SM Supplement to "Peak-Hour Road Congestion Pricing: Experimental Evidence and Equilibrium Implications"

Contents

SM.1 Departure Time Model Identification

In this section, I formally prove how identification of schedule costs and schedule heterogeneity in a departure time model depends on observing commuter reactions to congestion pricing. For analytical tractability, I proceed in a simplified model that maintains the key features of the full model: schedule preferences and a peak-hour (inverse U shaped) travel time profile. These results continue to hold when the travel time profile is endogenously determined in equilibrium. I use simulations to check a conjecture that the deadweight loss of peak-hour congetion in this model is decreasing in the schedule costs.

For intuition for the identification results, consider a commuter that we observe to leave at very different times on different days (as I document in Table SM.I). There are two ways this could arise. In the first scenario, the commuter has a unique ideal arrival time and high schedule flexibility. In this case, small idiosyncratic shocks have a large effect on departure times. In the second scenario, each day, the commuter draws an ideal arrival time from a dispersed distribution, but does not have much flexibility around that time.

These two cases are observationally equivalent for departure times, but they have different implications for how substitutable two departure times are to each other, on any given day. The key intuition for how congestion pricing leads to identification is that we can measure cross-price elasticities: how the probability of choosing departure time h depends on infinitesimal pricing of departure time h'.

SM.1.1 Simplified Departure Time Model

I assume that commuters have preferences directly over (continuous) departure times $h \in \mathbb{R}$. Unlike the main model where commuters have ideal *arrival* times, this assumption eliminates expectations over travel time uncertainty and greatly simplify the algebra.

Travel time is a (possibly degenerate) quadratic function of departure time. This captures the key shape of how travel time varies across the peak hour.[28] In most of the results below, schedule costs are quadratic and the ideal departure time is normally distributed. These assumptions rule out asymmetric (early/late) schedule costs yet deliver analytical tractability.

Given the focus on identification, I drop individual i and time t subscripts and assume that infinite data for a single individual is available. The utility for departure time h is

$$\underbrace{-\alpha T(h) - v(h - h^D)}_{u(h|h^D)} + \epsilon^D(h).$$

Here $v(\cdot)$ is the schedule penalty as a function of the deviation between departure time and the ideal departure time h^D. $\epsilon^D(h)$ are idiosyncratic shocks with scale β^{-1} that give rise to continuous logit choice probabilities. The ideal departure time h^D is distributed according to a cumulative distribution function F.

I assume that the value of travel time α is known and normalize it to $\alpha = 100$ INR/hour. Note, if travel time is not constant, this rules out a trivial source of non-identification due to scale.

The conditional probability density of choosing departure time h is given by the continuous logit density, and the unconditional density is given by integrating over F,

$$\pi(h|h^D) = \frac{\exp(\beta u(h|h^D))}{\int_{h'} \exp(\beta u(h'|h^D))dh'}, \text{ and } \quad \pi(h) = \int \pi(h|h^D)dF(h^D).$$

SM.1.2 Two Non-Identification Results with Observational Data

Before outlining the main results, I prove a general non-identification result in a simple setting where travel time is a constant (later, I will assume quadratic) and the ideal departure time distribution is unrestricted.

In this case, we can write the observed departure time as the sum of two independent random variables, corresponding to the ideal departure time, and the optimal departure time conditional on the ideal departure time. This exact decomposition helps clarify the source of non-identification.

Proposition 1. *Assume that travel time T is a constant (does not depend on departure time h). Normalize $\beta = 1$. Consider any family V of schedule delay functions $v \in V$, with at least two elements $v_1, v_2 \in V$ that differ on a non-zero measure set. Then, the schedule delay cost function $v(\cdot)$ is not identified given data on $\pi(h)$.*

Proof. If T does not depend on h, then $u(h|h^D)$ is only a function of the difference $h - h^D$. Hence, the

[28] The quadratic shape implies unrealistic negative travel time for very early or very late departure time. I later assume that schedule costs rise faster so that, on net, these departure times are unattractive.

optimal departure time random variable h^* can be written as the sum of two independent random variables, $h^* = h^D + \underbrace{h^* - h^D}_{h^E}$, where the pdf of h^E is

$$G(h^E) = \frac{\exp(-v(h^E))}{\int_h \exp(-v(h))dh}.$$

(Note: if v is quadratic then h^E is normally distributed.)

Consider two different schedule delay functions $v_1(\cdot)$ and $v_2(\cdot)$ and let h_1^E and h_2^E denote two independent random variables that have the corresponding pdfs G_1 and G_2.

Setting the ideal departure time distributions $h_1^D \sim G_2$ and $h_2^D \sim G_1$ (note that indices are switched) implies that the observed optimal departure time random variables $h_1^D + h_1^E$ and $h_2^D + h_2^E$ have the same distribution. Hence, the schedule cost function $v(\cdot)$ is not identified. □

The identification failure does not depend on constant travel time. I next prove the main non-identification result, in a model that is more strongly parametrized and where travel time is hump-shaped, which captures the peak-hour travel time profile. I make three functional form assumptions.

Assumption 1. $T(h)$ is quadratic, $T(h) = \tau_0 - \tau_1 h^2$ with $\tau_1 > 0$. Without loss of generality and for convenience I will set $\tau_0 = 0$.

Assumption 2. Schedule costs are quadratic, $v(h - h^D) = s(h - h^D)^2$ with $s > \tau_1$.

($s > \tau_1$ means that schedule costs dominate, and it implies that the commuter chooses departure times with negative travel time–very early or very late departure time–with very low probability.)

Assumption 3. The ideal departure time is normally distributed, $h^D \sim N(0, \sigma)$.

Proposition 2. *Fix the shape of the travel time profile τ_1 and maintain the VOTT normalization $\alpha = 100$ INR/hour. Under assumptions 1–3, the demand model parameters (β, s, σ) are not identified with data on observed departure times.*

This is not a trivial non-identification result due to scale, because VOTT α is normalized, and travel time is not constant.

The proof will show that it is possible to explain the same observed distribution of departure times by increasing schedule costs and increasing the spread of the ideal departure time distribution.

Proof of Propostion 2. I show that $\pi(h)$ is a normal distribution centered at zero. Its mean and variance depend on three variables (β, s, σ). Hence, the model is under-identified with two degrees of freedom.

The utility functions is (recall that the value of time spent driving α is normalized)

$$u(h|h^D) = \alpha\tau_1 h^2 - s(h - h^D)^2 + \epsilon^D(h).$$

Choice probabilities are given by

$$\begin{aligned}\pi(h) &= \int \pi(h|h^D)dF(h^D) \\ &= \int \frac{e^{-\beta\left(-\alpha\tau_1 h^2 + s(h-h^D)^2\right)}}{\int_{-\infty}^{\infty} e^{-\beta(-\alpha\tau_1(h')^2 + s(h'-h^D)^2)}\,dh'} \cdot \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{h^D}{\sigma}\right)^2} dh^D \\ &= \frac{1}{\sqrt{2\pi}\sqrt{\frac{s^2\sigma^2}{(s-\alpha\tau_1)^2} + \frac{1}{2\beta(s-\alpha\tau_1)}}} \exp\left(-\frac{1}{2}\frac{h^2}{\frac{s^2\sigma^2}{(s-\alpha\tau_1)^2} + \frac{1}{2\beta(s-\alpha\tau_1)}}\right).\end{aligned}$$

This is a normal distribution with mean zero and variance $\frac{s^2\sigma^2}{(s-\alpha\tau_1)^2} + \frac{1}{2\beta(s-\alpha\tau_1)}$. □

SM.1.3 Identification with Congestion Pricing Variation

I now study identification when we also observe choice probability distributions $\pi(\cdot|p(\cdot))$ in response to any possible pricing function $p(h)$.

Observing responses to pricing helps identify the cross-price elasticities for different departure times. This helps resolve the ambiguity discussed in the previous section, because different combinations of departure time distributions and conditional choice probabilities have different implications for cross-price elasticities.

The key object of interest is the impact of an "impulse" price function on choice probabilities. Slightly abusing notation (skipping a formal limit argument), we study "Kronecker delta" impulse pricing functions at h given by $p(x; h, \lambda) = \lambda 1(x = h)$ and study the effect of increasing λ around $\lambda = 0$ for given $h \neq h'$:

$$\left.\frac{d\pi(h'|p(\cdot;h,\lambda))}{d\lambda}\right|_{\lambda=0} = \frac{d}{d\lambda}\int \frac{\exp(\beta u(h'|h^D))}{\int_{h''}\exp(\beta u(h''|h^D) - \beta p(h'';h,\lambda))} dF(h^D).$$

For $h \neq h'$ and evaluating at $\lambda = 0$ this simplifies to

$$\beta \int \pi(h'|h^D)\pi(h|h^D)dF(h^D),$$

where $\pi(\cdot|h^D)$ denotes the conditional probability in the absence of pricing ($\lambda = 0$).

This expression shows that, for fixed $h - h'$, when conditional probabilities are concentrated (e.g. when β is high and/or the schedule cost function is steep around the ideal departure time), the cross-elasticities are close to zero. Intuitively, this suggests that knowing cross-elasticities for all h and h' solves the identification problem.

I now formally prove identification in the particular case considered in Result 2.

Proposition 3. *Fix the shape of the travel time profile τ_1. Under assumptions 1–3, the model parameters (β, s, σ) are identified with data on observed departure times and cross-elasticities for $h \neq h'$.*

Proof. Substituting the utility function and normal distribution for h^D in the expression for cross-elasticity, and computing integrals using Mathematica, yields

$$\beta \int \pi(h'|h^D)\pi(h|h^D)dF(h^D) = \beta^2(s-\alpha\tau_1)^{\frac{1}{2}}\left(s-\alpha\tau_1+4\beta s^2\sigma^2\right)^{-\frac{1}{2}} \exp\left(\frac{(s-\alpha\tau_1)2\beta^2 s^2\sigma^2}{(s-\alpha\tau_1+4\beta s^2\sigma^2)}(h'+h)^2-(s-\alpha\tau_1)\beta\left((h')^2+h^2\right)\right).$$

Taking log and grouping terms in a polynomial of h and h' gives

$$\begin{aligned}\ln\left(\int \pi(h'|h^D)\pi(h|h^D)dF(h^D)\right) = &-\beta(s-\alpha\tau_1)((h')^2+h^2)+ \\ &\frac{2\beta^2 s^2\sigma^2(s-\alpha\tau_1)}{s-\alpha\tau_1+4\beta s^2\sigma^2}(h'+h)^2+ \\ &\frac{1}{2}\ln\left(\frac{\beta^2(s-\alpha\tau_1)}{s-\alpha\tau_1+4\beta s^2\sigma^2}\right).\end{aligned}$$

By varying h and h', we have three identified coefficients and three unknowns (β, s, and σ). It is straightforward to check that this system of equations has a unique solution. □

SM.1.4 Equilibrium with Endogenous Congestion

I now show that the quadratic travel time profile assumed so far is consistent with equilibrium.

Assume that the travel time $T(h)$ is given by

$$T(h) = \lambda_0 + \lambda_1 \log(V(h)), \tag{7}$$

where $V(h)$ a measure of volume of travel around h. To construct V, assume that any trip at h affects the travel times of all other departure times (trips leaving both before and after h), with a weight given by a normal distribution pdf with standard deviation σ_V. That is, V is given by

$$V(h) = \int_{-\infty}^{\infty} \pi(h')\phi(h';h,\sigma_V)dh',$$

where $\phi(x;\mu,\sigma)$ is the normal pdf with mean μ and standard deviation σ, evaluated at x.

Given that π is a normal pdf, so will V, and hence travel time given by (7) will be quadratic in h.

Proposition 4. *This model has a unique equilibrium, where travel time is quadratic and choice probabilities follow a normal pdf. The following equilibrium equation holds:*

$$\frac{s^2\sigma^2}{(s-\alpha\tau_1)^2} + \frac{1}{2\beta(s-\alpha\tau_1)} + \sigma_E^2 = \frac{\lambda}{2\alpha\tau_1}.$$

SM.1.5 The Deadweight Loss of Congestion is Decreasing in Schedule Costs

Consider an equilibrium as described above. Based on Proposition 2, the observed choice probabilities and travel time profile are consistent with various combinations of schedule cost s and dispersion of ideal departure times σ.

Conjecture 1. *Holding fixed the equilibrium choice probabilities $\pi(h)$ and the profile of travel time $T(h)$, the deadweight loss of congestion (in absolute terms) is decreasing in schedule costs s.*

The deadweight loss does not appear to have a closed form solution. I use numerical simulations for 1,000 randomly chosen parameter vectors. I maintain the normalization $\alpha = 100$ INR/hour, and draw the following parameters uniformly and independently: $s \in [25, 125]$ INR/hour^2, $\sigma \in [0.05, 0.55]$ hours, $\beta \in [0.25, 0.75]$, $\sigma_V \in [0.5, 1, 5]$ hours, and $\lambda_1 \in [1.5, 2.5]$. In each simulation, I choose 10 alternate possible values of s' and solve for the implied σ' that leads to the same equilibrium as with the initial s, σ, and compute deadweight loss. In all 1,000 simulations, deadweight loss is decreasing in s'.

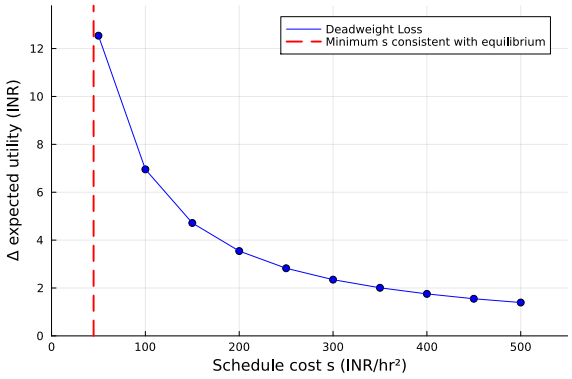


Figure SM1. Example deadweight loss versus schedule cost (holding observed equilibrium fixed)

SM.2 Route Choice Model Identification

To provide intuition for how VOTT and the route switching cost are separately identified using data from the route choice experiment, I analyze a version of the dynamic route choice model without departure time from section 4.2. I further assume no time discounting ($\delta = 0$).

Consider three time periods. At $t = 0$ the model is in steady state. At $t = 1$ the short route ($r = 0$) is unexpectedly charged p. At $t = 2$ the route is no longer charged. Denote $\pi_t(r_{t-1} \to r)$ the probability to use route r at time t if the $t-1$ route was r_{t-1} when there is no pricing, and $\pi_t(r_{t-1} \to r|p)$ with pricing p. Because there is no discounting, we have the following expressions for relative transition probabilities:

$$\frac{\pi_0(0 \to 0)}{1 - \pi_0(0 \to 0)} = \frac{\exp(0)}{\exp(\frac{-\gamma - \alpha\Delta T}{\mu})} \qquad \frac{\pi_0(1 \to 0)}{1 - \pi_0(1 \to 0)} = \frac{\exp(\frac{-\gamma}{\mu})}{\exp(\frac{-\alpha\Delta T}{\mu})}$$

$$\frac{\pi_1(0 \to 0|p)}{1 - \pi_1(0 \to 0|p)} = \frac{\exp(\frac{-p}{\mu})}{\exp(\frac{-\gamma - \alpha\Delta T}{\mu})} \qquad \frac{\pi_1(1 \to 0|p)}{1 - \pi_1(1 \to 0|p)} = \frac{\exp(\frac{-p-\gamma}{\mu})}{\exp(\frac{-\alpha\Delta T}{\mu})}.$$

It is easy to solve for the parameters α, γ, μ if these transition probabilities are known. Next, I show that these parameters are also unique determined by the detour route usage rates S_t in periods $t = 0, 1, 2$. These numbers satisfy the following equations (note that $t = 0$ and $t = 2$ have the same

transition probabilities)

$$\begin{aligned} S_0\pi_0(0 \to 1) &= (1 - S_0)\pi_0(1 \to 0) \\ S_1 &= S_0\pi_1(0 \to 0|p) + (1 - S_0)\pi_1(1 \to 0|p) \\ S_2 &= S_1\pi_0(0 \to 0) + (1 - S_1)\pi_0(1 \to 0). \end{aligned}$$

It is tedious but straightforward to show that these three equations uniquely determine α, γ, μ.

SM.3 Route Charge Treatment Regression Analysis

For the regression analysis of the route experiment, I focus on the early treatment group and the period before the experiment and the first two weeks during the experiment. I use the following specification:

$$y_{it} = \gamma^A \cdot T_i^{Early} W_t^1 + \gamma^{A,P} \cdot T_i^{Early} W_t^2 + \mu_t + \alpha_i + \varepsilon_{it}. \tag{8}$$

The coefficients of interest are γ^A and $\gamma^{A,P}$, which measure the impact of route congestion charges in the early charges group, and the persistence effect one week later, relative to similar commuters who anticipate that they will be treated in the fourth week of the experiment.

Panel A of Table SM.VII shows the impact of route charges on detour usage at the trip level. The sample is all trips between home and work. The results show a large increase of 27 percentage points during the first week in the experiment among the early treatment group, who faced charges that week. By comparison, only 11% of participants in the late group chose the detour that week. The second column shows that more than a third of this effect size persists one week later. Charges do not have a significant effect on the number of trips per day (columns 3 and 4). This means that there is no evidence that commuters reduce the number of trips to avoid route congestion charges, and the previous effects are driven by route switching.

I next analyze how baseline experience with detour routes affects the impact of charges. In Panel B, I restrict to commuters who use a detour route between home and work (or between work and home) at least once before the experiment. In general, the results from Panel A are amplified in this sample. Baseline usage is higher, as are the impact of charges (41 percentage points) and the persistence effect.

SM.4 Travel Demand Estimation

SM.4.1 Choice Probabilities

In the benchmark model with dynamic route choice and departure time choice, the departure time choice probabilities conditional on the chosen route (with $p_{it}(h, r) = 0$) is given by

$$\pi_i(h|r, h_{it}^A) = \frac{\exp((\sigma^{DT})^{-1}\mathbb{E}v(h, \mathcal{T}_i(h, r), h_{it}^A))}{\sum_{h'} \exp((\sigma^{DT})^{-1}\mathbb{E}v(h', \mathcal{T}_i(h', r), h_{it}^A))}.$$

These expressions show that the full model collapses to the single-route departure time choice model given by (2) when we condition on route and ideal arrival time. Similar expressions apply when we

include pricing $p_{it}(h,r)$.

In the full model, the expected utility of choosing route r is

$$\mathbb{E}u_{it}(r|h_{it}^A, r_{it-1}) = \sigma^{DT} \log\left(\sum_h \exp\left((\sigma^{DT})^{-1}\mathbb{E}v(h, \mathcal{T}_i(h,r), h_{it}^A)\right)\right) - \gamma \mathbf{1}(r \neq r_{it-1}) + \delta V_{it+1}(r).$$

This includes the "logsum" or 'inclusive value" term over departure times. In the upper nest, this leads to route choice probabilities (conditional on h_{it}^A)

$$\pi_{it}(r|h_{it}^A, r_{it-1}) = \frac{\exp((\sigma^R)^{-1}\mathbb{E}u_{it}(r|h_{it}^A, r_{it-1}))}{\exp((\sigma^R)^{-1}\mathbb{E}u_{it}(0|h_{it}^A, r_{it-1})) + \exp((\sigma^R)^{-1}\mathbb{E}u_{it}(1|h_{it}^A, r_{it-1}))}.$$

Unconditional probabilities follow by integrating over the ideal arrival time distribution f_i^A.

SM.4.2 GMM Moments That Exploit Experimental Variation

The two-step optimal GMM estimation finds the parameter vector $\theta = (\alpha, \beta_E, \beta_L, \gamma, \sigma^{DT}, \sigma^R, \eta_{\text{early}})$ that solves $\min_\theta \hat{g}(\theta)'\hat{W}\hat{g}(\theta)$ where the moment function $g(\theta)$ is described below, and $\hat{W}$ is the estimated optimal weighting matrix from the second step. (For the first step I use $\hat{W} = I$.)

Departure Time Moments. The first 49 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. Let k index the 5-minute-step departure time grid between -120 and $+120$ minutes relative to the rate profile peak. Denote $P_{ik}^{DT}(\theta, p_{it})$ the probability that the kth departure time is optimal when departure time and route pricing is p_{it}. In the data, define $\widetilde{P}_{ik}^{DT}(pre)$ and $\widetilde{P}_{ik}^{DT}(post)$ the fractions of trips starting in a 5-minute bin around the kth departure time for i in pre- and post- periods, respectively. The k-th moment is:

$$g_i^k(\theta, p_{it}) = (-1)^{1-T_i^{DT}}\left[\left(\widetilde{P}_{ik}^{DT}(post) - \widetilde{P}_{ik}^{DT}(pre)\right) - \left(P_{ik}^{DT}(\theta, p_{it}) - P_{ik}^{DT}(\theta, 0)\right)\right],$$

where T_i^{DT} is an indicator for departure time charges.

Route Moments. Ten moments match route choice market shares during five periods (before the experiment, and four weeks during the experiment) indexed by $t = 1, \ldots, 5$ and in two treatment groups (early and late charges).

Denote $P_{it}^A(\theta, p_{it})$ the probability to take the detour route (not intersect the congestion area) in time period t when pricing is p_{it}. In the data, define $\widetilde{P}_{it}^A$ the fraction of days when commuter home-work trips do not intersect the congestion area for individual i, which depends on i's treatment group. For $t = 1, \ldots, 5$, the route moments are:

$$g_i^{49+t}(\theta, p_{it}) = T_i^{Early} \cdot \left[\widetilde{P}_{it}^A - P_{it}^A(\theta, p_{it})\right]$$
$$g_i^{54+t}(\theta, p_{it}) = (1 - T_i^{Early}) \cdot \left[\widetilde{P}_{it}^A - P_{it}^A(\theta, p_{it})\right].$$

SM.5 Parameter Sensitivity Measure

Table SM4 reports the estimated sensitivity measure Λ from Andrews et al. (2017), scaled by the 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 measure is $\hat{\Lambda} = \left(\hat{S}'\hat{W}\hat{S}\right)^{-1}\hat{S}'\hat{W}\,\mathrm{diag}\,(\hat{\sigma})$ where $\hat{S}$ is the Jacobian evaluated at the estimated parameters, $\hat{W}$ is the optimal weighting matrix, and $\hat{\sigma}$ is the vector of bootstrap standard deviation of moment j.

SM.6 Road Technology Invariance Result

Conditional on the relationship (6) estimated on a representative sample, the impact of an additional trip on total driving time in Bangalore is invariant to the aggregate volume of traffic in Banglore, and it is invariant to the sample size used to estimate the road technology relationship.

The key intuition is that equation (6) depends on normalized density, so it is invariant to the true aggregate volume of traffic. Then, imagine that the aggregate volume is twice as large as initially believed. Then the impact of a single trip on travel delay will be twice as small. However, it will affect twice as many other commuters, so the impact on total time is not affected.

Using the notation from section 4.3, let $Q = (q(h,K))_{h,k}$ denote the pattern of departures, where $q(h,K)$ is the mass of trips of length K starting at h, based on a sample of N trips. Let $x = (x(h))_h$ denote the instantaneous travel delay profile, and $d = (d(h))_h$ the density profile. Similar to equation (6), assume that instantaneous delay satisfies $x(h) = \lambda_0 + \lambda_1 d(h)/N$ where N is the number of trips in the sample.

Proposition 5. *The marginal effect of an additional trip on total travel time does not depend on the sample size used to construct Q.*

Proof. Let $d(h', Q)$ denote density at time h' as a function of the pattern of departures, and $d(Q) = (d(h', Q))_{h'}$.

Travel times are uniquely determined by the instantaneous travel delay profile, which depends on normalized density. Hence, we can write average travel times as a function $\overline{T}\left(\frac{d(Q)}{N}\right)$. Note that total travel time in the city is $N\overline{T}$.

For every h', $d(h', Q)$ is homogeneous of degree 1 in Q. Consequently, the partial derivative $d_{h,K}(h', Q)$ with respect to the mass of trips with length K starting at h is degree 0 in Q, i.e. it does not depend on the sample size used to compute Q.

Consider adding a trip of length K that starts at h and denote the pattern of departures by $Q + \mathbf{1}(h,K)$. The change in *total* travel time is

$$N\left(\overline{T}\left(\frac{d(Q+\mathbf{1}(h,K))}{N}\right) - \overline{T}\left(\frac{d(Q)}{N}\right)\right) \approx N\frac{\partial \overline{T}}{\partial \mathbf{1}(h,K)} = N\sum_{h'}\frac{\partial \overline{T}}{\partial h'}\frac{d_{h,K}(h',Q)}{N}$$

The last term does not depend on N because neither $\frac{\partial \overline{T}}{\partial h'}$ nor $d_{h,K}(h', Q)$ depend on N. □

SM.7 Policy Simulations

For policy simulations I use a 5-minute departure time grid from 5am to 2pm. Each simulation has 3040 agents, with each real study participant replicated with 10 independent random draws of ideal arrival times from the distribution recovered with non-negative least squares (section 7.3). The vector of ideal arrival times is re-sampled during bootstrapping. Thus, the confidence intervals include uncertainty due to numerical simulation. Benchmark results are robust to using 10× more agents.

For the two-route equilibrium model, I assume double the volume of trips, so that on average the volume of trips per route remains the same.

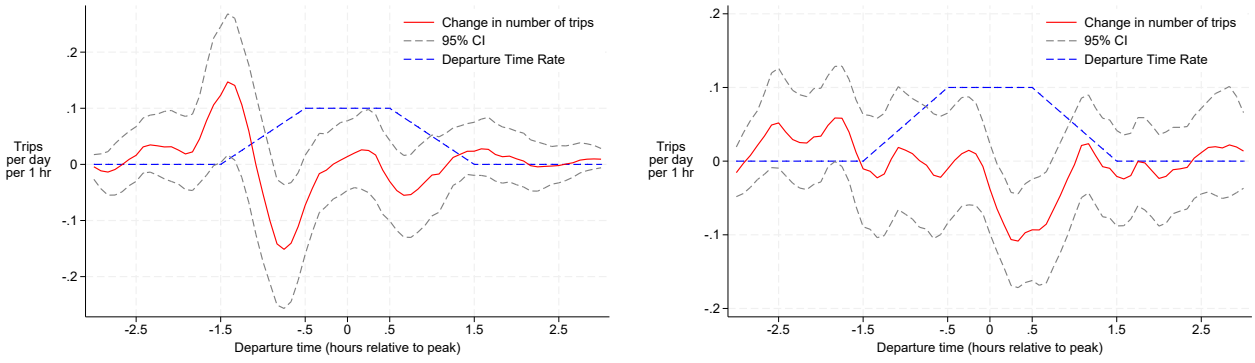
I use a nested logit model for the equilibrium model with an extensive margin decision. The outer nest has two options, taking the trip ($z = 1$) and not taking the trip ($z = 0$). Trips are valuable: a commuter not making a trip incurs a cost proportional to trip length $\omega_i = \omega \cdot \mathrm{K}_i/\overline{\mathrm{K}}$. Expected utility is given by:

$$\mathbb{E}u_i(x, h, h_i^A) = \begin{cases} \mathbb{E}v(h, \mathcal{T}_i(h), h_{it}^A) - p_{it}(h) + \varepsilon_{it}(1, h) & z = 1 \\ -\omega_i + \varepsilon_i(0, h) & z = 0 \end{cases}$$

where $\varepsilon_i(z, h)$ follow a type-1 extreme value distribution with correlation within each value of z, with logit scale parameter η for the trip (upper) nest. The congestion pricing experiment was not designed to estimate the extensive margin trip elasticity.

SM.8 Supplementary Material: Figures

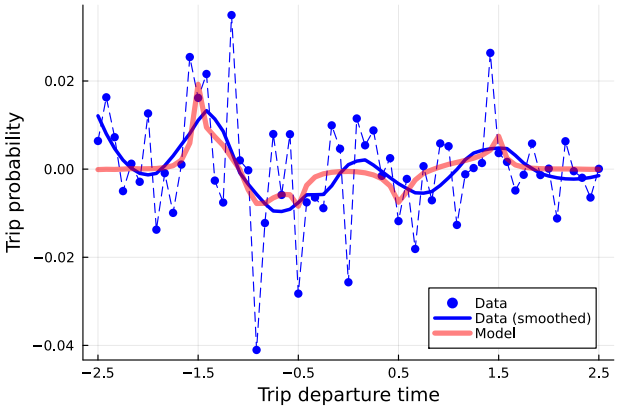
Figure SM2. Impact of Departure Time Charges on Departure Times (Commuting Trips)



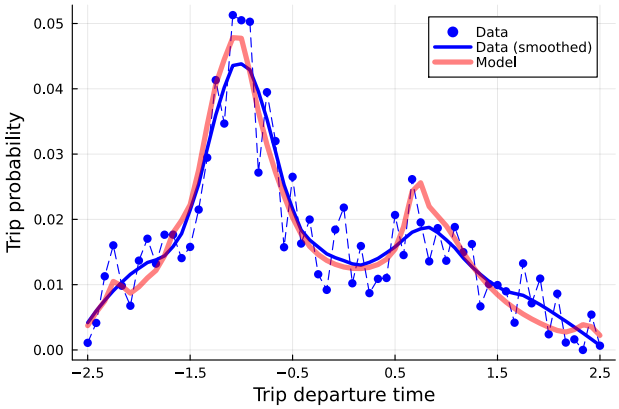
Panel (A) Morning Peak Panel (B) Evening Peak

Notes: Version of Figure 2 restricting to regular commuters and trips between home and work (both ways).

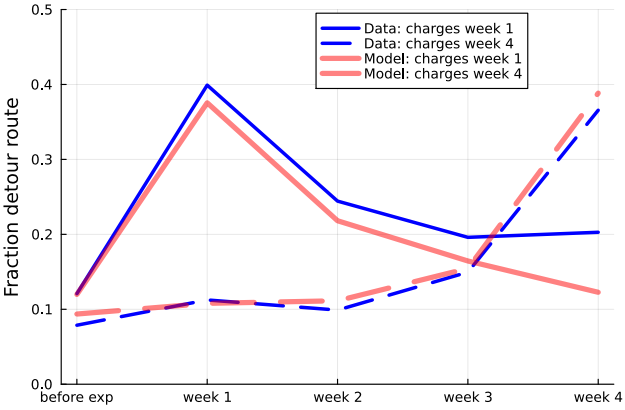
Figure SM3. Travel Demand Model Fit



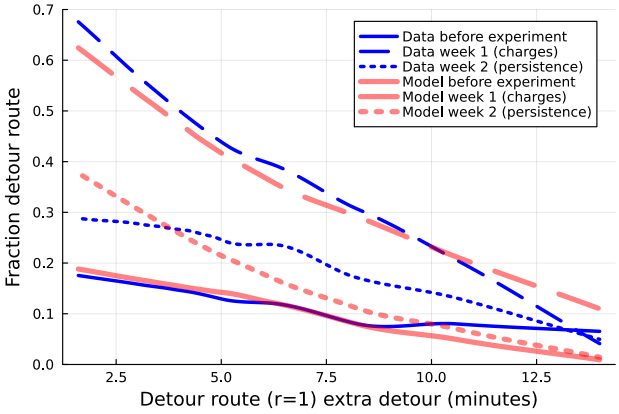
Panel (A) Departure Time Difference in Differences



Panel (B) Departure Time Control Post



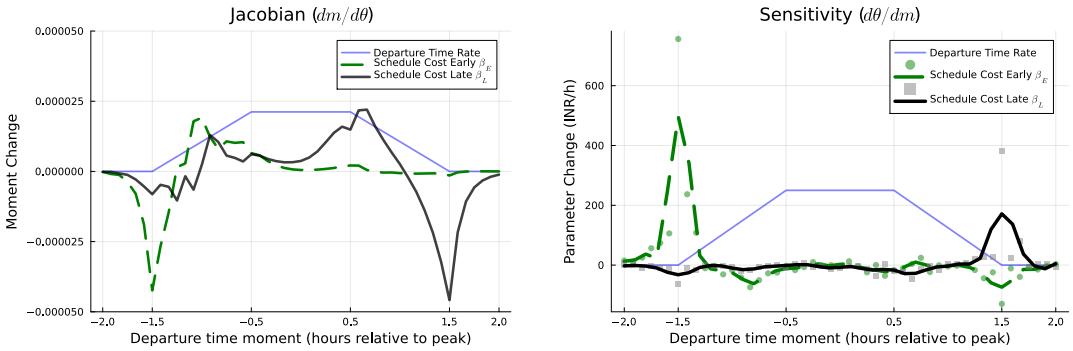
Panel (C) Detour Route Usage



Panel (D) Detour Route Usage Heterogeneity

Notes: This figure shows in- an out-of-sample fit for the estimated travel demand model. Panel A plots the departure time moments that correspond to the difference-in-differences (treated vs. control, during vs. before), the analogue of Figure SM2. Panel B shows the probability density of departure time in the control group during the experiment (Post). These moments are not directly targeted in the estimation (however, the ideal departure time distribution inversion routine depends on the distribution of departure time *before* the experiment). Panel C shows the dynamic route choice moments, the analogue of Figure 3. Panel D shows detour route choice heterogeneity by the amount of detour (in minutes), for the "early" treatment group, which receives charges in week 1. This is the analogue of Kreindler (2023), Figure A.4., and these moments are not targeted in estimation. For all graphs, the model is indicated by thicker, red lines.

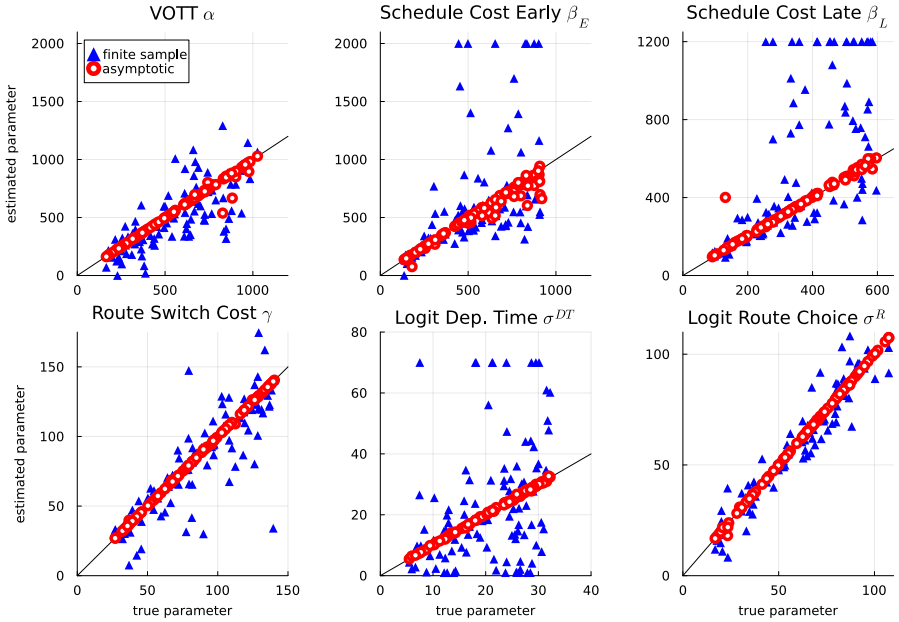
Figure SM4. Travel Demand Model: Understanding Identification



Panel (A) Jacobian: d moment/d parameter Panel (B) Sensitivity: d parameter/d moment

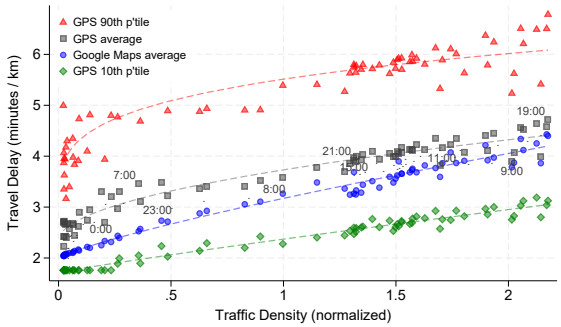
Notes: Moments are defined as $g_j(\theta) = s_j(\theta) - s_{j,data}$. Panel A plots the partial derivatives $dg(h,\theta)/d\beta_E$ and $dg(h,\theta)/d\beta_L$ for each departure time moment $g(h,\theta)$. Panel B plots the scaled sensitivity measure from Andrews et al. (2017) quantifying the change in the estimated early and late schedule cost parameters $\widehat{\beta}_E$ and $\widehat{\beta}_L$ given by one standard deviation change in each of the 49 departure time moments, as well as the LOESS fit. See Supplementary Material SM.5 for definitions.

Figure SM5. Travel Demand Model Numerical Identification Check and Finite Sample Properties

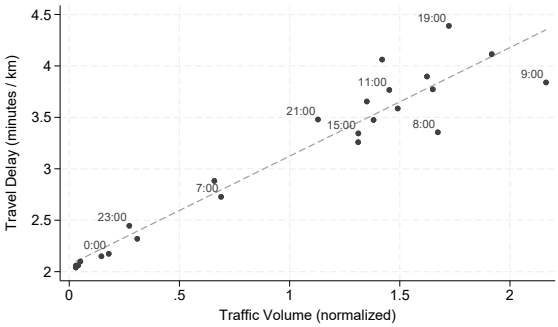


Notes: This figure compares true random parameters and the estimated parameters from simulated data, under two scenarios. In the "asymptotic" scenario (red circles) the simulated data has exact (route and departure time) choice probabilities. In the "finite sample" scenario (blue triangles) the simulated data has random choices and I use exactly the same data set size as in the real data (the number of observations per commuter). Simulations are based on 100 random parameters independently drawn between 25% and 175% of the benchmark estimated values. For each set of parameters, I first invert the f_i^A distributions from pre-experiment (real) data, then use it to simulate data. I then estimate the model on the simulated data using one random starting condition that is independent of the parameters used to simulate the model. Each graph shows the estimated parameter on the Y axis, and the true parameter on the X axis. Outlier values are censored. The diagonal line is identity. See also Table SM.XII.

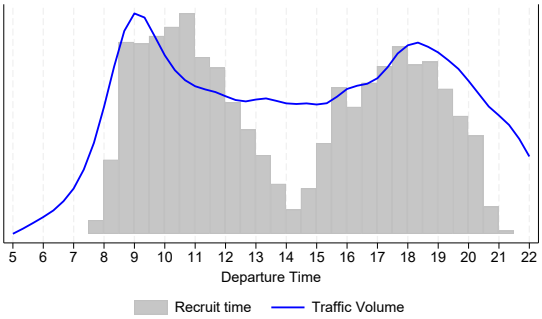
Figure SM6. Road Technology Estimation Robustness Checks



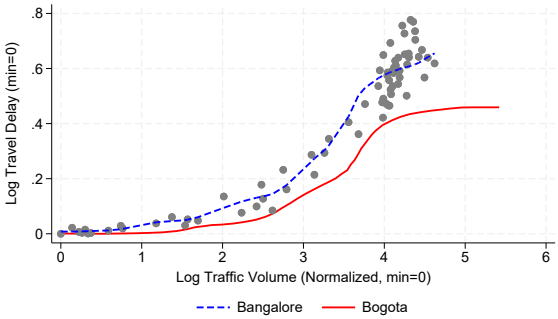
Panel (A) Travel Delay from GPS Data and Google Maps



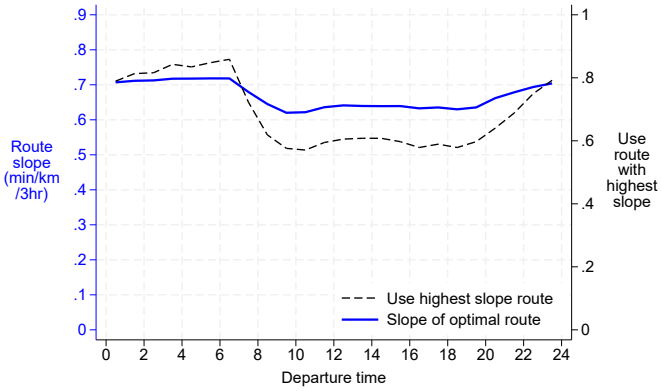
Panel (B) Travel Delay and Traffic Volume



Panel (C) Recruitment Time and Trip Time Distributions



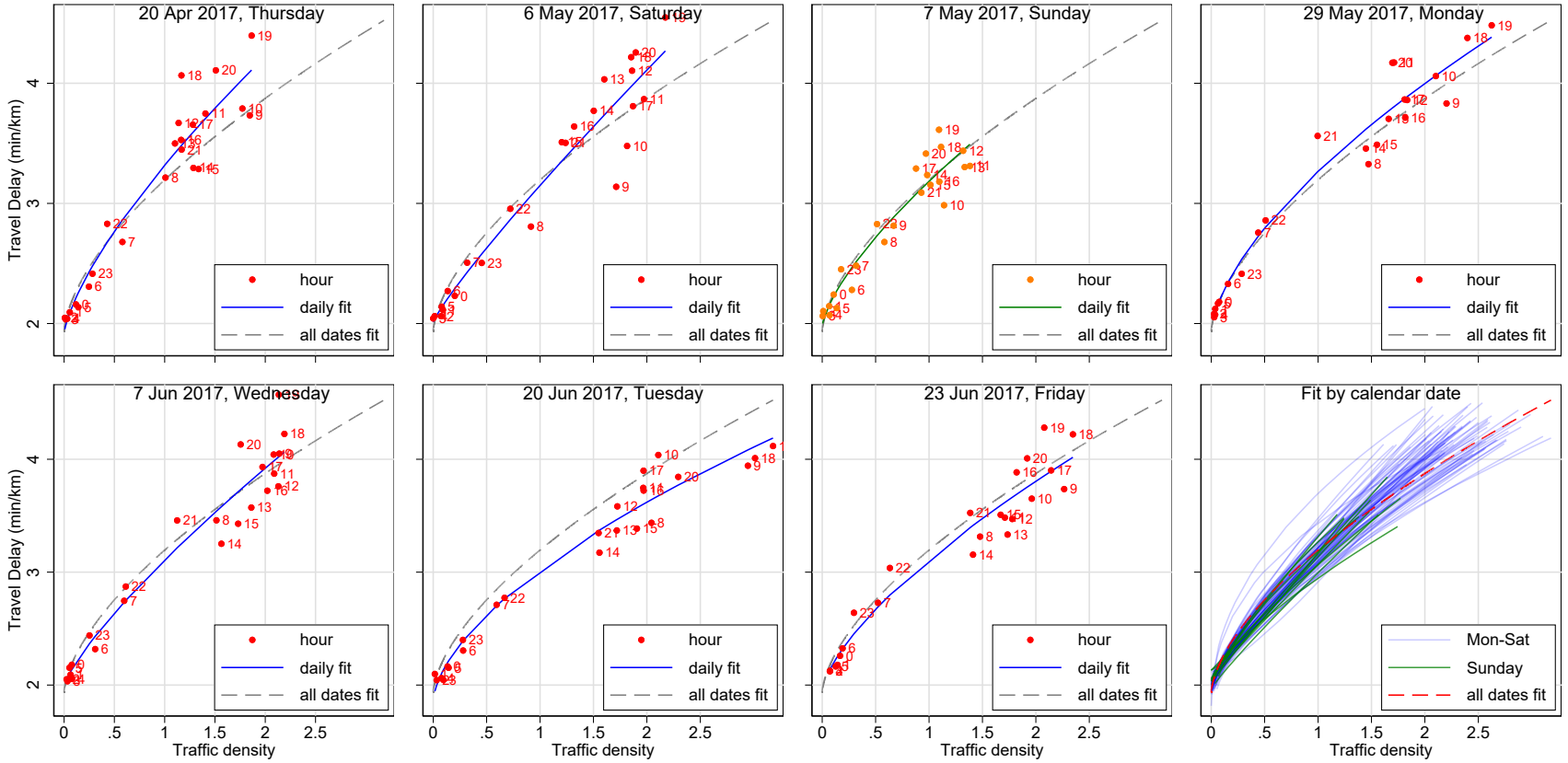
Panel (D) Comparison with (Akbar and Duranton, 2017)



Panel (E) Peak-hour Use of Lower Externality Routes

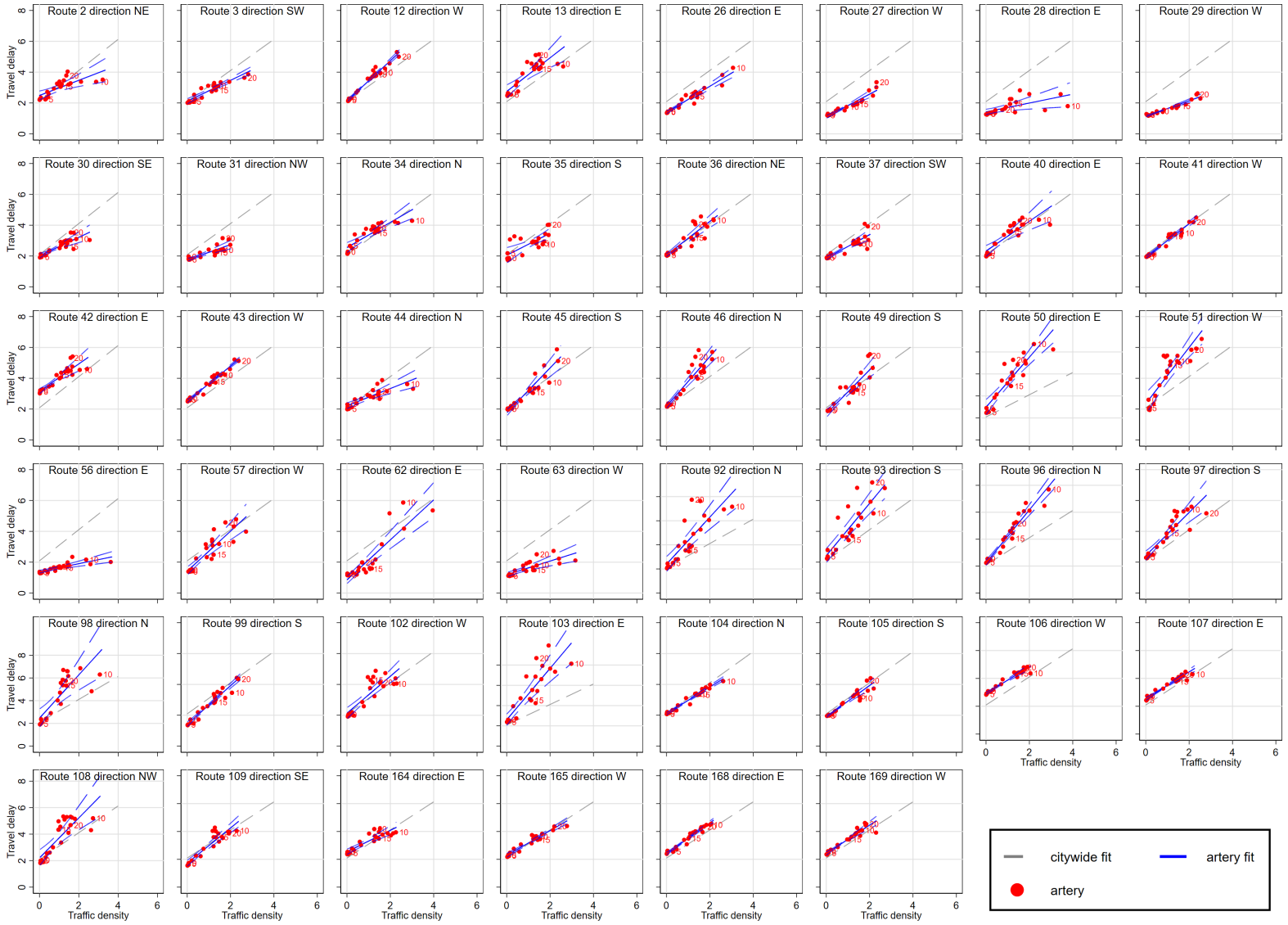
Notes: Panel A uses travel delay from GPS trips to replicate Figure 4, including percentiles. The sample is all weekday trips more than 2 kilometers long, without stops along the way, and with a trip diameter to total length ratio above 0.6 (the 25th percentile). For each hour-day I compute the average delay over trips starting in that interval. Panel B replicates Figure 4 with "volume," the normalized number of trips starting each hour on the X axis. 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.) Panel E describes peak-hour substitution towards routes with less steep travel time profiles. For each commuter in the experimental sample, I query from Google Maps the entire travel time profile for every route that is optimal at some departure time. For each route I compute its slope, the change in travel delay between 6:30 and 9:30 am. The right axis (black dashed line) plots the fraction of commuters for whom their highest slope route is fastest at departure time h. The left axis (blue solid line) plots the average slope of the optimal route at h.

Figure SM7. Road Technology at the Daily Level



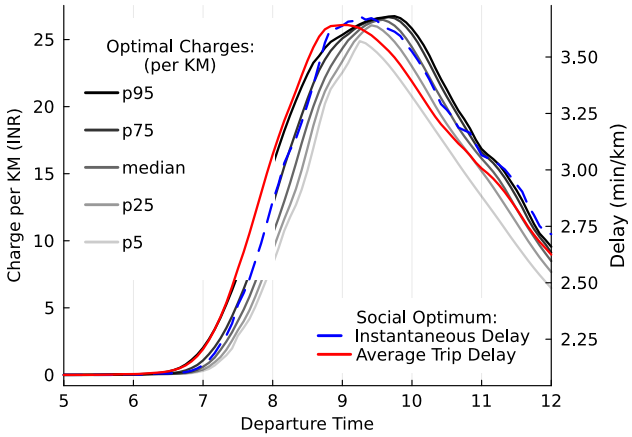
Notes. These graphs replicate Figure 4 panel A by date. The first 7 panels show the relationship between hourly GPS traffic volume and Google Maps travel delay for 7 randomly chosen calendar dates (one for each day of the week). The last panel overlays the predicted fit for all calendar dates in the sample. The sample is calendar dates with above-median number of GPS trips (at least 571 trips per day). Travel delay and traffic density at the day d and hour h level correspond to column 3 in Table III. Each fit is a power fit as in column 2 in Table III.

Figure SM8. Road Technology on Major Arteries

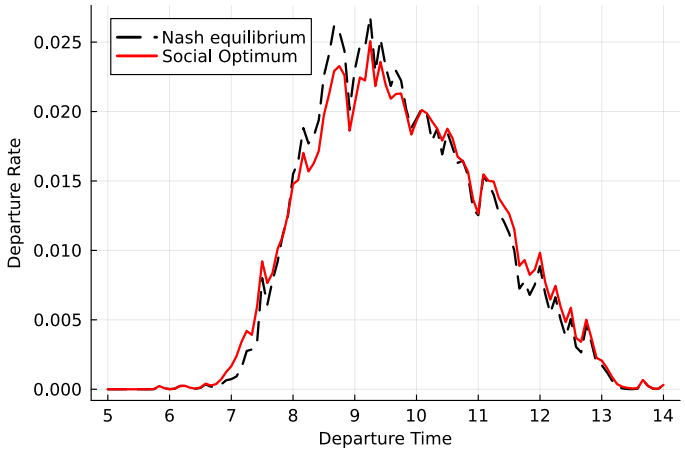


Notes. These graphs replicate Figure 4 for major arteries depicted in Kreindler (2023), Figure A.1., separately by direction. The Y axis is average Google Maps travel delay for that road segment. To compute traffic density at the artery level, I define a buffer area around each artery. I then count the number of GPS trips that travel along the artery in each direction for each time of day, excluding short trips that intersect the artery for less than 200 meters (which I assume correspond to cross-traffic). I obtain 268,292 trip segments on the 46 arteries. 95% confidence intervals based on Newey-West standard errors with a 3-hour lag also reported.

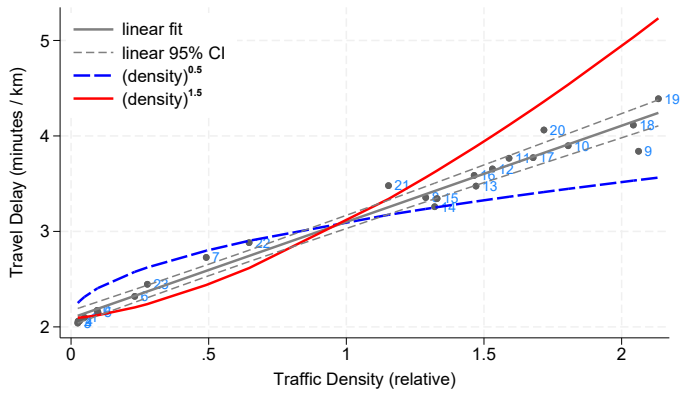
Figure SM9. Policy Counterfactual Additional Results



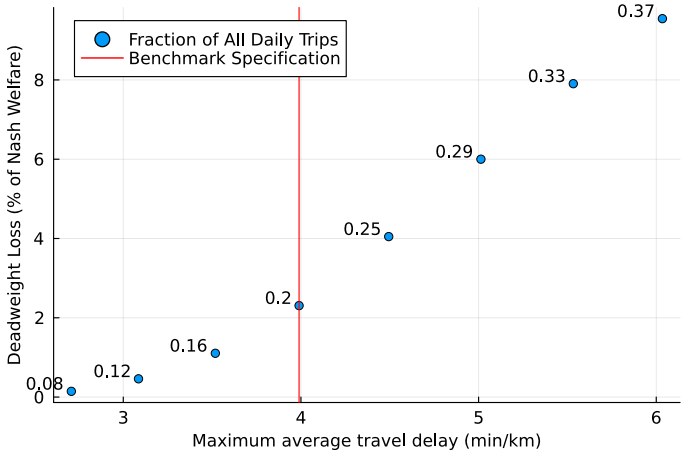
Panel (A) Optimal Congestion Charges



Panel (B) Departure Volume



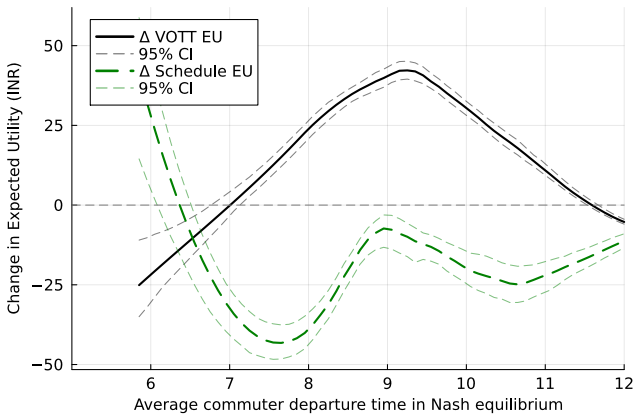
Panel (C) Non-linear Road Technologies



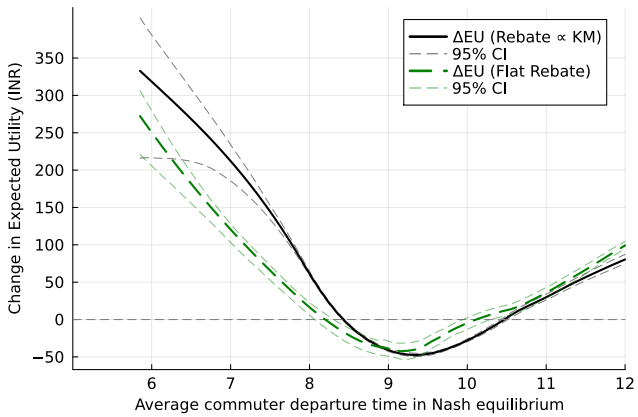
Panel (D) Varying the Total Volume of Trips

Notes. Panel A plots percentiles of the optimal charges (equal to marginal social cost) around the social optimum. For comparison, I plot the average trip delay (red, solid line) as in Figure 5, and the instantaneous travel delay (blue, dashed line). Panel B plots the rates of trip departure rates in the Nash equilibrium and in the social optimum. Panel C overlays the alternate road technologies used in panel D of Table IV, over the benchmark road technology (Figure 4). I use the estimated λ_0 and λ_1 from the benchmark linear equation 6 and only vary ν. Panel D shows equilibrium peak average travel delay (X axis) and welfare gain from optimal pricing (Y axis) when varying the total volume of trips used in the simulation. The number near each point is the assumed fraction of all daily trips that happen during peak time. In the benchmark specification, I assume this is 0.2.

Figure SM10. Decomposing Gains and Losses in the Social Optimum



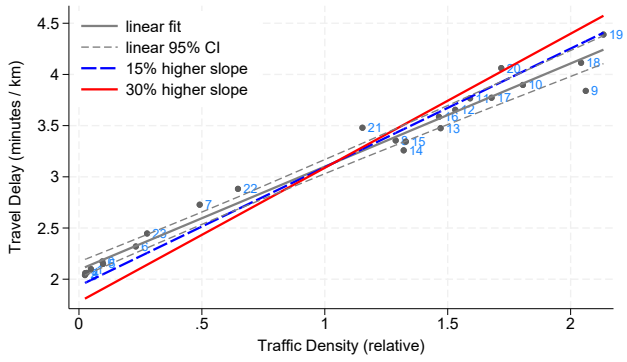
Panel (A) Real Changes: Travel Time and Schedule Costs



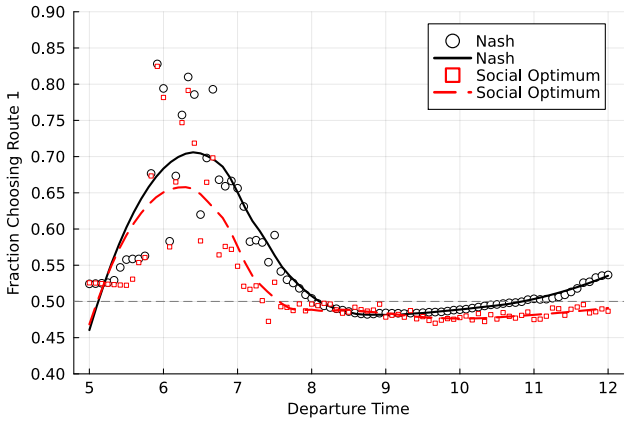
Panel (B) Net of Charges and Rebates

Notes. For each commuter and ideal arrival time h_i^A, the X axis is the average departure time $h_i = \mathbb{E}h(h_i^A)$ in Nash. Panel A plots the Nash–social optimum difference in $-\mathbb{E}\alpha T(h_i)$ vs h_i (black, solid line) and in $-\mathbb{E}\beta_E|h_i + T(h_i) - h_i^A|_- + \beta_L|h_i + T(h_i) - h_i^A|_+$ vs h_i (green, dashed line). Panel B plots average expected utility change vs h_i when commuters receive a rebate that is proportional to trip length (black, solid line) or constant (green, dashed line).

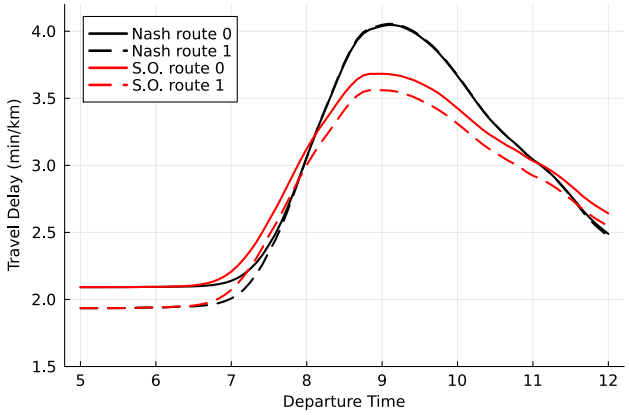
Figure SM11. Policy Counterfactual in Two Route Equilibrium Model



Panel (A) Road Technology for Two Routes



Panel (B) High-Externality Route Choice Prob.



Panel (C) Average Travel Delay by Route

Notes: This Figure describes the two-route equilibrium (panel E of Table IV). Panel A overlays the alternate high-externality route road technologies over the benchmark road technology (Figure 4). Panel B plots the probability of taking the high-externality route by departure time, in the Nash equilibrium and in the social optimum, in the two-route model where one route has 15% higher slope. Panel C replicates Figure 5 by route for the two-route model where one route has 15% higher slope.

SM.9 Supplementary Material: Tables

Table SM.I. Descriptive Statistics about Travel Behavior

Panel A. Trip Characteristics						
	Median	Mean	Std. Dev.	10 Perc.	90 Perc.	Obs.
Total Number of Trips	1.00	1.00	[0.00]	1.00	1.00	51,424
Number of Trips per Day	2.86	3.15	[1.16]	1.90	4.86	497
Median trip duration (minutes)	24.50	27.47	[12.82]	15.20	42.60	497
Median trip length (Km.)	5.93	7.19	[4.67]	2.92	13.36	497
Panel B. Commute Destination Variability						
	Median	Mean	Std. Dev.	10 Perc.	90 Perc.	Obs.
Regular Commuter	1.00	0.76	[0.43]	0.00	1.00	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	0.92	0.86	[0.16]	0.60	1.00	378
Panel C. Departure Time Variability						
(Standard Deviation of the Departure Time in hours)						
	Median	Mean	Std. Dev.	10 Perc.	90 Perc.	Obs.
First Trip (AM)	1.27	1.24	[0.50]	0.54	1.82	496
Last Trip (PM)	1.72	1.71	[0.50]	1.04	2.36	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	0.95	[0.63]	0.28	1.78	322

Notes: This table reports summary travel behavior statistics for the experimental sample of 497 commuters. See section 5.1 for the definition of home and work locations and of regular commuter. In panel C, I compute the within-commuter variation in departure times for different classes of trips.

Table SM.II. Experimental Design

Panel A. Treatment Strata

Strata			*Departure Time Sub-treatment*			
Route Eligibility	Car or Moto	Daily KM	High Rate	Low Rate	Info	Control
Eligible	Car	Low	3/8	1/8	2/8	2/8
Eligible	Car	High	1/8	3/8	2/8	2/8
Eligible	Moto	Low	3/8	1/8	2/8	2/8
Eligible	Moto	High	1/8	3/8	2/8	2/8
Ineligible	Car	Low	1/12	3/12	4/12	4/12
Ineligible	Car	High	3/12	1/12	4/12	4/12
Ineligible	Moto	Low	1/12	3/12	4/12	4/12
Ineligible	Moto	High	3/12	1/12	4/12	4/12

Panel B. Treatment Timing

			Treatment by Week in Experiment			
Route Eligibility	*Dep. Time Timing*	*Dep. Time Sub-Treatment*	1	2	3	4
Eligible	Late	High rate	R	H	H	H
		Low rate	R	L	L	L
		Information	R	I	I	I
		Control	R	C	C	C
Eligible	Early	High rate	H	H	H	R
		Low rate	L	L	L	R
		Information	I	I	I	R
		Control	C	C	C	R
Ineligible	Late	High rate	I	H	H	H
		Low rate	I	L	L	L
		Information	I	I	I	I
		Control	C	C	C	C
Ineligible	Early	High rate	H	H	H	I
		Low rate	L	L	L	I
		Information	I	I	I	I
		Control	C	C	C	C

Notes. There were eight strata in the experiment, all combinations of participants eligible or ineligible for the route charge, car or non-car (motorcycle or scooter) users, and participants with high or low daily travel distance in the baseline period. Departure time sub-treatment probabilities are given in panel A. There are eight route sub-treatments: all combinations of high/low charges, short/long detour, and early/late. All have equal probabilities. Sub-treatment are cross-randomized (see Kreindler, 2023, section A.6.). Treatment timing is presented in panel B. The letter R corresponds to the route treatment. The letters H, L, I and C respectively correspond to high-rate, low-rate, information and control in the departure time treatment.

Table SM.III. Experimental Participant Sample Representativeness

	(1)	(2)	(3)	(4)	(5)	(6)
	In Experiment		Not in Experiment		Difference	
	Mean	[SD]	Mean	[SD]	in SD units	N
Panel A. All Respondents Approached						
Male respondent	0.98	[0.13]	0.97	[0.17]	0.09**	8,231
Age	33.3	[8.2]	35.2	[8.7]	-0.21***	8,231
Car driver	0.30	[0.46]	0.41	[0.49]	-0.24***	8,227
Log vehicle price (residual)	10.5	[0.4]	10.5	[0.4]	-0.00	7,188
Panel B. Survey Respondents						
Log income	9.96	[0.71]	9.91	[0.73]	0.07	2,656
Stated Daily Travel (Km/day)	47.1	[24.0]	45.1	[25.1]	0.08*	4,427
Stated Value of Time (Rs/hr)	206.0	[138.9]	189.0	[151.3]	0.11*	1,001
Stated Schedule Flexibility (min)	20.0	[10.9]	18.7	[12.0]	0.11*	952

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Business owner or manager	Accountant, Teacher, Doctor	Software and IT	Engineers, Technical	Office staff	Manual jobs	Mobile professions	Student	Others, Retired	Total
Panel C. Survey Respondents										
In Experiment	16.7	7.5	10.3	14.3	15.4	8.4	15.6	9.0	2.9	455
Not in Experiment	15.6	6.2	10.1	11.2	18.1	9.5	12.0	13.4	3.9	2,458

Notes. These results describe respondent selection into experiment by comparing the experimental sample (497 respondents) to the entire sample of eligible commuters approached in gas stations by the survey team (panel A) and to the full survey sample (panels B and C). The sample in Panel A is all respondents approached in gas stations, excluding ineligible respondents. Weights are used to (a) account for missing data for each variable, and (b) to adjust for the estimated $\sim 52\%$ ineligible respondents among survey refusals (for refusals, $7,218$ respondents did not complete the eligibility filter, and I assume the same proportion were ineligible). Gender, age and car driver variables are visually assessed by the surveyor for all respondents. Vehicle value (residual) is imputed based on vehicle type (car/motorcycle), make and model, using pricing data scrapped from a used-vehicles website in Bangalore, residualized on a "car" dummy. Monthly income is self-reported during the recruitment survey (the respondent is handed the tablet to enter the amount confidentially – the surveyor never sees the amount), truncated at $100,000$ INR ($\sim 1,300$ USD). Occupation is self-reported during the recruitment survey. Value of time and schedule flexibility are based on choices in hypothetical scenarios in a follow-up phone survey; for details, see Kreindler (2023), section A.4.2. The difference in SD units includes significance levels from a (weighted) regression of the row outcome variable on an indicator for being in the experiment. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table SM.IV. Experimental Balance Checks

		Departure Time Treatments								Route Treatment			
		Information		Low Rate		High Rate		Obs.	Control Mean	Route Early		Obs.	Control Mean
			(S.E.)		(S.E.)		(S.E.)				(S.E.)		
(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	-0.93	(0.93)	1.26	(1.01)	-0.09	(1.07)	497	33.20	-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	0.01	(0.10)	-0.02	(0.14)	-0.07	(0.14)	411	10.11	-0.09	(0.12)	211	10.24
(6)	Frac days with good GPS data	-0.00	(0.03)	-0.02	(0.03)	-0.00	(0.03)	497	0.41	0.00	(0.03)	254	0.42
(7)	Frac days present at work	0.02	(0.03)	0.00	(0.03)	-0.03	(0.04)	497	0.69	-0.03	(0.03)	254	0.79
(8)	Number of trips per day	-0.11	(0.11)	-0.03	(0.13)	0.01	(0.14)	497	1.24	-0.00	(0.12)	254	1.15
(9)	Total distance per day (Km.)	-0.43	(0.69)	-0.19	(0.83)	0.34	(0.88)	497	8.26	0.19	(0.83)	254	8.79
(10)	Total duration per day (min)	-2.73	(3.03)	-1.25	(3.55)	1.43	(3.84)	497	35.27	0.49	(3.50)	254	35.49
(11)	Total D.T. hypothetical rate per day	-0.40	(3.82)	-0.40	(3.89)	0.03	(4.20)	497	38.38	-0.51	(3.78)	254	37.90
(12)	Total Route hypothetical rate per day	-2.27	(3.18)	-2.83	(4.23)	-0.40	(4.88)	497	23.83	0.70	(5.61)	254	50.73
(13)	Joint Significance Test F-stat	0.25								0.02			
(14)	Joint Significance Test P-value	0.86								0.90			

Notes. This table shows experimental balance checks for the departure time and route 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 "Route Treatment" columns, the sample is restricted to 254 participants who receive the route treatment, and the dependent variable is whether the respondent was assigned to the "early" route sub-treatment (to receive the route 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. Robust standard errors are shown in parentheses. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table SM.V. GPS Data Quality at Daily Level (Attrition Check)

	(1)	(2)
Treatment	*Departure Time*	*Route*
Commuter FE	X	X
High Rate × Post	0.02	
	(0.05)	
Low Rate × Post	0.00	
	(0.05)	
Information × Post	0.00	
	(0.04)	
Route Charges		0.02
		(0.04)
Post	0.08**	0.14***
	(0.03)	(0.04)
Observations	24,779	9,809
Control Mean	0.76	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, the sample in column 1 is restricted to the departure time treatment period, either the first or the last three weeks. The sample in column 2 is restricted to the first week in the experiment. All specifications include respondent and study cycle fixed effects. Standard errors are clustered at the respondent level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table SM.VI. Impact of Departure Time Charges on Daily Total Hypothetical Rate: Commuting Trips

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Time of Day	AM & PM		AM			PM	
		all	pre peak	post peak	all	pre peak	post peak
Commuter FE	X	X	X	X	X	X	X
Sample:	*Regular Commuters, Home-Work and Work-Home Trips*						
Charges × Post	-7.95***	-3.77**	-3.00*	-0.76	-4.18**	-0.88	-3.30***
	(2.89)	(1.90)	(1.56)	(1.20)	(1.67)	(1.23)	(1.08)
Post	-1.74	-0.74	-1.29	0.55	-1.00	-0.69	-0.31
	(2.65)	(1.74)	(1.30)	(1.36)	(1.61)	(1.18)	(1.06)
Observations	12,116	12,116	12,116	12,116	12,116	12,116	12,116
Control Mean	40.81	23.37	14.27	9.10	17.44	9.15	8.29

Notes: This table reports the impact of departure time charges on daily total hypothetical rates for regular commuters and commuting trips, separately by time interval. The sample of users and days, and the specifications, are the same as in Table I, panel B, further restricted to regular commuters and direct trips between their home and work locations (in either direction). Columns (3) and (6) restrict to trips before the peak, i.e. the mid-point of the rate profile. Columns (4) and (7) restrict to trips after the peak. Kreindler (2023), Table A.3., reports these results for 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 SM.VII. Impact of Route Charges on Detour Route Usage

	(1)	(2)	(3)	(4)
Outcome	*Use Detour Route*		*Number of Trips Today*	
Commuter FE	X	X	X	X
Panel A. All Commuters				
Treatment: Early × week 1	0.27***	0.26***	0.02	0.02
	(0.05)	(0.05)	(0.07)	(0.07)
Persistence: Early × week 2		0.09**		-0.09
		(0.04)		(0.08)
Observations	5,235	6,038	9,809	11,016
Control Mean (week 1)	0.11	0.11	0.73	0.73
Panel B. Commuters Who Used Detour at Baseline				
Treatment: Early × week 1	0.41***	0.41***	0.03	0.02
	(0.08)	(0.08)	(0.13)	(0.13)
Persistence: Early × week 2		0.13*		-0.01
		(0.07)		(0.13)
Observations	2,369	2,718	3,508	3,940
Control Mean (week 1)	0.18	0.18	0.87	0.87

Notes: This table reports difference-in-differences impacts of the route treatment on trip and daily outcomes. In the first two columns, an observation is a commuting trip between home and work, and the outcome is whether the commuting trip used a detour route (defined as any route that avoids the congestion area). The last two columns, an observation is a commuter, day combination, and the outcome is the total 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. The sample is restricted to 243 participants in the route treatment. In the first two columns, only frequent commuters are included. In panel B, the sample is restricted to commuters who used a detour route between home and work at least once before the experiment. All specifications include respondent and study cycle fixed effects. The mean of the outcome variable in the control (late) group in week 1 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 SM.VIII. Impact of Route Charge Sub-Treatments on Daily Outcomes

	Hypothetical Route Charges	
	(1)	(2)
Treated × High Rate	-41.1***	
	(13.1)	
Treated × Low Rate	-21.5	
	(13.3)	
Treated × Short Detour		-43.0***
		(15.4)
Treated × Long Detour		-26.2
		(17.8)
Observations	6,129	3,693
Commuters	243	148
Control Mean	117.1	122.7
P-val Equal Sub-treatment Effects	0.30	0.48

Standard errors in parentheses
* p<.10, ** p<.05, *** p<.01

Notes: This table reports difference-in-differences impacts of route sub-treatments on daily total hypothetical route charges. The sample in column 1 is the same as in Table SM.VII, covering the period before and during the first week in the experiment. In column 2 the sample is restricted to 148 route treatment 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). The outcome is total daily hypothetical route charges; higher values indicate lower detour usage. Standard errors in parentheses are clustered at the respondent level. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$

Table SM.IX. Travel Demand Estimates: Additional Results

	(1) Static Route Choice	(2) Asymmetric switching cost	(3) Time FE	(4) Half attention	(5) Parameters prop. to wage	(6) Single ideal arrival time
β_E: Schedule cost early (INR/hour)	647 [369, 2514]	525 [235, 3000]	488 [244, 2939]	244 [126, 541]	445 [121, 3000]	856 [267, 3000]
β_L: Schedule cost late (INR/hour)	499 [232, 3000]	343 [195, 1324]	358 [197, 1501]	1340 [601, 1966]	630 [188, 2555]	239 [112, 2491]
α: Value of travel time (INR/hour)	2200 [1798, 2772]	552 [198, 998]	494 [173, 1851]	28.6 [0, 161]	388 [121, 956]	645 [224, 1247]
γ: Route switching cost (INR)		55 [34.6, 69.9]	107 [43.7, 146]	39.3 [19.6, 61.8]	80.7 [43.1, 117.9]	81.6 [46.2, 105]
σ^{DT}: Logit departure time	15.6 [1, 142]	18.9 [1, 146.2]	18.2 [1, 134.4]	12 [1, 65.6]	35.2 [1, 333.5]	17.1 [1, 159.9]
σ^{R}: Logit route nest	95.9 [72.0, 130.3]	63.5 [49.4, 84.4]	62.9 [44.3, 96.1]	17.1 [1.11, 42.0]	58.3 [40.2, 85.8]	62.5 [47.0, 86.1]
Obs.	304	304	304	304	304	304
Model Components:						
Route choice model	Static	Dynamic	Dynamic	Dynamic	Dynamic	Dynamic
Fixed discount factor (δ)	-	0.90	0.90	0.90	0.90	0.90
Asymmetric switch cost ($\gamma_{01} = 2\gamma_{10}$)	-	Yes	-	-	-	-
Route Choice Time FE	-	-	Yes	-	-	-
50% share attentive to RCT	-	-	-	Yes	-	-
Moments:						
Departure Time (49)	Yes	Yes	Yes	Yes	Yes	Yes
Dynamic route choice (10)	-	Yes	Yes	Yes	Yes	Yes
Static route choice (2)	Yes	-	-	-	-	-

Notes: Column 1 fits a model with static route choice ($\delta = \gamma = 0$) using only two route choice moments: the fraction using route 1 when not charged during the experiment, and when charged. Column 2 modifies the benchmark model to include asymmetric switching costs parametrized by $\gamma_{01} = \gamma_{10} = 2\gamma$. Column 3 estimates time fixed effects $\eta_1, \eta_2, \eta_3, \eta_4$ that enter route 1 utility on the corresponding weeks during the experiment. Column 4 imposes that each commuter ignores experimental congestion charges with independent probability $p = 0.5$. In column 5, all preference parameters are proportional to w_i, commuter i's self-reported hourly wage. (Note that logit parameters are proportional to w_i and to normalized trip length, i.e. $\sigma_i^{DT} = \sigma \frac{w_i}{\overline{w}} \frac{K_i}{\overline{K}}$). In column 6, I assume that all commuters have the same ideal arrival time that does not vary over time, $h_{it}^A = h^A$. The optimization routine restricts $\alpha, \beta_E, \beta_L \leq 3000$ INR/hour (145 USD/ hour PPP). 95% confidence intervals from 500 Bayesian bootstrap iterations are reported in parentheses.

Table SM.X. Travel Demand Estimation: Discount Factor Robustness

	(1)	(2)	(3)	(4)	(5)
		Varying Discount Factor δ			Estimate δ
β_E: Schedule cost early (INR/hour)	339 [203, 813]	385 [214, 1179]	534 [255, 3000]	542 [255, 3000]	540 [258, 3000]
β_L: Schedule cost late (INR/hour)	3000 [1227, 3000]	3000 [889, 3000]	346 [197, 1426]	345 [195, 1594]	346 [196, 1463]
α: Value of travel time (INR/hour)	79.6 [0, 724]	81.6 [0, 598]	595 [249, 996]	618 [286, 1004]	612 [308, 953]
γ: Route switching cost (INR)	44.5 [24.1, 99.0]	50 [31.5, 100]	81.8 [53.1, 105]	77.6 [48.7, 101]	80.5 [54.6, 105]
σ^{DT}: Logit departure time	16.1 [1, 188]	16.3 [1, 157]	18.5 [1, 182]	17.2 [1, 138]	17.7 [1, 147]
σ^{R}: Logit route (upper nest)	23.9 [4.88, 78.78]	34.1 [9.71, 76.7]	63.4 [49.1, 85.2]	64.4 [51.1, 84.7]	64.3 [52.1, 82.8]
δ: discount factor					0.895 [0.457, 0.990]
Obs.	304	304	304	304	304
Model:					
Dynamic route choice model	Dynamic	Dynamic	Dynamic	Dynamic	Dynamic
Fixed discount factor (δ)	0.0	0.50	0.90	0.99	-
Moments:					
Departure time (49)	Yes	Yes	Yes	Yes	Yes
Dynamic route choice (10)	Yes	Yes	Yes	Yes	Yes
Route choice transition (1)	-	-	-	-	Yes

Notes: Columns 1-4 replicate column 1 in Table II with different assumptions on δ. In column 5 I estimate δ, using an additional moment. This moment measures the transition probability between route 0 and route 1, on average, between weeks 1-2, 2-3, and 3-4 during the experiment. In the data, I define that the commuter uses route 0 if the average weekly route choice of route 0 is strictly below 0.5. The optimization routine restricts $\alpha, \beta_E, \beta_L \leq 3000$ INR/hour (145 USD/ hour PPP). 95% confidence intervals from 500 Bayesian bootstrap iterations are reported in parentheses.

Table SM.XI. Dynamic Route Choice Model Identification

	(1) Full Model			(2) No departure time			(3) Simple Model ($\delta = 0$)		
	α	γ	σ^R	α	γ	σ^R	α	γ	σ^R
Estimated values	594.5	81.8	63.4	561.7	86.4	57.6	808.8	98.9	62.8
Jacobian: Change in Route 1 take-up Due to Change in Parameter									
Before Experiment	-0.09	-0.09	0.19	-0.09	-0.09	0.18	-0.19	-0.06	0.24
Week 1 (Charges)	-0.22	-0.39	0.2	-0.23	-0.4	0.21	-0.4	-0.35	0.25
Week 2 (After Charges)	-0.17	-0.13	0.18	-0.18	-0.14	0.18	-0.38	-0.06	0.19

Notes: This table reports the Jacobian matrix for three moments with respect to three route choice parameters (VOTT α, switch cost γ and logit scale σ^R). The three moments are the route treatment "early" group average detour route usage (1) before the experiment, (2) during week 1 in the experiment (when charges were in effect), and (3) in week 2 (after charges had ended). The first group of columns uses the benchmark model, and the next group uses the dynamic route choice model without departure time (column 4 in Table II). In the last group of columns, I estimate a simple model where a single agent faces the average detour (6.4 minutes) and the average route charge (144 INR), and I assume $\delta = 0$ (see Supplementary Material SM.2). Jacobian entries are divided by the value of the parameter, so they represent the semi-elasticity of the moment with respect to a proportional change in the parameter.

Table SM.XII. Travel Demand Model Finite Sample Properties Check

	(1)	(2)	(3)	(4)	(5)	(6)
	Estimated Parameter					
	$\widehat{\alpha}$	$\widehat{\beta}_E$	$\widehat{\beta}_L$	$\widehat{\gamma}$	$\widehat{\sigma^{DT}}$	$\widehat{\sigma^R}$
(True) Value of time α	1.02***	0.05	0.01	-0.01	0.00	0.00
	(0.11)	(0.09)	(0.18)	(0.01)	(0.01)	(0.00)
(True) Penalty early β_E	0.03	0.78***	-0.05	0.00	-0.01	0.00
	(0.10)	(0.12)	(0.14)	(0.01)	(0.01)	(0.00)
(True) Penalty late β_L	-0.07	0.02	1.65***	-0.01	0.01	0.01*
	(0.15)	(0.18)	(0.34)	(0.01)	(0.01)	(0.01)
(True) Switch Cost γ	-0.17	0.26	0.79	0.95***	-0.10*	0.03
	(0.69)	(0.66)	(1.07)	(0.03)	(0.06)	(0.03)
(True) Logit departure time σ^{DT}	0.48	-0.82	-2.41	-0.02	0.76***	0.21**
	(2.73)	(2.56)	(4.96)	(0.15)	(0.26)	(0.10)
(True) Logit route σ^R	-0.73	-0.41	0.48	-0.01	0.07	1.05***
	(0.84)	(0.77)	(1.83)	(0.05)	(0.05)	(0.03)
Observations	100	100	100	100	100	100

Notes: This table uses simulated data of exactly the same size as the data used in estimation to describe the finite sample properties of the estimation procedure. See notes for Figure SM5. Each column reports results from a quantile (median) regression of the estimated parameter on the vector of true parameters. $^{*}p \leq 0.10$, $^{**}p \leq 0.05$, $^{***}p \leq 0.01$