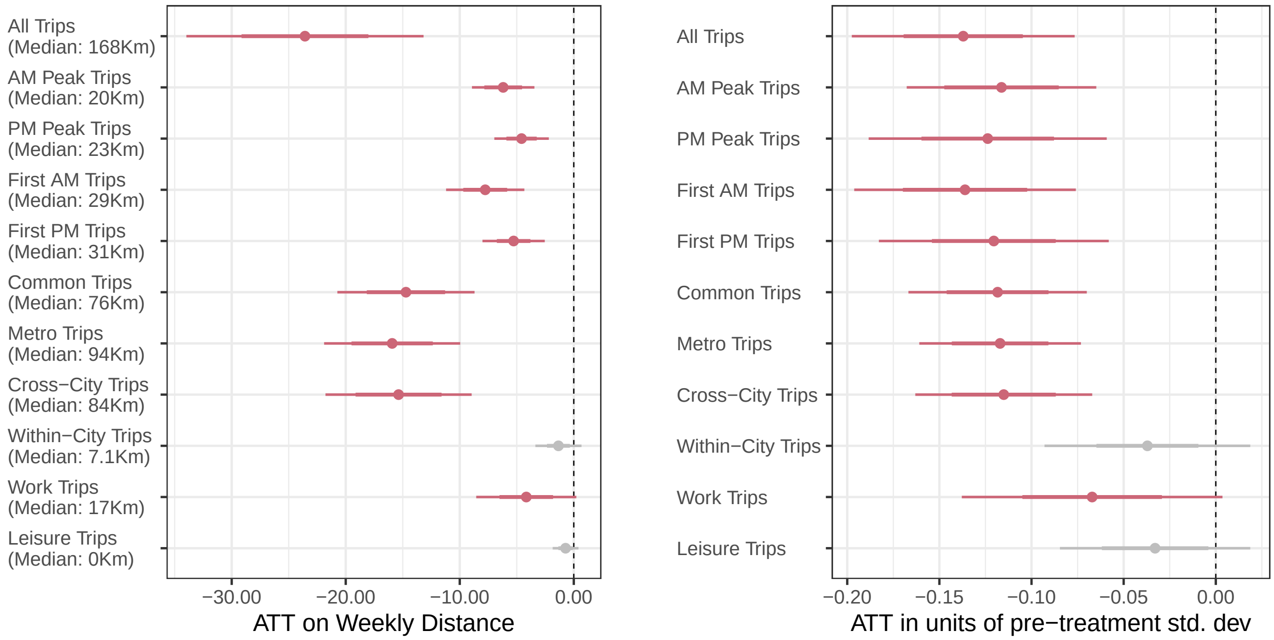


Figure 6: Average effects over week 4 to 16 post-treatment by trip category for total weekly distance driven

forming a portion of the direct environmental effect of our intervention through reduced driving and reduced CO2 emissions. We plot estimated effects on distance in Figure 6. Our estimated ATT for total weekly driving distance is a 24 kilometer decline, which translates to about 0.14 pre-treatment standard deviations. Within-city trips constitute a small portion of all driving in terms of distance, with a pre-treatment weekly median of 7 kilometers, whereas the weekly median for metro trips is 94 kilometers. We see a similar pattern with our effect estimates on total price, where the highest impacts are concentrated on non-leisure and longer trips.

To understand our impact on the intensive margin, we look at estimated effect on price per trip in Figure 7, and price per kilometer in Figure 8. Our drivers tend to take cheaper trips across all categories during the experiment period, paying 0.3 NIS less per trip or 0.02 less per kilometer driven. AM and PM peak hour trips see the highest decline among our categories measured in raw units, although the relative effect magnitudes in terms of pre-treatment standard deviations are about the same across all trips. The exception again are leisure trips where we see null effects, which constitute a very small portion of our trips with a pre-treatment median of 0 kilometers, and 0 NIS spent per week.

Our estimated effects on the intensive margin suggest either a departure time shift (to times where our drivers won't be charged) or a route shift (to routes where our drivers won't be charged) for trips that are still taken during the experimental period. While we can't directly address the route shift question since we don't have granular data on routes taken by our drivers, we have precise measurements of departure times.

For each trip's departure time, we compute the distance to the congestion pricing window in hours. This helps us understand whether our drivers adjust their behavior to avoid congested times when

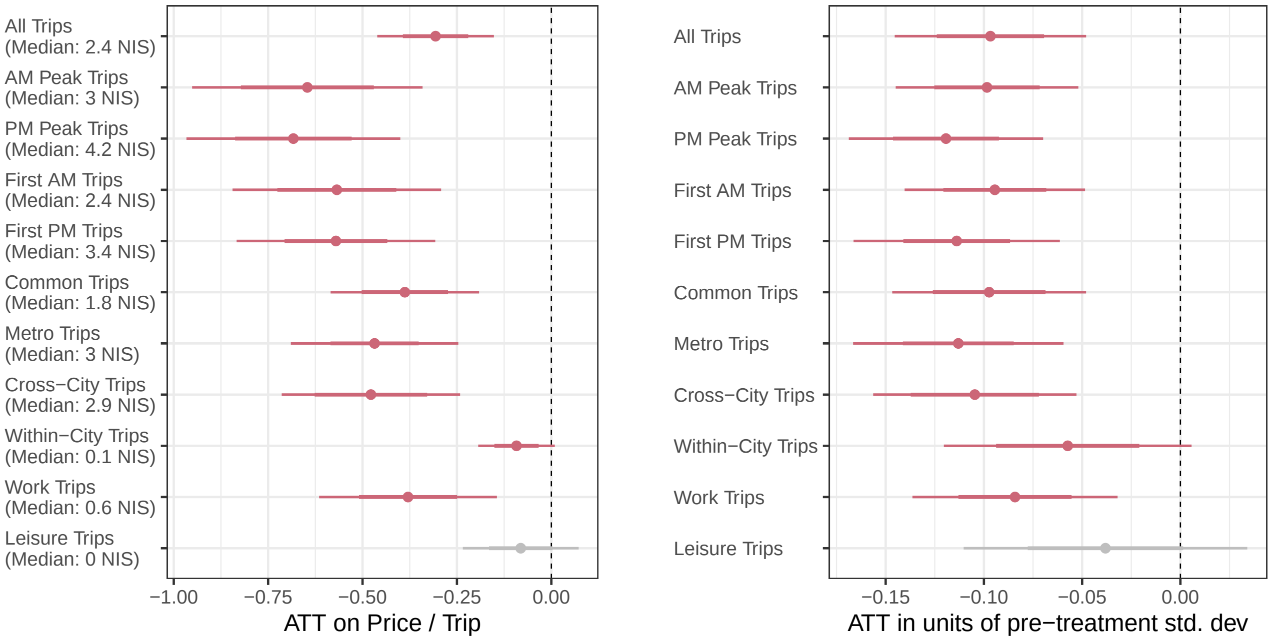


Figure 7: Average effects over week 4 to 16 post-treatment by trip category for price per trip

our pricing is active. We measure this distance symmetrically around two directions of the pricing window: a positive sign implies trips starting closer to the congestion pricing window, and a negative sign implies trips starting further away from the congestion pricing window [12].

In Figure 9, we report our estimates on departure time shifts. We see that our impacts are significant for evening trips, where drivers tend to move their trips away from congestion pricing window by about 3 minutes on average. The effect is mostly concentrated on metro trips and cross-city trips, and specifically the first PM trip taken by our drivers on a given day. This suggests that time shift behavior mostly impacts long-distance evening commutes between different cities, or between the suburbs and metro areas. Null effects on morning trips also suggest our drivers have less flexibility on average to change their departure times for morning commutes, compared to their evening commutes. Prior evidence from congestion pricing experiments in different cities around the world also shows departure time shift is a common response to time-restricted pricing schemes, which is consistent with what we see in our results (Li and Hensher, 2012; Karlström and Franklin, 2009).

[12] For AM peak: let $[l, u]$ be the peak time interval measured in hours since midnight, let $m := (l + u)/2$ be the midpoint, and let x be the relevant trip's departure time. If $x < m$, then our measure is is x - l , so negative times mean the trip started before peak hours and positive times mean the trip started after; if $x >= m$, then our measure is $-(x - u)$, so negative times mean the trip started after peak hours and positive times mean the trip started before peak hours. We define the distance to PM peak symmetrically, so that negative signs imply trips starting outside the peak window, and positive sign implying trips starting inside the peak window.

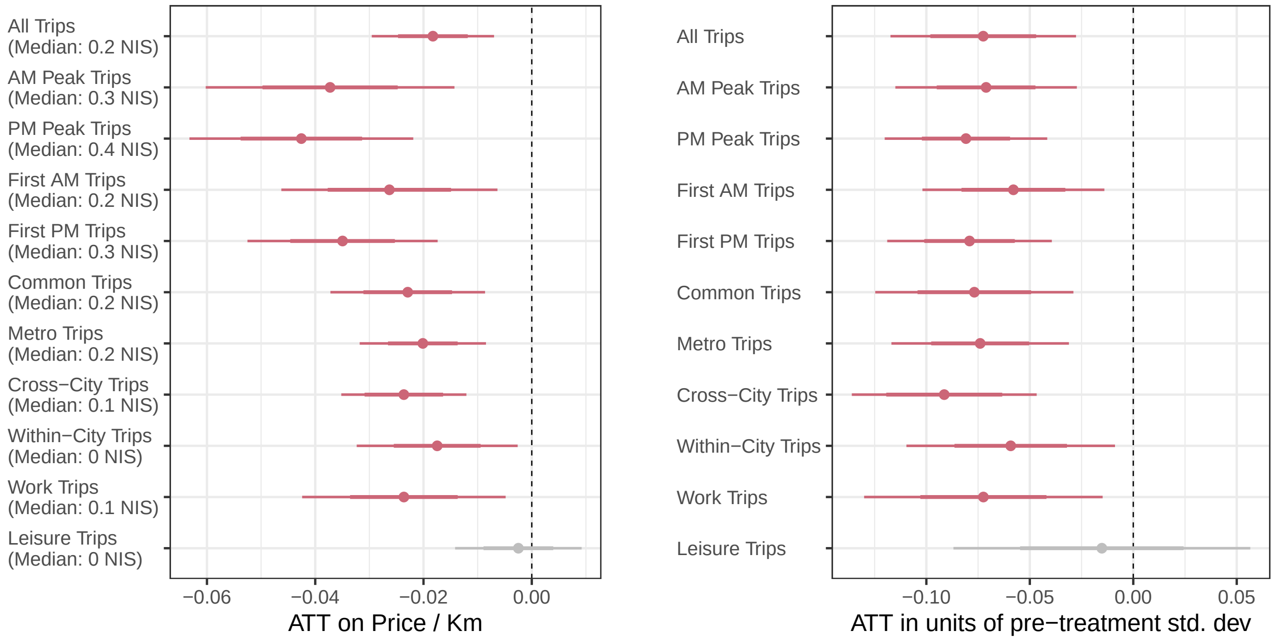


Figure 8: Average effects over week 4 to 16 post-treatment by trip category for price per kilometer

5 Heterogeneity in Responses to Usage-Based Pricing

In the prior sections, we have taken an in depth look at estimated responses to congestion pricing on average, over the treated subgroup in our experiment. We've also looked at how these average responses manifested themselves on different kinds of trips taken by our drivers. In this section, we focus on understanding heterogeneity in effects across three broad categories of driver variables: flexibility, congestion-inducing driving, and socioeconomic status.

5.1 Attributing Effect Heterogeneity to Driver Covariates

Recall that in our treatment effect analysis in Section 4, the control potential outcome model (4.1) we imposed to identify treatment effects via difference-in-differences did not restrict heterogeneity in treatment effects across drivers and weeks. As such, to assess whether or not significant treatment effect heterogeneity can be attributed to differences in driver characteristics, we will construct sufficiently accurate estimates of the effect of our treatment on each driver, predict those estimated driver treatment effects with driver characteristics, and then characterize how much those predictions vary across the population of drivers via the Sorted Effects method developed in Chernozhukov et al. (2018).

In particular, we first construct asymptotically unbiased estimates $\hat{\tau}_{it}$ of driver i's individual treatment effect $\tau_{it} := Y_{it}(C_i) - Y_{it}(0)$ in week $t \geq C_i$ as discussed at the end of Section 4.1. We then linearly project averages of these driver-specific treatment effect estimates in the fourth through 16th weeks post-treatment onto rich vectors X_i of interacted driver characteristics, yielding predicted driver-

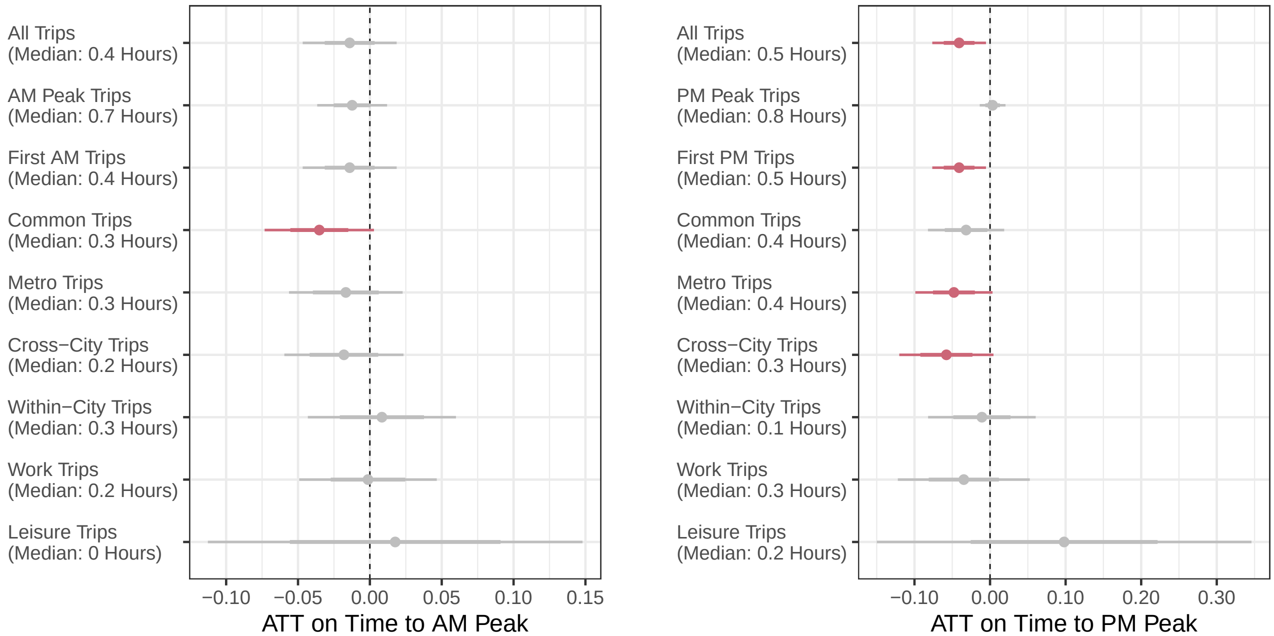


Figure 9: Average effects over week 4 to 16 post-treatment by trip category for distance to peak time for AM and PM trips

specific treatment effect estimates $\hat{\Delta}_i$ and their population analogues Δ_i:[13]

$$\hat{\Delta}_i := X_i \left(\sum_{i=1}^{N} X_i X_i' \right)^{-1} \sum_{i=1}^{N} X_i \cdot \frac{1}{12} \sum_{e=1}^{T} \hat{\tau}_{i(C_i+e)}, \quad \Delta_i := X_i \mathbb{E}[X_i X_i']^{-1} \mathbb{E}\left[X_i \cdot \frac{1}{12} \sum_{e=1}^{T} \tau_{i(C_i+e)} \right].$$

To assess whether or not driver covariates X_i can predict heterogeneous effects across drivers, we can test whether the distribution of Δ_i is non-degenerate. To do so, we estimate the the distribution of Δ_i via a fine grid of quantiles $\{\hat{Q}_{\Delta_i}(u)\}_{u \in [0.05, 0.06, \ldots, 0.95]}$ of the distribution of $\hat{\Delta}_i$ across drivers, and we construct 90% uniform confidence bands via a Bayesian bootstrap of the entire estimation procedure from the fixed effects regression (4.1) through the projection to compute $\hat{\Delta}_i$ (Chernozhukov et al., 2018). We plot the results of this procedure when the outcome of interest is a driver's total weekly price paid in Figure 10, and note that there is a roughly 30 NIS difference between the smallest and largest covariate-predicted effects in our sample of drivers, relative to a 35 NIS median pre-treatment weekly total price. Further, since the uniform confidence band does not contain any horizontal lines,[14] we can reject the null hypothesis that all quantiles of the distribution of Δ_i are the same. As such, we can conclude that there is substantial treatment effect heterogeneity that can be attributed to driver covariates.

[13]In particular, we let X_i be a fully-interacted quadratic basis for our covariates of interest.

[14]We can see that no horizontal lines pass through the uniform confidence band since the dashed line in Figure 10 corresponding to the upper bound of the confidence interval for smallest quantile $\hat{Q}_{\Delta_i}(0.05)$ of the distribution of Δ_i is below the dashed line corresponding to the lower bound of the confidence interval for the largest quantile $\hat{Q}_{\Delta_i}(0.95)$ of the distribution of Δ_i.

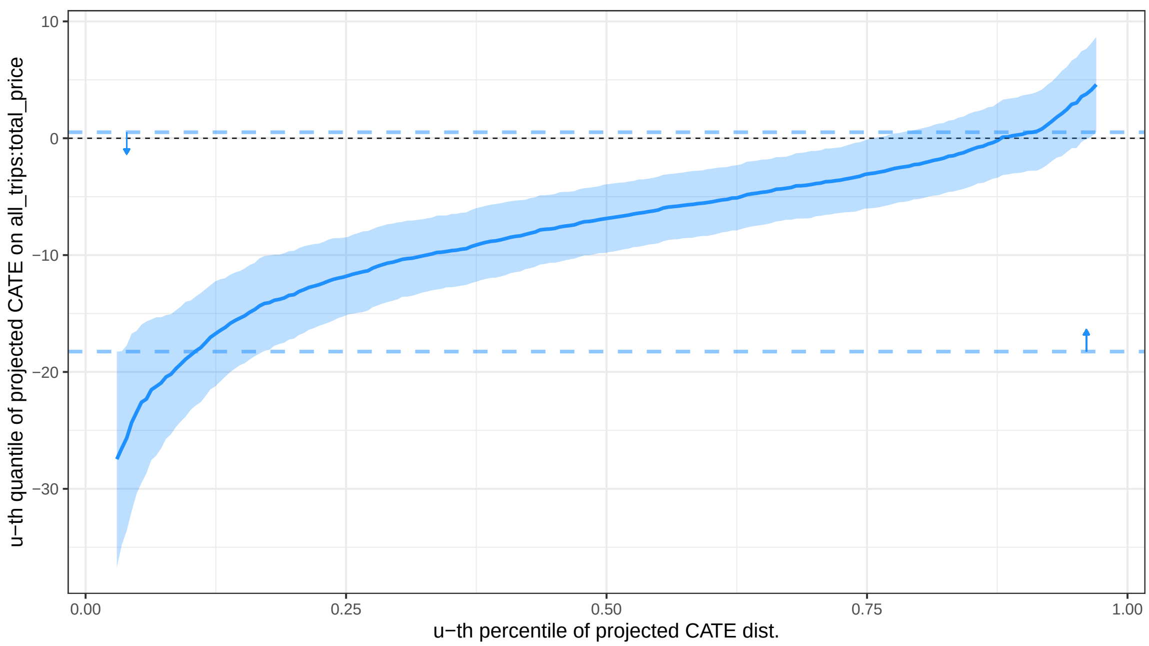


Figure 10: Estimated quantiles of the distribution of covariate-predicted driver-specific treatment effects, as well as a 90% uniform confidence band computed according to Chernozhukov et al. (2018). The dashed line with a downward-pointing arrow attached corresponds to the lower bound of the confidence interval for the largest quantile of the distribution of Δ_i, and the dashed line with an upward-pointing arrow attached corresponds to the upper bound of the confidence interval for the smallest quantile of the distribution of Δ_i; if the dashed line with the downward-pointing arrow attached is above the dashed line with the upward-pointing arrow attached, we can reject the null hypothesis of no heterogeneity in driver-specific treatment effects predicted by our covariates at the 10% level.

5.2 Covariate Differences Between the Least and Most Affected Drivers

Having established that there is significant covariate-predicted treatment effect heterogeneity across drivers, we now turn to characterizing how the most and least affected drivers differ in their observed characteristics. To do so, we again leverage methods developed in Chernozhukov et al. (2018) to estimate the vector μ of differences in average characteristics between the 20% of drivers with the least negative covariate-predicted treatment effects and the 20% of drivers with most negative covariate-predicted treatment effects:

$$\mu := \mathbb{E}[X_i \mid \Delta_i \geq Q_{\Delta_i}(0.8)] - \mathbb{E}[X_i \mid \Delta_i \leq Q_{\Delta_i}(0.2)].$$

We plot our estimates of these covariate differences between the least and most affected 20% shares of drivers in figures like Figure 11, where we normalize each covariate difference by the population standard deviation of that covariate to be able to compare differences across covariates with different units. In addition to each covariate difference point estimate denoted by a dot, we again use thick lines to denote confidence intervals uncorrected for multiple testing problems computed via a Bayesian bootstrap of the entire estimation procedure, thin lines to denote Bayesian bootstrap-based confidence intervals that control the FWER within each plot at 10% (Romano et al., 2010), and a non-gray color for a point estimate and its corresponding confidence intervals to indicate that the stepdown procedure developed in Romano and Wolf (2005) rejects the null hypothesis of zero effect size while controlling

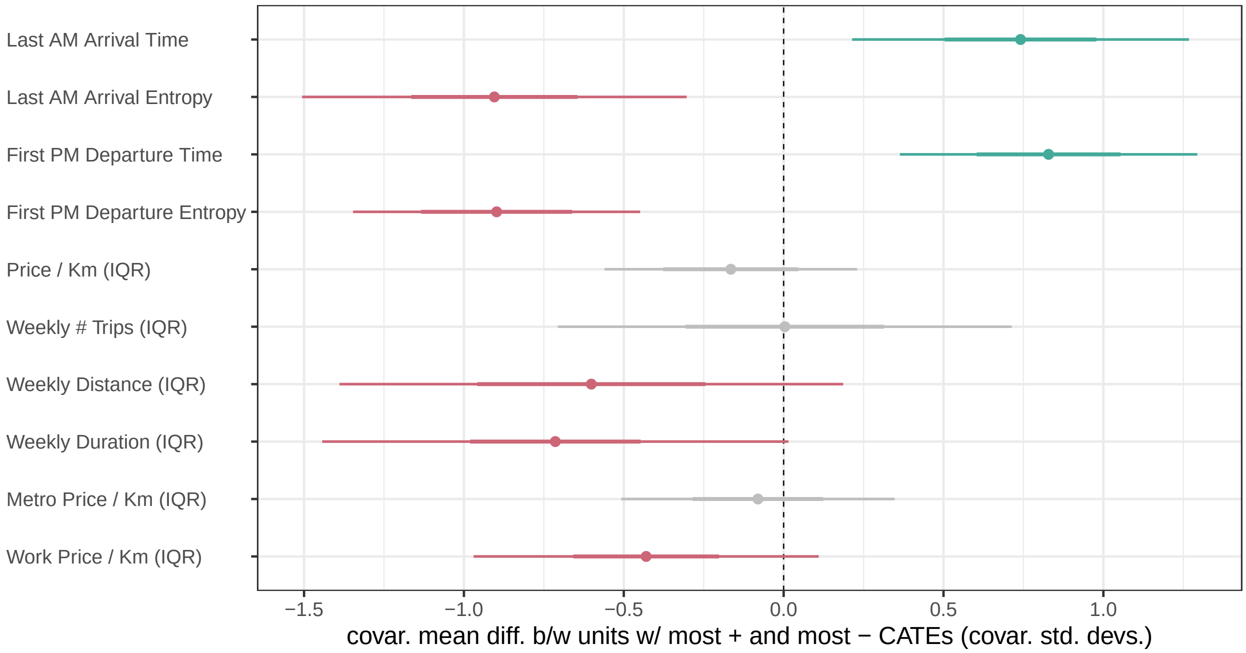


Figure 11: Sorted effects analysis results for flexibility covariates. A negative value implies the least affected drivers had lower covariate values, or conversely, the most affected drivers had higher covariate values. A positive value implies the least affected drivers had higher covariate values, or conversely, the most affected drivers had lower covariate values.

per-plot FWER at 10%.

Throughout the heterogeneity analysis presented in this section, for simplicity, we focus on our estimated effects for a single outcome, which is total weekly price paid. The effect of our intervention is generally in the negative direction, therefore the most positive covariate-predicted treatment effects in the plots belong to the *least* impacted drivers, and the most negative covariate-predicted treatment effects come from the *most* impacted drivers.

We begin by considering the results from our sorted effects heterogeneity exercise with respect to covariates that proxy for driver flexibility in Figure 11. Looking at 'commute-like' trips which we take to be the first PM departures and last AM arrivals for our drivers, we see that the least affected drivers take these trips later, by about 0.75 standard deviations. In other words, the most affected drivers tend to arrive earlier in the morning and leave earlier in the evening. Turning to our preferred variability measure for within-driver distributions of departure and arrival times, we compare the entropy values for the most affected and least affected drivers. We find that the least affected drivers have lower entropy. This means the within-driver distribution of PM departure times and AM arrival times are more concentrated and less dispersed for the least affected drivers. On the other hand, the most affected drivers have more dispersed distributions of PM departure times and AM arrival times, which suggests that they have more flexibility and less rigid schedules.

We also use the interquartile range of weekly driving metrics to measure the flexibility or variability in driving patterns over a longer time window: the within-driver IQR on the number of trips, total distance, and duration of trips. We generally see weaker evidence of these metrics being associated with differences in estimated treatment effects, compared to differences in entropy for departure times. Our point estimates suggest least affected drivers have smaller IQR's for their weekly durations and weekly distances, suggesting their driving patterns are more stable and don't change much week by week. After our FWER correction procedure, we fail to reject the null of no difference for the IQR's on weekly distance. We construct the within-driver IQR values for price per kilometer driven for different categories of trips to measure variability in prices paid. These also suggest that again, the least affected drivers have a smaller IQR values in the price they pay per km for their work trips.

Taken together, our analysis suggests most affected drivers showed more flexibility in their pre-treatment driving patterns - as measured by the metrics we defined above. We see evidence that they have more widespread within-driver distributions for AM arrival times and PM departure times, weekly driving durations, and prices paid per km for work trips.

Next, we consider whether the drivers most affected by our congestion pricing scheme are the drivers who contribute the most to traffic congestion, where we measure each driver's tendency to take part in congestion inducing driving behavior through the prices of their pre-treatment trips. By doing do, we're implicitly adopting a definition of "congestion inducing driving" based on the trips that Ayalon Highways chooses to specifically impose a price on - meaning trips that take place during peak hours as defined by their fee schedule, on the roads that they chose to price. These variables can also can be thought of as a measure of 'exposure' or 'intensity' of our treatment. Drivers that have more expensive trips pre-treatment will be exposed to our usage-based pricing intervention to a higher degree, assuming their travel needs don't change post-treatment. Figure 12 plots sorted effect analysis results with respect to these covariates.

We first look at this through the lens of cross-city and within-city trips. Within city trips tend to be much shorter and they are much less likely to cost anything, whereas cross city trips are longer and cost more. Unsurprisingly, the least affected drivers tend to have fewer number of cross city trips and spend less time on cross city trips. We see a significant difference for both duration and number of cross-city trips by about 1.2 covariate standard deviations. Consequently, we also observe large differences in average pre-treatment prices per trip and price per km: our least affected drivers are those that took less expensive trips pre-treatment, a difference of about 2 standard deviations on price per trip. These observations all point toward a perhaps not-so-surprising conclusion: drivers who changed their behavior the most tended to take the most expensive trips and/or more cross-city trips during the pre-treatment period.

Finally, we consider whether or not the least affected drivers differ significantly along socioeconomic dimensions from the most affected drivers. In Figure 13, we plot the average differences in several measures of driver socioeconomics between the least and most affected drivers. First and foremost, we can see that the most and least affected drivers do not differ in the socioeconomic indices of their home TAZs. While our confidence intervals are wide enough to not rule out some differences, our analysis suggests that, contrary to popular belief, usage-based congestion pricing does not disproportionately affect socioeconomically disadvantaged drivers, at least given the distribution of drivers' home locations and trip destinations. We also note that we were not able to detect any differences in other home TAZ demographic characteristics, which we do not report in Figure 13 for parsimony.

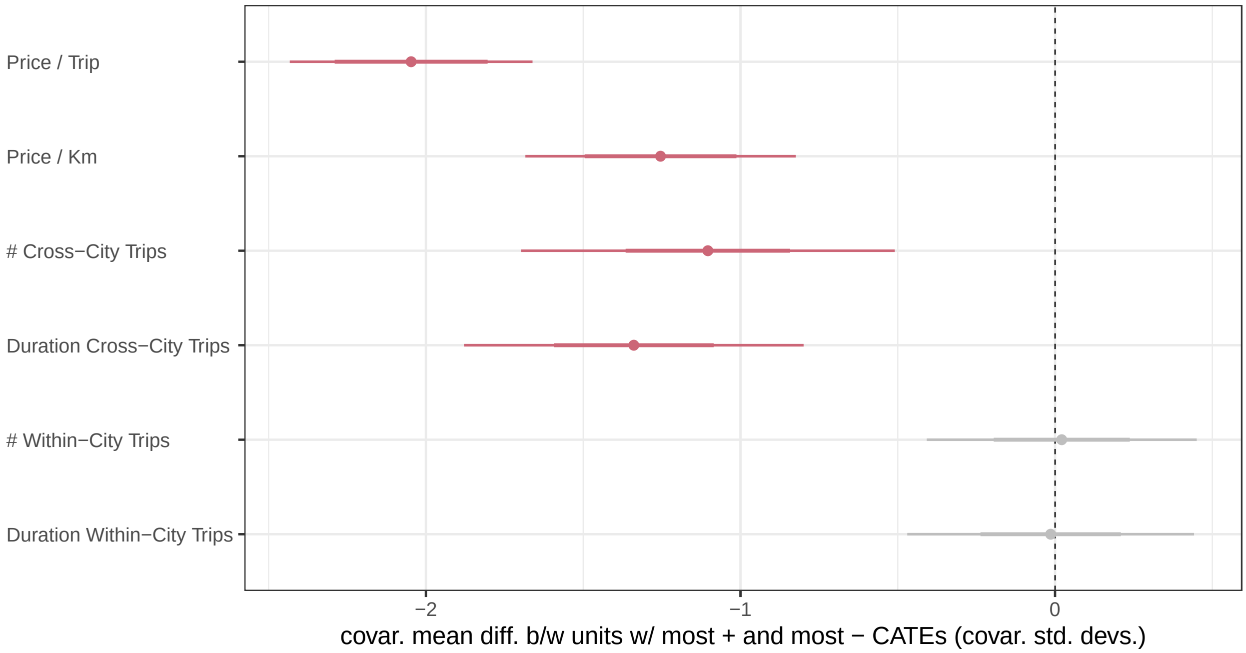


Figure 12: Sorted effects analysis results for congestion covariates. A negative value implies the least affected drivers had lower covariate values, or conversely, the most affected drivers had higher covariate values. A positive value implies the least affected drivers had higher covariate values, or conversely, the most affected drivers had lower covariate values.

The second important takeaway from Figure 13 is that public transit travel alternatives are much more accessible for the most affected drivers compared to the least affected drivers. While we cannot directly observe substitution behavior in our data, this result suggests that the existence of desirable public transit options could be an effective complement to our congestion pricing policy. Finally, we note that the most and least affected drivers do not meaningfully differ in how much their vehicles pollute, which could be another proxy for socioeconomic status given that lower polluting vehicles tend to be more expensive than higher polluting ones.

6 Impacts for Congestion

Thus far, we have established that our usage-based congestion pricing scheme meaningfully decreases driving during the most congested times and in the most congested areas. However, policymakers are most interested in the potential for usage-based congestion pricing to meaningfully increase road speeds. As such, to assess whether our treatment effects on driving behavior translate into meaningful impacts on road speeds, we also need to understand how reducing the number of drivers on each road affects the speed at which cars can travel along that road.

To estimate how traffic density on a road affects the speed at which cars travel along that road, we need high-frequency data on contemporaneous measurements of traffic densities and speeds, which are typically hard to come by on most roads besides the largest highways. Luckily, as described in Section 2, we have access to one year of contemporaneous road density and speed observations at 5

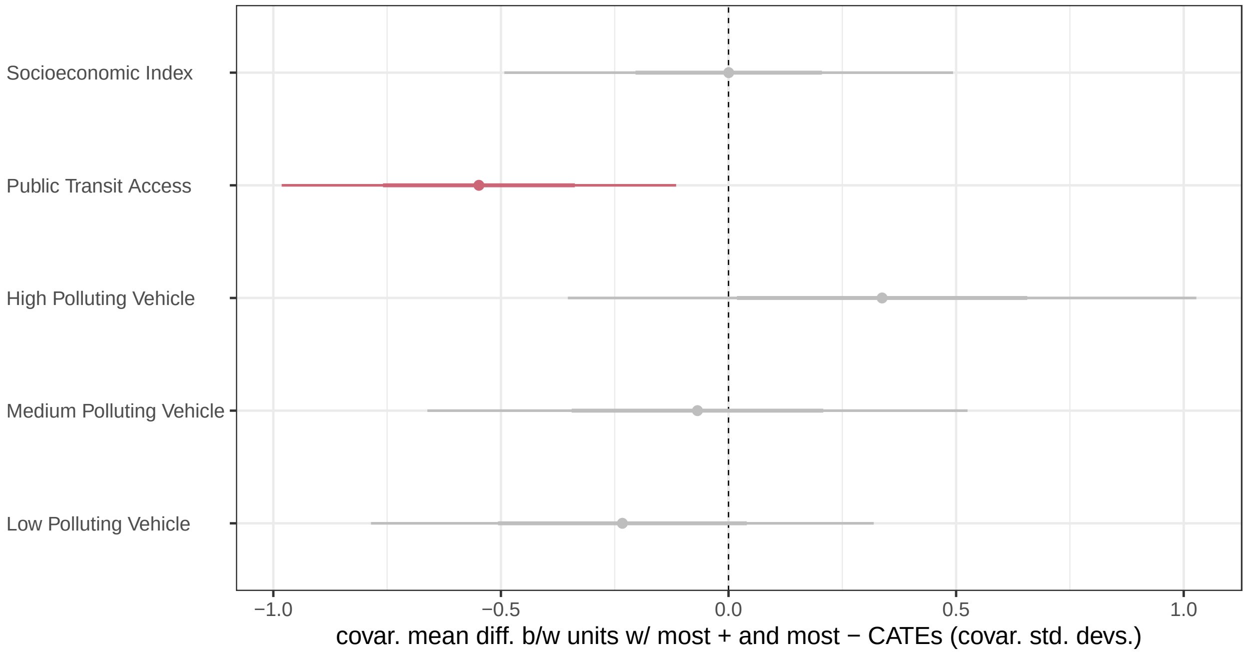
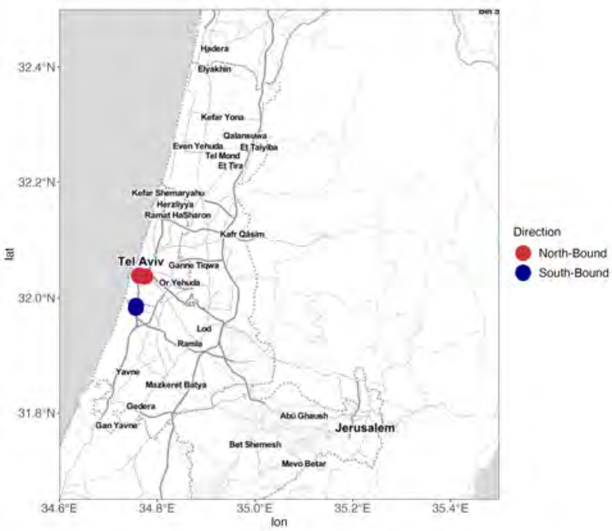


Figure 13: Sorted effects analysis results for socioeconomic covariates. A negative value implies the least affected drivers had lower covariate values, or conversely, the most affected drivers had higher covariate values. A positive value implies the least affected drivers had higher covariate values, or conversely, the most affected drivers had lower covariate values.

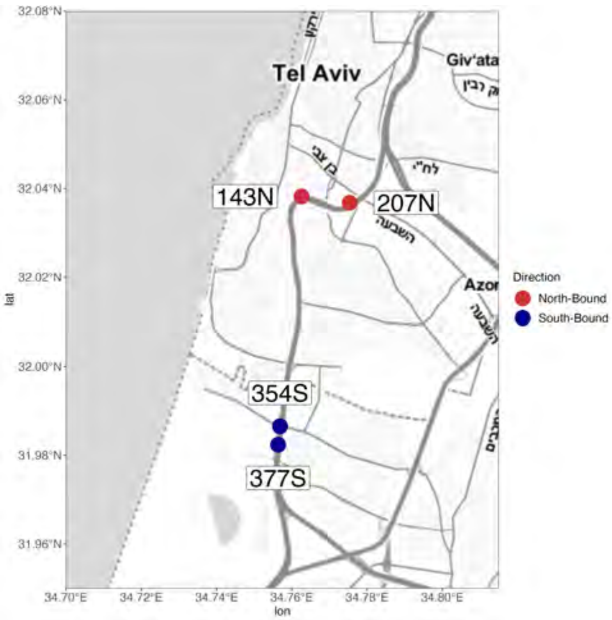
minute intervals from several sensors along a segment of the Ayalon, the largest highway in Israel, that is close to Tel Aviv, as illustrated in Figure 14. Figure 15 provides a scatter plot of our speed and density data. Since we infer that 18.5% of all trips in our data likely pass through highways like the Ayalon, and the median weekly share of drivers' pre-treatment trips that pass through highways like the Ayalon is 49.7%, understanding how our usage-based congestion pricing scheme can positively impact car speeds is of first-order importance. As such, we will focus on estimating the structural relationship between traffic densities and speeds on the Ayalon.

6.1 Structural Relationship Between Traffic Speeds and Densities on the Ayalon

To estimate the effect of traffic density on speeds, there are two important challenges to address. The first is that the shape of the function $h\colon \mathbb{R} \to \mathbb{R}$ mapping traffic densities D_t to speeds V_t is not known, but is important for determining the impact of our congestion pricing scheme on highway speeds. In particular, if the most frequent traffic densities occur in a region where h is flat and our congestion pricing scheme does not decrease the density of cars on the highway enough to leave h's flat region, then even if the effect of our treatment were seemingly substantial, it might still not be enough to have a meaningful impact on traffic speeds. Conversely, if the most frequent traffic densities occur in a region just above a region where h has a steep drop, then even a small decrease in traffic density could have large positive impacts on traffic speeds. To remain agnostic to the shape of h, we estimate h nonparametrically, unlike several recent papers like Yang et al. (2020) and Kreindler (2022) that impose strong parametric functional forms on h, and thus may over or understate the impacts of



(a) *General location of focal sensors*



(b) *Specific locations of lead and lag sensors.*

Figure 14: Map of the lead and lag Ayalon sensor locations of interest.

decreasing traffic density.

The second challenge is that traffic speed V_t and density D_t are simultaneously determined in equilibrium, and thus their relationship is endogenous (Small and Chu, 2003). To see why, note that drivers typically select the fastest route to get from their intended origin to their intended destination, so traffic density D_t on the Ayalon is determined by traffic speed V_t, or at least drivers' expectations about future speeds given the information available at the time of their departure. Additionally, traffic speed V_t and density D_t adjust to set the volume N_t of cars flowing over a segment of the highway equal to $V_t \cdot D_t$ in equilibrium. Further, because we only observe average speed V_t and density D_t in 5 minute intervals but underlying V_t and D_t are set essentially instantaneously, D_t is subject to measurement error.

To overcome this endogeneity issue, we instrument for traffic density D_t, as in Couture et al. (2018) and Yang et al. (2020), for example. In particular, we use the traffic density $\tilde{D}_{t-1}$ at one sensor further in the direction of traffic than our focal sensor 5 minutes prior as our instrument; in Figure 14, we use lagged densities from sensor 207N as our instruments for densities from our focal sensor 143N, and we use lagged densities from sensor 377S as our instruments for densities from our focal sensor 354S. Lagged "downstream" sensor traffic density $\tilde{D}_{t-1}$ is a relevant instrument for focal sensor density because sharp increases in $\tilde{D}_{t-1}$ can cause pile-ups upstream at the focal sensor in short order. $\tilde{D}_{t-1}$ is also excluded because the focal sensor's traffic speed V_t is only determined in equilibrium directly by D_t, which is our independent variable of interest, and N_t, which cannot adjust quickly enough to changes in $\tilde{D}_{t-1}$ to cause an exclusion restriction violation. In future work, we are also exploring alternative instruments based on measures of exposure to traffic accidents, about which we have exhaustive records for over a year lining up with our sample of Ayalon traffic speeds and densities.

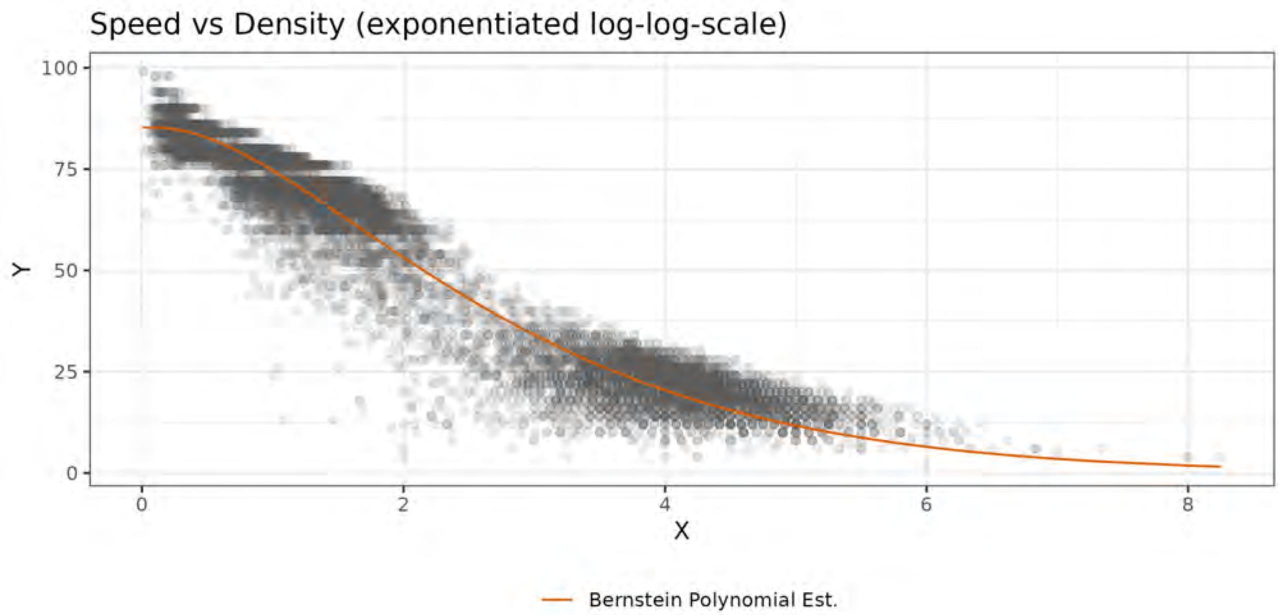


Figure 15: Scatter plot of contemporaneous traffic speed and density observations along with estimate of the structural relationship between traffic speed and density.

Concretely, to identify and estimate h while taking the challenges described above seriously, we impose the nonparametric instrumental variables (NPIV) model of Newey and Powell (2003) on the relationship between traffic speeds V_t, densities D_t, and lagged downstream sensor densities $\tilde{D}_{t-1}$:

$$\mathbb{E}[\log(V_t) - h(D_t) \mid \tilde{D}_{t-1}] = 0.$$

To estimate h under the NPIV model (6.1), we use a shape-constrained sieve two-stage least squares procedure studied in Chetverikov and Wilhelm (2017). In particular, we require that h be monotonically decreasing by approximating h via an appropriately constrained linear combination $\theta \in \mathcal{C}_J \subset \mathbb{R}^J$ of a J-dimensional Bernstein polynomial basis functions $B_J(D_t)$.[15] We construct unconditional moment conditions implied by (6.1) by interacting structural residuals $\log(V_t) - \theta' B_J(D_t)$ with a cubic, $K \geq J$-dimensional B-spline basis expansion $\Psi_K(\tilde{D}_{t-1})$ of $\tilde{D}_{t-1}$. In summary, we estimate h as follows:

$$\hat{h}(d) := \hat{\theta}' B_J(d),$$
$$\hat{\theta} := \arg\min_{\theta \in \mathcal{C}_J} \sum_{t=2}^{T} (\log(V_t) - \theta' B_J(D_t))^2 \Psi_K(\tilde{D}_{t-1})' \left(\sum_{s=2}^{T} \Psi_K(\tilde{D}_{s-1}) \Psi_K(\tilde{D}_{s-1})' \right)^{-1} \Psi_K(\tilde{D}_{t-1}).$$

We plot the estimate $\hat{h}$ of h resulting from the estimation procedure in (6.1) in Figure 15. Interestingly, h only appears to display mild thresholding behavior as discussed in the literature on hyper-congestion, in which traffic speed stays relatively flat as a function of traffic density as long as traffic density is low enough to sustain a free flow speed, and then above a certain density threshold,

[15] In particular, the constrained coefficient space $\mathcal{C}_J$ is such that $\theta_j \leq \theta_{j-1}$ for all entries j; see Appendix C.2 of Mogstad et al. (2018) for a clear reference.

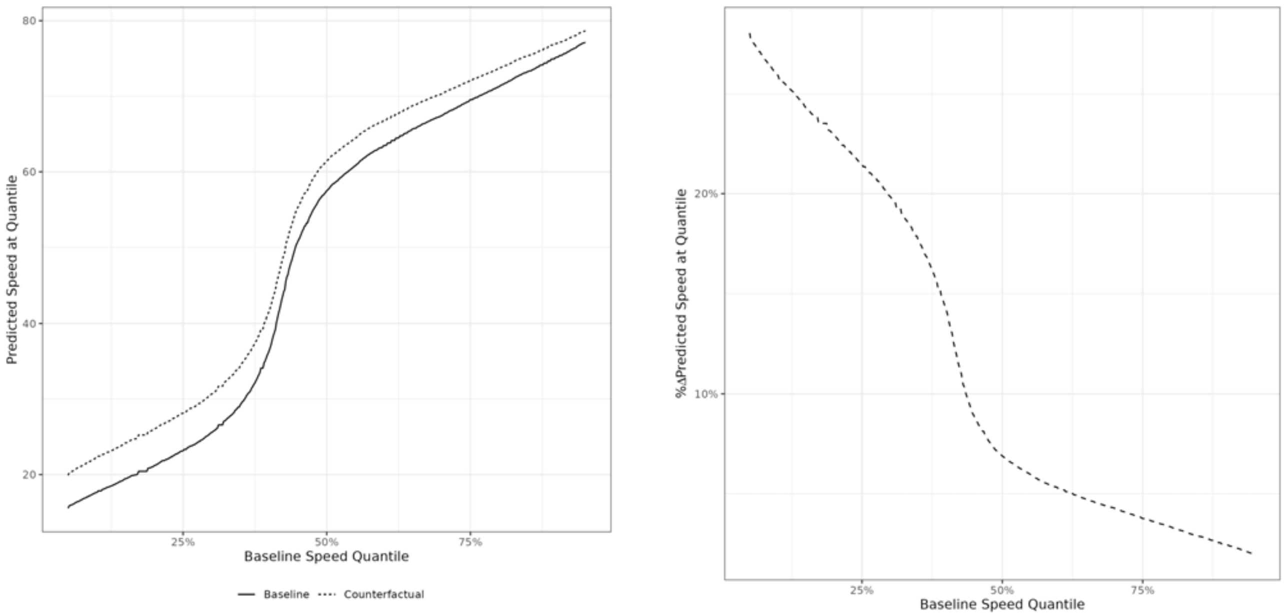
speed drops precipitously as a function of density (see e.g. Anderson and Davis (2020) for a review of the literature). Instead, after a short flat region at almost-zero densities, speed decreases in a more gradual, sigmoid-like shape as density increases.

6.2 The Distribution of Congestion Pricing Effects on Ayalon Traffic Speeds

Having estimated the structural function mapping traffic density on the Ayalon to traffic speed, we now turn to translating the effects of our congestion pricing scheme on driving behavior into its effects on the Ayalon's traffic speeds. An important caveat we must make before continuing is that our treatment effect estimates only measure changes in driver behavior in an equilibrium in which a negligible fraction of all of the drivers on the Ayalon are affected by our congestion pricing scheme. As such, if our congestion pricing scheme had non-trivial impacts on traffic speeds when scaled up to the whole population, drivers might change their behavior differently in response to having their driving priced. As such, we view the following extrapolation exercise as a study of the aggregate effects of our congestion pricing scheme in partial equilibrium; see Section 7 for a discussion of how we plan to conduct a more general equilibrium assessment of our congestion pricing scheme's effects. That being said, we still view our analysis below as informative about the scope for usage-based congestion pricing to improve traffic speeds during congested times of day on the Ayalon and other highways.

To assess the effect of our congestion pricing scheme on the distribution of traffic speeds on the Ayalon, we first estimate its effect on the number of trips taken by drivers on highways during peak hours via the methodology described in Section 4. As discussed in Section 2.2, we infer that a trip is likely to have taken a route through a highway if the fastest route as determined by speed limits from Open Street Map uses a highway road segment. We find that on average, drivers who take any highway trips at all decrease the number of highway trips they take during peak hours in response to our congestion pricing scheme by approximately 10% relative to their pre-treatment median number of weekly highway trips during peak hours. Such an effect estimate is in line with the magnitudes of our other treatment effect estimates reported in Section 4. For each observed traffic speed V_t and density D_t, we predict the counterfactual traffic speed that would be induced by our congestion pricing scheme $\hat{V}_t^* := \hat{h}((1 - 0.1) \cdot D_t)$.

Figure 16 illustrates our estimates of the effect of our congestion pricing scheme on the distribution of Ayalon traffic speeds in two ways. In Figure 16a, the solid line denotes quantiles of the baseline, observed distribution of Ayalon traffic speeds V_t, while the dashed line denotes quantiles of the counterfactual distribution of Ayalon traffic speeds $\hat{V}_t^*$. It appears that our treatment effects are large enough to increase traffic speeds by more than five kilometers per hour during the most congested times with the 25% lowest traffic speeds and several kilometers per hour in the least congested times with the 50% highest traffic speeds. In addition, we plot the percent change in each quantile of the distribution of Ayalon traffic speeds when moving from the baseline distribution of observed traffic speeds V_t to the distribution of counterfactual traffic speeds $\hat{V}_t^*$ in Figure 16b. Given that traffic speeds are already quite low in the 25% most congested times, a five kilometers per hour increase in traffic speeds is a large relative speed increase.



(a) Quantiles of Observed, Counterfact. Ayalon Traffic Speeds

(b) % Changes in Quantiles of Ayalon Traffic Speeds

Figure 16: This figure illustrates the effect of our usage-based congestion pricing scheme on the distribution of traffic speeds on the Ayalon.

7 Next Steps

Recall from Section 6.2 that our treatment effect estimates only reflect changes in driver behavior in an equilibrium in which a negligible fraction of all of the drivers on the Ayalon are affected by our congestion pricing scheme. As such, if our congestion pricing scheme has non-trivial impacts on traffic speeds when scaled up to the whole population, drivers might change their behavior differently in response to having their driving priced than we estimated in Section 4.

To assess the impacts of our usage-based congestion pricing scheme under general equilibrium, we have developed and will estimate a model of drivers' decisions around which trips to take, when to take them, and what route through the Israeli road network to take, allowing those decisions to depend on road speeds that are set endogenously in equilibrium. Below, we sketch out our model and how one could identify and estimate it, as well as the data work we intend to complete to assemble the ingredients needed to estimate the model credibly.

7.1 Driver Choice Model and Traffic Equilibrium

For simplicity, we focus on the time-of-day and route choices for weekday morning trips taken by drivers in our sample. In our model, each driver is endowed with a "most common trip" that they can take or not take each morning. Crucially, we allow for the possibility that exogenously, a driver may not need to take their common trip on a given day (e.g. because they work from home some days, or they only take their children to daycare some days each week), and in our model, we distinguish between the lack of a common trip on a given day due to a missing exogenous choice opportunity versus an active decision by a driver to not take a trip. Not taking this "zero inflation" into account as in other papers in this literature could lead to an overstatement of sensitivity to congestion prices, since a lack of common trips could be due to the lack of opportunity to choose to take them, not due to the active choice not to take them.

In particular, on each day t, a driver i with covariates $X_i \in \mathbb{R}^K$ receives the opportunity to take their most common trip from TAZ A_i to TAZ B_i with probability π_{it}; the indicator $D_{it} = 1$ if driver i received a choice opportunity on day t and $D_{it} = 0$ otherwise.[16] If $D_{it} = 1$, then driver i actively chooses a "route" J_{it} from an exogenously given discrete set $\mathcal{J}_i$ of "routes" to take to complete the trip, or to not take the trip at all, which we denote $J_{it} = 0$. We define a route $j \in \mathcal{J}_i$ as a particular path through the road network taken at a particular time, which is endowed with a congestion price P_{ijt} that is individual-specific, as well as a travel time T_{jt}. Driver i chooses J_{it} to maximize their indirect utility, which is given as follows:

$$U_{ijt} := \underbrace{\xi_{jt} - \gamma_0 T_{jt}}_{\delta_{jt}} - \alpha_0 P_{ijt} - \sum_{k=1}^{K} X_{ik} \left(\alpha_k P_{ijt} + \gamma_k T_{jt}\right) + \varepsilon_{ijt}, \quad U_{i0t} := \varepsilon_{i0t},$$

where ξ_{jt} is a route-and-day specific unobservable that affects the desirability of taking route j on day t for all drivers, and ε_{ijt} are type-I extreme value distributed idiosyncratic utility shocks that are independent across drivers i, routes j, and days t. Using standard arguments, drivers' choice probabilities conditional on everything except the logit shocks ε_{ijt} are given as follows:

$$\mathbb{P}(J_{it} = j \mid \pi_{it}, \xi_{jt}, X_i, P_{ijt}, T_{jt}) = \pi_{it} \frac{\exp\left(\delta_{jt} - \alpha_0 P_{ijt} - \sum_{k=1}^{K} X_{ik}\left(\alpha_k P_{ijt} + \gamma_k T_{jt}\right)\right)}{1 + \sum_{j' \in \mathcal{J}_i} \exp\left(\delta_{j't} - \alpha_0 P_{ijt} - \sum_{k=1}^{K} X_{ik}\left(\alpha_k P_{ij't} + \gamma_k T_{j't}\right)\right)},$$
$$\mathbb{P}(J_{it} = j \mid \pi_{it}, \xi_{jt}, X_i, P_{ijt}, T_{jt}) = (1 - \pi_{it}) + \pi_{it} \frac{1}{1 + \sum_{j' \in \mathcal{J}_i} \exp\left(\delta_{j't} - \alpha_0 P_{ijt} - \sum_{k=1}^{K} X_{ik}\left(\alpha_k P_{ij't} + \gamma_k T_{j't}\right)\right)}.$$

Importantly, we allow the travel time T_{jt} to be determined endogenously by the route choices of the population of drivers. In particular, each route j is characterized by a sequence of road segments $\{r_{j\ell}\}_{\ell=1}^{L_j}$ that a driver must traverse to complete a trip via route j, as well as a window of time in which the driver begins and completes the trip w_j. Cars travel along road segment r with length D_r at traffic speed V_{rtw} during time window w, where V_{rtw} is determined as a function of the total share S_{rtw} of the driver population traveling along that road segment during time window w on day t:

$$V_{rtw} = h_r\left(S_{rtw}\right), \quad S_{rtw} := \sum_{j \in \mathcal{J}} \mathbb{1}\{w = w_j\} \sum_{\ell=1}^{L_j} \mathbb{1}\{r = r_{j\ell}\} \mathbb{P}(J_{it} = j),$$

where we let $\mathcal{J}$ denote the set of all possible routes across all drivers in the population. The travel

[16] In principle, the following model could apply separately to all of a driver's common trips, but for expositional simplicity, we focus on a driver's most common trip.

time T_{jt} is then determined by the sum of the traversal times across the road segments in the route:

$$T_{jt} = \sum_{\ell=1}^{L_j} V_{r_{j\ell}tw_j} \cdot D_{r_{j\ell}}.$$

The endogeneity of T_{jt} is evident given that it is a function of road segment speeds V_{rtw}, which are in turn a function of driver choices $\mathbb{P}(J_{it} = j)$, which are of course determined by T_{jt}. We define a traffic equilibrium on day t as a set of road segment speeds V_{rtw} across time windows such that the fixed point equation (7.1) is satisfied for all road segments in all time windows, and all drivers are maximizing their utilities. Computing equilibria under this model then reduces to solving the fixed point equations (7.1).

To identify the driver choice part of the model, we assume there are at least $|\mathcal{J}_i| \geq 2$ routes available for all drivers to choose from, an assumption which is certainly reasonable if we allow drivers to choose from a menu of potential leaving times. Then, a Poisson regression of outcomes $Y_{ijt} := \mathbb{1}\{J_{it} = j\}$ for $j > 0$ on P_{ijt}, $X_{ik}P_{ijt}$, and $X_{ik}T_{jt}$ as well as i-by-t and j-by-t fixed effects will yield consistent estimates of the parameters α_0, α_k, and γ_k. α_k and γ_k are identified from the exogenous variation of X_k across drivers, while α_0 is identified directly from this regression since P_{ijt} changes from zero to a positive price for each driver once the congestion pricing scheme has become active for them. We can also recover asymptotically unbiased estimates of π_{it} and δ_{jt} as simple byproducts of the Poisson regression results. To recover γ_0, we can then regress δ_{jt} on T_{jt}, instrumenting for T_{jt} with an appropriate instrument. We can also identify the congestion functions h_r across road segments r by instrumenting for the total population share on segment r in window w in (7.1) with the same instrument we used to identify γ_0.

7.2 Data Work

Estimating the model described in the prior section requires data on travel times T_{jt} across all possible routes, which are in turn determined by contemporaneous road speeds V_{rtw} and occupancies S_{rtw}. As mentioned in Section 6, such data are difficult to come by for all but the biggest roads, so we resort to combining several data sources to estimate them. In particular, we have acquired high frequency GPS pings for the entire fleet of public buses in Israel over the course of 2021, which we have used to estimate average road speeds over fine-grained time intervals across the large set of roads traversed by the buses. In addition, we have scraped Google Maps' predicted travel times twelve times per day during most of 2021 for a set of 300 routes representative of the trips taken by drivers in our sample. We are currently building a graph neural network trained on these sources of data as well as the characteristics of the road network to predict traffic speeds V_{rtw} for cars driving along the same roads. Given V_{rtw}s, we can then impute driver shares on each road segment S_{rtw} by simulating a non-stationary Markov jump process over road segments calibrated to our predicted road speeds V_{rtw} and the trips taken by drivers in our sample.

References

Anderson, M. L. and L. W. Davis (2020): "An empirical test of hypercongestion in highway bottlenecks," *Journal of Public Economics*, 187, 104197.

Barth, M. and K. Boriboonsomsin (2008): "Real-world carbon dioxide impacts of traffic congestion," *Transportation research record*, 2058, 163–171.

Borusyak, K., X. Jaravel, and J. Spiess (2021): "Revisiting event study designs: Robust and efficient estimation," *arXiv preprint arXiv:2108.12419*.

Callaway, B. and P. H. Sant'Anna (2021): "Difference-in-differences with multiple time periods," *Journal of econometrics*, 225, 200–230.

Chandler-Wilde, H. (2023): "These Are the World's Most Congested Cities," *Bloomberg*.

Chernozhukov, V., I. Fernández-Val, and Y. Luo (2018): "The sorted effects method: discovering heterogeneous effects beyond their averages," *Econometrica*, 86, 1911–1938.

Chetverikov, D. and D. Wilhelm (2017): "Nonparametric instrumental variable estimation under monotonicity," *Econometrica*, 85, 1303–1320.

Couture, V., G. Duranton, and M. A. Turner (2018): "Speed," *Review of Economics and Statistics*, 100, 725–739.

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

Kahneman, D., A. B. Krueger, D. Schkade, N. Schwarz, and A. A. Stone (2006): "Would you be happier if you were richer?" *science*, 312, 1908–1910.

Karlström, A. and J. P. Franklin (2009): "Behavioral adjustments and equity effects of congestion pricing: Analysis of morning commutes during the Stockholm Trial," *Transportation Research Part A: Policy and Practice*, 43, 283–296.

Kreindler, G. (2022): "Peak-hour road congestion pricing: Experimental evidence and equilibrium implications," *Working paper*.

Li, Z. and D. A. Hensher (2012): "Congestion charging and car use: A review of stated preference and opinion studies and market monitoring evidence," *Transport Policy*, 20, 47–61.

Martin, L. A. and S. Thornton (2017): "To Drive or Not to Drive: A Field Experiment in Road Pricing," Tech. rep., Working Paper.

Mogstad, M., A. Santos, and A. Torgovitsky (2018): "Using instrumental variables for inference about policy relevant treatment parameters," *Econometrica*, 86, 1589–1619.

Newey, W. K. and J. L. Powell (2003): "Instrumental variable estimation of nonparametric models," *Econometrica*, 71, 1565–1578.

OpenStreetMap contributors (2017): "Planet dump retrieved from https://planet.osm.org ," https://www.openstreetmap.org.

Romano, J. P., A. M. Shaikh, and M. Wolf (2010): "Hypothesis testing in econometrics," *Annu. Rev. Econ.*, 2, 75–104.

Romano, J. P. and M. Wolf (2005): "Stepwise multiple testing as formalized data snooping," *Econometrica*, 73, 1237–1282.

Shao, J. and D. Tu (2012): *The jackknife and bootstrap*, Springer Science & Business Media.

Small, K. A. and X. Chu (2003): "Hypercongestion," *Journal of Transport Economics and Policy (JTEP)*, 37, 319–352.

Yang, J., A.-O. Purevjav, and S. Li (2020): "The marginal cost of traffic congestion and road pricing: evidence from a natural experiment in Beijing," *American Economic Journal: Economic Policy*, 12, 418–453.

d