The Welfare Effect of Road Congestion Pricing: Experimental Evidence and Equilibrium Implications*

Gabriel E. Kreindler†

January 17, 2018

Abstract

The textbook policy response to traffic externalities is congestion pricing. However, quantifying the welfare consequences of pricing policies requires detailed knowledge of commuter preferences and of the road technology. I study the peak-hour traffic congestion equilibrium using rich travel behavior data and a field experiment grounded in theory. Using a newly developed smartphone app, I collected a panel data set with precise GPS coordinates for over 100,000 commuter trips in Bangalore, India. To identify the key preference parameters in my model – the value of time spent driving and schedule flexibility – I designed and implemented a randomized experiment with two realistic congestion charge policies. The policies penalize peak-hour departure times and driving through a small charged area, respectively. Structural estimates based on the experiment show that commuters exhibit moderate schedule flexibility and high value of time. In a separate analysis of the road technology, I find a moderate and linear effect of traffic volume on travel time. I combine the preference parameters and road technology using policy simulations of the equilibrium optimal congestion charge, which reveal notable travel time benefits, yet negligible welfare gains. Intuitively, the social value of the travel time saved by removing commuters from the peak-hour is not significantly larger than the costs to those commuters of traveling at different, inconvenient times.

*I am deeply grateful to Ben Olken, Esther Duflo, Frank Schilbach, and Edward Glaeser for their advice and generous support throughout this project. I especially thank Nikhil Agarwal, Vikas Dimble, Matt Lowe and Yuhei Miyauchi for useful conversations. I thank Alex Bartik, Moshe Ben-Akiva, Benjamin Faber, John Firth, Chishio Furukawa, Nick Hagerty, Rachel Glennerster, Tetsuya Kaji, Jing Li, Rachael Meager, Scott Nelson, Will Rafey, Otis Reid, Mahvish Shaukat, Dan Waldinger, and participants at the J-PAL Bangalore brown bag lunch for many helpful suggestions. Anupriya Khemka, Keerthana Jagadeesh, and Ashwin MB provided excellent research assistance. I also thank Mohannad Abunassar, Maryam Archie, Priya Chetri, Sasha Fleischman, Mahima Gupta, Aditi Sinha, Mamta Jat, Kristina Kelhofer, Michelle Nenciu, Sebastian Quinones, Sarvottam Salvi, Meghna Singh, Sahana Subramanyam, Tammy Tseng, Thuy Duong Vuong, Lantian Xiang, and Massieh Zare, who contributed valuable research assistance at various stages of the project. I gratefully acknowledge design and technical support for the smartphone app "Bangalore Traffic Research" from Adrian Drewett and Dharmendra Singh from Gridlocate Ltd. Funding for this project was generously provided by the Weiss Family Fund for Research in Development Economics, the IGC Cities Fund, the J-PAL Pilot Fund, and the J-PAL Urban Services Initiative Pilot Fund. This project has human subjects approval from MIT COUHES (protocol 1511312369A002) and IFMR (IRB00007107) and was registered in the AEA RCT Registry (AEARCTR-0002083).

†MIT Department of Economics. Email: gek@mit.edu. Website: http://economics.mit.edu/grad/gek

1 Introduction

Traffic congestion is a significant urban disamenity, especially in developing countries, where urban population and private vehicle ownership are growing rapidly.[1] Even holding fixed the number of vehicles in use, peak-hour traffic jams may be particularly inefficient, as large numbers of commuters driving at the same time cause longer travel times for everyone. Reflecting this concern, various urban traffic policies focus on reducing peak-hour congestion, either through pricing or quantity restrictions.[2]

However, it is challenging to evaluate the welfare impact of such policies or, more generally, to quantify the inefficiency in a decentralized unpriced equilibrium. These calculations depend critically on how drivers value the time they spend driving, as well as on their flexibility of changing the *timing* of their trips. Intuitively, the inefficiency may be small if commuters have sufficient schedule flexibility so as to eliminate congestion peaks in the decentralized equilibrium to begin with. Alternatively, if everyone is very inflexible, the distribution of departure times under the social optimum will look similar to the unpriced equilibrium. Apart from these preference parameters, the road technology mediates the externalities that drivers impose on each other. Engineering studies at the road or highway segment level, typically in developed countries, regularly find a convex impact of vehicular volume on travel time (Small et al., 2007), which suggests large social marginal costs when congestion is already high. However, we know little about this relationship at the commuter-trip level and in large cities in developing countries, where the infrastructure, types of vehicles driven, and driving styles differ considerably from the settings of existing studies.[3]

This paper analyzes the peak-hour traffic congestion equilibrium using a rich data set of travel behavior and a theory-grounded field experiment on congestion charge policies. The backbone of the study is a fine-grained panel data set of trips from a sample of around 2,000 car and motorcycle commuters from Bangalore, India. I collected this data using a novel smartphone app that passively logs precise GPS location data, which I helped design for this purpose. The data covers over 100,000 individual trips and almost one million kilometers of travel inside the city.

I outline the key preference parameters of interest in this setting using a simple model of commuting trip scheduling and route choice, building on classic models in transportation economics (Arnott et al., 1993; Noland and Small, 1995). In the model, commuters choose between two route options and decide their trip

[1] Between 2005 and 2015, new private vehicle registrations have grown at 15% and 43% per year in India and China, compared to 0% in the United States and Europe (OICA, 2016).

[2] The congestion charge policy in Stockholm and Singapore's Electronic Road Pricing (ERP) policy have higher fees during the morning and evening peak hours. Jakarta's former "3-in-1" and the current "odd-even" policies are in effect during morning and evening peak hours only. Similarly, Manila's Unified Vehicular Volume Reduction Program (UVVRP) only applies during peak hours in certain parts of the city.

[3] A notable exception is Akbar and Duranton (2017), who use a household travel survey in Bogotá together with Google Maps data collected several years later. They find a small elasticity and hypothesize that drivers are more likely to use side roads during peak hours.

departure time, taking into account the expectation and uncertainty in travel times for different departure times and for the two routes, their ideal arrival time (unobserved to the researcher), and their schedule flexibility.[4] The key parameters for welfare are the value of time spent driving, and the schedule costs of arriving earlier or later than desired. Intuitively, these parameters respectively measure the benefits and the costs of policies that aim to reduce peak-hour congestion by inducing commuters to travel before or after the peak. In an ideal experiment, we would observe choices for various vectors of prices over departure times and drive times. By contrast, observational data does not have sufficient independent variation in costs over these two dimensions. For example, an earlier departure time typically affects both the probability of arriving late and the mean driving time.

In order to identify the model parameters, I designed and implemented two realistic congestion charge policies as part of a field experiment with 497 commuters. The two policies introduce exogenous price variation in departure times and driving times, which identifies the value of time and schedule cost parameters. Under the "departure time" policy, trips are charged according to a pay-per-km rate that is higher during peak-hour departures. Under the "area" policy, commuters face a flat fee for driving through a small area, chosen such that there exists an alternate, untolled route with a longer driving time. Participants were randomized into treatment and control groups for the "departure time" policy. For the "area" policy, the timing of the treatment was randomized among participants. For both treatments, charges were calculated automatically on a daily basis, using the smartphone app travel data, and subtracted from a prepaid virtual account. In order to separate the effect of price incentives from that of information, daily SMS updates, and experimenter demand effects, participants in a separate "information" sub-treatment received daily SMS information notifications and flat weekly bank transfers.

Experimental results show that commuters have moderate flexibility to adjust trips away from typical work hours in order to save money. Under "departure time" charges, commuters leave earlier in the morning and later in the evening. These findings are consistent with working hours acting as constraints, at least in the short run. During the morning interval, participants advance their trips by around 4-6 minutes on average, an effect driven by a subset of commuters that responds more strongly. Responses in the low rate sub-treatment are roughly half of those in the high rate sub-treatment, although imprecisely estimated, and I do not find any impact of the information treatment, which suggests that commuters are responding to prices rather than to other aspects of the intervention. Under "area" charges, participants cross the congestion area around 20% less frequently, and switch to longer routes. Neither randomly doubling the congestion charge nor shortening the detour affects this fraction. These findings are consistent with considerable preference

[4] The model abstracts from the extensive margin travel decision to focus on the within-day distribution of travel.

heterogeneity, and the implied value of time for the marginal commuter is large relative to the average hourly wage for this sample.

I next use the experimental price variation to structurally estimate a model of route and departure time choice for the morning home to work commute. The model includes random utility shocks over routes and departure times, leading to a nested logit specification, and it accounts for the non-linear structure of incentives in the experiment. I construct individual-level choice sets using Google Maps driving time data collected for each driver's typical route and detour route, at all departure times, and I calibrate the driving time uncertainty. I simulate the model to compute choice probabilities, and perform an additional step to invert the individual-specific distribution of ideal arrival times from observed departure times in the pre-experimental period. I estimate the model using two-step GMM and moments chosen to exploit the variation induced by the congestion charge experiments. The estimated schedule cost of early arrival, at around Rs. 320 per hour (approximately $5), is roughly a quarter of the value of time spent driving, at Rs. 1,120 per hour. Late arrival schedule costs are large but imprecisely estimated. I also estimate the probability that a participant responds to the experiment, by matching the distributions of individual effects in the departure time and area treatments. Around half of participants respond to the experiment.[5]

The structural demand estimates show that commuters are moderately schedule flexible relative to how much they value time spent driving. However, to understand the equilibrium welfare costs of congestion, we need to combine these demand estimates with knowledge about the shape and size of the technological part of the externality.

I document a moderate and linear impact of traffic volumes on travel times. I use all the GPS trips data collected using the smartphone app to measure volumes, and Google Maps data on travel delay collected daily to measure driving times.[6] The average travel delay for trips starting at a certain time of the day is linearly increasing in the average volume of departures at that time. In particular, I do not find any convexity for high levels of traffic, unlike previous empirical estimates for highway road segments. Quantitatively, making an average length trip during peak hours increases aggregate driving time for everyone else by approximately 15 minutes, which is roughly half of the private trip duration.[7]

Finally, I compute the optimal equilibrium congestion charge profile, which implements the social optimum. I simulate the city-wide traffic equilibrium in an environment where agents have preferences drawn from those estimated from the data, and aggregate travel volumes determine the travel delay profile through

[5]The experimental results show stark heterogeneity in individual responses, which is not well explained by models with random coefficients. The binary "response probability" does a much better job at replicating this pattern.

[6]I validate the Google Maps driving time data using median driving times from the GPS data. The two measures co-vary with a slope close to 1.

[7]I show that such calculations depend on a semi-elasticity and do not require knowledge of the *total* number of vehicles.

the estimated road technology. This approach has the benefit of relying entirely on estimated parameters. However, I ignore extensive margin responses, and results may differ if long-term responses to congestion charges differ significantly, for example if in the long run firms can accommodate more flexible schedules. I compute the social optimal allocation by finding a Nash equilibrium with congestion charges with the following fixed point property: the charge for departure time h is equal to the marginal social cost of driving at time h.

The social optimum has notable travel time savings relative to the decentralized unpriced equilibrium. Travel is one minute faster from a base of 39 minutes, which is 7% of the travel time above free-flow speeds. However, welfare gains are negligible. This is due to the fact that the social benefits of travel time savings are almost fully offset by the scheduling costs incurred by drivers who now avoid the peak-hour. I conduct counterfactual simulations with alternate policy parameters and road technologies and show that the linear road technology is important for driving these findings.

This result implies that peak-hour congestion pricing and similar quantity-based restrictions are not warranted in Bangalore for the sole purpose of flattening peak-hour congestion by re-allocating drivers across departure times.

This project builds on and contributes to several literatures.

First, transportation economists have developed a rich theoretical literature that models traffic equilibria. Vickrey (1969) and Henderson (1974) introduced the inefficiency due to trip scheduling and their ideas were further formalized by Arnott et al. (1993) and Chu (1995).[8] I build on these early models and adapt them to make it easier to apply them to data on real travel behavior.

Second, the vast majority of transportation research uses survey methods (such as trip diaries) to measure travel behavior. In this project, I collected precise travel behavior data based on detailed GPS traces from around 2,000 participants using their own smartphones. This method circumvents misreporting and recall bias issues that affect survey methods, and makes it easier to collect longitudinal data. Early studies that collected GPS travel behavior data were typically limited to small samples and used special GPS devices that participants carried with them during the study (see, for example, Papinski et al., 2009). Zhao et al. (2015) use a smartphone app and a respondent-supervised machine learning trip classification algorithm to measure travel behavior in Singapore. In this project, I designed and calibrated an entirely automatic trip detection algorithm, which makes it easier to collect large quantities of travel behavior data.[9]

[8]Later on, this literature evolved towards more sophisticated models, for example the joint analysis of departure time and network routing models (Yang and Meng, 1998), and studies of the distributional impacts of pricing (van den Berg and Verhoef, 2011; Hall, 2016).

[9]The app used in this study is also more battery efficient – an important requirement in this setting – due to not collecting accelerometer data.

Third, most estimates of travel preferences in transportation research and planning are based on "stated preferences," whereby survey respondents make hypothetical choices between alternatives that involve trade-offs (Ben-Akiva et al., 2016). Despite their flexibility, stated preferences may bias results if respondents do not properly anticipate their own behavior, a problem the literature has attempted to minimize through careful survey design. In this paper, I measure revealed preferences (real behavior) after experimentally introducing congestion charges for some commuters.

There are relatively few studies that estimate commuter preferences using a revealed-preference approach. Small et al. (2005) analyze real-world driver decisions to use a faster tolled lane to estimate the value of time and of reliability, and Bento et al. (2017) estimate the value of urgency in a similar setting. Estimates of scheduling preferences are even rarer (Small, 1982). A separate group of papers analyzes reduced form impacts of road pricing experiments. Tillema et al. (2013) study a pilot offering rewards for avoiding peak-hour driving. In a contemporaneous study, Martin and Thornton (2017) analyze a randomized experiment that implemented several types of congestion charges in Melbourne, Australia. They report reduced-form effects and implied elasticities, and document that peak-hour distance charges reduced peak-hour travel, cordon charges reduced cordon entries, especially for commuters moderately close to public transit, while commuting to work was not affected. In this paper, I bring these two strands of the literature together by designing a randomized experiment in order to be able to recover the key commuter preference parameters in the model, value of time and scheduling preferences.

Fourth, a growing empirical literature documents the impact of traffic policies on traffic volumes, travel times and air pollution. Several papers analyze the aggregate impact of real-world congestion pricing policies, in London (TfL, 2006; Prud'homme and Bocarejo, 2005; Raux, 2005), Milan (Gibson and Carnovale, 2015) and Stockholm (Karlström and Franklin, 2009), while another strand studies non-price, vehicle quantity restrictions (Davis, 2008; Kreindler, 2016; Hanna et al., 2017; Gu et al., 2017). These papers measure impacts on aggregate outcomes, and either do not address the welfare implications of these policies, or, in a few cases, perform basic welfare calculations treating travel at any time of the day as a single good. Here, I combine estimated preferences with road technology estimates to run equilibrium policy simulations, which allows me to assess policy welfare impacts.

Fifth, empirical studies of the relationship between traffic density, speed and flows mostly focus on small road segments (Small et al., 2007). Geroliminis and Daganzo (2008) used data from fixed-loop detectors and taxi GPS data over a large area in Yokohama, Japan, to document that speed decreases strongly with vehicle density. We do not have similar estimates for cities in developing countries. Akbar and Duranton (2017) use trip data from a household travel survey in Bogotá, Colombia and travel times collected from Google Maps several years later, to estimate both the demand for travel and supply (or road technology). They find a small elasticity of travel time with respect to volume of travel, and I show that the results in Bangalore and Bogotá are very similar. In this paper, I use a large GPS data set with precise information on traffic volumes and driving times, as well as contemporaneous driving time data from Google Maps, and document a linear, moderate relationship between volume and travel times, both within day and across days.

I organize the rest of the paper as follows. Section 2 describes traffic congestion and travel behavior in Bangalore. Section 3 sets up a theoretical model of travel preferences and analyzes the experiment within the model. Section 4 describes the data collection and study sample, and section 5 describes the experimental design and reduced form results. Section 6 describes the structural estimation, section 7 quantifies the traffic congestion technology, section 8 reports the policy counterfactual simulations and quantifies the inefficiency in the decentralized equilibrium, and section 9 concludes.

2 Setting

Traffic Congestion and Travel Behavior in Bangalore, India

Similar to other large cities in developing countries, Bangalore's fast-growing population and economy put stress on its transportation network, which suffers from severe road traffic congestion. In Bangalore, commuters essentially depend on the road in order to reach their destinations. Nearly all motorized transport, both private and public, travels on urban roads, so, to a first approximation, congestion affects all commuters.[10]

Traffic congestion in Bangalore is extreme, and shows significant and predictable within-day variation. Figure 1 shows average predicted travel delay in minutes per kilometer, collected from the Google Maps API on 28 routes in the study area.[11] On average between 7 am and 10 pm on weekdays and across all routes, it takes 3.41 minutes to advance one kilometer. Travel delay is the inverse of speed, so this is equivalent to a speed of 10.9 miles per hour. This is extremely slow, but broadly in line with speeds in other heavily congested large cities in developing countries, such as downtown Jakarta, Indonesia (Hanna, Kreindler and Olken 2017) and Delhi (Kreindler, 2016). Bangalore is much slower compared to cities in the U.S. For example, Anderson (2014) finds an average travel delay of 0.7 minutes per kilometer on urban highways in Los Angeles.

Figure 1 also shows strong predictable within day variation in traffic congestion. Between 7 am and 9 am, travel delay increases by 0.75 minutes per kilometer, or 30%. In other words, a trip that would take

[10] The 2011 census reports that roads are used by 97% of all commuters – excluding those who do not travel, walk or use the bicycle. The main modal split is 33% using motorcycles, 15% cars, and 44% bus. Ridership on the Bangalore metro was below 100,000 per day in 2016, accounting for no more than 4% of all commuters.

[11] Results from 178 routes across Bangalore show a very similar shape and slightly lower travel delay levels.

an hour starting at 7 am would take 80 minutes starting at 9 am. Similarly large changes in average travel delay occur around the evening peak. Here we are taking an average over many different routes that cover all directions and that may have different temporal patterns. In smaller areas, the within-day variation in expected travel time is likely larger. In addition, these results ignore travel time uncertainty, which increases alongside expected travel time.

Assuming that commuters have some flexibility in their schedules, these results suggest that it may be more efficient if some people traveled at earlier or later times, in order to avoid the peak hours. The individual-level GPS data collected for this study using the smartphone app shows that commuters do indeed vary their departure times significantly from day to day. Table 1 reports descriptive statistics about travel behavior in the study sample. Panel C shows the within-person departure time variability in the morning and evening. For the first trip in the morning, the standard deviation for the median person is 1.3 hours, which implies a 95% confidence interval of five hours for the departure time! Even restricting to trips between home and work, the median commuter's departure time is covered by a 95% confidence interval of almost two hours for the morning and three and a half hours in the evening.

However, the daily variation in how commuters travel does not automatically mean that commuters have flexible schedules and that they would respond strongly to policies that give incentives for off-peak travel. It is possible that desired travel times change from day to day (based on changes in work or other constraints), yet commuters may be inflexible around those times on any particular day. Similarly, the existence of large, predictable travel time differentials between different times of the day is not by itself enough to understand the externality imposed by an additional commuter on the road at a given time. Overall, the facts discussed in this section suggest that peak-hour inefficiency is a possibility, yet they are not enough to quantify the welfare impact of policies that aim to reduce peak-hour traffic. In order to study these issues formally, I next introduce and analyze a model of within-day travel behavior and traffic congestion.

3 Theoretical Framework

The profile of traffic congestion within a day cannot be summarized effectively by a single aggregate statistic. Instead, commuters choose when to travel taking into consideration congestion at each time of the day, among other factors. Moreover, a commuter's impact on driving times experienced by others will also depend on their departure time. This is a departure from classic models, where externalities operate through a single aggregate measure (Beckmann et al., 1956; Diamond, 1973).

The model introduced here and elaborated in section 6 puts a specific structure on the demand substitution pattern between travel at different times of the day. I also use the model to explain why the key parameters cannot be identified from observational data alone, and to show that two specific congestion charge policies help solve this problem.

The model is based on the classic formulation of preferences over scheduling and time spent driving from (Arnott et al., 1993), modified to include travel time uncertainty and ideal arrival time variation. It abstracts from the extensive margin decision to travel. I first set up the model for a single route, then introduce the route choice problem briefly in section 3.2 and formally in section 6.

3.1 Model Setup

An atomistic commuter decides when to travel from a stable origin (home) to a stable destination (work), taking into account traffic conditions at different departure times. Define $u(h_D, T)$ the utility from departure time h_D and travel time T, and assume it is quasi-linear in money. This general formulation can include preferences to depart and arrive at specific times, the distaste for time spent traveling, and variations in travel times based on departure time. The commuter maximizes expected utility, namely solves $\max_{h_D} \mathrm{E}_T u(h_D, T(h_D))$, where travel time $T(h_D)$ is stochastic and realized only after departure.

I assume utility takes the following form:

$$u(h_D, T) = -\alpha T + -\beta_E \left|h_D + T - h_A^*\right|_- - \beta_L \left|h_D + T - h_A^*\right|_+$$

The commuter cares about travel time and the arrival time $h_A = h_D + T$. Travel time cost is linear, and α measures the value of time. The second and third terms measure scheduling preferences over arrival time (Arnott et al., 1993). The commuter has an ideal arrival time h_A^*, and constant per-unit of time costs of arriving early (β_E) and of arriving late (β_L). Here, $|x|_-$ and $|x|_+$ respectively denote the negative and positive parts of x (both defined as non-negative numbers). I assume that the ideal arrival time is known in advance but it can change from day to day.

The key parameters of interest in this model are α, β_E, and β_L. The first measures the benefits of any policy that improves expected travel times. The other two parameters capture the costs of a policy that attempts to push commuters away from the peak-hour, namely the costs of traveling at inconvenient times.

Under these assumptions, the commuter's problem becomes

$$\max_{h_D} -\mathrm{E}_T \left[\alpha T(h_D) + \beta_E \left|h_D + T(h_D) - h_A^*\right|_- + \beta_L \left|h_D + T(h_D) - h_A^*\right|_+\right] \tag{1}$$

While the commuter preferences over arrival times have a kink at the ideal arrival time, the uncertainty in travel time smoothes out the utility. The first order condition can be re-written as the following identity:

$$\pi(h_D^*) = \frac{\alpha \cdot d\mathrm{E}_T T / dh_D + \beta_L}{\beta_E + \beta_L} \tag{2}$$

Here $\pi(h_D) = \Pr\left(h_D + T\left(h_D\right) < h_A^*\right)$ is the probability of arriving early when departing at h_D, which depends on the distribution of travel time shock $T\left(h_D\right) - \mathrm{E}_T T\left(h_D\right)$. At the optimum departure time, the probability to arrive early depends on the following factors. If the slope of expected travel time with respect to departure time is positive, the commuter has an incentive to leave earlier to take advantage of faster travel, and this effect is increasing in the value of time, α. This is captured in the first term in the numerator. The commuter also chooses departure time to balance the costs of arriving early and arriving late. The more costly it is to arrive late, the earlier the optimal departure time will be. This effect is captured by the $\frac{\beta_L}{\beta_E+\beta_L}$ term.

3.2 Identifying preferences using congestion charges

The key parameters α (value of time spent driving) and β_E and β_L (schedule costs of arriving early and late) are not typically identified from observational data. There are two problems. First, a change in departure time leads to a change in the distribution of arrival times, and it may also lead to a change in expected travel time. This makes it difficult to disentangle the relative importance of schedule costs from the value of time, as shown in expression (2). For example, assuming we know h_A^*, if we observe someone leave early, we do not know if they do so in order to take advantage of faster travel times (assuming $d\mathrm{E}_T T/dh_D > 0$) or because the cost of arriving late is very high. The second problem is that it is difficult to learn anything from day to day variation in departure times. If we allow the ideal arrival time h_{At}^* to vary by day (t) – for example because the commuter needs to arrive earlier or later to work on some days – then the individual optimal departure time will for the most part track h_{At}^*. To see this, assume the travel time distribution is independent of departure time, then this relationship is exactly linear. In other words, in this model, day-to-day changes in observed departure time are not informative about the underlying parameters α, β_E, β_L.

I now introduce two congestion charge policies that create price variation that will help identify the required parameters. This procedure has the added benefit of providing monetized estimates of α, β_E and β_L. The first policy imposes a marginal cost of departure time, $m = p \cdot h_D$. Intuitively, this creates an independent incentive to change departure time and leave earlier. The first order condition with pricing changes to

$$\pi(h_D^*) = \frac{\alpha \cdot d\mathrm{E}_T T/dh_D + \beta_L + p}{\beta_E + \beta_L}$$

By observing the commuter's departure time behavior for various values of p, and given knowledge of the shape of the function π, we are able to identify the denominator and numerator in (2). However, β_L and β_E are not identified separately without knowing α, how the commuter values time spent commuting, except in the case when $d\mathrm{E}_T T/dh_D = 0$, that is when expected travel time does not depend on departure time.

Now consider a different congestion charge scheme to help identify the marginal value of time α. First, consider an extension of the model where the commuter also chooses one of two routes $j \in \{0, 1\}$. The route $j = 0$ is the shorter, direct route from home to work, while the $j = 1$ (detour) route takes more time. Travel time on route j at departure time h_D is denoted by $T_j(h_D)$, and satisfies $\mathrm{E}T_1(h_D) > \mathrm{E}T_0(h_D)$ for all h_D. Under the second congestion charge policy, the commuter has a choice between using the short route $j = 0$ and paying a flat fee m, or taking the detour route $j = 1$ for free. The fee m that makes the commuter indifferent between the two options is informative about the value of time α, although β_E and β_L also play a role by determining the optimal departure time for each route.

Taken together, the two congestion charge schemes jointly identify the value of time and schedule flexibility parameters, assuming we know the distributions of travel time and shape of idiosyncratic shocks. The field experiment is designed based on these insights.

3.3 Closing the model: road technology, equilibrium, and social optimum

To close the model, we need to specify how the distribution of travel time at each departure time depends on the volume of traffic at that and other departure times. I consider a simple relationship between the rate of departures at a given time h, and the expected travel time starting at that time. Assume $\mathrm{E}T(h) = F(Q(h))$, where Q is quantity or volume of traffic departing at time h, and $F(\cdot)$ is a function that describes the road technology. (In section 7 I will show that a linear F provides a very good fit to the data.[12]) I further assume that the travel time distribution at a given departure time is fully determined by the expected travel time. (In the empirical application, I show that a log-normal distribution with standard deviation following a quadratic function in the mean offers a good fit to the data.)

A Bayesian Nash equilibrium of this model is described by a pair of travel decisions $(h_{Di})_i$ for all commuters i in the population, and a travel profile $(T(h))_h$ such that commuters respond optimally to traffic conditions, and the travel profile is determined by the aggregate pattern of departures. It is also possible to compute a Nash equilibrium in the presence of departure time monetary charges $\tau(h)$. Simulations in section 8 identify a unique and stable equilibrium. Intuitively, commuters have well-defined desired arrival times, and congestion makes traveling at a given time strategic substitutes.

The social optimum can be implemented as a Nash equilibrium with "Pigou" charges, where the charge $\tau(h)$ at departure time h is exactly the marginal social cost of a commuter traveling at h. We can compute

[12] The bottleneck model is a classic alternative with useful theoretical characteristics (Arnott et al., 1993). In that model, traffic is modeled as a bottleneck with fixed flow capacity. If vehicles arrive at the bottleneck at a rate below capacity, they pass through without delay. As soon as the incoming flow exceeds capacity, a queue forms. The queue is cleared in a first-come-first-served order, at the bottleneck capacity rate per unit of time. The wait time (and queue length) depends on the entire distribution of departure times in the past. Unfortunately, this type of model does not fit the data in this setting.

the marginal social cost of a commuter i traveling at a given departure time h by computing the two equilibria when i leaves at h and when i does not travel at all, and comparing the (utilitarian) welfare for the other commuters.

Having estimated the demand parameters and equipped with a model of road technology, it is possible to compute the equilibrium, evaluate welfare under counterfactual congestion charge policies, compute the externalities imposed in the unpriced equilibrium, and compute the optimal charges.

4 Data Sources and Study Sample

The data backbone of the project is a data set of trips with precise GPS coordinates, collected using a newly developed smartphone app. This data was used both for measuring detailed travel behavior and for implementing the congestion charge policies in the experiment. This section describes how the app was designed, how the data was collected, and how it was automatically cleaned and classified. I also briefly describe several other data sources. The section ends with a description of the study participant recruitment procedure.

4.1 GPS trip-level data from smartphone app

GPS traces. Travel behavior data was collected using a smartphone app that works in the background of any GPS-enabled Android smartphone and passively collects phone location data, without requiring any user input. To conserve battery power, updates were collected at variable time intervals, between every 30 seconds while traveling and every 6 min when in stationary mode.[13] The phone location is identified by the phone operating system using GPS information, as well as cell phone network and WiFi information. (Henceforth, I will refer to this data simply as *GPS data*.) The app uploads data to a server at regular intervals using the phone's data connection. The app has a simple interface that shows a map with the user's current location, and users can receive notifications in the phone notification panel.

Measuring travel behavior using a smartphone-based app has several major advantages over previous data collection techniques. Most often, surveys collect self-reported behavior, which is affected by recall bias, rounding of departure times and trip duration, and tends to underestimate within-person temporal and route variation Zhao et al. (2015). The study app solved these issues by collecting the relevant information completely automatically, without any user input at the beginning or end of a trip, and without requiring participants to later review and validate their trips. Using a smartphone as sensing device also improves

[13] The app, called "Bangalore Traffic Research," was available from Play Store during the study period,. I worked together with GridLocate Ltd, a GPS tracking solutions company, to adapt one of their products to the specific needs of this project.

over previous studies that required participants to carry a separate GPS device.[14]

Trip Data Processing. The raw GPS data for each user-day was automatically cleaned and classified into *trips* and *locations*. I designed and implemented a sequence of algorithms that eliminates outliers and imprecise GPS data points, and segments each day into a sequence of trips and locations, as well as segments corresponding to missing data. Consecutive trips with short stops (at most 15 minutes) between them are linked together into *chains*, which is the unit of analysis. There is no direct way to distinguish the travel mode, including walking or public transport. However, short walking trips are automatically excluded from the sample of trips. The algorithm tags trips outside Bangalore, defined as more than 18km away from the city center. This algorithm was used during the experiment to compute congestion charges for participants in the treatment groups.

Missing GPS data was caused by technical as well as human factors. The app does not record location data if the phone or the location services are switched off, if app permissions are revoked, or if the phone is unable to determine its own location. The last situation may arise, for example, when the phone's 3G internet connection is switched off, because an active connection helps achieve a faster first GPS location. I classify data into three quality categories based on the total duration without location data, and the total distance traveled without precise route information: good quality data, insufficient data, and no data. During the experiment, around 75% of days are good quality, which is the category used for analysis.

Common Destinations and Regular Commuters. Travel behavior and preferences may differ on regular and non-regular trips. In order to be able to control for this important regularity in travel behavior, I identify common, recurring destinations at the commuter level (such as a workplace or school) using a clustering algorithm to group locations into groups, followed by manual review of the location groups most frequently visited.[15] The home location is easy to identify as the most common location group. I then classify one or at most two location groups as "work" destinations. Next, I compute the fraction of distance traveled between home and work, as well as the fraction of days present at work. Using these two variables, I classify participants into regular and variable commuters. Around 75% of study participants have a regular destination, and the median regular commuter visits work on 91% of weekdays (Table 1 Panel B).

Google Maps data. I collected two types of Google Maps data on travel times that include information on traffic congestion. The first data set collected "live" or real-time travel time on 178 routes across Bangalore, including 28 routes in the study area of South Bangalore, every 20 minutes throughout the day,

[14] In a phone survey performed after the experiment ended, only 2.5% of respondents said they left their phone at home "sometimes, for usual destinations."

[15] For grouping locations, I used the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm implemented in the `sklearn` package in python. I then define the top two groups as home and work candidates, respectively, and classify all trips based on whether they connect one or both of these locations.

for 207 days in 2017. This data will be used to calibrate the distribution of travel times, holding route and departure time fixed, and in order to measure the road technology impact of traffic volume on speeds. The second data set is individual-level data on typical travel times between their home and work locations, at all departure times during the day. This data will be used to understand the choice set faced by individual commuters.

4.2 Study sample and survey data

Study participants were recruited in a random sample of gas stations in South Bangalore.[16] Surveyors approached private vehicle drivers (commuters) who were using a car, motorcycle or scooter, excluding taxis and professional drivers, and invited them to participate in a study about understanding traffic congestion in Bangalore. Respondents first answered a short eligibility filter,[17] and if eligible the surveyor explained in broad terms the study purpose, mentioning monetary rewards for participation and the possibility to receive monetary incentives tied to changes in travel behavior. Respondents were invited to install the study smartphone app and answer a very short survey on the spot. All respondents received a study kit including a branded study flyer and consent form.[18] The recruitment survey collected basic contact data, demographic variables (age category, gender, self-reported income, occupation) as well as information on the vehicle (car, motorcycle or scooter, brand and model, odometer reading).[19]

In the weeks after recruitment, we collected travel data from the participant smartphone app.[20]

5 Congestion Charge Policies Experimental Design

I designed and implemented two congestion charge policies that capture the main dimensions of traffic congestion: time and location. The first policy, called "departure time" congestion charges, imposed a pay-

[16] Gas stations are ideal locations to meet commuters who regularly use their private vehicle. (During piloting our team attempted household visits, which suffered from a very low probability of finding respondents at home.) In gas stations, surveyors worked Monday–Saturday in one of two shifts, 8 am – 1 pm or 3 pm – 8 pm.

[17] A respondent was eligible if they reported being the owner or regular user of the private vehicle used on that day, traveling with it or another private vehicle at least 20 Km in total per day, at least three days per week, owning a smartphone and not planing to leave Bangalore for more than two weeks over the following two months. Smartphone usage is very high: 76% of participants eligible based on other conditions owned a GPS Android smartphone (an additional 12% owned an iphone and were not included).

[18] Out of 16,912 persons approached, 43% refused to be interviewed. A further 28% were ineligible. Out of eligible respondents, 27% or 2,300 accepted to install the app. This is calculated assuming the same fraction of ineligibles between those who answered the initial filter and those who refused.

[19] A few variables were collected for all respondents, including refusals: age category, vehicle type, brand and model. I also scraped vehicle prices from an online marketplace, and merged this data with the recruitment survey data for all respondents who were approached.

[20] The study team monitored quality and contacted respondents in case data quality problems arose. Participants were also offered an incentive worth Rs. 300 in phone recharge for providing one week of quality data.

per-kilometer congestion rate that was higher during morning and evening peak hours. Peak hours are a natural target for traffic policies, and congestion charge policies in Stockholm and Singapore have the same feature of higher fees for peak-hour travel. The second policy, called "area" congestion charges, imposed a flat fee for crossing or driving through a specific circular area. This is modeled after flat fee cordon pricing policies (such as those in London and Milan), with an additional special focus on detour route decisions.[21] In addition to emulating common congestion pricing policies, the two policies were designed such that commuter responses to these charges identify the key parameters of the travel demand model described in section 3.

The experimental sample was selected based on app data quality and a second eligibility check.[22] Participants were invited to meet with a surveyor to discuss the second part of the study (the experimental phase). Overall, 497 or 22% of all app participants were enrolled in the experiment on a rolling basis. After the meeting was scheduled and before it took place, participants were randomized into treatments; all participants (including the control group) met in person with a surveyor at a location convenient for the respondent. During the meeting, surveyors explained the treatment and (if applicable) how congestion charges function. Participants were told that the purpose of the study is to understand how commuters in Bangalore would react to the presence of charges, and the surveyors emphasized that there are no correct or incorrect behaviors in response to congestion charges.

During the experiment, charges were deducted from a pre-paid virtual account that was set up for each participant. The outstanding balance at the end of each week during the experiment was transferred to the participant's bank account. In addition to any charges due to their travel behavior, participants were charged a flat fee for no or severely incomplete GPS data, and in case they did not make any trips on a given weekday.[23] A maximum daily total charge and minimum account balance of Rs. 250 also applied. Account opening balances were chosen independently for each participant, based on a model that predicted expected charges given baseline travel behavior and a hypothesis of responsiveness to treatment. (The target final account balance was randomized to either Rs. 500 or Rs. 1,000 per week.) Charges were calculated automatically and participants received daily account balance updates through SMS and app notifications. In addition, weekly phone calls reminded participants about their treatment group details. Participants also received support materials such as a laminated rate card with information about congestion charges (see Appendix Figure A3). To establish trust, participants received a welcome bank transfer soon after the first meeting, and/or an external smartphone battery (power bank) as a gift during the meeting. A study call

[21]The diameters of the congestion areas in London and Milan are 6.5 and 3.5 kilometers, respectively, whereas in this experiment they range between 0.5 and 2 kilometers. Study participants never have a stable destination inside the congestion area, and always have a detour route that takes at most 14 minutes more than their usual route.

[22]Commuters with less than 5km of travel per day, and those who actually lived or spent considerable amount of time outside Bangalore, were dropped.

[23]The "no trip" fee was designed to dissuade incentive gaming by leaving one's smartphone at home for the entire day.

center was available if study participants had questions or complaints.

Experimenter demand effects are an important concern in this setting. Commuters in Bangalore generally care deeply about traffic congestion, and study participants may be motivated to avoid congested times or areas by a sense of civic duty. While these responses may in principle be real, it is also possible that they are specific to this (short-term) experiment, where their participation was voluntary and compensated. I took several steps to guard the experimental results against this possibility. First, during the meeting surveyors were trained to present the options in a neutral light, and to emphasize at least twice that the study does not have a preference over whether participants change or do not change their behavior. Importantly, the experimental design includes a departure time "information" treatment, where participants received flat payments as well as SMS and app notifications that concerned how they can change departure times to avoid traffic. Finally, both types of congestion charges had sub-treatments with price variation, in principle allowing the estimate price responsiveness controlling for overall responsiveness.[24]

Participants were added to the experiment on a rolling basis, and the allocation to treatments was pre-randomized for each stratum. There were eight strata in the experiment, all combinations of participants eligible or ineligible for the area charge, car or non-car (motorcycle or scooter) users, and participants with high or low daily travel distance in the baseline period. The strata, sub-treatments for each of the departure time and area treatments, and timing, is described in Appendix Tables A11 and A12. All departure time and area sub-treatments were cross-randomized within each stratum, and the sub-treatments in each main treatment were stratified in time, across blocks of 8 consecutive slots.

5.1 Congestion Charges

Departure Time Pay-per-Km Congestion Charge

Participants in this treatment were charged for each trip based on a per-km rate and the length of their trip. The rate was positive during a 3-hour interval during the morning and a 3-hour interval during the evening. Each charged interval had the same structure: a one hour increasing "shoulder" ramp when the rate grew linearly from zero to the peak rate, one hour of peak rate, and a one hour decreasing "shoulder" ramp when the rate fell linearly to 0. Appendix Figure A3 shows an example rate card (given to study participants) that illustrates the charges for the morning interval.[25]

[24] In addition, as argued in section 3, we are mainly interested in *relative* preferences over time spent driving and schedule costs, and these measures are more robust to experimenter effects than the absolute values, as there is no obvious reason for these effects to disproportionally affect one treatment over the other.

[25] The start time of the charged interval differed by at most ±30 minutes between commuters, and was designed to maximize the overlap between the shoulder periods and typical departure times for that commuter based on baseline data. This procedure was implemented for *all* commuters before randomizing them between the treatment groups.

Four sub-treatments were designed to separate the impact of prices from other features of the intervention. The sub-treatments were: control, information, low rate, and high rate (see Appendix Table A11). Participants in the control group were monitored for 5 weeks, received regular updates about their data quality, and received a flat Rs. 300 payment per week for participation. I included an information group in order to measure the bundle of experimenter demand, information and reminder provision, and other non-price features. Participants in the information group received daily messages about the trips they had completed the previous day, together with advice about quicker travel times outside the morning and evening peak hours. They also participated for 5 weeks. The low and high rate groups had a maximum (peak) congestion rate of Rs. 12/Km and Rs. 24/Km, respectively. These participants received this treatment for three consecutive weeks out of four in total, either the first three or the last three. (During the remaining week, they received the information group treatment.) Before the start of the congestion charge phase, participants underwent a three-day *trial phase* where they received congestion charge messages to understand how charging works. In total, low rate and high rate participants also were in the experiment for approximately 5 weeks.

Area Congestion Charge

Participants in this treatment were charged for driving through a congestion area that was chosen individually for each participant. The area was a disc with radius 250m, 500m or 1000m positioned along a route used frequently by the participant during the pre- period. The area induced an alternate non-intersecting detour route, which was between 3 and 14 minutes longer than the original route. If no area charge with this property was found, the participant was ineligible for the area congestion charge. (Roughly half of the experiment participants were are eligible.) The charge was in effect between 7 am and 9 pm, and applied at most once for the morning interval (7 am – 2 pm) and at most once for the evening interval (2 pm – 9 pm). The area congestion charge was implemented for one week (five weekdays). The area location, radius, boundaries, and induced detour were emphasized by the surveyor during the meeting before the experiment, and this information was repeated in each daily reminder SMS sent to study participants.[26]

The area treatment did not include a pure control group, due to the smaller size of participant pool. However, participants were randomized between being treated early (in the first week after the meeting) or late (in the last or forth week of the study), which is the basis for the experimental comparison.

The area sub-treatments were designed to identify the effect of price and detour time variation on choices. On two randomly chosen days, the congestion charge was 50% higher. The following sub-treatments were cross-randomized (see Appendix Table A11). Low rate participants were charged a baseline charge of Rs. 80

[26]The area location did not specifically target congested areas. Surveyor were instructed to not convey this idea to the participant during the in-person meeting, and if they were asked to reply that the area was selected by an algorithm.

(and Rs. 120 on the two days per week when the charge was higher), while High Rate participants were charged Rs. 160 (and Rs. 240 respectively). Long detour participants had an area location and radius that induced a predicted detour between 7 and 14 minutes above the usual route, if such an area existed. Short detour participants had an area that induced a predicted detour between 3 and 7 minutes above the usual route, if such an area existed.

5.2 Reduced-Form Responses to Congestion Charges

Reduced-Form Specification

The congestion charges described above may affect the number as well as the temporal and spatial distribution of trips. In order to capture unconditional effects, I first aggregate outcomes at the day level and run the following difference-in-difference specification:

$$y_{it} = \delta^I T_i^I + \delta^L T_i^L + \delta^H T_i^H + \gamma^I T_i^I \times Post_t + \gamma^L T_i^L \times Post_t + \gamma^H T_i^H \times Post_t + \mu_t + \alpha_i + \varepsilon_{it}, \qquad (3)$$

where y_{it} is an outcome of interest for commuter i on day t, such as the number of trips that day, $Post_t$ is a dummy for the period of the experiment, T_i^I, T_i^L and T_i^H are dummies for the information, low rate and high rate departure time sub-treatments, and α_i is a commuter fixed effect, μ_t is a study cycle fixed effect whose categories are the period before the experiment, and each week in the experiment. The coefficients of interest, γ^I, γ^L and γ^H, respectively measure the impact of information, low congestion rates and high rates relative to control, during the experiment relative to the period before.

The sample is all non-holiday weekdays when the respondent does not travel outside Bangalore. During the experiment, I include the three weeks when charges are in effect; in the control and information groups I also keep three weeks to make the timing in each sub-treatment comparable. Where necessary for the construction of the y_{it} variable, the sample is restricted to days with "good quality" GPS data, as defined above. Standard errors are clustered at the commuter level. For trip level outcomes, I use the same specification with outcome y_{jit} corresponding to trip j of commuter i on day t.

For the Area treatment, there is no pure control group. Instead, the empirical strategy is based on comparing commuters randomly assigned to be treated early or late. Specifically, in the first week I compare commuters treated early (treated group) to those treated late (control group). In the fourth week, these roles are reversed. The period before the experiment and the second and third week during the experiment are included to gain precision when estimating individual fixed effects. Specifically, define $Treated_{it} = (1 - T_i^{Late}) \times \mathbb{1}(t \in W_1) + T_i^{Late} \times \mathbb{1}(t \in W_4)$ where T_i^{Late} is an indicator for being treated late, and W_s is an indicator for week s. I run the following specification:

$$y_{it} = \gamma^A \cdot Treated_{it} + \mu_t + \alpha_i + \varepsilon_{it} \quad (4)$$

The coefficient of interest is γ^A, which measures how the outcome y_{it} differs as a result of being exposed to area congestion charges, relative to similar commuters who are not treated that week.

Experimental Integrity Checks

Table A2 reports the experimental balance check. The different treatment groups are similar along demographic and pre- period travel behavior variables. All coefficients are small, and join significance tests cannot reject the null of no effect.

Given that smartphone app data was used to implement the congestion charges, it is especially important to ensure that treated participants did not differentially tamper with their smartphones by switching the phone or the GPS sensor off during certain trips or on certain days. During the experiment, participants provided good quality GPS data on approximately 75% of weekdays. Appendix Table A1 shows that the departure time and area sub-treatments did not have any detectable differential impact on GPS data quality. This suggests that missing GPS data was mostly due to technical and human factors unrelated to gaming incentives.[27]

The Impact of Departure Time Pay-per-Km Charges

Commuters may respond to charges by canceling trips with departure times during the charged period, as well as by rescheduling these trips to departure times with lower charges. Figure 2 shows the causal impact of congestion charges on the distribution of trip departure times, for the morning and evening charges. It plots a locally linear difference-in-difference by departure bin. To construct Figure 2, for each commuter, day and departure time *relative* to the midpoint of the congestion charge for the commuter,[28] I compute the number of trips that start around that time, using an Epanechnikov kernel. Then, for each departure time I run a regression similar to (3) except that I compare the low rate and high rate charge groups (identified by T_i^{LH}) to the control and information groups combined. Figure 2 plots the coefficients on $T_i^{LH} \times Post_t$ as well as pointwise 95% confidence intervals.

Commuters substitute away from departure times with high charges towards departure times with lower charges. In the morning (panel A) there is strong substitution within the early ramp interval, when the

[27] The experiment was generally successful in terms of *retaining* study participants: around 5% of participants dropped out right after the meeting, and this figure rose to 10% on the last day of the study. Drop outs are 2 percentage point more frequent in the treatment group, yet this difference is not statistically significant (p-value 0.20).

[28] Recall that the congestion charge has the same shape for everyone, but its location varies by commuter, including for those in the information and control groups.

charge is linearly increasing. In this interval, there is a marginal incentive to advance one's departure time. The results suggest that study participants understood this feature and decided to leave earlier and take advantage of lower charges. There is suggestive evidence of an increase in the number of trips starting right after the end of the charged period; note that the exact position of this increase does not map cleanly to the predicted response given incentives, in the way that the early AM change does.

The results for the evening period are broadly mirrored, namely commuters substitute towards later departure times on the decreasing ramp of the congestion charge profile. However, the results are slightly weaker and less precise. In sum, Figure 2 shows that commuters responded to charges by advancing their departure times in the morning when this leads to lower charges, and delaying their departure times in the evening.

Table 2 shows results on daily outcomes. Panel A of Table 2 shows impacts on trip shadow rates. This outcome is computed in the same way for every trip in the data given its departure time and the commuter's own congestion charge rate profile (which is defined for everyone irrespective of treatment group) and using a normalized peak rate of 100. The rates are then summed over all trips in the day. This outcome is a summary statistic for whether the commuter changed their travel behavior to avoid charges, and includes intensive and extensive margin responses. The results show that the High Rate sub-treatment leads to a decrease of around 14 from a base of 97 in the control group. The Low Rate treatment also appears to lead to a decrease in rates, of roughly half the size of that of the High Rate group, yet these results are not significant. The information group does not seem to have any effect on charges. In panel B, the outcome is the total number of trips in the first column, and the total number of trips during the morning and evening in the other columns. The point estimates are negative, small, and far from statistical significance.

Running the same specification at the trip level leads to similar results. Daily charges are mechanically related to the number of trips per day that occur in the charged interval. Even in the absence of a treatment effect on the number of trips, chance variation in the number of trips per day between treatment groups reduces the precision of the estimates in panel A. Table 3 explores the impact of charges at the *trip* level instead of daily level. Panel A covers the entire sample of commuters and trips, panel B covers only regular commuters and trips between home and work or vice-versa, and panel C covers all trips belonging to the approximately 25% variable commuters. In addition to full day results in column (1) and results in the morning and evening in columns (2) and (4), the table also reports results restricted to the *early* morning interval (all departure times before the midpoint of the peak of the rate profile) in column (3), and restricted to the *late* evening interval in column (5).

Trips in the High Rate have on average lower rates by around $13 - 15\%$ relative to the control group (panel A), with a larger and precisely estimated effect in the morning. The coefficients for the Low Rate treatment are also negative, of roughly half the size, yet not statistically significant. The effects are more precisely estimated for regular commuters in panel B. In particular, the coefficients for early morning and late evening are negative and larger than for the entire day period, as suggested in Figure 2. In panel C there is no evidence that congestion charges changed the distribution of trip departure times for variable commuters. In the entire table, no discernible pattern emerges for the information group, suggesting that information alone did not shift travel behavior.

In Appendix Figure A1 panel A, I investigate the heterogeneity in individual responses to the departure time treatments. Pooling together the Low Rate and High Rate (as in Figure 2), the figure shows that treatment group respondents have a bi-modal distribution in the within-person change in shadow trip rates. This suggests that a certain group of commuters decided to change their behavior to take advantage of lower charges, while others did not make any changes.

The Impact of Area Charges

Following the discussion of the model in section 3, we are interested in the impact of Area charges on the choice probability of alternate routes that avoid the congestion area. Tables 4 and 5 report the results at the day and trip level, respectively.

Panel A of Table 4 reports the impact on total shadow charges due to crossings of the congestion area. These are calculated for every trip in the sample, and the charge for a crossing is normalized to 100. The results show a large, precisely estimated decrease in the probability to cross the congestion area. The decrease is around 23% of the control mean, significant at the 1% level. The impact is similar in the morning and evening intervals, and roughly similar for participants treated in the first or the last week (columns 4-6). Panel B shows the impact of being treated on the number of trips in the day. Being treated results in around $6-10\%$ more trips per day, with the effect concentrated in the morning and for participants treated in the last week. The increase in number of trips seems related to a small increase in data quality in the treatment group (both effects are concentrated in the 4th week). Note that a larger number of trips will tend to mechanically increase the coefficient on shadow rates, so the treatment impact may in reality be slightly more negative than the result in column (1).

Table 5 investigates whether the area charge induced commuters to take a longer detour, and the choice probability of alternate routes that avoid the congestion area. The table shows results at the trip level and restricts the sample to regular commuters and trips from home to work or vice-versa.[29] Panel A reports the impact on whether the trip intersects the congestion area, and shows a large reduction of 23 percentage

[29] 93% of Area treatment participants are regular commuters.

points on a base of 83%, or equivalently a 29% reduction in area crossings. The effect is very precisely estimated, and of similar magnitude in the morning and evening intervals.

Panel B uses trip duration as outcome variable and reports the experimental effect of being treated on trip duration. The point estimates are positive on average, yet small and not significant. This result is likely due to lack of power to detect a reduced-form effect on trip duration. Indeed, multiplying the treatment effect in panel A by the average difference in duration (4 minutes) we find an average increase of 1.45 minutes for Treated respondent. The point estimates are most often smaller, but we cannot reject this value either.

One concern would be that study participants identified alternate routes that avoided the congestion area that are quicker than what I estimated using Google Maps. To explore this hypothesis, Appendix Table A5 shows the non-experimental correlation between trip duration and whether a trip is charged, including commuter-level, directed route fixed effects. Charged trips are significantly shorter, by about 5 minutes. Moreover, this effect is significantly larger for respondents in the Long Detour Area sub-treatment (column 2), and the extra duration for avoiding trips closely tracks the Google Maps predicted detour of the quickest non-intersecting alternative (column 3). These results show that the Google Maps data accurately predicts the extra time detour incurred in real trips that avoid the congestion area.

Randomly varying the crossing charge and the detour length does not affect the response to the area treatment. Indeed, Table 6 shows that neither doubling the congestion charge (randomized across participants), nor having a 50% higher charge on a random day (randomized within participant), has any significant effect on shadow charges (columns 2 and 3). The last column shows that participants randomly assigned to a short detour ranging between 3 and 7 minutes (as opposed to the long detour, between 7 and 14 minutes) do not reduce their shadow charges more. These results are consistent with high levels of heterogeneity in the population, whereby some participants are easy to sway to change their routes (low values of time), while the others are much more difficult to convince (high values of time).

Individual level response heterogeneity is consistent with this story (Appendix Figure A1 panel B). For each area participant, I count the fraction of days crossing the congestion area, separately when treated and when in the control group. The distribution in the control group is concentrated near 1, as most commuters select the shortest route in the absence of charges (solid, gray bars). In the presence of charges, the distribution becomes bi-modal, with around 20 per cent of the population in the lowest bin, implying that some participants stopped crossing the congestion area at all (outline, red bars).

On the other hand, results on observable sources of heterogeneity are somewhat imprecise (Appendix Table A6). Regular commuters and self-employed commuters appear to respond more to the departure time treatment (columns 1 and 2, panel A), although these differences are not quite statistically significant. Surprisingly, commuters with more expensive vehicles seem to respond more to the departure time treatment, and there is also evidence that they reduce their number of trips (column 4, panels A and B). There is also suggestive evidence that older respondents respond more to both treatments (column 5). There is no evidence that stated preferences predict responses in the experiment (columns 6 and 7).

In summary, departure time pay-per-km charges caused commuters to change their departure times towards departure times with lower charges, especially towards earlier departures in the morning and later departures in the evening. This means that commuters have some flexibility to move trips away from typical work hours in order to save money. These results are driven by a subset of commuters who responds more strongly. Responses in the low rate sub-treatment are roughly half of those in the high rate sub-treatment, although imprecisely estimated, and I do not find any impact of the information and nudges treatment.

Area congestion charges lead to a precisely estimated shift to routes that avoid the congestion area. Participants intersect the congestion area around 20% fewer times when "area" charges are in effect. Doubling the congestion charge or shortening the implied detour experimentally do not affect this fraction. (Routes that do not intersect the congestion area are on average 5 minutes longer.) Consequently, the naive implied value of time for the marginal participant lies between Rs. 1,152 and Rs. 2,304 per hour, both of which are large. These findings are consistent with considerable preference heterogeneity.

However, in order to better quantify these results – especially in terms of interpreting and comparing the responses to the two treatments – I will next estimate a structural model where agents choose departure times and routes.

6 Structural Travel Demand Estimation

I now estimate the key parameters in a model of travel demand over routes and departure times, using experimental variation from the congestion charge treatments. This procedure will provide monetary measures of individual preferences over schedule inflexibility and mean driving time.

I first augment the model set up in section 3 with route choice, commuter heterogeneity and random utility shocks, and derive the choice probabilities. I then describe the Google Maps data used to construct individual-level choice sets, discuss the experimental moments, and finally I present discuss the results and robustness exercises.

Nested Logit Model over Routes and Departure Times

To make the model of the morning home to work commute introduced in section 3 easier to fit to real data, I add route choice, commuter heterogeneity and random utility shocks. On day t, a commuter i chooses a route type $j \in \{0, 1\}$, where $j = 0$ represents any route from home to work that intersects the congestion area, and departure time h_D, chosen from a discrete grid of departure times H_D. Utility is given by:

$$\begin{aligned} U_{it}\left(h_D, j, h^*_{Ait}\right) = & -\alpha \mathrm{E}T_i\left(j, h_D\right) - \beta_E \,\mathrm{E}\left|h_D + T_i\left(j, h_D\right) - h^*_{Ait}\right|_- - \beta_L \,\mathrm{E}\left|h_D + T_i\left(j, h_D\right) - h^*_{Ait}\right|_+ \\ & - m^{DT}_{it}\left(h_D\right) - m^A_{it}\left(j\right) + \varepsilon_{it}\left(j, h_D\right), \end{aligned} \tag{5}$$

where h^*_{Ait} is the ideal arrival time on day t, $T_i\left(j, h_D\right)$ is the (random) driving time on route j, $m^{DT}_{it}\left(h_D\right)$ represents the departure time congestion charge that may apply to the current trip, $m^A_{it}\left(j\right)$ is the area congestion charge for route j, and $\varepsilon_{it}\left(j, h_D\right)$ is a random utility shock for route j and departure time h_D on day t.[30] Both h^*_{Ait} and $\varepsilon_{it}\left(j, h_D\right)$ are drawn i.i.d. each day. Expectations are with respect to the random driving time T_i. The key preference parameters of interest are α, β_E and β_L, respectively the value of mean travel time, and the schedule costs of arriving early and late. Commuter heterogeneity is captured by different distributions of ideal arrival times h^*_{Ait} and different driving time profiles T_i.

In order to allow different patterns of substitution between departure times and between routes, I assume that the random utility shocks $\varepsilon_{it}\left(j, h_D\right)$ follow an extreme value distribution with correlation within each route. This leads to a nested logit structure over routes and departure times. The two route choices constitute the upper nest, while the choice over departure times is the within-nest component.[31,32]

The distribution of the random utility shocks depends on two parameters, σ and μ, which will be estimated from the data.[33] The probability to choose a given departure time and route can be decomposed as $\Pr\left(j, h_D \mid h^*_{Ait}\right) = \Pr\left(h_D \mid j, h^*_{Ait}\right) \Pr\left(j \mid h^*_{Ait}\right)$. Denote by $V_{it}\left(h_D, j, h^*_{Ait}\right)$ the constant part of utility in (5) (without the utility shock $\varepsilon_{it}\left(j, h_D\right)$), then the departure time choice conditional on route is

$$\Pr\left(h_D \mid j, h^*_{Ait}\right) = \frac{\exp\left(\frac{1}{\sigma_i} V_{it}\left(h_D, j, h^*_{Ai}\right)\right)}{\sum_h \exp\left(\frac{1}{\sigma_i} V_{it}\left(h, j, h^*_{Ai}\right)\right)} \tag{6}$$

Costs scale approximately linearly with route length, so I normalize the logit parameter for commuter i by i's route length, namely $\sigma_i = \frac{KM_i}{\overline{KM}}\sigma$ where $\overline{KM}$ is the sample average of KM_i. This means that all commuters have similar probabilities to choose non-optimal departure times, instead of commuters who

[30] When calculating departure time congestion charges, I ignore the trip distance dependence on route j. In the experiment, the area and departure time treatments never apply at the same time; trip distance still matters, by making route $j = 1$ relatively less attractive when departure time charges are in effect. However, this effect is of secondary importance.

[31] The assumption of independent utility shocks at several minute intervals along the departure time grid may not seem particularly attractive. However, the resulting choice probabilities have a familiar form. To see this, assume that utility is quadratic – which always holds as an approximation around the optimum h^*_D – then as the grid becomes finer the multinomial logit model becomes equivalent to choosing the optimum h^*_D plus a random noise term, with the standard deviation of the noise term related to the inverse curvature of the utility function at the optimum.

[32] It is possible to set up more detailed models over departure times and routes. For example, transportation researchers have developed route choice models that are considerably more sophisticated than the one used here (Ben-Akiva M., 2003). However, this model serves the primary purpose of understanding the margin of route choice highlighted in the experiment, namely the trade-off between taking a longer route and paying a higher congestion charge.

[33] The normalization used here is that utility is expressed in Rupees.

travel far having more precise choices, as would be implied by a constant $\sigma_i = \sigma$.

The route choice probability is given by

$$\Pr\left(j \mid h^*_{Ait}\right) = \frac{\exp\left(\frac{1}{\mu} V_{it}\left(j\right)\right)}{\exp\left(\frac{1}{\mu} V_{it}\left(0\right)\right) + \exp\left(\frac{1}{\mu} V_{it}\left(1\right)\right)} \tag{7}$$

where $V_{it}\left(j\right) = \sigma_i \log\left(\sum_h \exp\left(\frac{1}{\sigma_i} V_{it}\left(h, j, h^*_{Ait}\right)\right)\right)$ is the expected utility assuming i chooses route j, called the "logsum" term for route j. The parameters σ and μ measure the importance of utility shocks for departure time choice and route choice, respectively. Higher values correspond to more importance given to utility shocks (less precise choices).[34] Overall choice probabilities are obtained by integrating over the (individual-specific) distribution of ideal arrival times h^*_{Ait}.

To capture the stark heterogeneity documented in the reduced form experimental results, I assume that each participant responds to experimental congestion charges with some probability p, while with probability $1-p$ they behave as if there were no charges. This assumption has two possible interpretations: either this behavior reflects real preferences, that is, a fraction $1-p$ of the population is infra-marginal to the incentives offered in the experiment, or for other reasons these participants decided to ignore the experiment, in which case we do not know their true preferences.[35]

To summarize, agents choose routes and departure times according to nested logit, and they ignore monetary charges with some probability. The full vector of parameters to be estimated is $\theta = (\alpha, \beta_E, \beta_L, \sigma, \mu, p)$ as well as the individual specific distributions of ideal arrival times h^*_{Ait}.

Data Sources and Model Simulation

In addition to behavior data collected using the smartphone as part of the experiment, fitting this model requires knowledge of the counterfactual distribution of driving times. For average driving times $\mathrm{E}T_i\left(0, h_D\right)$ and the short route length KM_i, for each person I collected Google Maps predicted driving times on their home to work route at all departure times throughout the day. To calibrate the distribution of driving times conditional on route and departure time, I use live Google Maps data collected on a set of 178 routes across Bangalore. Conditional on route and departure time, driving time is approximately log-linearly distributed across the 146 weekdays in the data, with the standard deviation well explained by a quadratic in the average driving time (Appendix Figures A5 and A6). Thus, for each commuter and departure time, I assume that

[34] Commuters not in the area treatment only choose departure time, according to multinomial logit (there is no route choice). Their choice probabilities are given by $\Pr\left(h_D \mid h^*_{Ait}\right) \propto \exp\left(\frac{1}{\sigma_i} V_{it}\left(h_D, h^*_{Ait}\right)\right)$.

[35] A more traditional way to capture preference heterogeneity is by assuming random coefficients, that is that parameters α, β_E and β_L vary at the individual level according to some distribution (such as log normal). In this setting, estimating models with random coefficients fails to fully capture the heterogeneity documented in section 5.2 and Appendix Figure A1.

driving time follows such a distribution given the measured Google Maps average driving time. I calibrate driving times on the alternate route ($j = 1$) as an individual-specific constant multiple of driving times on the shortest route ($j = 1$).[36] I assume that the relevant variation in commuter beliefs is captured by the Google Maps travel time. In particular, if commuters systematically over- or under-estimate travel time differences, then the structural estimates from these procedure should be adjusted based on those beliefs.

The estimation sample covers the morning interval, covers all trips between home and work, and restricts to 308 regular commuters with at least two observed trips between home and work in the morning interval during the experiment.

To compute choice probabilities, I use formulas (6) and (7) given individual preference parameters α, β_E, β_L, the nested logit parameters σ and μ, the ideal arrival time h^*_{Ait}, and the travel time distributions for each route and departure time, denoted $\tau_i \equiv (T_i(j, h_D))_{j,h_D}$. During estimation, I assume candidate values for the first five preference parameters, while the travel times are taken from the Google Maps data together with the log-normal distribution assumption. The remaining difficulty is that the distribution of h^*_{Ait} is neither observed nor known *a priori*. To overcome this, I use the observed distribution of departure times in the pre period (before the experiment) to obtain the distribution of ideal arrival times conditional on other parameters. I then use this distribution for h^*_{Ait} to compute choice probabilities both before and during the experiment.[37] During the experiment, the terms $m^{DT}_{it}(h_D)$ and $m^A_{it}(j)$ are either zero or the congestion charges experienced by commuter i in that period, denoted by M^{DT}_{it} and M^A_{it}.

GMM Estimation and Moment Choice

To estimate the model parameters, I use the generalized method of moments (GMM).[38] I use four sets of moments: (1) difference in difference changes in departure time "market shares," (2) the variance of individual-level changes in shadow charges in the departure time treatment and control groups, (3) route choice "market shares" when treated and not treated with area charges, and (4) the 3-bin histograms for

[36]For area treatment participants, before the experiment, I obtained from Google Maps the driving time for the quickest route that does not intersect the congestion area, for a departure time of 9 am for all participants. I assume that the driving times on the detour route ($j = 1$) are a constant multiple of the driving times on the main (intersecting) route, namely $T_i(1, h_D) = \lambda_i T_i(0, h_D)$. The constant λ_i is chosen to match the alternate route travel time at 9 am for person i, as queried before the experiment.

[37]Specifically, I first fit a normal distribution on departure times during the pre period. This is done before estimation, and confidence intervals in Table 7 do not take into account that the departure time distributions are themselves estimated. Then, for given parameter values, I find the distribution of ideal arrival times that, under optimal behavior, would give rise to the normal fit on departure times. This inversion is computationally expensive to do precisely. Instead, I make the following approximations: (1) for each ideal arrival time h^*_{Ait} the optimal departure time is normally distributed around the utility maximizing departure time, with the standard deviation given by the curvature of the utility around the optimum (see footnote (31)), and (2) I assume that the standard deviation is constant for all h^*_{Ait} which allows me to obtain the distribution of optimal departure times by shrinking the distribution of departure times. I then invert the optimal departure time relationship to obtain the distribution of ideal arrival times.

[38]Nested logit has a closed form likelihood function, recommending maximum likelihood on efficiency grounds. Nevertheless, with GMM it is possible to choose moments such that parameters are essentially identified from experimental variation.

individual sample frequency of choosing the short route when treated and not treated with area charges.

The first 61 moments match the difference in difference in departure time market shares, between the departure time treatment and control groups, during the experiment relative to before. Formally, for each 5-minute departure time bin h^k between -2.5 and 2.5 hours relative to the rate profile peak, and for each participant i, I compute the probability that i leaves during h^k, conditional on a trip being made. In the model, for a day t, let $P_{itk}^{DT}\left(\theta,\tau_i,m_{it}^{DT}\right)=\Pr\left(h_i\left(\theta,\tau_i,m_{it}^{DT}\right)\in h^k\right)$ where the random departure time (relative to i's peak) h_i depends on preference parameters, the travel time profile τ_i and charges m_{it}^{DT}. In the data, define $\tilde{P}_{ik}^{DT}\left(pre\right)$ and $\tilde{P}_{ik}^{DT}\left(post\right)$ the fractions of trips starting in bin h^k for individual i in pre- and post- periods, respectively. Recall that M_{it}^{DT} denotes the charges assigned to i in the experiment, and for convenience make the dependence on θ and τ_i implicit. For $k\in\{1,\ldots,61\}$, the k-th moment is:

$$\begin{aligned}g_i^k\left(\theta\right)=&\left(\left(\tilde{P}_{ik}^{DT}\left(post\right)-\tilde{P}_{ik}^{DT}\left(pre\right)\right)\cdot T_i^{LH}-\left(\tilde{P}_{ik}^{DT}\left(post\right)-\tilde{P}_{ik}^{DT}\left(pre\right)\right)\cdot\left(1-T_i^{LH}\right)\right)\\&-p\cdot\left(\left(P_{itk}^{DT}\left(M_{it}^{DT}\right)-P_{itk}^{DT}\left(0\right)\right)\cdot T_i^{LH}-\left(P_{itk}^{DT}\left(M_{it}^{DT}\right)-P_{itk}^{DT}\left(0\right)\right)\cdot\left(1-T_i^{LH}\right)\right)\end{aligned}$$

where T_i^{LH} is an indicator for being in any of the departure time treatment groups (low or high rate), t is a day during the experiment, and the heterogeneity parameter p enters by attenuating the model term. Intuitively, for given p these moments help identify the schedule costs β_E and β_L, as well as the logit parameter σ. The magnitude of responses on the early and late ramps of the congestion rate profile identify the first two parameters, while the precision of these responses helps identify σ.

The departure time heterogeneity moments target the variance of the individual-level change in shadow charges for trips in the early morning, between the pre and post periods. For these moments, it is important to take sampling variation into account when simulating the model, so denote N_i^{pre} and N_i^{post} the number of days in the pre and post periods for i. Assume h_{it} for $t=\left\{1,\ldots,N_i^{pre}+N_i^{post}\right\}$ are independent random variables, the first N_i^{pre} distributed according to $h_i\left(\theta,\tau_i,0\right)$, and the rest according to $h_i\left(\theta,\tau_i,M_{it}^{DT}\right)$, in both cases conditional on departure times in the two hours before the rate profile peak, namely $h_i\in[-2,0]$. Define $ch\left(h\right)$ to the be the shadow charge of departure time h, and the random individual effect as

$$ch_i^{DT}=\frac{1}{N_i^{post}}\sum_{t=N_i^{pre}+1}^{N_i^{pre}+N_i^{post}}ch\left(h_{it}\right)-\frac{1}{N_i^{pre}}\sum_{t=1}^{N_i^{pre}}ch\left(h_{it}\right)$$

Denote the individual effect in the data by $\tilde{ch}_i^{DT}$. The two departure time heterogeneity moments match the variance of ch_i^{DT} in the treatment and control groups. The expressions for full response ($p = 1$) are:[39]

$$g_i^{62} = \left(var\left(ch_i^{DT}\right) - \widehat{var}\left(\tilde{ch}_i^{DT}\right) \right) \cdot T_i^{DT}$$
$$g_i^{63} = \left(var\left(ch_i^{DT}\right) - \widehat{var}\left(\tilde{ch}_i^{DT}\right) \right) \cdot \left(1 - T_i^{DT}\right)$$

The first moment helps identify the probability p that a study participant responds to the treatment. Indeed, given other parameter values, p affects the variance of the individual effect, by splitting the sample between participants who respond and those who do not respond. The second moment helps ensure that the model is able to replicate the sampling variation in individual effects.

The next two moments match route choice market shares, namely the probability to intersect the congestion area when treated and when not treated for commuters in the area congestion charge treatment. Formally, define $P_i^A\left(\theta, \tau_i, m_{it}^A\right) = \Pr\left(j\left(\theta, \tau_i, m_{it}^A\right) = 0\right)$ the probability to take the short route (intersect the congestion area), where the random route choice j depends on preference parameters, the travel time profile τ_i and charges m_{it}^A. In the data, define $\tilde{P}_i^A\left(treat\right)$ and $\tilde{P}_i^A\left(control\right)$ the fraction of days (mornings) when the commuter intersects the congestion area, when treated and when not treated, respectively. Recall that M_{it}^A denotes the area charge assigned to i in the experiment, and for convenience make the dependence on θ and τ_i implicit. The area moments are:

$$g_i^{64}\left(\theta\right) = \left(p \cdot P_i^A\left(M_{it}^A\right) - (1-p) \cdot P_i^A\left(0\right) - \tilde{P}_i^A\left(treat\right)\right) \cdot T_i^A$$
$$g_i^{65}\left(\theta\right) = \left(P_i^A\left(0\right) - \tilde{P}_i^A\left(control\right)\right) \cdot T_i^A$$

where T_i^A is an indicator for being in the area treatment. Without area charges, a commuter will only choose the detour route ($j = 1$) due to large utility shocks that offsets the driving time penalty. For a given value of time α, this helps identify the outer nest logit parameter μ. With area charges, there is an additional monetary benefit to choosing the detour, and for given p these moments together help identify α.

The area heterogeneity moments target the distribution of individual-level sample frequency of intersecting the area. Once again, it is important to take sampling variation into account, so define N_i^{treat} and $N_i^{control}$ the number of days when i is treated and not treated, respectively. Assume j_{it} for $t = \left\{1, ..., N_i^{control} + N_i^{treat}\right\}$ are independent random variables, the first $N_i^{control}$ distributed according to $j\left(\theta, \tau_i, 0\right)$, and the rest according to $j\left(\theta, \tau_i, M_i^A\right)$. Define the (random) sample average of intersecting

[39] Note that ch_i^{DT} is random for a given commuter i, and its distribution also differs between commuters. We are interested in the overall variance, both between commuters and within commuter, as this is what we see in the data. Hence, I use the following shorthand notation: $var\left(ch_i^{DT}\right) \equiv \mathrm{E}\left(ch_i^{DT} - \mathrm{E}\frac{1}{N}\sum_{j=1}^{N} ch_j^{DT}\right)^2$ and $\widehat{var}\left(\tilde{ch}_i^{DT}\right) \equiv \left(\tilde{ch}_i^{DT} - \frac{1}{N}\sum_{j=1}^{N} \tilde{ch}_j^{DT}\right)^2$.

the area in control and treatment as

$$ch_i^{A,control} = \frac{1}{N_i^{control}} \sum_{t=1}^{N_i^{control}} j_{it} \quad \text{and} \quad ch_i^{A,treat} = \frac{1}{N_i^{treat}} \sum_{t=N_i^{control}+1}^{N_i^{control}+N_i^{treat}} j_{it}$$

In the data, denote the corresponding quantities by $\tilde{ch}_i^{A,control}$ and $\tilde{ch}_i^{A,treat}$, respectively. We are interested in the distribution of these variables. The four area heterogeneity moments match the probability that these variables are the middle or top third of the unit interval (the moment for the bottom third is omitted because it is colinear with the others), when treated and not treated. The expressions for full response ($p = 1$) are:

$$\begin{aligned}
g_i^{66} &= \left(\Pr\left(ch_i^{A,treat} \in [1/3, 2/3)\right) - \mathbb{1}\left(\tilde{ch}_i^{A,treat} \in [1/3, 2/3)\right)\right) \cdot T_i^A \\
g_i^{67} &= \left(\Pr\left(ch_i^{A,treat} \in [2/3, 1]\right) - \mathbb{1}\left(\tilde{ch}_i^{A,treat} \in [2/3, 1]\right)\right) \cdot T_i^A \\
g_i^{68} &= \left(\Pr\left(ch_i^{A,control} \in [1/3, 2/3)\right) - \mathbb{1}\left(\tilde{ch}_i^{A,control} \in [1/3, 2/3)\right)\right) \cdot T_i^A \\
g_i^{69} &= \left(\Pr\left(ch_i^{A,control} \in [2/3, 1]\right) - \mathbb{1}\left(\tilde{ch}_i^{A,control} \in [2/3, 1]\right)\right) \cdot T_i^A
\end{aligned}$$

Intuitively, the first set of moments will help identify the probability p that a study participant responds to the treatment, by matching the empirical histogram in the treated group with an average between the model treated and the model control histograms.

6.1 Structural Estimation Results

Table 7 shows the estimation results from two-step GMM, using 100 random parameter starting values to ensure convergence to the global minimum of the objective function.

Commuters value time spent driving at Rs. 1,122, and the estimated schedule cost of arriving earlier than ideal is Rs. 320. Commuters are thus relatively schedule flexible to leave earlier in the morning. To put these values in context, a commuter with these preferences would be indifferent between leaving one hour earlier if the driving time from leaving early was 15 minutes lower (this back of the envelope example ignores uncertainty). In particular, this means that commuters have some ability to "self-insure" against congestion, in the sense that commuters will tend to change departure times in response to a *localized* increase in congestion, which will reduce the welfare impact of the shock. It is important to note that the estimated value of time is significantly larger than the average self-reported monthly income of Rs. 270 per hour (Rs. 39,000 per month).[40]

[40] It is possible that this estimate of α also includes a fixed cost of switching routes. The field experiment was designed to separate the fixed cost of route change and marginal costs of travel time, through the low and high rate sub-treatments in the area treatment. Given that I do not find any reduced form effect of increasing the area congestion charge, I model the route

The late arrival cost β_L cannot be estimated precisely from the data. The underlying reason is that in Figure 2 there is no reduced form impact on late departures. This tells us that β_L is large; however, it is not clear how large. For the estimation in Table 7, this parameter is fixed at β_L = Rs. 4,000. Appendix Figure A8 (Panel A) shows that the GMM objective function is mostly flat above this value. In Appendix Table A7, I show that using β_L = Rs. 1,000 or β_L = Rs. 8,000 instead has no detectable effect on the other estimated parameters. This inflexibility of leaving later is consistent with work requirements acting as a firm constraint.

Around half of all study participants responded to congestion charges ($\hat{p} = 0.46$). Intuitively, this value maximizes the variance of individual responses, emphasizing the stark response heterogeneity in the data.

Both logit parameters are estimated to be approximately Rs. 37, indicating a small or moderate amount of noise in choices. The fact that these terms are equal means that I cannot reject the multinomial logit model over the entire decision space. The inner nest logit parameter σ, corresponding to departure time choice, is estimated with significantly more noise than the outer nest parameter μ, which corresponds to route choice. This is related to commuter heterogeneity in terms of ideal arrival time. In principle, both a wide distribution of h^*_{Ait} and a large σ will imply a wide observed distribution of departure times. The logit parameter is separately identified from the shape of the *experimental* response to departure time congestion charges. As σ becomes smaller, the impact concentrates around the kinks of the congestion ramp. In practice, and with the available data, I can only estimate σ somewhat imprecisely.

Appendix Figure A7 shows the model fit graphically by plotting the data and model prediction for the moments used in estimation. The model generally fits the data well, and in particular it does a good job of replicating the variance in individual effects for departure times and in route choices (panels B, C, and E).

I use two empirical methods to shed light on how model parameters are identified. The first is to show numerically that the estimation procedure can recover the parameters using simulated data for various sets of random parameter. Appendix Table A8 shows that estimated parameters track the true parameters closely. The estimated slope between the underlying parameter and the GMM estimate is close to 1, and the R^2 is very high, in all cases except for the inner nest (departure time) logit parameter σ, for which the slope is above 1 and statistically significantly different from zero, yet noisier.

The second exercise is to compute the sensitivity measure from Andrews et al. (2017). The (scaled) sensitivity matrix Λ captures how estimated parameters depend on the different moments of the data. Specifically, each entry $\Lambda_{\gamma k}$ measures the impact of a standard deviation increase in moment g^k, $1 \leq k \leq 69$, on estimated parameter $\gamma \in \{\alpha, \beta_E, \beta_L, \sigma, \mu, p\}$. The results generally confirm the intuitions described earlier,

choice decision in this parsimonious way.

while also emphasizing that parameters are jointly estimated, with contributions from several moments. As expected, the early schedule cost β_E depends most strongly on departure time moments in the early ramp part of the departure time problem (departure times between -1.5 and -0.5 in Appendix Figure A8, Panel B). However, the area moments also have important contributions (column 2 in Appendix Table A9). As expected, the value of time driving α is most strongly identified by the area moments (column 1 in Appendix Table A9). The probability to respond, p, is affected more strongly by the area moments than by the departure time heterogeneity moments (column 5 in Appendix Table A9).

Overall, the structural model offers a good fit to how commuters responded to the congestion charge experiments. The results indicate that commuters are fairly flexible to change their schedules by leaving earlier locally around their ideal departure time, relative to how much they value time spent driving. However, in order to quantify the externalities involved in peak-hour traffic congestion, and the welfare impacts of congestion mitigating policies, it is also necessary to know how traffic responds to aggregate changes in driving patterns.

7 The Road Traffic Congestion Technology

Each additional vehicle on the road leads to slower road speeds. I now quantify this external cost using all the GPS trip data collected during the study, and real-time Google Maps driving time data collected during the same period on a set of routes in Bangalore.[41]

The traditional approach to studying this relationship in transportation engineering has been to analyze road or highway segments in developed countries. Empirical estimates vary considerably, in part due to the variation in the specific roads considered.[42] From an economic perspective, we are interested in full trips, not only road segments or small areas. Indeed, commuters make decisions over trips, and trips cover large areas and different types of roads. There are few empirical studies that measure travel time costs and external costs at the trip level. Geroliminis and Daganzo (2008) use GPS taxi trip data from 140 taxis for one month in Yokohama, Japan, and show that average trip speed declines strongly at times of the day when many trips are taking place. Akbar and Duranton (2017) measure road traffic volume from around 20,000 motorized trips recorded in a household transportation survey in Bogotá, Colombia, and travel times from

[41] Driving also imposes other external costs, such as increases in pollution emissions, pollution exposure (which is related to traffic speeds), and accidents. Here I am only considering the impact on higher (and less reliable) driving times.

[42] A commonly used functional form to describe travel time T as a function of incoming flow V is given by $T = T_f \cdot \left(1 + a \cdot (V/V_k)^b\right)$, where T_f is time under free-flow, and V_k is the maximum road capacity. The parameter values for a and b vary considerably. For example, the Bureau of Public Roads (BPR) and the updated BPR functions use $a = 0.15$, $b = 4$ and $a \in [0.05, 0.2]$, $b = 10$, respectively. See section 3.3.2 in Small et al. (2007) for a review of estimated and postulated functional forms.

real-time Google Maps data collected several years later. They establish a much smaller elasticity of travel time with respect to the volume of traffic. Their results suggest that there are fundamental differences in city-wide road technology in Bogotá relative to cities in richer countries. One potential concern with their approach is attenuation bias due to survey survey recall bias, which can lead to mis-measurement in the traffic volume measure. For example, survey respondents may omit trips or only report imprecise departure and arrival times. In this paper, I use precise GPS data, contemporaneous real-time Google Maps data, and a larger sample of trips than in the two previous papers. I show at the end of this section that the elasticity in Bogotá estimated by Akbar and Duranton (2017) is very similar and slightly smaller than what I find in Bangalore.

To measure the *quantity* of driving, I rely on 117,527 trips coded from GPS data from 1,747 app users, covering 185 calendar dates and 44,034 user-days with travel information.[43] (This sample includes the experimental sample, as well as other study participants who used the smartphone app for shorter periods of time and were not included in the experiment.)

For road *speeds*, I use two different data sources that give very similar results. My main data source is Google Maps travel delay data collected on 28 routes in the study area over the same calendar period, at 20 minute intervals.[44] I also compute trip-level travel delay directly from the GPS data.[45] The two measures track each other exceptionally well at the level of departure time (column 3 in Table 8 and panel A in Appendix Figure A9).

I use a simple empirical specification to measure the impact of traffic volumes on travel delay. (I discuss possible threats to causal inference below.) In order to reduce measurement error in the dependent variable, I summarize traffic volume and travel delay along two dimensions: trip departure time, and calendar date. In the first case, the average departure rate at a given time of the day measures *inflows* into the urban road network, so this approach is similar to classic transport engineering estimates, except that here I consider results for a large urban area.[46] The second approach considers the total travel on a given calendar date. In

[43] I restrict the sample to trips longer than 2 km. Shorter trips have higher travel delay, possibly because of higher likelihood of walking trips. Results are almost identical including all trips.

[44] Travel delay is the inverse of driving speed, measuring the number of minutes necessary to cover 1 kilometer, on average. To obtain it, for each route I divide driving time (in minutes) by the route path length (in kilometers).

[45] Trips in the GPS data may have considerably more noise, for example due to short stops along the way, or errors in trip classification. I use medians to summarize this data in order to limit the influence of outliers. The sample is all weekday trips shorter than 2km, without stops along the way. To avoid circuitous trips, I restrict to trips with diameter to total length ratio above 0.6 (the 25th percentile). For each departure time, I compute the median delay of all trips starting around that departure time (weighting each trip using an Epanechnikov kernel with bandwidth 20 minutes around the reference departure time). In addition, Appendix Table A10 reports results from quantile (median) regressions.

[46] It is plausible that travel conditions depend on the history of inflows, not only on contemporaneous inflow. One way to model this is to measure the number or *density* of vehicles on the road at any given time. This approach gives very similar results, see panel B in Appendix Figure A9. Intuitively, the two variables are strongly correlated, because trips are short relative to the scale of peak/off-peak fluctuation. In addition, models that includes lags will fit the data marginally better. Indeed, in panel A of Figure 3 the travel delay at 9 am – following a large inflow – is slightly lower than predicted by the

both cases, I normalize the dependent (volume) variable to mean 1, so the results are directly comparable. I am not able to distinguish between the impact of motorcycles and cars, and cannot account for vehicle occupancy.[47] Results should be interpreted as the average effect along these dimensions.

Travel delay is well explained by a linear function of traffic volume, and results are similar using variation within day and across calendar dates. Figure 3 shows the main results in graphical form. Panel A shows results at the departure time level (collapsing over all weekdays in the data), and plots the results for all departure times at 30 minute intervals, while panel B shows results by calendar date. Columns 1, 2 and 4 in Table 8 show the same results in regression format. For departure times, an increase in the number of vehicles equal to 10% of the mean is associated with an increase of 0.106 minutes per kilometer higher travel times (column 1). The relationship is close to linear, and I can reject at the 95% level an exponent of 1.18 (column 2). The relationship is similar and slightly shallower across calendar dates at 0.097 minutes per kilometer (column 2). The difference may partly reflect attenuation bias due to measurement error in traffic volumes along calendar dates.

These results imply that every additional (average length) trip departing increases the aggregate driving time of everyone else on average by approximately 4.2 minutes for a 7 am departure time, and by approximately 17 minutes for trips departing at the morning peak (9 am) or evening peak (7 pm). (For reference, the average trip duration is 33 minutes.) To derive this result, note that the impact on aggregate driving time is equal to the traffic volume, times the marginal impact on travel delay, and times the average trip length, namely $\tilde{Q}(h) \frac{\partial T(h)}{\partial \tilde{Q}(h)} \cdot \overline{KM}$. The first two terms form a semi-elasticity, so this social cost calculation does not depend on the scaling of traffic volume $\tilde{Q}$. In other words, for a representative sample the in-sample calculation is consistent for the population calculation. The empirical results show that $\frac{\partial T(h)}{\partial \tilde{Q}(h)} = 1.06\,\mathrm{min/km}$ (for any h), and the average trip length is 8.0 kilometers, which gives an effect of $8.5 \cdot \tilde{Q}(h)$ minutes, where $\tilde{Q}(h)$ is the relative traffic volume at h. Figure 3 shows that $\tilde{Q}(7\,\mathrm{am}) \approx 0.5$ and $\tilde{Q}(9\,\mathrm{am}) \approx \tilde{Q}(7\,\mathrm{pm}) \approx 2$, which gives the figures cited above.

One potential concern is whether the data used here is representative for Bangalore. It is reassuring that the results from two completely different data sets on speeds (GPS data and Google Maps data) give similar results: column 3 in Table 8 shows that the slope between the two variables is very close to one. The other concern is whether the traffic quantity measure is not representative in a way that is correlated with the pattern of congestion. In principle, it is possible that the survey team recruited disproportionately more (or fewer) respondents during peak hours, which may bias the results. However, the link between recruitment

linear relationship, yet delay continues to rise after 9 am despite slightly decreasing inflows. Here, I use the more parsimonious functional relationship.

[47]The share of trips made by car is roughly constant throughout the day, at around a third.

time and average departure time is very weak. Indeed, the R squared of a regression of trip departure time in the morning on morning recruiting time is below 4%, and below 2% for the evening.[48]

Interpreting these results as the causal impact of driving on external driving time costs raises several potential concerns. One issue arises if different types of drivers systematically travel at different times, for example if inherently slower drivers are more likely to travel during peak hours. A related concern is if peak-hour and off-peak trips differ in some dimension correlated with speed, such as trip length. In principle, these issues could even affect the Google Maps travel delay estimates, if Google's algorithms do not correct for such biases. To address these concerns, in Appendix Table A10 I run trip-level quantile (median) regressions of trip delay on the traffic volume at the trip departure time, where I control for trip length and commuter fixed effects. The results are broadly similar and somewhat smaller than those in Table 8. More generally, any factor correlated with the within-day or across-date distribution of traffic volume, which also directly impacts driving times, is a potential omitted variable. For example, anticipated weather and road network shocks (e.g. construction, closures) may bias the estimates downwards. These factors are unlikely to be a major concern in this setting. First, weather during the study period was very stable, and there were no major road network shocks. Secondly, the within-day results are less likely to be significantly biased by this type of factors, because weather and road network shocks tend to last longer. Higher pedestrians flows during peak hours may bias our results *upwards*, if pedestrians interact with and slow down incoming traffic. Anecdotally, drivers in Bangalore tend to not slow down considerably when pedestrians cross the road.

The results in Bangalore are very similar to those reported by Akbar and Duranton (2017) in Bogotá. Appendix Figure A9 panel D compares the log-log curves in the two cities. The curve for Bogotá is slightly lower and the maximum elasticities in Bangalore and Bogotá are 0.33 and 0.25, respectively. The curve in Bogotá also becomes flat for high values of traffic volume. There may be two reasons for this. First, note that the local linear fit in Bangalore also has a slightly lower slope for high volumes; this may be due to the linear contemporaneous road technology specification, which omits traffic volume lags. In particular, traffic volume rises quickly in the morning, and speed grows slightly slower (only to continue to grow even past the peak in traffic volume); this tends to attenuate the relationship between traffic volume and travel times. Another potential reason for the more pronounced flat region in Akbar and Duranton (2017) is that survey respondents are likely to give typical departure times that underestimate the variability in departure times, which tends to overestimate peak-hour volumes. Indeed, Zhao et al. (2015) document exactly this phenomenon by comparing survey data with precise GPS travel data collected with a smartphone app on the same sample in Singapore.

[48]Panel C in Appendix Figure A9 shows graphically that the recruit time and trip departure time distributions are very different.

Previous engineering studies on road segments show that travel time responds strongly and convexly to traffic inflows. Intuitively, one would expect this relationship to be even stronger in a high congested city, such as Bangalore. In fact, I provided evidence for a shallower and linear relationship. The slope is several times shallower than the slope identified based on taxi trips in Geroliminis and Daganzo (2008). There are several potential reasons why the road technology may be different in Bangalore. The high ratio of motorcycles may render traffic more fluid; however, using the GPS data I find that motorcycles are faster only during the night, and have a similar speed as cars during the day. Another hypothesis is that drivers switch to side streets during peak hours, thus avoiding traffic build-ups on main thoroughfares (Akbar and Duranton, 2017). Further, the driving style in cities like Bangalore, where anecdotally vehicles are driven close to each other, may attenuate traffic jams. (However, in principle this type of driving could also make jams worse.) Another structural difference is the smaller number of automatic traffic signals compared to cities like Yokohama, Japan, and potentially higher reliance on traffic police agents, which may also affect the bottleneck properties at certain key junctures. I consider average travel times over several months, which are likely the relevant measure to measure the expected externality. This is similar to Akbar and Duranton (2017) and unlike Geroliminis and Daganzo (2008), who use instantaneous relationships.[49]

In this section, I provided new evidence that despite high levels of traffic congestion in Bangalore, the shape of the road technology externality is moderate and linear throughout the distribution of traffic volume. Equipped with this estimate, we are now in a position to quantify the inefficiency involved in the peak-hour traffic equilibrium.

8 Policy Simulations

In this section I quantify welfare and the inefficiency in the no-toll equilibrium, and explore how these numbers depend on preferences and on the road technology.

Commuters make departure time decisions based on their own travel time and schedule costs, and have some flexibility to adjust departure times to avoid congestion, as shown in the experimental results. However, their decisions also affect the other traffic participants by increasing delays at the times when they travel, an effect mediated through the road technology that was quantified in the previous section. Moreover, other commuters adjust to the increase in congestion, and this has either positive or negative first order impacts on welfare, as the envelope theorem does not hold for welfare at an inefficient equilibrium. For example, for the same level of congestion, traveling after the peak-hour may have a higher externality because it induces

[49] In principle, it is possible that the instantaneous relationship is convex, and the peak-hour is realized at slightly different times on different days, which would smooth out the relationship. However, the relationship in Figure 3 looks similar when using a single day of data (not shown).

other commuters to switch to earlier (more congested) travel times.

I now study these interactions and their welfare consequences using a simulation model of the city-wide road traffic equilibrium. I use the model to solve for decentralized Nash equilibria without or with departure time charges, I then compute the marginal social cost of departing at a certain time around a given Nash equilibrium, and I finally solve for the social optimum and compare the improvement relative to the decentralized equilibrium using various benchmarks. In any equilibrium with charges, I assume that tax revenue is transferred back to commuters lump-sum.

The model parameters are derived entirely from demand and road technology estimates presented in previous sections. At the same time, there are several important limitations of this approach. I continue to abstract from the extensive margin decision of whether to travel using a private vehicle. Thus, the analysis here pertains specifically to the within-day inefficiency due to commuters wanting to travel at similar times. I also do not take into account longer term preferences and adjustments, which may be different from the short-term responses measured in the experiment. As in the road technology estimation, I do not distinguish between the externalities generated by motorcycles and cars, although in practice the latter is likely to be higher. Finally, this analysis ignores other traffic, including trips that are not between home and work, bus passengers (who would also benefit from reductions in travel times), taxis, bus and truck traffic (which may respond differently to similar congestion charges and may affect traffic differently). I also do not measure in these calculation other important social costs of congestion, such as pollution generation and pollution exposure.

The simulation environment is populated with N agents. Each agent has a single route and chooses a morning departure time according to multinomial logit probabilities, using $\hat{\sigma}$ as estimated previously. Each simulation agent is a copy of a real study participant, with the same route length and preferences as estimated in section 6, with a fixed ideal arrival time randomly drawn from the distribution estimated for that agent. In practice, I replicate each real commuter and draw 120 ideal arrival times for each copy, for a total of $N = 36,960$ simulation agents.

I use an asynchronous logit best-response dynamic to compute Nash equilibria. Given (fixed) congestion charges that depend on the departure time, and an initial travel time profile, each period a 1% random sample of agents re-compute their choice probabilities,[50] and then the travel time is updated given the aggregate volume at each departure time (integrated choice probability over all simulation agents). The simulation stops when every agent is close to best-responding, namely when the ℓ^2-norm of changes in choice probability, averaged over the entire population, is below a certain threshold. This procedure leads to fast convergence to

[50] Travel time uncertainty is parametrized based on the mean travel time, as was done for structural estimation. Travel time is log-normal distributed, and travel delay standard deviation is quadratic in the mean. See Appendix Figure A6.

equilibrium; indeed, the choice probability norm roughly halves after half of the population updates (every 50 periods), and it takes around 7 revisions per capita to reach equilibrium. Moreover, this dynamic has a natural interpretation in terms of commuters revising their actions periodically. In practice, the simulation finds a unique equilibrium independently of starting conditions, which is consistent with travel at various times being strategic substitutes due to congestion.

The marginal social cost imposed by a commuter leaving at a certain departure time h_D should be calculated allowing other commuters to adjust.[51] Indeed, other commuters may change their departure times in response to the increase in congestion, which will decrease their costs and may have either positive or negative spillovers.[52] This effect is quantitatively meaningful; for example, the partial equilibrium social welfare cost of an additional departure at 9:40 am, starting from the Nash equilibrium, is Rs. -408.6, compared to Rs. -352.0 after recomputing the equilibrium with the fixed departure at 9:40 am. Moreover, for the same level of congestion, the marginal social cost depends on the slope of congestion around that point. Figure 4 shows in blue (right axis) that the marginal social cost is higher after the peak. This happens because displaced commuters tend to leave earlier (because $\beta_E < \beta_L$) and when the additional commuter departs after the peak, this switching leads to even more congestion at earlier times.

The social optimum is a Nash equilibrium with departure time Pigou charges.[53] It has the following fixed point property: charges at departure time h_D equal the marginal social cost of an additional commuter at h_D. To find this fixed point, I use a lazy adjustment dynamic for charges. The starting point is the Nash equilibrium, and for each iteration I compute the marginal social cost and update charges at each departure time with a $1/3$ weight on the new marginal social cost and $2/3$ on the current charge. This procedure converges in around 15 iterations with precision Rs. 0.1 for welfare.

The social optimum leads to small but notable improvements in travel times. Figure 4 shows the travel delay under the decentralized unpriced equilibrium and under the social optimum. The social optimum has a lower peak, and more commuters departing early, between 5:30 am and 7:45 am. However, the distance between the two travel delay profiles is not very large; at the peak, travel delay improves by 0.14 minutes per kilometer, which translates to 0.9 and 2.3 minutes faster travel time for the commuters at the 25th and 75th percentile of route distance. One reason for this moderate difference is that moving some people away from the peak does not have an outsized effect on congestion under the linear road technology. The social marginal

[51] Arnott et al. (1993) make the same point in their model with identical agents, where the stark implication is that MSC does not depend on departure time (as in equilibrium all agents are indifferent). The more general point also applies in this setting, where agents differ.

[52] Computing the marginal social cost thus requires computing a new Nash equilibrium for each departure time.

[53] I assume that the social planner knows individual preferences but does not observe the exact realization of the random utility shocks. In this case, where also all commuters have the same externality conditional on departure time, the planner can implement the social optimum with departure time congestion charges.

cost function is also drawn (right Y axis). For the same level of congestion, social marginal cost is higher after the peak, and yet there is almost no aggregate difference between the Nash and social optimum on that side of the graph. This happens because the cost of departing later β_L is very high. Despite large congestion charges, individual changes under the social optimum are small. The average change in average departure time (conditional on ideal arrival time) is leaving 3 minutes earlier, and the 25th and 75th percentiles are 4.1 minutes earlier and half a minute later. These number are within the range of experimental responses to the departure time policy. Hence, these counterfactual results do not rely on extrapolating based on the functional form in preferences.

Table 9 quantifies the effects on travel time and on welfare. The social optimum leads to a reduction of 1.04 minutes in expected travel time, from a base under Nash of 38.7 minutes.[54] This represents a 2.7% improvement relative to the Nash equilibrium, or a 6.8% improvement when considering only travel time above free-flow. (Free-flow is defined as a speed of 2.14 minutes per kilometer, which is the intercept in Table 8.)

The improvement in welfare under the social optimum are an order of magnitude smaller. In other words, the travel time benefits are nearly offset by schedule costs incurred by commuters who are induced to travel at privately inconvenient times. Welfare is Rs. 4.5 per commuter per morning higher under the optimum (7 US cents), from a total trip cost of around Rs. 773, which represents a roughly half percentage improvement. Relative to free-flow, the improvement is only 1.3%. Moreover, to achieve the social optimum, commuters pay on average Rs. 267.3 in charges, or 35% of their average private cost. For this exercise, I assumed that charges are a costless transfer, whereas in reality policy enforcement and attention costs may be important. It is likely that real-world, more forceful policies that attempt to cap peak-hour congestion may lower welfare. This framework can be used to quantify these effects.

The road technology plays a key role for these results. Indeed, the welfare gain would be higher and would depend more on preferences if travel time was convex in traffic volume. Figure 5 shows this by plotting the improvement of going from the unpriced equilibrium to the social optimum, for travel times (panel A) and for welfare (panel B). The black and red lines denote the current (linear) and counterfactual (third power) road technology, while preferences vary along the X axis. The welfare gains are an order of magnitude higher with the power road technology, and range between 3.5% and 5.5%, relative to $0 - 0.5\%$ with the linear technology.

Overall, I have shown using policy simulations grounded in demand and road technology estimates

[54] Note that the average route length is not evenly distributed across departure time, with commuters who travel far slightly more likely to depart early. Moreover, under the social optimum this effect is slightly stronger, which contributes to lower average travel time than suggested by Figure 4 alone.

that the social inefficiency due to departure times is likely small. This result highlights the importance of measuring and considering the (schedule) costs of a policy that attempts to clear up the congestion peak-hour. An important reason for these findings is the size of the road externality, and especially its linearity, which implies that even for high levels of congestion the travel time benefit of removing a commuter from the peak is the same as for lower levels. By consequence, road traffic congestion does not warrant intervention through corrective taxation solely for the reason of departure time inefficiency.

9 Conclusion

Reducing traffic congestion has significant benefits in terms of the value that commuters put on the time they spend driving; it can also lead to improved subjective well-being (Anderson et al., 2016). This makes it tempting to only consider these benefits when thinking about traffic policies, and in particular about policies designed to reduce peak-hour congestion. However, it is also important to take into account the costs of disruption to commuter schedules.

In this paper, I collected new data on travel behavior and implemented a field experiment motivated by a model of travel demand to study both sides of the peak-hour traffic equilibrium. Estimating a model of the morning commute decision using experimental variation, I find that the cost of arriving earlier than ideally desired is around 4 times smaller than the value of time spend driving. To put this in context, as a first approximation this means that a commuter facing a one hour expected drive time would prefer to leave one hour earlier, if this reduced the expected drive time to less than 45 minutes.

Surprisingly, given high levels of traffic congestion in Bangalore, I find a moderate road traffic externality, and a *linear* effect of traffic volume of travel times, including for high values of traffic volume. This result is robust to using different data sources and analyzing this relationship across days or within day and across departure times. These estimates are smaller than a previous study in Yokohama, Japan (Geroliminis and Daganzo, 2008), and they rule out hyper-congestion. Differences in driving style and the density of traffic control measures (traffic lights) are possible explanations for the discrepancy between the two settings.

Putting demand and road technology estimates together, I calculate the equilibrium optimal congestion charge for the morning peak-hour. I find that relative to the decentralized equilibrium without charges, the social optimum allocation leads to notable travel time benefits of around 1.2 minutes per trip from a base of 36.5 minutes (which is 30% of the improvement that can be achieved by spreading all commuters evenly between 5 am and 12 pm). However, the welfare gains from optimal charging are small, as the travel time benefits are almost fully offset by the schedule costs of making commuters travel at inconvenient times.

The elasticities of travel at different times may differ from those estimated here in the long-run and in the presence of a city-wide policy. For example, firms may adjust by providing more flexible schedules, allowing commuters to more easily change their travel timings. As in other large cities marred by high congestion, some large companies in Bangalore are already implementing this type of flexible work-hours policies (Merugu et al., 2009). Moreover, around 20% of the sample in this study are self-employed and may already have higher autonomy in deciding their own schedules. Finally, it is worth noting that the welfare impacts of firm-level work-hours changes is ambiguous, due to complementarities *between* firms of having similar work hours (Henderson, 1981).

The cost and road technology estimates in this paper are also useful for thinking about the extensive margin of making trips with private vehicles, which was held constant throughout this analysis. Indeed, the same methods can be used to compute the social marginal cost of adding a commuter at a certain departure time,[55] and thus calculating the optimal congestion charge. The results will depend on the elasticity of making a trip with respect to generalized travel cost. Given the small share of metro travel in Bangalore, this effect would likely mostly go through canceling non-essential trips, working from home, or switching to bus travel.[56] Moreover, the same framework can be used to include additional externalities, such as the air pollution that drivers generate, and the impact that longer driving times have on exposure to air pollution, which is high on urban roads.

This paper argues that the peak-hour traffic congestion equilibrium is close to efficiency from the point of view of travel speed externalities. This does not imply that there are no welfare enhancing policies to ease traffic congestion. Pricing the extensive margin may be a viable policy, either done directly or through taxes on gasoline or private vehicle ownership. Road infrastructure investment – including investments to make road network flows more efficient – may currently be at an inefficient level. Of course, developing viable public transport options may also contribute to lower congestion and more convenient travel. However, the results in this paper do put into focus the welfare costs that well-intended traffic control policies may have on commuters affected by such charges or restrictions.

[55] The marginal social cost will not be exactly the same, because some commuters who are at the margin may cancel their trip in response to an increase in congestion.

[56] There are two reasons why this margin was not studied experimentally here. The first is a measurement issue, as commuters would have strong incentives to leave their smartphones at home given monetary incentives to reduce the number of trips. It is possible to solve this problem by installing a GPS device in the private vehicle – as Singapore plans to implement its next generation Electronic Road Pricing policy, and as in the experiment studied in Martin and Thornton (2017). The second reasons is that extensive margin changes will plausibly take longer, as commuters need to find alternate travel arrangements or substitutes for canceling unessential trips. Indeed, Martin and Thornton (2017) do not find any reduced-form impact of distance-based congestion charges on trip extensive margin, even after two months.

References

Akbar, P. A. and Duranton, G. (2017). Measuring the cost of congestion in highly congested city: Bogotá. *CAF Development Bank of Latin America Working Paper 2017/04*.

Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review*, 104(9):2763–96.

Anderson, M. L., Lu, F., Zhang, Y., Yang, J., and Qin, P. (2016). Superstitions, street traffic, and subjective well-being. *Journal of Public Economics*, 142:1 – 10.

Andrews, I., Gentzkow, M., and Shapiro, J. M. (2017). Measuring the sensitivity of parameter estimates to estimation moments*. *The Quarterly Journal of Economics*, 132(4):1553–1592.

Arnott, R., de Palma, A., and Lindsey, R. (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *The American Economic Review*, 83(1):161–179.

Beckmann, M. J., McGuire, C. B.Winsten, C. B., and Yale University. Cowles Foundation for Research in, E. (1956). *Studies in the economics of transportation.* New Haven, Published for the Cowles Commission for Research in Economics by Yale University Press, 1956 [cl955].

Ben-Akiva, M., McFadden, D., and Train, K. (2016). Foundations of stated preference elicitation consumer behavior and choice-based conjoint analysis. *Working Paper*.

Ben-Akiva M., B. M. (2003). *Handbook of Transportation Science*, chapter Discrete Choice Models with Applications to Departure Time and Route Choice. Springer.

Bento, A., Roth, K., and Waxman, A. (2017). Avoiding traffic congestion externalities? the value of urgency. *Working Paper*.

Chu, X. (1995). Endogenous trip scheduling: The henderson approach reformulated and compared with the vickrey approach. *Journal of Urban Economics*, 37(3):324 – 343.

Davis, L. (2008). The effect of driving restrictions on air quality in mexico city. *Journal of Political Economy*, 116(1):38–81.

Diamond, P. A. (1973). Consumption externalities and imperfect corrective pricing. *The Bell Journal of Economics and Management Science*, 4(2):526–538.

Geroliminis, N. and Daganzo, C. F. (2008). Existence of urban-scale macroscopic fundamental diagrams: Some experimental findings. *Transportation Research Part B: Methodological*, 42(9):759 – 770.

Gibson, M. and Carnovale, M. (2015). The effects of road pricing on driver behavior and air pollution. *Journal of Urban Economics*, 89(Supplement C):62 – 73.

Gu, Y., Deakin, E., and Long, Y. (2017). The effects of driving restrictions on travel behavior evidence from beijing. *Journal of Urban Economics*, 102(Supplement C):106 – 122.

Hall, J. D. (2016). Pareto improvements from lexus lanes: the effects of pricing a portion of the lanes on congested highways. *Working Paper*.

Hanna, R., Kreindler, G., and Olken, B. A. (2017). Citywide effects of high-occupancy vehicle restrictions: Evidence from three-in-one in jakarta. *Science*, 357(6346):89–93.

Henderson, J. (1974). Road congestion: A reconsideration of pricing theory. *Journal of Urban Economics*, 1(3):346 – 365.

Henderson, J. (1981). The economics of staggered work hours. *Journal of Urban Economics*, 9(3):349 – 364.

Karlström, A. and Franklin, J. P. (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(3):283 – 296. Stockholm Congestion Charging Trial.

Kreindler, G. (2016). Driving delhi? behavioural responses to driving restrictions. *Working Paper*.

Martin, L. A. and Thornton, S. (2017). Can road charges alleviate congestion? *Working Paper SSRN*.

Merugu, D., Prabhakar, B. S., and Rama, N. (2009). An incentive mechanism for decongesting the roads: A pilot program in bangalore. In *Proc. of ACM NetEcon Workshop*.

Noland, R. B. and Small, K. A. (1995). Travel-time uncertainty, departure time choice, and the cost of the morning commute. *working paper Institute of Transportation Studies, University of California, Irvine*.

OICA (2016). Vehicle sales in the bric countries from 2005 to 2015 (in units). Technical report, International Organization of Motor Vehicle Manufacturers.

Papinski, D., Scott, D. M., and Doherty, S. T. (2009). Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based gps. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(4):347 – 358.

Prud'homme, R. and Bocarejo, J. P. (2005). The london congestion charge: a tentative economic appraisal. *Transport Policy*, 12(3):279 – 287.

Raux, C. (2005). Comments on "The London congestion charge: a tentative economic appraisal"; (Prud'homme and Bocajero, 2005). *Transport Policy*, 12(halshs-00067920).

Small, K. (1982). The scheduling of consumer activities: Work trips. *American Economic Review*, 72(3):467–79.

Small, K., Verhoef, E., and Lindsey, R. (2007). *The Economics of Urban Transportation*. Taylor & Francis.

Small, K. A., Winston, C., and Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4):1367–1382.

TfL (2006). Central london congestion charging impacts monitoring fourth annual report. Technical report, Transport for London.

Tillema, T., Ben-Elia, E., Ettema, D., and van Delden, J. (2013). Charging versus rewarding: A comparison of road-pricing and rewarding peak avoidance in the netherlands. *Transport Policy*, 26(Supplement C):4 – 14.

van den Berg, V. and Verhoef, E. T. (2011). Winning or losing from dynamic bottleneck congestion pricing?: The distributional effects of road pricing with heterogeneity in values of time and schedule delay. *Journal of Public Economics*, 95(7):983 – 992.

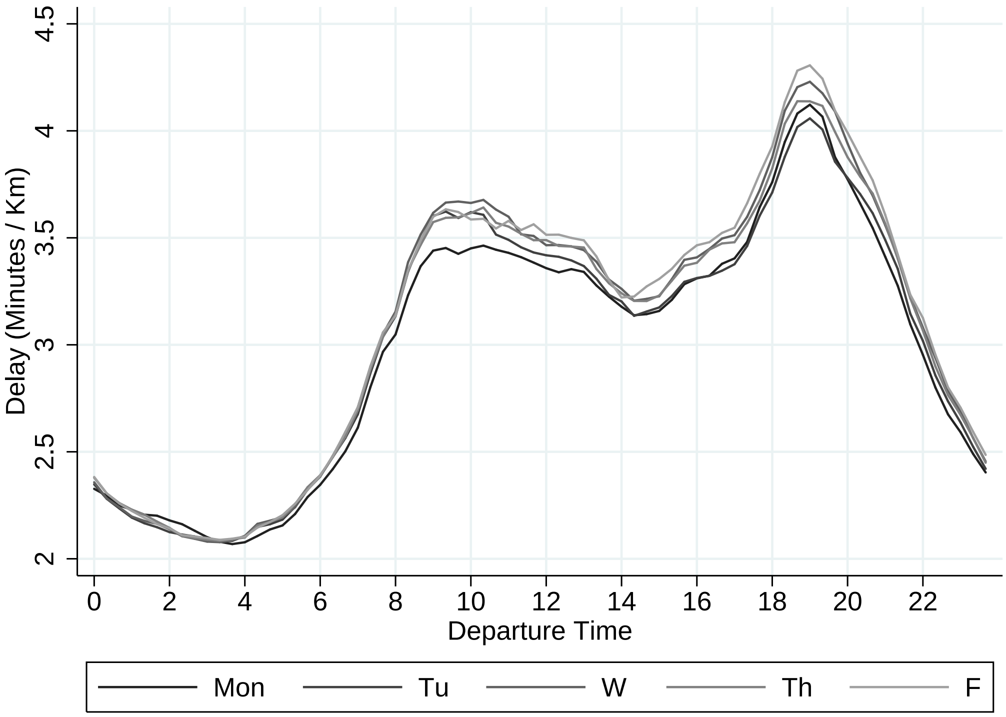
Vickrey, W. S. (1969). Congestion theory and transport investment. *The American Economic Review*, 59(2):251–260.

Yang, H. and Meng, Q. (1998). Departure time, route choice and congestion toll in a queuing network with elastic demand. *Transportation Research Part B: Methodological*, 32(4):247 – 260.

Zhao, F., Pereira, F. C., Ball, R., Kim, Y., Han, Y., Zegras, C., and Ben-Akiva, M. (2015). Exploratory analysis of a smartphone-based travel survey in singapore. *Transportation Research Record: Journal of the Transportation Research Board*, 2494:45–56.

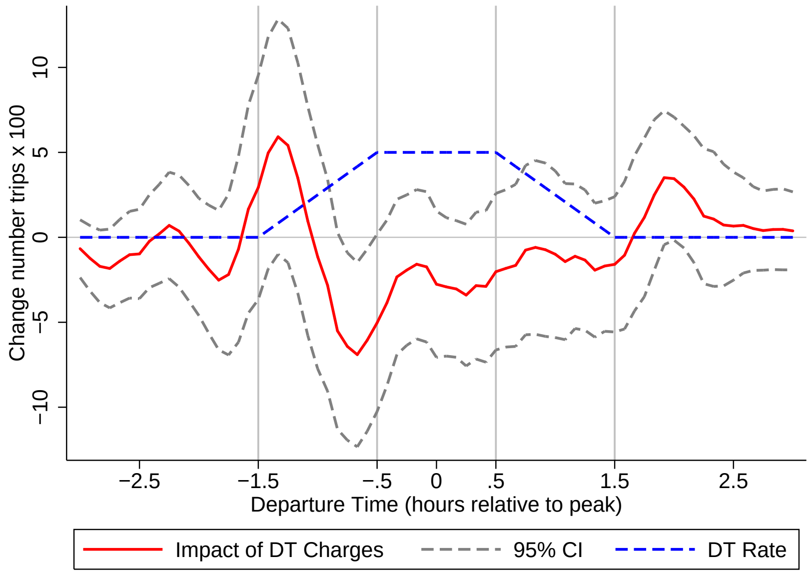
Figures

Figure 1: Average Predicted Travel Delay in the Study Region in Bangalore

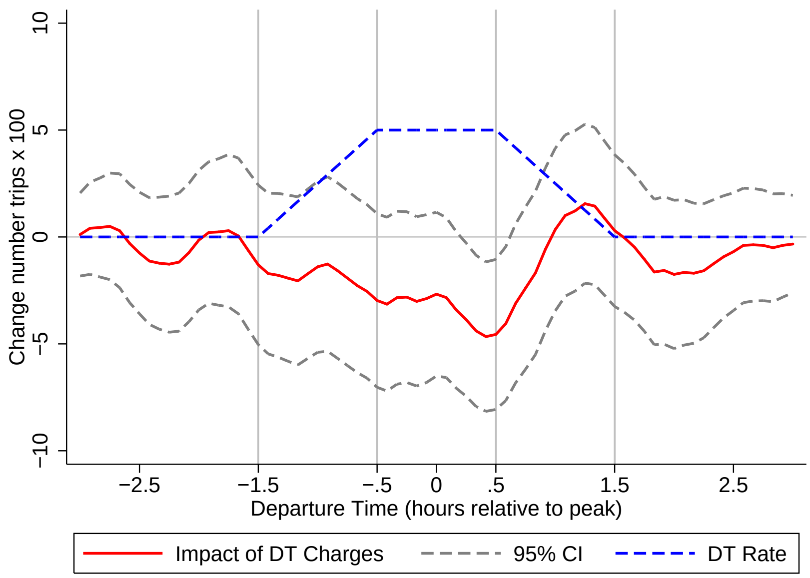


Notes: This graphs plots the average predicted travel delay on 28 major routes across the study area of South Bangalore, by day of the week. *Travel delay* is defined as the number of minutes to cover one kilometer, i.e. the inverse of speed. (A travel delay of 2 minutes per kilometer corresponds to 18.6 miles per hour.) The travel time and route length data is obtained from the Google Maps API. For each route, weekday and departure time (at 20 minute frequency) I queried the typical travel time under normal conditions, as predicted by Google.

Figure 2: Impact of Departure Time Charges on the Distribution of Departure Times



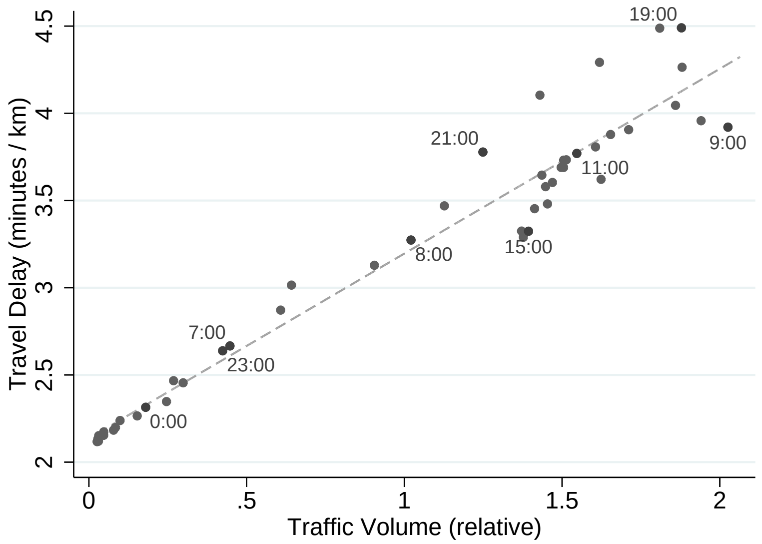
Panel (A) Morning Peak



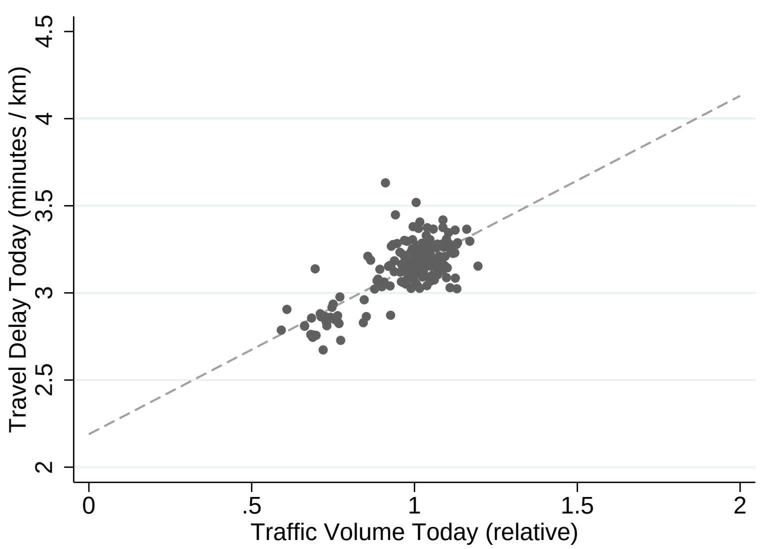
Panel (B) Evening Peak

Notes: These graphs plot the impact of departure time charges on the distribution of departure times, in the morning and evening. To construct this figure, for each commuter, day and departure time relative to the commuter's mid-point of the congestion charge rate profile, I compute the number of trips that start approximately at that time, using an Epanechnikov kernel with bandwidth 20 or 30 minutes for AM and PM, respectively. Then, for each relative departure time I run a difference-in-difference regression that identifies the impact of being in any of the charged sub-treatments (either High Rate or Low Rate) relative to being in the control or information sub-treatments. Each figure plots the charged sub-treatment times *Post* interaction coefficients, as well as pointwise 95% confidence intervals clustered at the individual level. The dimension of the Y axis is the number of trips at a given departure time, divided by 100. The sample is all non-holiday weekdays with good quality GPS data, excluding days outside Bangalore. In the post period, only the first or the last three weeks are included.

Figure 3: Road Technology: Travel Delay Linear in Traffic Volume



Panel (A) By Trip Departure Time



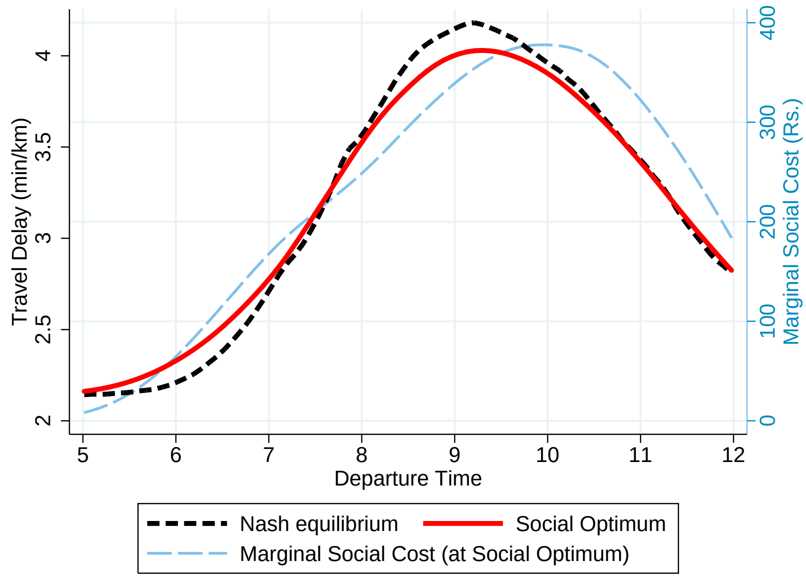
Panel (B) By Trip Calendar Date

Notes: These graphs show that travel delay is approximately linear in the volume of traffic.

Data. The volume measures are based on GPS data covering 117,527 trips from 1,747 app users across 185 days (including weekends). In panel A, all weekday trip departure times are aggregated at the departure time minute level, then smoothed using a local linear regression with Epanechnikov kernel with 10 minutes bandwidth, and finally normalized to have mean 1. In panel B, for each date I compute the number of trips per capita (using the number of app users that day), and again normalize this variable to have mean 1. The travel delay measures use Google Maps data collected over 28 routes in South Bangalore, every 20 minutes daily for 185 weekdays. In panel A, I compute the average delay over all weekdays and routes for each departure time, interpolating at the minute level. In panel B, I compute the average delay over all routes and departure times, for each day in the data.

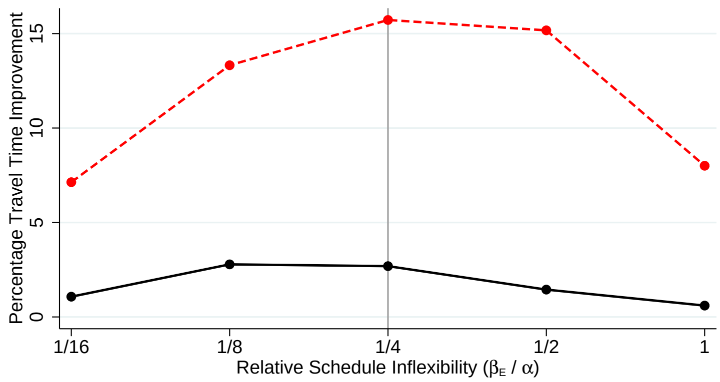
The OLS slopes for the two panels are 1.06 (0.06) and 0.97 (0.04), respectively. Table 8 reports the regression version of these relationships.

Figure 4: Unpriced Nash Equilibrium and Social Optimum (Policy Simulation)

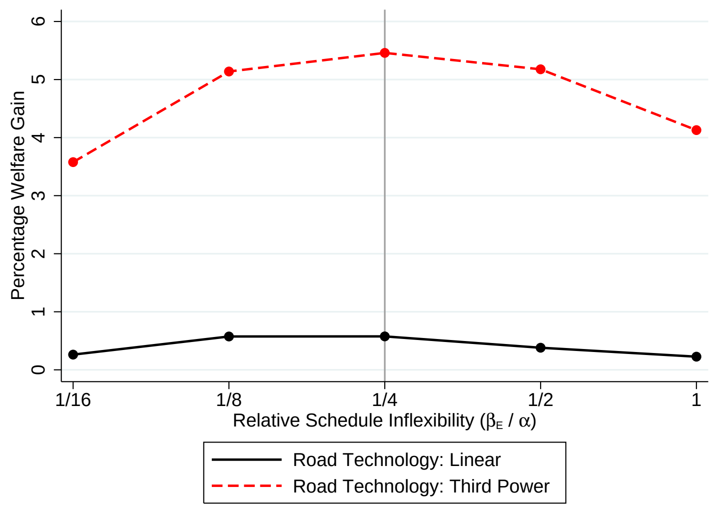


Notes: This graph shows the profile of travel delay under the simulated Nash equilibrium for morning departures (black, dashed line, left axis) and under the social optimum (red, solid line, left axis). The social optimum is a Nash equilibrium implemented with (equilibrium-consistent) social marginal charges in Rupees (blue, long dashed line, right axis). Under the Nash equilibrium, departure time choice probabilities are given by multinomial logit based on the travel time profile, and the profile itself is determined based on the road technology formula and aggregate traffic volume at each departure time. It is the end point of an asynchronous logit "best-response" dynamic whereby 1% of the population updates their choices each period (and travel delay updates in response). To compute the marginal social cost of adding a commuter at departure time h_D, I compute the new Nash equilibrium with that (fixed) addition and compute the change in total expected utility. The social optimum is a Nash equilibrium with the following fixed point property: departure time charges are exactly the marginal social cost of adding a commuter at that departure time. I compute the social optimum by updating departure time charges towards the marginal social costs until convergence.

Figure 5: Policy Simulations with other Preference and Road Technology Parameters



Panel (A) Change in Travel Time



Panel (B) Change in Welfare

Notes: This graph plots the improvement in average travel times and welfare of going from the no-toll equilibrium to the social optimum (as a percentage of the value in the no-toll equilibrium), for various assumptions on preference parameters and road technology. The black solid line corresponds to the linear road technology from Table 8 ($Delay = \lambda_0 + \lambda_1 Volume$ where $Volume$ is relative volume), while the red dashed line corresponds to road technology given by a thrid power ($Delay = \lambda_0 + \lambda_1 Volume^3$). The X axis reports the approximate ratio of early schedule cost (β_E) to value of time spent driving (α); at the center I report results using the estimated value of $\hat{\beta}_E/\hat{\alpha} = 0.28 \approx 1/4$; the other estimates vary this by a factor of 1/4, 1/2, 2 and 4. For each point, I compute the Nash (to toll) equilibrium and the social optimum for that road technology and preference parameters as described in Table 4, and plot the percentage impovement in the social optimum relative to the Nash.
.

Tables

Table 1: Descriptive Statistics about Travel Behavior

Panel A. Trip Characteristics	Median	Mean	Std. Dev.	10 Perc.	90 Perc.	Obs.
Total Number of Trips						51,164
Number of Trips per Day	2.85	3.15	[1.16]	1.90	4.85	497
Median trip duration (minutes)	24.50	27.38	[12.77]	15.05	42.60	497
Median trip length (Km.)	5.91	7.17	[4.66]	2.90	13.36	497
Panel B. Commute Destination Variability						
Regular Commuter		0.76				497
Frac. trips Home-Work, Work-Home	0.38	0.39	[0.21]	0.13	0.67	378
Frac. of trips Work-Work	0.03	0.06	[0.08]	0.00	0.15	378
Frac. of days present at Work	0.91	0.86	[0.16]	0.61	1.00	378
Panel C. Departure Time Variability						
(Standard Deviation of the Departure Time in hours)						
First Trip (AM)	1.27	1.24	[0.50]	0.52	1.85	496
Last Trip (PM)	1.72	1.71	[0.50]	1.06	2.34	497
First Home to Work Trip (AM)	0.48	0.62	[0.52]	0.15	1.28	332
Last Work to Home Trip (PM)	0.80	0.94	[0.62]	0.28	1.78	321

Notes: This table reports summary travel behavior statistics for the experimental sample of 497 commuters and 51,164 trips. For panel B, I classify each commuter as “regular” or “variable” based on a hybrid automatic and manual algorithm to identify common destinations (home or nighttime and work or daytime). In panel C, I compute the within-commuter variation in departure times for different classes of trips.

Table 2: Impact of Departure Time Charges on Daily Outcomes

	(1)	(2)	(3)
Time of Day	AM & PM	AM	PM
Commuter FE	X	X	X
Panel A. Total Shadow Rates Today			
High Rate × Post	-13.9**	-7.8**	-6.1*
	(6.1)	(3.8)	(3.4)
Low Rate × Post	-7.4	-2.8	-4.6
	(6.3)	(3.7)	(3.8)
Information × Post	-0.3	-0.2	-0.0
	(5.4)	(3.3)	(3.3)
Post	1.1	-0.9	2.1
	(4.9)	(2.9)	(3.1)
Observations	15,610	15,610	15,610
Control Mean	96.5	48.3	48.2
Panel B. Number of Trips Today			
High Rate × Post	-0.11	-0.04	-0.06
	(0.14)	(0.07)	(0.07)
Low Rate × Post	-0.06	-0.00	-0.07
	(0.14)	(0.07)	(0.07)
Information × Post	0.08	0.05	0.03
	(0.13)	(0.06)	(0.07)
Post	0.04	-0.01	0.06
	(0.11)	(0.06)	(0.06)
Observations	15,610	15,610	15,610
Control Mean	3.05	1.16	1.30

Notes: This table reports difference-in-difference impacts of the departure time sub-treatments on daily total shadow (per-Km) rates and total number of trips. In panel A, the outcome is the sum over all trips that day of the trip shadow rate. The shadow rate for a given trip is between 0 and 100 and is computed based on the trip departure time, the respondent's rate profile, and a peak rate of 100 for all respondents. (See Appendix Figure A3 for an example of rate profile.) In panel B, the outcome is the number of trips that day. The sample is all non-holiday weekdays with good quality GPS data, excluding days outside Bangalore. In the post period, only the first or the last three weeks are included. Column (2) and (3) restrict to the morning interval (7am-1pm) and to the evening interval (4-10pm), respectively. All specifications include respondent and study cycle fixed effects, and *Post* is an indicator for days during the experiment. The mean of the outcome variable in the control group during the experiment is reported for each specification. Standard errors in parentheses are clustered at the respondent level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 3: Impact of Departure Time Charges on Trip Shadow Rate

	(1)	(2)	(3)	(4)	(5)
Time of Day	AM & PM	AM	AM pre peak	PM	PM post peak
Commuter FE	X	X	X	X	X
Panel A. Full Sample					
High Rate × Post	-3.99***	-6.23***	-5.60	-3.12*	-5.40*
	(1.34)	(2.25)	(3.50)	(1.84)	(2.80)
Low Rate × Post	-1.85	-2.71	-3.59	-1.33	1.07
	(1.41)	(2.24)	(3.49)	(1.96)	(3.20)
Information × Post	-1.00	-2.32	-0.13	-0.41	0.40
	(1.06)	(1.81)	(2.70)	(1.63)	(2.53)
Post	-0.36	-0.81	-0.95	-0.35	-3.28
	(1.09)	(1.70)	(2.38)	(1.76)	(2.45)
Observations	43,776	16,764	7,592	18,468	7,899
Control Mean	31.64	41.81	46.98	37.21	44.29
Panel B. Regular Commuters, Home-Work and Work-Home Trips					
High Rate × Post	-4.97*	-7.48**	-10.12**	-1.54	-8.97
	(2.68)	(3.38)	(4.50)	(3.85)	(6.15)
Low Rate × Post	-4.02	-2.97	-9.12**	-5.38	-9.67
	(3.18)	(3.98)	(4.46)	(4.90)	(7.07)
Information × Post	-0.25	0.85	-2.73	-1.15	-1.97
	(2.06)	(2.96)	(3.32)	(3.42)	(4.60)
Post	-2.07	-2.73	-0.94	-3.86	-4.79
	(1.97)	(2.73)	(3.17)	(3.40)	(5.05)
Observations	11,895	5,789	3,782	4,862	2,113
Control Mean	37.08	44.59	44.91	38.16	42.28
Panel C. Variable Commuters, All Trips					
High Rate × Post	-1.99	-1.00	-4.41	-4.73	-2.10
	(2.90)	(5.05)	(10.32)	(4.26)	(7.64)
Low Rate × Post	0.47	-4.28	-6.15	-0.19	11.16
	(2.30)	(5.35)	(9.83)	(4.68)	(8.67)
Information × Post	-1.37	-3.38	-1.07	-1.22	-0.99
	(2.21)	(4.10)	(8.51)	(3.41)	(6.18)
Post	-1.05	-1.27	-0.17	-1.33	-2.45
	(2.22)	(3.72)	(8.07)	(3.55)	(5.81)
Observations	8,177	2,826	961	3,432	1,439
Control Mean	27.37	34.91	41.51	36.18	41.67

Notes: This table reports difference-in-difference impacts of the departure time sub-treatments on (per-Km) trip shadow rates. The shadow rate for a given trip is between 0 and 100 and is computed based on the trip departure time, the respondent's rate profile, and a peak rate of 100 for all respondents. The sample of users and days, and the specifications, are the same as in Table 2. In addition, columns (3) and (5) respectively restrict to trips before the morning peak (between 7 am and the mid-point of the AM rate profile), and after the evening peak (between the mid-point of the PM rate profile and 10 pm). Panel B restricts to regular commuters and direct trips between their home and work locations, and panel C restricts to variable commuters. Standard errors in parentheses are clustered at the respondent level . $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 4: Impact of Area Charges on Daily Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Time of Day	AM & PM	AM	PM	AM & PM	AM	PM
Commuter FE	X	X	X	X	X	X
Panel A. Total Shadow Charges Today						
Treated	-22.8***	-13.2***	-9.6***			
	(5.5)	(3.4)	(3.3)			
Treated in Week 1				-26.2***	-16.3***	-9.9*
				(8.3)	(5.0)	(5.3)
Treated in Week 4				-19.2*	-9.9*	-9.3*
				(10.1)	(6.0)	(5.5)
Observations	8,878	8,878	8,878	8,878	8,878	8,878
Control Mean	107.7	54.4	53.3	107.7	54.4	53.3
Panel B. Number of Trips Today						
Treated	0.17**	0.12**	0.06			
	(0.08)	(0.05)	(0.05)			
Treated in Week 1				0.06	0.04	0.01
				(0.13)	(0.07)	(0.09)
Treated in Week 4				0.30*	0.20**	0.10
				(0.16)	(0.08)	(0.10)
Observations	8,878	8,878	8,878	8,878	8,878	8,878
Control Mean	2.50	1.13	1.37	2.50	1.13	1.37

Notes: This table reports difference-in-difference impacts of the Area treatment on daily total shadow charges and total number of trips. In panel A, the outcome is the sum over all trips that day of the trip shadow charge. The shadow charge of a trip is equal to 100 if the trip intersects the respondent's congestion area, and 0 otherwise. In panel B, the outcome is the number of trips that day. The sample is all non-holiday weekdays with good quality GPS data, excluding days outside Bangalore. In the post period, all days except trial days are included. Column (2) and (5) restrict to the morning interval (7am-2pm), and columns (3) and (6) to the evening interval (2-10pm). The sample is restricted to the 243 participants in the Area treatment. The Treated dummy is equal to one in the week when the individual is treated (first and fourth week of the experiment for "early area" and "late are" sub-treatment commuters, respectively) and zero otherwise. All specifications include respondent and study cycle fixed effects. The mean of the outcome variable in the control group in weeks one and four of the experiment is reported for each specification. Standard errors in parentheses are clustered at the respondent level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 5: Impact of Area Charges on Trip Duration and Trip Shadow Charge

	(1)	(2)	(3)	(4)	(5)	(6)
Time of Day	AM & PM	AM	PM	AM & PM	AM	PM
Route FE	X	X	X	X	X	X
Panel A. Trip Shadow Charge						
Treated	-22.5***	-25.9***	-19.0***			
	(3.4)	(3.8)	(4.1)			
Treated in Week 1				-23.6***	-26.1***	-21.5***
				(4.9)	(5.0)	(6.3)
Treated in Week 4				-21.3***	-25.6***	-16.1**
				(6.4)	(7.4)	(7.0)
Observations	7,455	4,108	3,347	7,455	4,108	3,347
Control Mean	83.4	85.1	81.3	83.4	85.1	81.3
Panel B. Trip Duration (minutes)						
Treated	0.52	0.66	0.40			
	(0.72)	(0.74)	(1.29)			
Treated in Week 1				-1.14	-0.53	-1.55
				(0.97)	(1.04)	(1.74)
Treated in Week 4				2.49**	2.05	2.72
				(1.09)	(1.25)	(1.80)
Observations	7,455	4,108	3,347	7,455	4,108	3,347
Control Mean	40.81	39.15	42.73	40.81	39.15	42.73

Notes: This table reports difference-in-difference impacts of the Area treatment on trip shadow charge (panel A) and on trip duration (panel B). The shadow charge of a trip is equal to 100 if the trip intersects the respondent's congestion area, and 0 otherwise. The sample of users and days are the same as in Table 4, except that we restrict to regular commuters and direct home to work or work to home trips. All specifications include route fixed effects. Standard errors in parentheses are clustered at the respondent level . $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 6: Impact of Area Charge Sub-Treatments on Daily Outcomes

	(1)	(2)	(3)	(4)
Commuter FE	X	X	X	X
Panel A. Total Shadow Charges Today				
Treated	-22.8***	-21.4***	-25.0***	-23.6**
	(5.5)	(7.4)	(5.9)	(11.1)
Treated × High Rate		-3.0		
		(9.6)		
Treated × High Rate Day			5.4	
			(5.4)	
Treated × Short Detour				3.2
				(12.0)
Observations	8,878	8,878	8,878	5,417
Control Mean	107.7	107.7	107.7	110.6
Panel B. Number of Trips Today				
Treated	0.17**	0.09	0.24**	0.19
	(0.08)	(0.09)	(0.10)	(0.13)
Treated × High Rate		0.17		
		(0.14)		
Treated × High Rate Day			-0.16*	
			(0.10)	
Treated × Short Detour				-0.07
				(0.16)
Observations	8,878	8,878	8,878	5,417
Control Mean	2.50	2.50	2.50	2.53

Notes: This table reports difference-in-difference impacts of Area sub-treatments on daily total shadow charges and total number of trips. For outcome definitions and specifications see the notes for Table 4. The sample is the same as in Column (1) in Table 4. In column (4) the sample consists of the 148 Area participants for whom candidate areas included at least one with short detour (3-7 minutes) and at least one with long detour (7-14 minutes). (See section 4 for more details.) The specification in column (4) includes fixed effects for each day in the experiment. Standard errors in parentheses are clustered at the respondent level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table 7: Structural Parameter Estimates

(1)	(2)	(3)	(4)	(5)	(6)
Value of time α (Rs/hr)	Schedule cost early β_E (Rs/hr)	Logit inner σ (dep. time.)	Logit outer μ (route)	Probability to respond p	Ratio α/β_E
1,121.9	319.4	36.5	36.9	0.46	3.51
(318.7)	(134.5)	(65.4)	(9.3)	(0.13)	(1.11)

Notes: This table reports structural estimates of model parameters using two-step GMM with 69 moments. The first set of moments match the difference in difference average number of trips in each of 61 departure time bins (between -2.5 and $+2.5$ hours around the peak-hour, in 5 minute increments). Two moments match the variances of individual changes in shadow charges for trips between the morning rate profile peak and two hours earlier, in treatment and control. Two moments match the probability to intersect the congestion area with and without area charges. Four moments match the fraction of commuters whose sample frequency to intersect the congestion area lies in the middle third and top third of the unit interval, with and without area charges. Data on the distributions of travel times at different departure times and routes was collected from Google Maps. Model simulation details are described in Section 6. The two-step GMM is estimated with 100 random initial conditions. The cost of late arrival is held fixed at $\beta_L =$ Rs. $4,000$ (Appendix Figure A8 shows that the objective function is mostly flat for $\beta_L \geq$ Rs. $4,000$. Appendix Table A7 shows that results are essentially unchanged by using $\beta_L =$ Rs. $1,000$ or $\beta_L =$ Rs. $8,000$). Standard errors from 100 bootstrap runs are shown in parentheses.

Table 8: Road Technology: Travel Delay Linear in Traffic Volume

	(1)	(2)	(3)	(4)
Dependent Variable:	Google Maps Travel Delay (min/km)			
Sample:	Departure Time			Dates
Traffic Volume	1.06***	1.15***		0.97***
	(0.06)	(0.12)		(0.04)
Traffic Volume Exponent γ		0.89***		
		(0.15)		
GPS Travel Delay (min/km)			1.00***	
			(0.04)	
Constant	2.14***	2.08***	-0.10	2.19***
	(0.03)	(0.06)	(0.12)	(0.04)
Observations	1,440	1,440	1,440	185
Traffic Volume Std.Dev.	0.69	0.69		0.16
R^2	0.94	0.94	0.95	0.56

Notes: Table version of Figure 3 and Appendix Figure A9. This table shows that travel delay – measured either using Google Maps or GPS data – is approximately linear in the volume of traffic.

Data. The volume measures are based on 117,527 trips from 1,747 app users across 185 days (including weekends). Google Maps travel delay was collected over 28 routes in South Bangalore, every 20 minutes daily for 207 days (including weekends).

Variables and Sample. In the first and second columns, all weekday trip departure times are aggregated at the departure time minute level, then smoothed using a local linear regression with Epanechnikov kernel with 10 minutes bandwidth, and finally normalized to have mean 1; I compute the average delay over all weekdays and routes for each departure time, interpolating at the minute level. In the last column, for each date I compute the number of trips per capita (using the number of app users that day), and again normalize this variable to have mean 1; I compute the average delay over all routes and departure times, for each day in the data. GPS travel delay in column 3 is computed based on the GPS trips data. The sample is all weekday trips without any stops along the way, and with a trip diameter to total length ratio above 0.6 (the 25th percentile). For each departure time that is a multiple of 20 minutes, I compute the *median* delay of all trips starting around that departure time (weighting each trip using an Epanechnikov kernel with bandwidth 20 minutes around the reference departure time), then interpolate the result at the minute level.

Specifications. Columns 1, 3, and 4 report OLS regression with Google Maps Delay as outcome, with Newey-West standard errors, with three-hour lag in columns (1) and (3), and 10 day lag in column (4). Column 2 reports results from a nonlinear regression $Delay_h = \lambda_0 + \lambda_1 Volume_h^{\gamma}$ with HAC standard errors with Newey-West kernel and three-hour lag.$^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

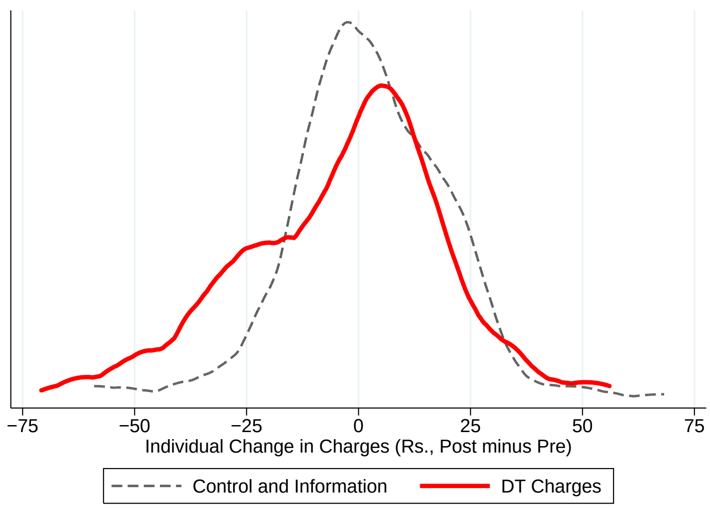
Table 9: Travel Times and Welfare in the Unpriced Nash Equilibrium and in the Social Optimum

	(1) Nash	(2) Social Optimum	(3) Improvement	(4) Improvement (% of Nash)
Panel A. Benefits and Welfare				
Travel Time (minutes)	38.7	37.7	-1.04	-2.69%
Welfare (Rupees)	-773.4	-769.0	4.46	-0.58%
Panel B. Benefits and Welfare Relative to Free-flow				
Travel Time (minutes)	15.4	14.4	-1.04	-6.77%
Welfare (Rupees)	-337.8	-333.3	4.46	-1.32%

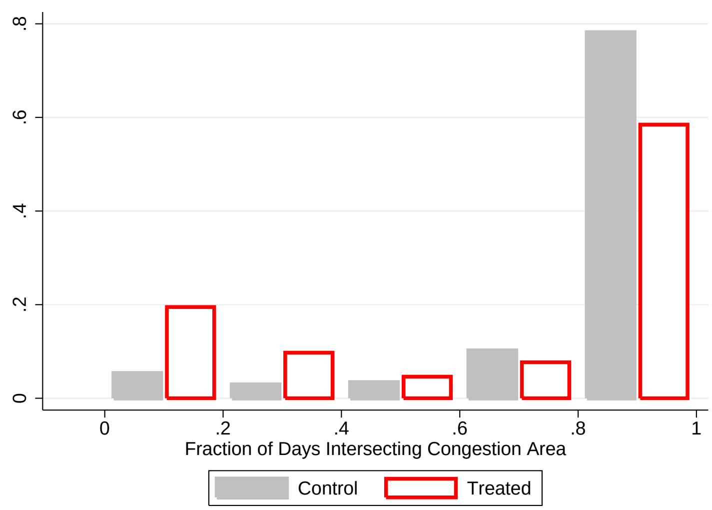
Notes: This table reports average travel times and welfare under the decentralized unpriced Nash equilibrium and under the social optimum. In panel B travel times and welfare are computed relative to the "free-flow" benchmark, where delay is constant at 2.14 minutes per kilometer regardless of traffic volume. (The average trip length is 10.9 Km.) Travel times are calculated taking individual route length into account, and welfare is the sum over all simulation agents of expected utility, including travel time and scheduling costs, and assuming charges are transferred lump-sum back to commuters. Columns 3 and 4 report the improvement from the unpriced Nash to the social optimum, in levels and as a fraction of the baseline (Nash) value.

A Appendix Figures

Figure A1: Treatment Heterogeneity for Departure Time and Area Treatments



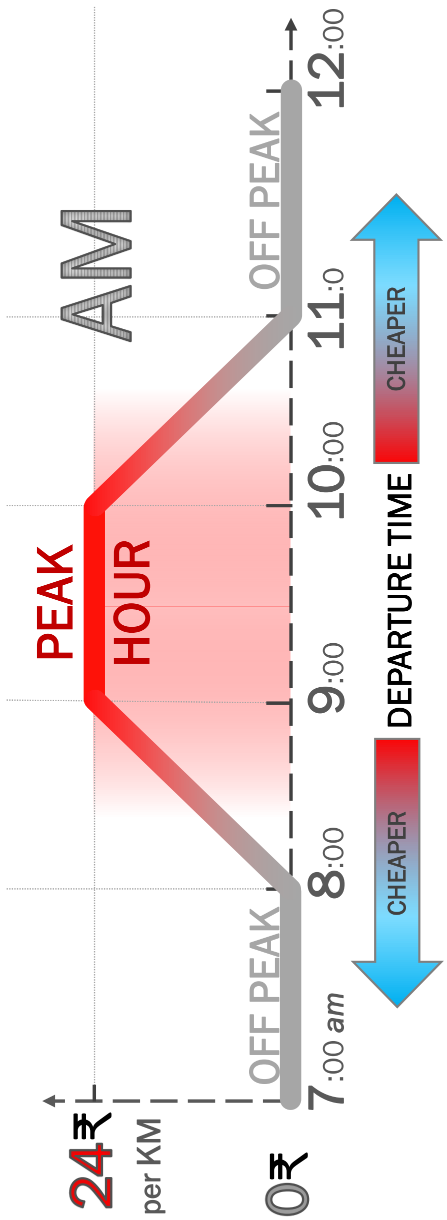
Panel (A) Individual-Level Change in Shadow Trip Rate, by Departure Time Treatment



Panel (B) Frequency of Avoiding the Congestion Area, by Area Treatment Status

Notes. These figures show the distribution of individual changes in shadow charges in the departure time treatment and control (panel A) and the distribution of individual frequency of intersecting the area when treated and not treated in the area treatment (panel B). In panel A, the sample is all regular commuters and all trips between the morning profile peak and 2 hours earlier. The graph plots the pre-post change in shadow trip rates for each participant, separately for participants with charges (Low Rate and High Rate treatments) versus those without charges (the control and information treatments). (The graph with shrunk distributions using empirical Bayes shows a similar pattern.) The sample for panel B is all days with trips in the morning between home and work for regular commuters (see Table 5 column 2). The graph shows the histogram of the fraction of days when a participant intersects the congestion area, separately for treated and not treated participants. Both graphs suggest a stark form of heterogeneity in how commuters respond to charges.

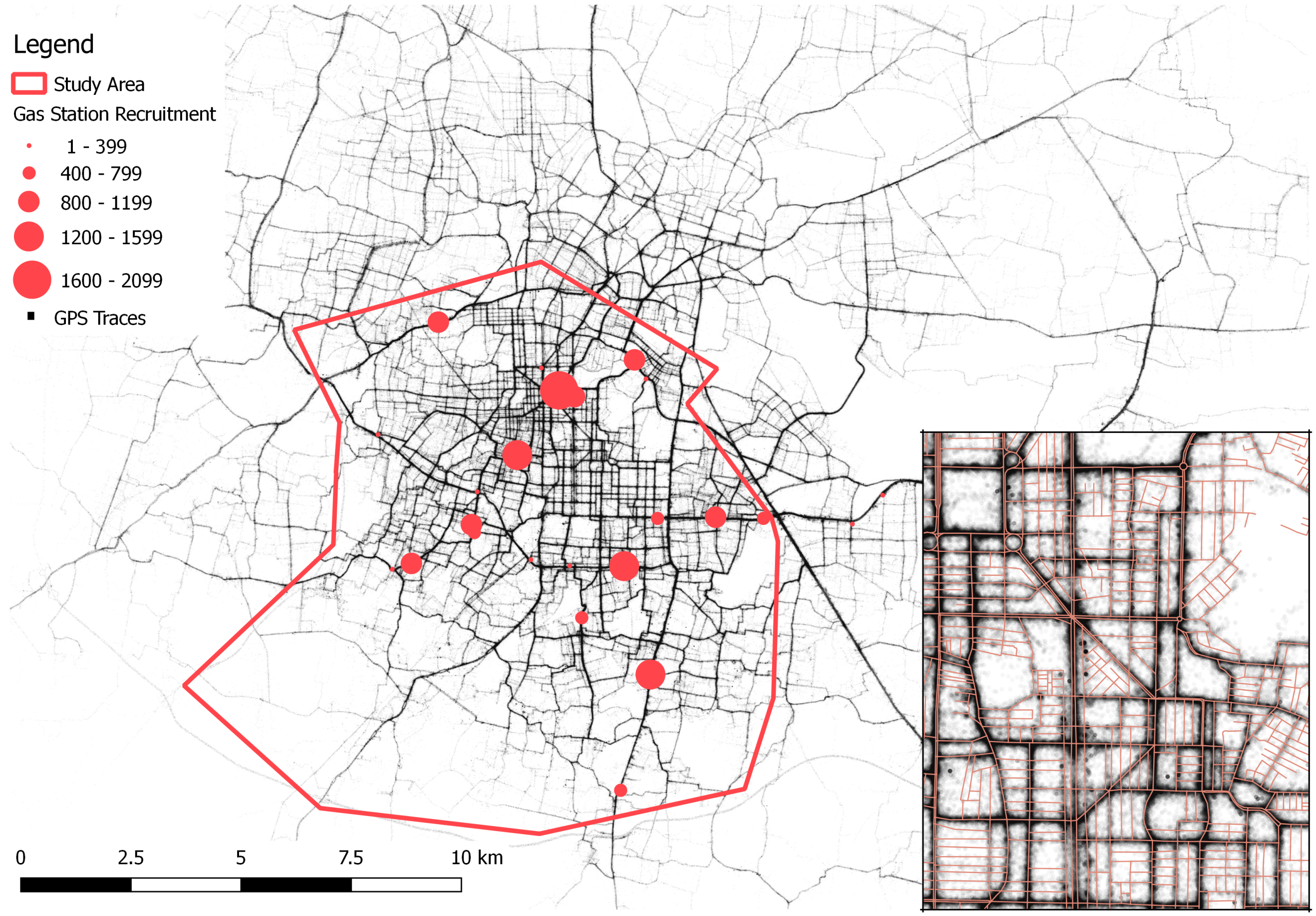
Figure A3: Departure Time Congestion Charge (AM) Rate Profile Card Example



	MORNING AM	
	Departure time	₹perKM
OFF-PEAK	7:00	0₹ per KM
	7:10	
	7:20	
	7:30	
	7:40	
	7:50	
SHOULDER	8:00	0
	8:10	4
	8:20	8
	8:30	12
	8:40	16
	8:50	20
PEAK	9:00	24₹ per KM
	9:10	
	9:20	
	9:30	
	9:40	
	9:50	
	10:00	
SHOULDER	10:10	20
	10:20	16
	10:30	12
	10:40	8
	10:50	4
	11:00	0
OFF-PEAK	11:10	0₹ per KM
	11:20	
	11:30	
	11:40	
	11:50	
	12:00	

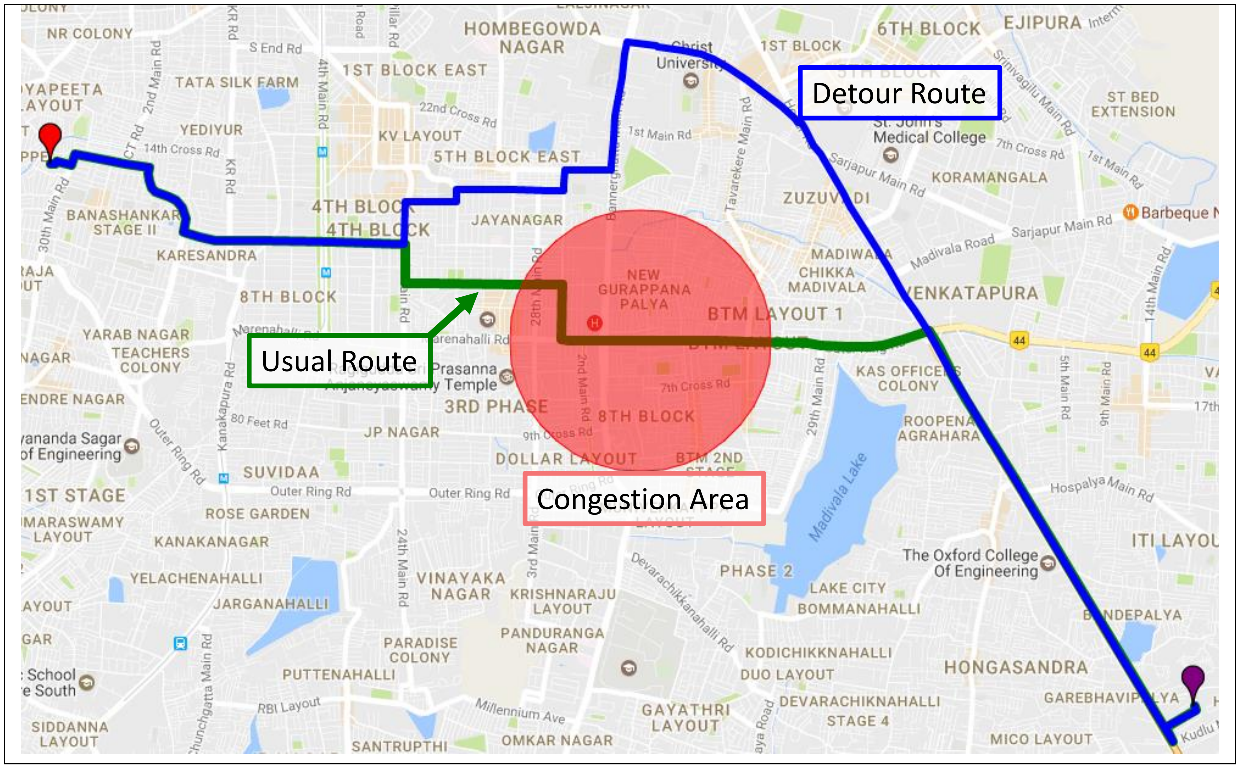
Notes: This figure shows an example of Rate Profile card that study participants in the departure time charge sub-treatments received. The cards for different participants differed in the value of the *peak rate* (Rs. 12/Km and Rs. 24/Km in the Low Rate and High Rate sub-treatments, respectively), and in the starting time of the profile (between 8 am and 9 am for the morning profile, and between 5 pm and 6 pm for the evening profile).

Figure A2: Study Area and Recruitment Locations



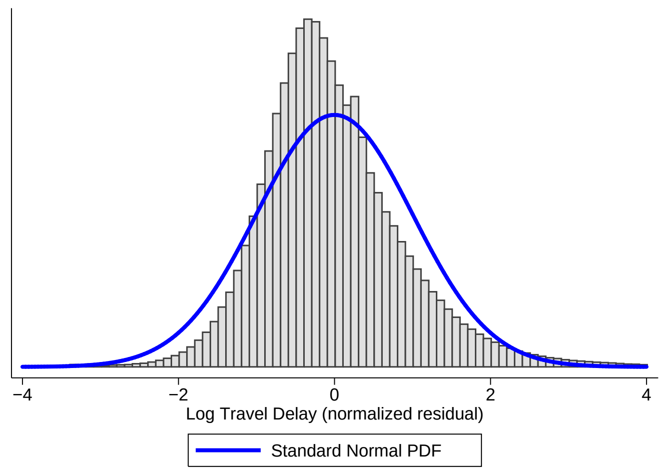
Notes. This figure shows the area of South Bangalore where the study was conducted. The red discs represent the randomly chosen gas stations where study participants were recruited (the diameter indicates the number of commuters approached). The black points represent all the GPS data collected during the study. (In the inset the Bangalore Open Street Map road network is overimposed for comparison.)

Figure A4: Area Congestion Charge Example



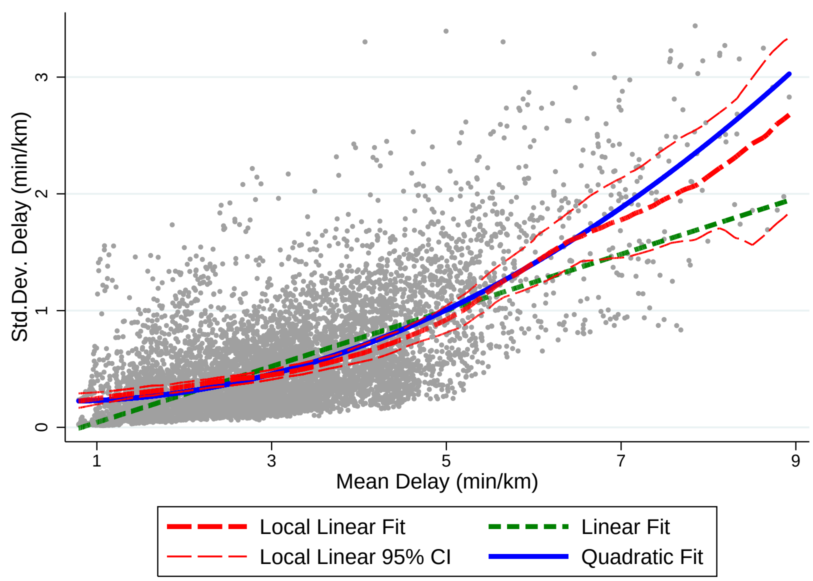
Notes: This figure shows an example of congestion area. Congestion areas were selected as follows. Given a regular route between home and work for a participant (in green), several "candidate" areas were selected along the route, with a radius of 250m, 500m or 1000m. For each candidate area, I found the quickest detour route that avoids the congestion area, based on a custom algorithm using multiple queries to the Google Maps API. Candidate areas with detours between 3 and 14 minutes longer were manually reviewed, and the final area was randomly selected from within this group.

Figure A5: Google Maps Travel Time is Approximately Log-Linear Distributed



Notes: This figure shows the shape of the day-to-day variation of log normalized travel time. For each route and departure time cell, I consider the distribution of travel times over 147 weekdays. Within each cell, I compute the normalized residual by subtracting the mean and dividing by the standard deviation for that cell. The graph shows the distribution of the log residuals for all cells, and a standard normal (solid blue line).

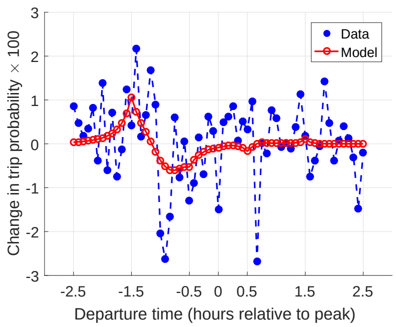
Figure A6: Travel Time Standard Deviation is Approximately Quadratic in Travel Time Mean



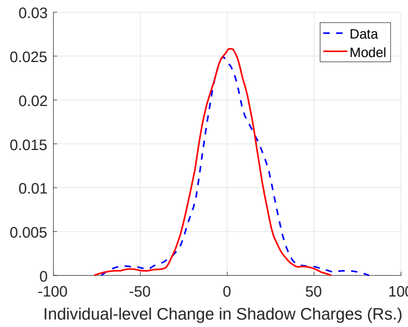
Notes: This figure shows the relationship between the mean and standard deviation of travel time. Each dot represents a route and departure time cell, and the two axes measure the mean and standard deviation in that cell over over 147 weekdays. The local linear, linear and quadratic fits are respectively shown in red (long dash), green (dash) and blue (solid). The local linear fit uses and Epanechnikov kernel with 0.5 minute per kilometer bandwidth, and 95% confidence intervals, bootstrapped by route, are also shown (thin red dashed line). The estimated quadratic equation is:

$$\text{StdDevDelay} = \underset{(0.02)}{0.24} - \underset{(0.01)}{0.05} \cdot \text{MeanDelay} + \underset{(0.002)}{0.04} \cdot \text{MeanDelay}^2$$

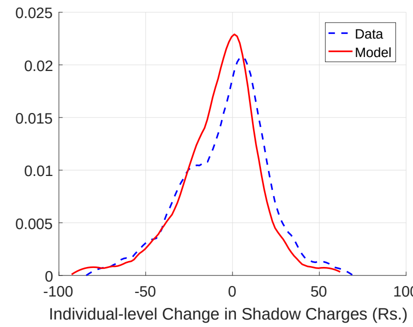
Figure A7: Structural Model Fit



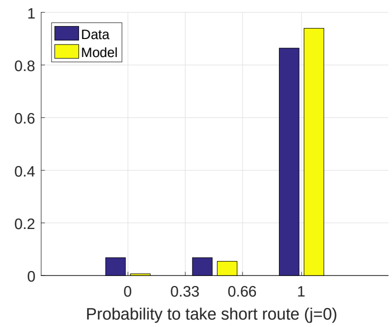
Panel (A) Departure Time



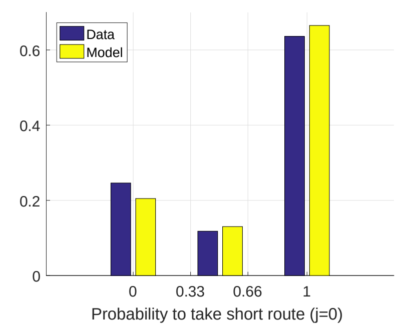
Panel (B) DT Individual (Control)



Panel (C) DT Individual (Treatment)



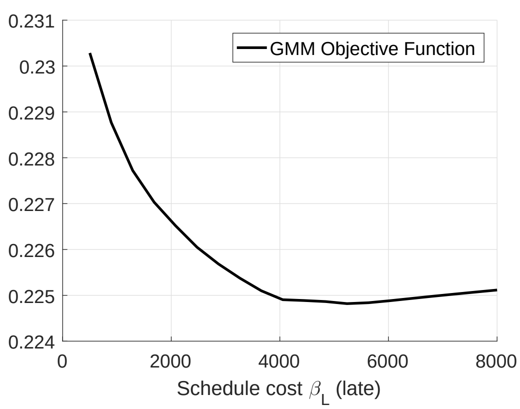
Panel (D) Area Individual (Control)



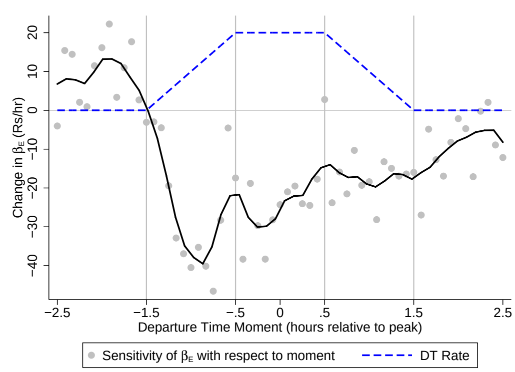
Panel (E) Area Individual (Treated)

Notes: This figure shows the fit of the estimated structural model. Panel A shows the 61 moments that target the difference in difference in number of trips by departure time bin. Panels B and C show the distributions of individual effects in the departure time treatments (changes in shadow charges between pre- and post-). Panels D and E show the distributions of individual effects in the area treatment (fraction of days intersecting the congestion area when treated and not treated).

Figure A8: Structural Model Diagnostics



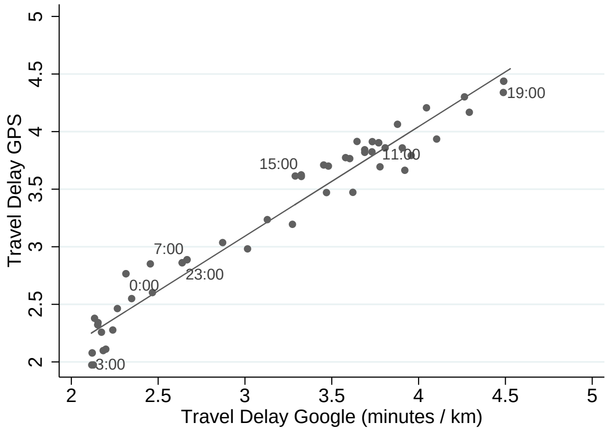
Panel (A) GMM Objective as Function of Late Schedule Cost β_L



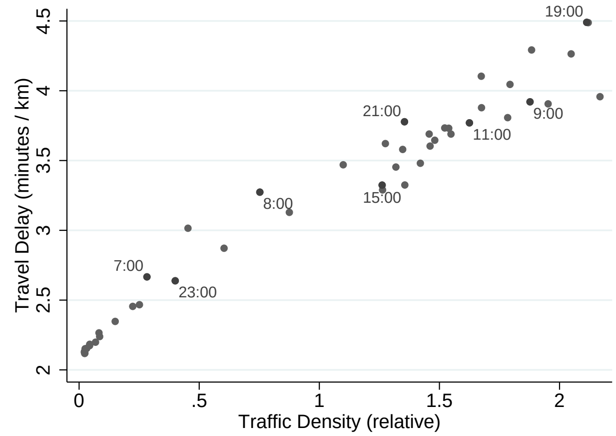
Panel (B) Sensitivity of Early Schedule Cost β_E with Respect to Departure Time Moments

Notes: Panel A shows that the objective function is mostly flat for values of the late schedule cost β_L above Rs. 4,000. It is evaluated at the estimated parameters, using the optimal weighting matrix. Panel B plots the scaled sensitivity measure from Andrews et al. (2017), quantifying the change in the estimated early schedule cost parameter β_E given by one standard deviation change in each of the 61 departure time moments (see Appendix Table A9 for the full definition of the sensitivity measure).

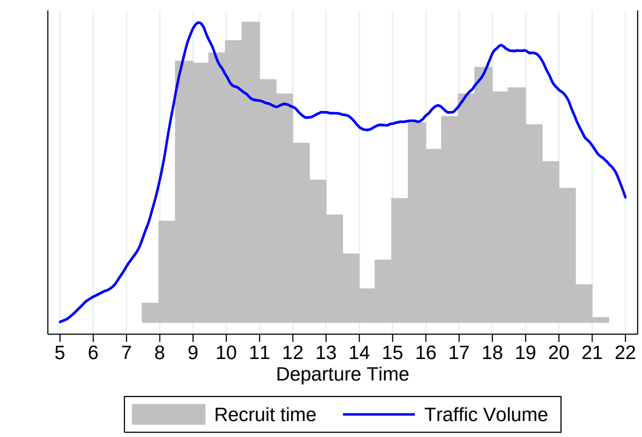
Figure A9: Road Technology Estimation Robustness Checks



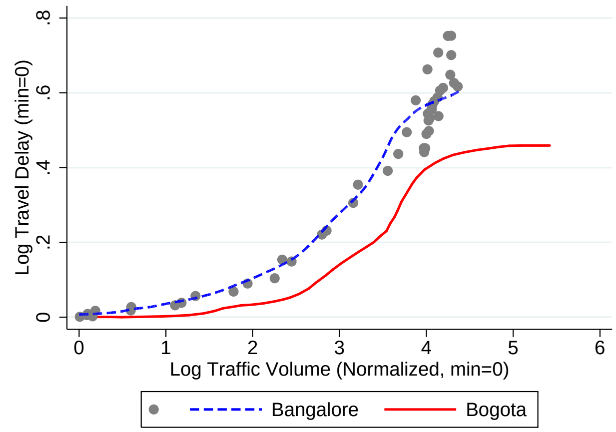
Panel (A) Travel Delay from GPS Data and Google Maps



Panel (B) Travel Delay and Traffic Density



Panel (C) Recruitment Time and Trip Time Distributions



Panel (D) Comparison with (Akbar and Duranton, 2017)

Notes: Panel A shows the relationship between travel delay measured using Google Maps and travel delay measured using GPS trips from smartphone app users, at the level of departure time. The notes for Table 8 describe the samples and variable construction. Each dot represents the average delay from Google Maps (X axis) over all weekdays in the sample, and median delay from GPS data (Y axis) over all weekday trips in the sample. The OLS fit with slope 1.00 (0.04) is also shown.

Panel B replicates Figure 3 with traffic density instead of volume of departures on the X axis. Road density at a certain time is defined as the number of ongoing trips.

Panel C plots the distribution of participant recruitment times (histogram in solid gray) and the distribution of trip departure times (kernel density plot in solid blue line). Both Y axes start at zero.

Panel D compares log-log road technology estimates from this paper (gray dots, dashed blue line) with those from Akbar and Duranton (2017) in Bogotá (red solid line). (Their estimate is computed from Figure 4 panel C.) Akbar and Duranton (2017) use a transportation survey to measure traffic volume at different times of the day. Zhao et al. (2015) document that in Singapore survey respondents report more concentrated departure times in the morning and evening, compared to real departure times as measured with a GPS smartphone app; this leads to overestimating peak-hour volumes. A similar effect may explain the slightly flatter region for high traffic volumes in Bogotá.

B Appendix Tables

Table A1: GPS Data Quality at Daily Level (Attrition Check)

	(1)	(2)	(3)	(4)
Commuter FE	X	X	X	X
Panel A. Departure Time Treatment				
High Rate × Post	0.01			
	(0.05)			
Low Rate × Post	-0.01			
	(0.05)			
Information × Post	-0.01			
	(0.04)			
Post	0.09***			
	(0.04)			
Observations	24,827			
Control Mean	0.76			
Panel B. Area Treatment and Sub-treatments				
Treated	0.05**	0.04	0.05**	0.05
	(0.02)	(0.03)	(0.02)	(0.03)
Post	0.06*	0.06*	0.03	0.07**
	(0.03)	(0.03)	(0.03)	(0.04)
Treated × High Rate		0.01		
		(0.04)		
Treated × High Rate Day			-0.00	
			(0.02)	
Treated × Short Detour				-0.05
				(0.05)
Observations	13,479	13,479	13,479	8,032
Control Mean	0.73	0.73	0.73	0.76

Notes. This table shows experimental impacts on the quality of the GPS data received from study participants. The outcome is a dummy for good quality GPS data on a given day. The sample covers all non-holiday weekdays for all experiment participants, excluding days outside Bangalore. In the post period, in panel A only the first or the last three weeks are included, and in panel B only the first and the last week are included. Panel B restricts to 243 participants in the Area treatment, except in column (4) where the sample consists of the 148 Area participants for whom candidate areas included at least one with short detour (3-7 minutes) and at least one with long detour (7-14 minutes). (See section 4 for more details on the candidate area selection process.) All specifications include respondent and study cycle fixed effects; column (4) includes fixed effects for each day in the experiment. The mean of the outcome variable in the control group during the experiment is reported for each specification. Standard errors in parentheses are clustered at the respondent level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A3: Impact of Departure Time Charges on Daily Shadow Charges

	(1)	(2)	(3)
Time of Day	AM & PM	AM	PM
Commuter FE	X	X	X
Panel A. Total Shadow Charges Today			
High Rate × Post	-22.7**	-16.5**	-6.2
	(11.1)	(7.4)	(6.3)
Low Rate × Post	-11.7	-3.5	-8.2
	(13.5)	(8.5)	(7.9)
Information × Post	9.5	2.9	6.6
	(10.2)	(6.7)	(6.2)
Post	0.6	-1.2	1.9
	(9.1)	(5.6)	(6.0)
Observations	15,610	15,610	15,610
Control Mean	151.0	81.7	69.3

Notes: This table replicates panel A in Table 2 using shadow *charges* instead of shadow *rates*. The shadow charge for a trip is equal to the shadow rate multiplied by the trip length in kilometers. Shadow charges are expressed in Rupees and are calculated based on a peak rate of Rs. 24/Km for for all respondents.

Table A2: Experimental Balance Checks

		Departure Time Treatments								Area Treatment			
		Information	(S.E.)	Low rate	(S.E.)	High rate	(S.E.)	Obs.	Control Mean	Area Early	(S.E.)	Obs.	Control Mean
(1)	Car user	0.01	(0.02)	0.01	(0.01)	0.01	(0.02)	497	0.28	-0.01	(0.01)	254	0.28
(2)	Regular destination	-0.05	(0.05)	0.00	(0.05)	-0.09*	(0.05)	497	0.77	-0.05	(0.03)	254	0.95
(3)	Age	-0.85	(0.93)	1.34	(1.01)	-0.03	(1.07)	497	33.13	-1.35	(0.94)	254	34.30
(4)	Log vehicle price	0.11**	(0.05)	0.09	(0.05)	0.03	(0.06)	453	11.06	0.00	(0.05)	231	11.17
(5)	Log income	-0.00	(0.10)	-0.03	(0.14)	-0.08	(0.14)	410	10.13	-0.09	(0.12)	211	10.24
(6)	Frac days with good GPS data	0.00	(0.03)	0.01	(0.04)	-0.02	(0.04)	497	0.62	0.02	(0.03)	254	0.64
(7)	Frac days present at work	0.01	(0.03)	0.00	(0.03)	-0.03	(0.04)	497	0.70	-0.03	(0.03)	254	0.79
(8)	Number of trips per day	-0.19	(0.14)	-0.01	(0.16)	-0.15	(0.15)	497	1.91	-0.00	(0.13)	254	1.69
(9)	Total distance per day (Km.)	-0.95	(0.86)	0.49	(1.14)	-0.50	(1.04)	497	12.95	0.43	(0.98)	254	13.19
(10)	Total duration per day (min)	-5.09	(3.71)	0.48	(4.50)	-2.36	(4.33)	497	54.82	1.59	(3.96)	254	52.48
(11)	Total D.T. shadow rate per day	-1.25	(4.68)	1.18	(5.00)	-3.82	(4.88)	497	59.07	-1.81	(4.59)	254	57.23
(12)	Total Area shadow rate per day	-3.35	(4.05)	-1.85	(5.22)	-4.29	(5.51)	497	36.07	-0.09	(6.79)	254	76.83
(13)	Joint Significance Test F stat	0.61		0.13		0.68				0.00			
(14)	Joint Significance Test P-value	0.44		0.72		0.41				0.99			

Notes. This table shows experimental balance checks for the departure time and area treatments. Variables 1,3,4, and 5 are from the recruitment survey, while the remaining eight variables are calculated from the GPS trips data before the experiment. Each row and group of columns combination reports coefficients from a regressions with the row header as outcome. In the "Area Treatment" columns, the sample is restricted to 254 participants who receive the area treatment, and the dependent variable is whether the respondent was assigned to the "early area" sub-treatment (to receive the area charges in week 1 as opposed to week 4). All regressions include randomization strata dummies. Rows 13 and 14 report the F-statistic and p-value from column-wise joint significance tests. The joint significance test for the all the "Departure Time" regressions has F-statistic of 0.32 and a p-value of 0.81. Robust standard errors are shown in parentheses. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A4: Impact of Departure Time Charges on Trip Shadow Charge

	(1)	(2)	(3)	(4)	(5)
Time of Day	AM & PM	AM	AM pre peak	PM	PM post peak
Commuter FE	X	X	X	X	X
Panel A. Full Sample					
High Rate × Post	-6.00*	-12.98**	-19.60*	-2.39	-10.55
	(3.10)	(5.44)	(10.07)	(4.53)	(6.60)
Low Rate × Post	-3.43	-4.86	-13.31	-4.68	-4.29
	(3.93)	(6.92)	(11.39)	(5.55)	(8.55)
Information × Post	1.66	-2.55	-5.92	4.19	6.28
	(2.67)	(4.59)	(7.07)	(4.06)	(6.40)
Observations	43,776	16,764	7,592	18,468	7,899
Control Mean	49.49	70.74	83.87	53.44	59.88
Panel B. Regular Commuters, Home-Work and Work-Home Trips					
High Rate × Post	-14.29*	-27.00**	-44.46**	2.67	-10.17
	(8.07)	(11.22)	(17.39)	(10.71)	(13.40)
Low Rate × Post	-10.56	-10.71	-30.68*	-10.76	-34.80
	(10.53)	(12.66)	(16.20)	(16.79)	(26.84)
Information × Post	1.34	-0.63	-7.17	8.43	3.16
	(5.32)	(7.30)	(8.59)	(9.82)	(11.43)
Observations	11,895	5,789	3,782	4,862	2,113
Control Mean	68.87	83.39	85.20	70.19	76.48
Panel C. Variable Commuters, All Trips					
High Rate × Post	-7.98	-5.25	1.04	-16.16*	-25.36*
	(6.05)	(11.06)	(20.67)	(8.58)	(14.65)
Low Rate × Post	0.69	-5.54	10.28	-4.08	27.66
	(8.22)	(17.12)	(27.75)	(11.76)	(19.59)
Information × Post	-1.73	-3.96	-1.10	-3.13	-3.07
	(5.92)	(9.42)	(19.31)	(7.40)	(11.17)
Observations	8,177	2,826	961	3,432	1,439
Control Mean	37.09	49.88	61.41	46.82	49.64

Notes: This table replicates Table 3 using shadow *charges* instead of shadow *rates*. The shadow charge for a trip is equal to the shadow rate multiplied by the trip length in kilometers. Shadow charges are expressed in Rupees and are calculated based on a peak rate of Rs. 24/Km for for all respondents.

Table A5: Trip Duration for Trips that Intersect or Do Not Intersect the Congestion Area

	(1)	(2)	(3)
	Trip Duration (minutes)		
Route FE	X	X	X
Trip Charged	-4.81***	-3.16**	-0.28
	(0.98)	(1.34)	(1.79)
Trip Charged × Long Detour		-3.66*	
		(1.91)	
Trip Charged × Predicted Detour			-0.83***
			(0.29)
Observations	7,455	7,455	7,455
Control Mean	38.51	38.51	38.51

Notes: This table compares the trip duration (in minutes) of trips that intersect and trips that do not intersect the congestion area. The sample is home to work or work to home trips of area participants on non-holiday weekdays. Each specification includes route fixed effects. "Trip Charged" is a dummy for whether the trip intersects the congestion area. Column (2) includes the interaction with the "Long Detour" Area sub-treatment. Column (3) includes the interaction with detour duration (in minutes) as predicted from Google Maps.

Table A6: Treatment Heterogeneity

Heterogeneity Dummy Variable K	(1) Regular Destination	(2) Self Employed	(3) Car Driver	(4) Small Log Vehicle Value	(5) Older	(6) Small Stated α	(7) Small Stated β	(8) Short Route	(9) Seldom Avoid Area
Panel A. *Departure Time Treatment: Trip Rate*									
Charges×Post×($K=0$)	-1.25	-2.74**	-2.89**	-5.81***	-1.06	-3.41**	-5.04***	-2.85*	
	(2.17)	(1.30)	(1.35)	(1.63)	(1.90)	(1.52)	(1.92)	(1.47)	
Charges×Post×($K=1$)	-4.11***	-7.01***	-4.69**	-0.85	-4.70***	-4.26**	-2.68	-3.95**	
	(1.37)	(2.68)	(2.26)	(1.59)	(1.47)	(1.96)	(1.66)	(1.77)	
Observations	43,776	43,170	43,776	43,776	43,776	40,783	39,639	43,776	
Participants $K=0$	119	407	350	280	175	252	218	249	
Participants $K=1$	378	82	147	217	322	205	228	248	
Control Mean $K=0$	29.71	32.34	32.16	32.57	30.90	32.25	32.43	30.73	
Control Mean $K=1$	33.34	32.73	33.06	32.24	33.32	33.11	32.68	34.59	
P-value interaction	0.27	0.15	0.50	0.03	0.13	0.73	0.35	0.63	
Panel B. *Departure Time Treatment: Number of Trips Today*									
Charges×Post×($K=0$)	-0.40*	-0.10	-0.14	-0.36**	-0.09	0.03	-0.15	-0.10	
	(0.24)	(0.13)	(0.15)	(0.15)	(0.20)	(0.19)	(0.15)	(0.20)	
Charges×Post×($K=1$)	-0.09	-0.34	-0.20	0.11	-0.19	-0.20	-0.16	-0.28*	
	(0.14)	(0.35)	(0.19)	(0.19)	(0.15)	(0.16)	(0.18)	(0.15)	
Observations	15,610	15,367	15,610	15,610	15,610	14,416	14,073	15,610	
Participants $K=0$	119	407	350	280	175	252	218	249	
Participants $K=1$	378	82	147	217	322	205	228	248	
Control Mean $K=0$	2.98	2.78	3.01	2.84	2.87	2.93	2.79	3.20	
Control Mean $K=1$	2.94	3.70	2.82	3.10	3.00	3.02	3.11	2.68	
P-value interaction	0.26	0.52	0.80	0.06	0.69	0.37	0.94	0.46	
Panel C. *Area Treatment: Trip Shadow Rate*									
Treated×($K=0$)		-11.91***	-11.54***	-11.29***	-7.04**	-12.92***	-9.65**	-11.46***	-9.43***
		(2.49)	(2.56)	(2.80)	(3.56)	(2.97)	(4.04)	(2.81)	(2.74)
Treated×($K=1$)		-7.94**	-12.73***	-12.54***	-14.18***	-10.19***	-13.07***	-12.66***	-14.19***
		(3.58)	(3.95)	(3.38)	(2.66)	(3.36)	(2.73)	(3.38)	(3.26)
Observations		20,367	20,594	20,594	20,594	18,741	18,260	20,594	20,594
Participants $K=0$		190	163	133	73	100	90	123	110
Participants $K=1$		32	63	93	153	104	109	103	116
Control Mean $K=0$		47.03	44.10	46.14	39.79	41.99	43.13	46.99	34.43
Control Mean $K=1$		37.31	46.75	42.84	47.35	46.56	45.21	42.07	53.00
P-value interaction		0.36	0.80	0.78	0.11	0.54	0.48	0.79	0.27
Panel D. *Area Treatment: Number of Trips Today*									
Treated×($K=0$)		0.21**	0.08	0.15	0.15	0.19	0.18	0.16	0.32**
		(0.09)	(0.10)	(0.10)	(0.15)	(0.13)	(0.12)	(0.13)	(0.13)
Treated×($K=1$)		-0.07	0.40**	0.20	0.18*	0.14	0.19	0.21*	-0.00
		(0.24)	(0.16)	(0.14)	(0.10)	(0.12)	(0.12)	(0.11)	(0.10)
Observations		8,745	8,878	8,878	8,878	8,056	7,874	8,878	8,878
Participants $K=0$		204	174	141	79	108	95	132	121
Participants $K=1$		35	69	102	164	110	118	111	122
Control Mean $K=0$		2.28	2.55	2.44	2.43	2.51	2.32	2.80	2.38
Control Mean $K=1$		3.80	2.37	2.58	2.53	2.45	2.58	2.17	2.61
P-value interaction		0.29	0.09	0.74	0.88	0.79	0.96	0.76	0.05

Notes: This table reports heterogeneous experimental response by observable characteristics. All heterogeneity variables K are dummy variables. They are: whether the commuter has a stable destination in column 1, whether the commuter's vehicle value is below median in column 4, whether the commuter is at least 35 years old in column 5, whether the stated preference value of time (α) is below median in column 6, whether the stated preference schedule cost (β) is below median in column 7, whether the daily average kilometers travelled pre-experiment is below median in column 8, and whether the frequency of intersecting the congestion area pre-experiment is below median in column 9.

Data. Vehicle values are scrapped from a global online marketplace and matched by vehicle type, brand and model. Stated preferences are from a phone survey. Value of time is measured by asking for the fee that would make commuters indifferent between their usual travel time plus the fee, or a longer travel time. The measure of schedule costs is computed asking by how much commuters would advance (or delay) their departure time if each minute leaving earlier (or later) was less expensive. A commuter has "small stated β" if the absolute change in departure time is above median. The stated preference values are residuals after controlling for morning or evening and cheaper earlier or cheapear later effects.

Specification. Each regression includes commuter fixed effects, study period fixed effects interacted with each group. The last line in each panel reports the p-value from the test of whether the two groups ($K=0$ and $K=1$) responded identically to the experiment.

Table A7: Structural Estimation Robustness Checks

(1) Value of time α (Rs/hr)	(2) Schedule cost early β_E (Rs/hr)	(3) Schedule cost late β_L (Rs/hr)	(4) Logit inner σ (dep. time.)	(5) Logit outer μ (route)	(6) Probability to respond p
1,187.2	345.2	1,000	27.3	37.3	0.47
1,092.3	322.9	8,000	31.9	36.8	0.47

Notes: This table replicates Table 7 using different assumptions for the late schedule cost. In Table 7 the late cost is fixed at $\beta_L =$ Rs. $4,000$; here it is fixed at $\beta_L =$ Rs. $1,000$ and $\beta_L =$ Rs. $8,000$.

Table A8: Numerical Model Identification Check

	(1) $\hat{\alpha}$	(2) $\hat{\beta}_E$	(3) $\hat{\sigma}$	(4) $\hat{\mu}$	(5) $\hat{p}$
Value of time α	1.12***				
	(0.04)				
Penalty early β_E		1.00***			
		(0.17)			
Logit inner σ			1.41**		
			(0.55)		
Logit outer μ				1.08***	
				(0.03)	
Probability to respond p					1.05***
					(0.03)
Observations	90	90	90	90	90
R^2	0.92	0.43	0.06	0.92	0.90

Notes: This table shows numerically that the GMM estimation is able to recover model parameters. To construct it, I drew 100 random sets of model parameters, and for each set I simulated the model to generate choice data corresponding to those parameters, and estimated the structural model on the simulated data. Each column in this table reports the results from a regression of the estimated parameter on the original parameter. Each random parameter $\theta_0 \in \{\alpha, \beta_E, \beta_L, \sigma, \mu, p\}$ was drawn independently from a uniform distribution $U\left(0.3 \cdot \hat{\theta}, 2 \cdot \hat{\theta}\right)$ where $\hat{\theta}$ is the GMM parameter estimate from Table 7. The simulated data covered five times more commuters than the real data. When running GMM, the random initial conditions did not depend on the original parameters, and the late schedule cost was fixed at $\beta_L =$ Rs. $4,000$ as in Table 7. Ten observations where the estimated parameters had extreme values were dropped; results are essentially unchanged with all 100 observations, except for column (3) where the coefficient becomes 0.53 (0.78). Robust standard errors in parentheses. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A10: Road Technology Trip Level Regressions

	(1)	(2)	(3)	(4)
Dependent Variable:	Trip Delay (min/km)			
Commuter FE			X	X
Traffic Volume at Trip Departure Time	0.87***	0.84***	0.70***	0.70***
	(0.04)	(0.03)	(0.04)	(0.04)
Trip Length (km)		-0.05***		-0.01**
		(0.00)		(0.00)
Constant	2.47***	2.97***	3.11***	3.17***
	(0.06)	(0.05)	(0.06)	(0.06)
Observations	61,234	61,234	61,234	61,234

Notes: This table reports trip-level quantile (median) regressions of the trip delay, defined as trip duration in minutes divided by trip length in kilometers, on the average traffic volume at the trip departure time and trip length. The sample is all weekday trips more than 2km in length, without any stops along the way, and with a trip diameter to total length ratio above 0.6 (the 25th percentile). Columns 3 and 4 first residualize the trip delay on commuter fixed effects. Standard errors in parentheses are clustered at the commuter level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table A11: Experimental Design (strata, sub-treatments, timing)

Strata	*Congestion Charge Treatments* Departure Time (DT) Sub-treatments	 Area Sub-treatments	*Timing*
(Area eligible Area ineligible) × (Car Motorcycle) × (High Daily KM Low Daily KM)	(High Rate Low Rate Information Control)	(Low Rate High Rate) × (Long Detour Short Detour)	(DT First DT Last)

Table A9: Structural Estimation Sensitivity Measure

	(1) Value of time α (Rs/hr)	(2) Schedule cost early β_E (Rs/hr)	(3) Logit inner σ (dep. time.)	(4) Logit outer μ (route)	(5) Probability to respond p
Panel A. Departure Time Moments					
(1-61) Average absolute value for departure time bin moments	28.7	17.3	6.32	0.32	0.005
(62) Variance individual effects departure time treatment	167.1	1.5	-5.16	0.59	-0.017
(63) Variance individual effects departure time control	61.7	9.1	9.88	0.08	-0.016
Panel B. Area moments					
(64) With charge: probability to intersect area	-170.8	-23.9	16.82	0.37	0.039
(66) With charge: sample frequency to intersect area $\in [1/3, 2/3)$	170.7	34.5	-4.41	3.47	0.029
(67) With charge: sample frequency to intersect area $\in [2/3, 1]$	238.9	26.4	-20.69	0.97	-0.060
(65) Without charge: probability to intersect area	187.2	49.2	2.38	-0.89	0.080
(68) Without charge: sample frequency to intersect area $\in [1/3, 2/3)$	-121.0	-23.3	3.70	1.16	-0.026
(69) Without charge: sample frequency to intersect area $\in [2/3, 1]$	-220.1	-48.0	1.51	0.60	-0.065

Notes: This table reports the estimated sensitivity measure Λ from Andrews et al. (2017), scaled by the (bootstrap) standard deviation of each moment. Each entry Λ_{pj} measures the change in estimated parameter θ_p due to a one standard deviation change in moment m_j. The matrix Λ is estimated by $\hat{\Lambda} = \left(\hat{S}'\hat{W}\hat{S}\right)^{-1}\hat{S}'\hat{W}\mathrm{diag}\left(\hat{\sigma}_j\right)$ where $\hat{S}$ is the Jacobian of the moments with respect to parameters evaluated at the estimated parameters, $\hat{W}$ is the estimated optimal weighting matrix, and $\mathrm{diag}\left(\hat{\sigma}_j\right)$ is the diagonal matrix with jth entry given by the (bootstrap) estimated standard error of moment j.

Table A12: Subtreatment Probabilities by Stratum

Strata				*Departure Time (DT) Sub-treatments*					*Area Sub-treatments*					*Timing*		
									High Rate		Low Rate					
No.	Area	Car or Moto	Daily KM	High Rate	Low Rate	Info	Control	*Total*	Short Detour	Long Detour	Short Detour	Long Detour	*Total*	DT Weeks 1-3	DT Weeks 2-4	*Total*
1	Eligible	Car	Low	3/8	1/8	2/8	2/8	1	1/4	1/4	1/4	1/4	1	1/2	1/2	1
2	Eligible	Car	High	1/8	3/8	2/8	2/8	1	1/4	1/4	1/4	1/4	1	1/2	1/2	1
3	Eligible	Moto	Low	3/8	1/8	2/8	2/8	1	1/4	1/4	1/4	1/4	1	1/2	1/2	1
4	Eligible	Moto	High	1/8	3/8	2/8	2/8	1	1/4	1/4	1/4	1/4	1	1/2	1/2	1
5	Ineligible	Car	Low	1/12	3/12	4/12	4/12	1					0	1/2	1/2	1
6	Ineligible	Car	High	3/12	1/12	4/12	4/12	1					0	1/2	1/2	1
7	Ineligible	Moto	Low	1/12	3/12	4/12	4/12	1					0	1/2	1/2	1
8	Ineligible	Moto	High	3/12	1/12	4/12	4/12	1					0	1/2	1/2	1