When Should Carpools in HOV Lanes be Encouraged?

Jonathan E. Hughes[1] and Daniel Kaffine[1*]

Abstract

A variety of policies to encourage carpooling in high-occupancy vehicle (HOV) lanes have been adopted by state and local governments in the US to lower congestion and reduce driving related externalities such as air pollution. We analytically model highway congestion and other vehicle-related externalities. Encouraging carpooling decreases total costs when congestion relief in mainline lanes outweighs increased HOV lane congestion due to carpooling. Importantly, entry of new drivers via induced travel demand can negate the congestion and air pollution benefits of increased carpooling. We explore our theoretical predictions using 10 years of detailed traffic data from Los Angeles to estimate time and route specific marginal external costs. Because costs vary substantially across routes, hours and days, current policies to promote carpooling will often increase social costs.

*The authors thank Tania Barham, Harrison Fell, Craig McIntosh, Edward Morey and seminar participants at the Association of Environmental and Resource Economists summer meeting, LAMETA Montpellier, the NBER Summer Institute, Resources for the Future, Stanford University, the Triangle Resource and Environmental Economics Seminar Series and the University of Pittsburgh for helpful comments.

[1]Department of Economics, University of Colorado at Boulder, Box 0256 UCB, Boulder, CO 80309. Email: jonathan.e.hughes@colorado.edu, Telephone: (303) 735-0220, Fax: (303) 492-8960

1 Introduction

External costs from traffic congestion in the US exceed $120 billion per year (Schrank, Lomax, and Eisele, 2012). While a number of cities have implemented some limited form of congestion pricing, public opposition to tolls has prevented widespread comprehensive adoption of these policies. Instead, policy makers have largely relied on a variety of policy alternatives that center on explicit or implicit subsidies for reduced driving. Among these policies are programs to encourage carpooling in "carpool" or "high-occupancy vehicle" (HOV) lanes.[1]

While a survey of the policy space suggests that encouraging HOV lane use is viewed as a desirable policy goal, it is not clear a priori if increasing HOV lane use reduces total congestion costs. First, forming new carpools decreases mainline congestion, but it also increases HOV lane congestion. On the one hand, HOV lanes are less congested than neighboring mainline lanes, implying a reduction in congestion costs, but on the other hand, there are more people per vehicle in the HOV lane who are affected by additional carpools. Second, reducing congestion in the mainline may entice new drivers to begin driving, eroding congestion relief in the mainline and increasing costs from other vehicle-related externalities. Here, we use a simple economic framework to investigate when and where policies to increase HOV lane use lower *total* congestion and other vehicle related costs. We use our model to study traffic on Los Angeles freeways. Because the wisdom of increasing HOV lane use critically depends on congestion cost, we exploit detailed data on traffic conditions to estimate time-varying marginal external costs of traffic congestion in mainline and HOV lanes.

Similar classes of problems have been well-studied in urban economics. A large literature has explored the properties of second-best congestion tolls

[1] High-occupancy vehicle (HOV) or "carpool" lanes are highway lanes where access is limited to vehicles carrying a minimum number of riders. The Federal Highway administration estimates there are over 150 highways with HOV lanes and over 1,000 HOV lane miles in US metropolitan areas (http://ops.fhwa.dot.gov/freewaymgmt/faq.htm). Policies to encourage carpooling include subsidized or preferential parking, "guaranteed ride home" programs, and informational campaigns (Federal Highway Administration, 2016).

where pricing is only available on a portion of the road network. In particular, the classic two-route problem focuses on two substitute routes and derives optimal policies when policy makers can only encourage or discourage travel on one of the routes (Lévy-Lambert, 1968; Marchand, 1968; Verhoef, Nijkamp, and Rietveld, 1996; Liu and McDonald, 1998). More generally, Hughes and Kaffine (2017) apply a similar approach to a broad class of second-best resource problems where users face congestion costs and substitute across resources with different access costs. A smaller literature has examined issues of HOV lane use. Yang and Huang (1999) is the most related to our approach in that they consider highways with mainline and HOV lanes, where HOV lane commuters face both congestion costs and costs of forming a carpool, and derive optimal lane-differentiated congestion tolls.

This work highlights the critical importance of understanding congestion costs and travel demand on substitute routes throughout the network.[2] For instance, Verhoef, Nijkamp, and Rietveld (1996) and Liu and McDonald (1998) find the optimal second-best policy can be a tax or subsidy depending on the relative congestion costs across routes and the overall elasticity of demand for transportation.[3] Despite the large number of theoretical papers and simulation studies highlighting the importance of these issues, there is relatively little empirical evidence quantifying and comparing marginal external costs of traffic congestion on substitute routes.

We make several contributions to this literature. First, following Verhoef, Nijkamp, and Rietveld (1996) and Yang and Huang (1999) we develop a straight-forward extension of the static two-route model for congestion on mainline and HOV lanes. We consider external costs due to congestion and other vehicle related externalities such as air pollution and traffic accidents. Importantly, we derive sufficient statistics for determining whether HOV lane

[2]As noted by De Palma, Kilani, and Lindsey (2005) "A drawback of second-best tolling is that it requires extensive information on speed-flow curves and demand elasticities throughout the network."

[3]Further, these quantities can vary throughout the network and over time even in a static model of traffic patterns during peak and off-peak periods. For a detailed explanation see Verhoef (1999).

use should be encouraged or discouraged. Further, we highlight situations when encouraging HOV lane use may actually *increase* emissions.

Second, we exploit 10 years of detailed data on traffic flows in Los Angeles to estimate the marginal external costs of congestion for mainline and HOV lanes. We show these relationships are identified from plausibly exogenous variation in travel demand as opposed to shocks to highway capacity, for instance weather or traffic accidents.

Third, we document substantial variability in marginal external congestion costs along mainline and HOV lanes across routes and hours of the day. During peak hours, mainline marginal external congestion costs can reach \$1.18 per mile, while HOV lane costs can reach \$1.39 per mile. By contrast, external congestion costs are zero during the overnight and early morning hours when freeways are uncongested. Moreover, for a single peak hour, *i.e.* 8 am or 5 pm, costs also vary substantially from day to day on any given route. For instance, marginal net external HOV lane congestion costs, taking into account mainline congestion and moderate induced demand, on I-10E vary from -\$1.95 at the 10th percentile to zero at the 90th percentile during the 5pm peak hour. On I-210E with moderate induced demand, marginal net external cost estimates imply HOV lanes should be subsidized at the 75th percentile but taxed at the 90th percentile. Further, while these estimates are essential for understanding the costs and benefits of encouraging HOV lane use, they also apply more generally to second-best toll policies.

Fourth, given our estimates we investigate when and where encouraging HOV lane use would decrease *total* congestion based on the true (unobserved) level of induced travel demand. We find that encouraging HOV lane use increases total peak hour congestion costs for median induced demand ranging from 0.33 to 0.84. Our results underscore the need to quantify congestion costs in realtime when setting optimal policies rather than relying on average levels or rules of thumb.

2 Conceptual framework

We model HOV lane policies using a static two-route model for mainline and HOV lane congestion (Verhoef, Nijkamp, and Rietveld, 1996; Yang and Huang, 1999). Specifically, we adapt the framework developed in Hughes and Kaffine (2017) to HOV lanes where increased carpooling may remove two or three solo occupancy vehicles from the mainline depending on HOV lane restrictions. Users may also choose an uncongestible alternative outside option. In Section 2.1, we analyze the decentralized equilibrium where users minimize costs across their alternatives. Section 2.2 describes the allocations of a cost-minimizing social planner who can influence the allocations of users across options, but cannot limit the total number of highway users.[4] We next extend the model along two dimensions. Section 2.3 considers other vehicle related externalities, such as air pollution, that depend on total vehicle use. Section 2.4 considers the cases where the *outside option* is congestible and when users of the outside option generate use-externalities (e.g. emissions).

Consider the case of $\bar{N}$ total cost-minimizing commuters who make a single trip and choose between a mainline freeway and carpooling in a high-occupancy vehicle (HOV) lane. Additionally, commuters can select an uncongestible alternative outside option, for instance telecommuting or public transportation. The number of commuters who choose the mainline lane is n_l, which is equivalent to the number of mainline *vehicles*. The number of commuters who chose the outside option is n_a. Since it will be convenient below to write the congestion cost functions in terms of the number of *vehicles*, we assume a binding 2+ HOV lane restriction such that the number of HOV *vehicles* is n_h and the number of HOV commuters is $2 \times n_h$.[5] Finally, each commuter must be allocated such that $\bar{N} = n_l + 2n_h + n_a$.

Commuters who choose the mainline lane face (common) congestion costs

[4] Alternatively we could frame the problem in terms of the total benefits for travel and costs as in Verhoef, Nijkamp, and Rietveld (1996). Hughes and Kaffine (2017) show that approach yields results identical to those presented here.

[5] In our numerical analysis the exception to this rule is I-10, which has a 3+ HOV lane restriction during peak hours. In that case, we assume the number of HOV vehicles is n_h and the number of HOV commuters is $3 \times n_h$.

that depend on the total number of mainline drivers, denoted by $T(n_l)$. Commuters in the HOV lane face congestion costs $T(n_h)$ and an access cost $\tau(n_h)$, *i.e.* the transaction cost of forming a carpool. Users of the outside option face heterogenous costs $A(n_a)$, for instance the cost of public transportation or lost productivity from telecommuting. Both access costs $\tau(n_h)$ and outside option costs $A(n_a)$ vary by commuter.[6] Finally, we assume users of each mode receive the same benefit from the trip such that cost-minimization implies utility maximization.

We can understand the net benefit or costs of increasing HOV lane consumption by comparing how commuters would allocate themselves in the decentralized equilibrium with the allocations of a cost-minimizing social planner.

2.1 Decentralized equilibrium

In the decentralized equilibrium commuters sort across the three commute options, the mainline, HOV lane and outside option, until no user can lower her costs by choosing another option such that:

$$\begin{aligned} T_l(n_l) &= T_h(n_h) + \tau(n_h) \qquad (1) \\ T_l(n_l) &= A(n_a) \end{aligned}$$

The first condition equates the commute cost in the mainline with the commute cost in the HOV lane, *i.e* the marginal user is indifferent between the mainline and HOV lanes. The second relation equates congestion costs in the mainline to the outside option cost. This condition states the entry criterion, namely that the marginal commuter is indifferent between commuting in the mainline lane and choosing the outside option.

Note that each user considers her private cost of commuting but fails to consider the effect of her choice on other users. In particular if travel time increases in the number of mainline and HOV lane commuters, *i.e.* there

[6] The relationships $\tau(n_h)$ and $A(n_a)$ can each be thought of as an ordering of users from lowest to highest cost. However, since each user pays her individual cost, no rents are generated for users of the outside option.

is congestion, driving creates a negative externality and allocations in the decentralized equilibrium will be inefficient.

2.2 Constrained social planner

Next consider the allocations of a social planner who is unable to limit the total number of commuters but can choose how commuters are allocated across the three options.[7] We use the term "constrained social-planner" to describe the case where the planner is unable to limit entry.

The social planner chooses n_h, n_l, n_a to minimize total costs, subject to the constraint that all users must be allocated, and the fact that congestion relief in the mainline will induce users from the outside option until congestion costs in the mainline are equal to the cost of the outside option. Since the social planner is concerned with welfare effects across *users*, each term containing n_h and τ in (2) is multiplied by two to reflect the number of commuters per vehicle. Formally, the planner solves:

$$\min_{n_h, n_l, n_a} \quad 2T_h(n_h)n_h + \int_0^{n_h} 2\tau(n)dn + T_l(n_l)n_l + \int_0^{n_a} A(n)dn \tag{2}$$

$$s.t. \quad \bar{N} - 2n_h - n_l - n_a = 0$$

$$T_l(n_l) = A(n_a)$$

Iterative substitution of the first order conditions of (2) yields an expression analogous to (1) above, namely:

$$2T_l - 2T_h = 2\tau + 2n_h T_h' - 2n_l T_l' (\frac{A'}{A' + T_l'}). \tag{3}$$

Comparing this relationship with the decentralized equilibrium (1), note the additional final terms that capture the marginal external cost of an additional HOV lane user $2n_h T_h'$ and the marginal external benefit of reduced mainline congestion $2n_l T_l'(\frac{A'}{A'+T_l'})$. Equation 3 also makes clear the role of 2+ HOV

[7] This assumption holds in most transportation markets where congestion pricing or other policies to limit driving are unavailable.

lane restrictions requiring two riders per car. On one hand, forming a carpool from two mainline drivers replaces two vehicles with one vehicle. This lowers congestion in the mainline. On the other hand, every HOV lane vehicle carries two users who are each affected by increased congestion. Since the marginal effects offset one another, we can rewrite (3) more simply as:

$$T_l - T_h = \tau + n_h T_h' - n_l T_l'(1 - \alpha) \tag{4}$$

where, $\alpha = 1 - \frac{A'}{A'+T_l'}$. The parameter α is the induced demand effect and captures commuters' trade-offs between mainline congestion and outside option costs. Formally, α is a function of the slopes of the mainline congestion and outside option cost functions. Intuitively, $\alpha = 1$ implies that each user shifted from the mainline lane to the HOV lane is replaced with a new entrant from the outside option. Similarly, $\alpha = 0$ means zero entry from the outside option, *i.e.* the total number of drivers is fixed.[8]

We use the term "marginal net external cost" to describe the combined effects of HOV lane congestion and mainline congestion relief. From (4) we see that whether the planner would encourage or discourage HOV lane use depends not only on the marginal congestion costs in the mainline and HOV lanes, but also entry due to induced demand. If the level of induced demand is known, the optimal subsidy or tax to encourage or discourage HOV lane use can be calculated from the marginal net external cost.

Alternatively, we derive a relationship for the "critical" level of induced demand α^* at which the decentralized equilibrium would be optimal. Intuitively, there is some level of induced demand at which the marginal net external cost is zero and the social planner would not wish to reallocate commuters across lanes. Specifically, the marginal net external cost is zero when:

$$\alpha^\star = (1 - \frac{n_h T_h'}{n_l T_l'}). \tag{5}$$

Therefore, the planner would increase HOV lane consumption for $\alpha < \alpha^*$ and

[8] Similarly, $\alpha = \frac{1}{2}$ implies one new entrant for every two users shifted out of the mainline.

decrease consumption for $\alpha > \alpha^*$.

2.3 Other vehicle related externalities

Up to this point we have only considered externalities related to congestion. However, other vehicle related externalities such as local air pollution, climate change and accidents likely contribute to policy makers' desire to promote carpooling. We adapt the result in Hughes and Kaffine (2017) to the case where additional use externalities depend on the total number of *vehicles*. In this case, the constrained social planner's problem is given by:

$$\min_{n_h, n_l, n_a} \quad 2T_h(n_h)n_h + \int_0^{n_h} 2\tau(n)dn + T_l(n_l)n_l + \int_0^{n_a} A(n)dn + E(n_h + n_l) \quad (6)$$

$$s.t. \quad \bar{N} - 2n_h - n_l - n_a = 0$$

$$T_l(n_l) = A(n_a)$$

Iterative substitution of the first order conditions yields an expression analogous to (4) above, namely:

$$T_l - T_h = \tau + n_h T_h' - n_l T_l'(1 - \alpha) + E(\alpha - \frac{1}{2}) \quad (7)$$

where E represents the additional external cost per vehicle and where we assume a 2+ HOV lane restriction as before. The final externality term has an intuitive interpretation. Since the additional use externalities depend on the total number of vehicles, when induced demand is less than 0.5 ($\alpha < 0.5$), carpools formed from two mainline drivers are replaced by, on average, less than one new entrant. Therefore, the total number of vehicles and additional vehicle related externalities decreases. If induced demand is greater than 0.5 ($\alpha > 0.5$), then external costs are *increasing* with HOV lane use.[9] In other words, if induced demand is greater than 0.5, policies that aim to reduce

[9]Formally, the critical level of induced demand with additional use externalities E can be written as: $\alpha^\star = \frac{E - 2n_h T_h' + 2n_l T_l'}{2(E + n_l T_l')}$. Differentiating gives $\frac{d\alpha^\star}{dE} = \frac{1}{2(E+n_l T_l')} - \frac{E - 2n_h T_h' + 2n_l T_l'}{2(E+n_l T_l')^2} = \frac{1}{2(E+n_l T_l')} - \frac{\alpha^\star}{(E+n_l T_l')}$. If $\alpha^\star < 0.5$ then $\frac{d\alpha^\star}{dE} > 0$, and if $\alpha^\star > 0.5$ then $\frac{d\alpha^\star}{dE} < 0$

pollution by promoting carpooling in HOV lanes would have the perverse effect of *increasing* pollution.

Therefore, $\alpha = 0.5$ emerges as a "magic number" for determining the extent to which the planner would wish to encourage or discourage HOV lane use. To see this, we can find the critical level of induced demand (α^*) that sets the marginal net external costs in (7) to zero or:

$$\alpha^\star = 1 - \frac{n_h T_h' + \frac{1}{2}E}{n_l T_l' + E}. \tag{8}$$

Thus, at low levels of induced demand, including use-externalities provides a stronger case for increasing HOV use. By contrast, at higher levels of induced demand, including use-externalities provides a stronger case for decreasing HOV use. Similarly, the critical level of induced demand will tend toward 0.5 as use-externalities increase.[10] The intuition is that if use-externalities are sufficiently large relative to congestion concerns, then the key question is whether or not increasing HOV use leads to an overall increase or decrease in the number of vehicles, with 0.5 as the cutoff between those two outcomes.

2.4 Congestion and use externalities in the outside option

The assumptions above characterize the choices of the constrained social planner if the outside option mainly consists of non-congestible options, *i.e.* telecommuting or public transit. In reality, outside option users may cause congestion or produce other vehicle related externalities. First, consider the case where the outside option may include another congestible route, for instance a surface street or alternate highway. In this case, users in the decentralized equilibrium can sort across four options, the mainline highway, the HOV lane, an uncongestible outside option and the alternate route, until no user can lower her costs by choosing another option. Similarly, the social planner now also con-

[10] We see from the expression above, as the use-externality increases relative to congestion costs, the critical level of induced demand converges in the limit to $\alpha^\star = 0.5$

siders congestion costs on the alternate route in addition to the other options. In this case, it is straightforward to show that whether encouraging HOV lane use increases total congestion depends on marginal external costs on the alternate route. The closer marginal costs are to those on the mainline the more likely it is that encouraging HOV lane use reduces total costs. If the alternate route is a surface street, the marginal costs of congestion are likely much lower compared with the mainline highway.[11] If the alternate route is another highway, the congestion costs could be similar. Since we think the majority of alternate routes in Los Angeles are surface streets, and because marginal congestion costs of these alternatives are unobserved, our numerical analysis below abstracts away from congestion effects on substitute routes.

The case where the outside option produces other vehicle related externalities is more nuanced. Assume E_A captures the external cost of an additional outside option user. Considering use externalities in the mainline, HOV lane and outside option, the constrained social planner problem becomes:

$$\min_{n_h,n_l,n_a} \quad 2T_h(n_h)n_h+\int_0^{n_h} 2\tau(n)dn+T_l(n_l)n_l+\int_0^{n_a} A(n)dn+E(n_h+n_m)+E_A(n_a) \tag{9}$$

$$s.t. \quad \bar{N}-2n_h-n_l-n_a=0$$

$$T_l(n_l)=A(n_a)$$

Iterative substitution of the first order conditions of (9) yields:

$$2T_l-2T_h=2\tau+2n_hT_h'-2n_lT_l'(1-\alpha)+2E(\alpha-\frac{1}{2})-2E_A\alpha$$

We simplify Equation (9) by dividing by two noting, as before, each carpool replaces two mainline vehicles, but each HOV lane carpool carriers two users.

$$T_l-T_h=\tau+n_hT_h'-n_lT_l'(1-\alpha)+E(\alpha-\frac{1}{2})-E_A\alpha \tag{10}$$

[11]While the marginal effect of an additional vehicle on travel time may be similar in size or somewhat larger on the surface street, highway vehicle densities are much higher on a multilane expressway. Therefore, the marginal external cost is likely much smaller on the surface street. See Appendix A for further discussion on this point.

To understand how externalities in the outside option affects the planner's choice to encourage or discourage carpooling, we solve for the critical level of induced demand (α^*) that sets marginal net external costs to zero.

$$\alpha^\star = 1 - \frac{n_h T_h' + \frac{1}{2}E - E_A}{n_l T_l' + E - E_A}. \tag{11}$$

Let $E_A = \gamma E$ where $0 \leq \gamma \leq 1$. Intuitively, this assumes some fraction δ of outside option users, *e.g* surface street drivers, generate vehicle related use externalities. In this case, $\alpha^\star = 1 - \frac{n_h T_h' + E(\frac{1}{2} - \gamma)}{n_l T_l' + E(1-\gamma)}$. To illustrate how other use externalities affect the critical level of induced demand, consider the case where E is large. Specifically, the $\lim_{E_F \to \infty} \alpha^\star = \frac{1}{2(1-\gamma)}$ and the critical level of induced demand depends only on γ. When $\gamma = 0$, *i.e.* there are no use externalities in the outside option, $\lim_{E \to \infty} \alpha^\star = \frac{1}{2}$, which replicates our result in Section 2.3. Larger outside option use externalities increase the critical level of induced demand, *e.g.* if $\gamma = \frac{1}{4}$, $\lim_{E_F \to \infty} \alpha^\star = \frac{2}{3}$, if $\gamma = \frac{1}{2}$, and if $\lim_{E_F \to \infty} \alpha^\star = 1$. If every outside option user is a driver, *i.e.* $\gamma = 1$, $\lim_{E_F \to \infty} \alpha^\star = \infty$, then forming a carpool always reduces the number of vehicles and encouraging carpooling is always optimal in the case where other use externalities are large. Because γ is unobserved, our numerical simulations below abstract from use-externalities in the outside option. However to the extent they exist, encouraging HOV lane use is more likely to reduce total costs.

Overall, these results make it clear that whether encouraging carpooling in HOV lanes increases or decreases total congestion critically depends on the congestion levels in mainline and HOV lanes and induced demand from the outside option. In the following sections, we apply our framework to detailed traffic data from California freeways. We focus on the critical levels of induced demand and marginal net external costs that make encouraging HOV lane use optimal.

3 Data

Our analysis focuses on 12 highway routes in Los Angeles, California from 2002 through 2011. Los Angeles is well-known for high levels of traffic congestion and for having one of the most extensive networks of HOV lanes in the nation.[12] Figure 1 shows the locations of the routes in our study, which include both north-south and east-west highways in and around downtown Los Angeles. We exploit detailed traffic data from the Freeway Performance Measurement Systems (PeMS). From PeMS we observe average vehicle speed and hourly flow rates at nearly 600 locations on the citys major highways. We aggregate the individual detector-level data to route-level data to capture traffic patterns and representative commutes.

Following the traffic engineering literature we use vehicle density, defined as the number of vehicles per mile, as our measure of consumption. This choice seems appropriate, since we assume each traveler makes at most one peak hour trip, but chooses the mode. In this case, an increase (decrease) in demand for a given mode translates into a greater (lesser) number of users on the road or outside option at a given time. The day-to-day variation in observed density and travel time define the travel time density curves from which we calculate the marginal external cost of congestion for each route and each time period in our sample.

We impose both spatial and temporal restrictions on our data to focus on congested periods and locations. Our routes include all freeways that have both mainline and HOV lanes subject to these restrictions. First, from all the possible highway routes for which we have PeMS data, we identify congested locations by looking at average vehicle speeds at various points along each freeway during the morning and evening commute periods. When congestion occurs, average speeds drop below the free flow traffic speed. These areas of reduced speed define the post-mile ranges for the congested routes. In most cases, the congested sections of highway are bounded by features of the road network, typically interchanges. In some cases we are limited by the locations

[12] Los Angeles county alone has 36% of the HOV lanes miles in the state of California.

of PeMS detectors. Second, we restrict our sample to weekdays and drop observations for Federal holidays, and the weeks of Christmas, Thanksgiving and Easter. This results in 239 daily observations per route per year. Third, some of our results focus on two commute hours, 8 am for the morning peak and 5 pm for the evening peak period. We classify each route as an am-peak or pm-peak route based on whether the observed average congestion is more severe in the morning or evening. Our analysis of average vehicle speeds confirms that these hours accurately reflect peak commuting times.

Because we are interested in travel time and speed differentials between the mainline and HOV lanes, we match PeMS detectors by type at the post-mile level. We limit our analysis to only those detectors where mainline and HOV traffic are monitored at the same location.[13] Speeds and flows are measured at between 10 and 40 locations along each route. We drop any routes for which we observe traffic conditions at fewer than 10 locations. Following the above criteria, we select the routes shown in Table 1.

For each route we estimate the average travel time and density for the mainline and HOV lanes for each hour in the sample using the detector-level data. To do this we replicate the procedure traffic engineers term "walking the vector." Beginning at the start of each route, we calculate the route-level travel time as:

$$T_{it} = \sum_{s=1}^{S} \left(\frac{1}{\overline{speed}_{j,j+1}}\right)(PM_{j+1} - PM_j) \tag{12}$$

for detector j along route i with S total detectors, where $\overline{speed}_{j,j+1}$ is the average speed between detectors j and $j+1$ and PM_j is the recorded postmile for detector j (for notational convenience, the route i subscripts are suppressed).

To estimate the average density for each route, day and lane type, we first calculate the mile-weighted average hourly flow ($\bar{F}_{it}$) and speed ($\bar{S}_{it}$). Density ($\bar{n}_{it}$) for route i and time t is then calculated using the identity $n_{it} = \bar{F}_{it}/\bar{S}_{it}$.[14].

[13] While this restriction is not necessary, it helps to ensure consistency in route distances, average speeds and flows across the lane types.

[14] Alternatively, one could define consumption in terms of traffic flow rather than density. However, density seems most appropriate in our setting as discussed above. For an excellent discussion of the relationships between travel time, speeds, flow and density in the context

The difference in travel time between the mainline and HOV lanes equals the transaction cost of carpool formation net of fuel savings and other non-congestion related differences between modes.

Summary statistics for each route during off-peak (2am) and peak periods (8am/5pm) are shown in Table 1. We see that routes in our sample vary from five to nearly twenty miles in length. During the off-peak hour, average vehicle densities in both mainline and HOV lanes are low, less than 5 cars per lane-mile on average, and average speeds are at free flow speeds of approximately 60 to 65 miles per hour. We see the effects of congestion during peak hours where average mainline vehicle densities are substantially higher ranging from 37 to 47 cars per lane-mile in the mainline and average speeds fall to between 29 and 43 miles per hour. Vehicle densities are lower and average speeds are higher in the HOV lanes, between 16 and 39 cars per lane-mile and between 40 and 55 miles per hour during peak hours. In all cases, the average speed on each route is greater in the HOV lane than the adjacent mainline lane. This is consistent with our conceptual model where consumers weigh the transaction costs of carpool formation (τ) against the differences in travel time in the mainline and HOV lanes. We explore this issue further in our numerical results below.

Finally, in our results below including other driving related externalities such as air pollution, carbon emissions and accidents, we use an average value of \$0.06 per vehicle-mile (Parry, Walls, and Harrington, 2007). For the value of time we assume \$21.46 per hour for Southern California drivers (Small, Winston, and Yan, 2005) for both HOV and mainline commuters.[15]

4 Numerical strategy

Our goal is to determine when increasing HOV lane use would lower total congestion costs on each of the routes we study. We utilize two approaches based on the conceptual framework in Section 2. First, we calculate the critical levels

of traffic congestion see Verhoef (1999).

[15] Analysis of the 2009 National Household Travel Survey data for California suggests carpoolers have incomes similar to other drivers.

of induced demand for each route and hour in the sample with and without additional vehicle related externalities. Whether encouraging carpooling in HOV lanes lowers or raises costs depends on whether the critical values are above or below the actual level of induced demand. Second, we calculate the marginal net external costs for HOV lanes at various levels of induced demand. These values are analogous to the second-best congestion tolls and indicate when carpooling should be encouraged (subsidized) or discouraged (taxed). To do this we first describe our approach to estimating congestion costs.

Our static model assumes the observed mainline and HOV lane vehicle densities are unbiased estimates of the decentralized equilibrium values. Formally, $n_{ht} = n_{ht}^{DE} + \epsilon_{ht}$ and $n_{lt} = n_{lt}^{DE} + \epsilon_{lt}$, where n_{ht} and n_{lt} are the observed vehicle densities in the HOV and mainline lanes, n_{ht}^{DE} and n_{lt}^{DE} are the decentralized equilibrium densities and ϵ_{ht} and ϵ_{lt} are well-behaved optimization errors. Throughout our analysis we assume the number of users (occupants) per HOV lane vehicle equals the minimum requirement, either two or three users for 2+ and 3+ HOV lanes, respectively.[16]

Next, we need to understand how marginal changes in vehicle densities affect travel times on the routes we study. We use observed travel times and vehicle densities to estimate congestion cost functions. Figure 2 plots hourly average route travel time and density for route I-605 N during 2011. Several features common to all our routes are worth illustrating with this specific example. First, during periods of low density (cars/lane-mile), increasing the number of vehicles in either the mainline of HOV lane does not increase travel time. In other words, external costs of traffic congestion are zero during these periods, typically during the early morning hours. Second, once vehicle density reaches a critical level, approximately 20 cars per mile in the mainline and 10 cars per mile in the HOV lane, increasing vehicle density increases travel time for all drivers within the lane, *i.e.* there is traffic congestion. Therefore, the marginal external cost of increased density during these periods is the product of the slope of the travel time density curve, vehicle density, and

[16]Further, increasing HOV lane use by one vehicle removes two or three vehicles from the mainline lane depending on the HOV lane occupancy requirement.

drivers' value of time.[17] Third, the onset of congestion occurs earlier for a given density on HOV lanes compared with mainline lanes of the same route. This difference could be due to the relative ease with which mainline drivers can pass slower vehicles on a multilane highway. Our routes have between three and four mainline lanes and usually only one HOV lane. Further, HOV lanes are often separated from adjacent mainline lines by a physical barrier or a double solid line that prevents passing into or out of the mainline lane except in designated areas. Therefore, a single slower vehicle in the HOV lane would have a larger effect on trailing vehicles. Fourth, while our estimates below allow the travel time density relationships to vary from year to year, we find the data, specifically the shapes of the travel time density curves, remarkably consistent across the years of our sample.

To account for both congested and uncontested periods as well as differences across routes and lane types, we adopt a flexible modeling approach. For each route, lane type and year of the sample we estimate the piecewise model below using non-linear least squares:

$$T_{it} = \beta_0(n_{it} < c) + \beta_1(n_{it} - c) + \beta_2(n_{it}^2 - c^2) * (n_{it} \geq c) + \epsilon_{it} \qquad (13)$$

where T_{it} is travel time, n_{it} is vehicle density on route lane-type i at time t and the β's are parameters to be estimated. We assume the slope of the linear uncongested period is zero and estimate the density c associated with the onset of congestion. We use a quadratic model for the congested periods both because it fits the observed data well, for example see Figure 2, and because it enables us to easily recover the marginal cost $n_i T_i'$.

In estimating Equation 13 we assume the variation in observed traffic densities is exogenous, *i.e.* shifts in travel demand, such that changes in density trace out the travel time density curve. This assumption seems reasonable for highway travel where road capacities are fixed in the short-run. The relationships between vehicle density and travel times are largely determined

[17] More formally the marginal external costs are $n_l T_l' V$ and $n_h T_h' V$ in the mainline and HOV lanes, where V is the value of time spent commuting.

by physical characteristics of the roadway including the number of lanes, lane width, grade, and curvature. Our estimates are unbiased if changes in density come from variation in travel demand or accidents on alternate routes.[18]

On the other hand, our estimates are biased if factors affecting vehicle density also alter the relationship between density and travel time, *i.e.* shifts in highway capacity. For example, accidents, weather, or changes in the types of vehicles and drivers on the roadway may alter both densities and the travel time density curve (Kockelman, 1998). We investigate our assumption two ways. First, we illustrate how vehicle densities shift throughout the day tracing out a travel time density curve consistent with our assumptions. Second, we show the estimated travel time density relationships are not substantially affected by factors that could lead to shifts in highway capacity.

Figure 3 plots vehicle densities and travel times for a representative route, again I-605 N in 2011, for the mainline and HOV lanes. Observations from different hours are indicated by the different shading patterns and plotted against all hours from 2011. We see for both the mainline and HOV lanes, moving through the hours of the day traces out the travel time density curve from low vehicle density in the early morning hours to high density during peak hours. This pattern is consistent with our identifying assumption that demand shifts account for the changes in observed travel times and densities. The other routes in our sample exhibit qualitatively similar behavior.

Next, we check whether weekends or holidays, seasons or weather conditions, data imputation or "incidents" such as lane obstructions, stalled vehicles and accidents, yield substantially different relationships for travel time and density compared to the full sample of observations. We discuss these results here and present representative figures in the appendix.

Holidays and weekends show lower vehicle densities and travel times on

[18]There is a fair amount of variation in aggregate travel demand even within weekday peak hours. For example, hourly mean vehicle miles traveled (VMT) in 2011 for all PeMS locations in Los Angeles and Ventura counties at 8 am excluding holidays and weekends is approximately 5.81 million miles. The 10th and 90th percentiles of aggregate VMT in 2011 are 5.68 and 5.94 million miles with minimum and maximum values of 4.97 and 6.07 million miles.

both the mainline and HOV lanes. However, the overall travel time density relationships during these off-peak periods follow the same patterns as our preferred weekday sample. One might expect weekend drivers to exhibit different driving behavior or have different values of time relative to weekday commuters. This could show up as a steepening of flattening of the travel time density curve. However, we see no evidence of this in either the representative route or other routes we study. For instance, Appendix Figure 1 plots observations from holidays and weekends against all weekday observations from I-605 N in 2011. For both the mainline and HOV lanes we see the off-peak data fall along the lower density portion of the travel time density curves for this route.

Next, we investigate potential effects by season or due to rainfall. Surprisingly, traffic patterns across seasons are quite similar. For instance, Appendix Figure 2 plots mainline and HOV lane travel times and densities by quarter. Similarly, Appendix Figure 3 plots mainline and HOV lane travel times for days with rainfall (precipitation > 0 in.) and heavy rain (precipitation > 0.5 in.). While one may expect heavy rainfall to increase the effect of increased vehicle density on travel time, we do not find strong evidence of this in our data.

PeMS engineers are forced to impute missing observations when highway detectors malfunction. We find the overall travel time density relationships do not vary substantially with the proportion of imputed observations in the data. Appendix Figure 4 compares the full sample on I-605 N in 2011 with hours where the number of actual (non-imputed) observations is less than 50% and less than 25%.[19] We see little qualitative evidence highly imputed observations are systematically different than the rest of the sample.

Finally, we compare travel time density relationships estimated using all observations with those estimated after dropping *all* observations that occur on a day with a recorded incident. Appendix Figure 5 plots the restricted

[19]PeMs classifies observations based on the fraction of observations within an hour that are imputed. We average the reported percent observed across all the detectors along a route for each our of the sample.

sample against all observations. The black curve is fitted to the data excluding incidents. The dark gray curve reproduces our main estimates. While dropping days with incidents eliminates over 90 percent of our observations, the estimated travel time density curves are very similar to those estimated with the full sample. In fact, for the mainline it is difficult to distinguish the restricted sample curve as it is almost completely covered by the full sample estimate. The HOV lane curve is somewhat flatter when incidents are excluded, though still quite similar in shape to the full sample. Further, we show below eliminating incident days does not materially affect our estimates of the critical levels of induced demand or marginal net external cost.

5 Critical levels of induced demand

We calculate $\alpha^\star$ by evaluating the derivatives of the fitted functions, T_h' and T_l', at the observed densities and substitute these values into Equation 5 or 8. To allow for changes in highway characteristics over time, we repeat this procedure separately for each year of our sample. Figure 4 plots the distributions of critical levels of induced demand levels with and without additional vehicle-related externalities. Recall induced demand less than $\alpha^\star$ implies encouraging HOV lane use reduces total congestion costs and induced demand greater than $\alpha^\star$ implies encouraging HOV lane use increases total congestion costs. Panel *a* plots α^* for all hours and panel *b* focuses on the peak congested hour. Note that encouraging HOV lane use is beneficial over a large range of induced demand. During peak hours, panel *b*, the distribution shifts to the left indicating *encouraging* HOV lane use increases total congestion costs for more modest levels of induced demand. Notice when other use externalities are considered, the distribution contracts inward toward a critical value of 0.5 consistent with our conceptual model.

Table 2 summarizes points on the distributions of the critical level of induced demand $\alpha^\star$ for each of the routes in our sample. We focus on the peak congestion periods of 8 am and 5 pm for morning and afternoon peaking routes. Panel *a* presents results ignoring non-congestion related use-externalities, while panel b includes other use-externalities. Beginning with panel a we see the mean $\alpha^\star$ ranges from 0.29 for I-210 E to 0.81 for I-10 W. The median values are slightly larger ranging from 0.33 to 0.84.

When use-externalities are added, in panel b, we see the distributions of $\alpha^\star$ move toward the central value of 0.50. Median values range from 0.34 to 0.82. The critical $\alpha^\star$ decreases on routes with median values greater than 0.50 when use-externalities are considered. The critical $\alpha^\star$ increases on routes with median values less than 0.50. The relatively small movement in $\alpha^\star$ when non-congestion use-externalities are included is consistent with previous literature that finds congestion costs tend to dominate calculations of external costs (Parry and Small, 2005; Parry, Walls, and Harrington, 2007; Bento et al., 2014).

The values reported in Table 2 suggest on average, encouraging HOV lane use is beneficial for moderate levels of induced demand. However, there is substantial variation across and within routes. At the mean values, encouraging consumption of the I-210E HOV lane would *increase* total congestion costs if induced demand is greater than 0.30. For five of the routes, encouraging HOV lane use increases congestion if induced demand is greater than 0.30 for at least 10 percent of the peak hours observed.[20] For ten of the routes, encouraging HOV lane use raises total congestion for at least 10 percent of the peak hours observed if induced demand is greater than 0.50. Finally, looking at the maximum for each route in panel b we note $\alpha^\star$ is never equal to 1.0. This implies that with full induced demand (Duranton and Turner, 2011), increasing HOV lane use *always* increases total congestion costs.

6 Optimal HOV subsidies and tolls

Table 3 shows the mean calculated marginal external costs of congestion for mainline and HOV lanes across the twelve routes in our sample during the peak am and pm commute periods. The rightmost columns translate congestion costs into marginal net external costs under various levels of induced demand

[20]Specifically, routes I-105E, I-105W, I-210E, and the southern portion of I-405S.

taking into account mainline congestion relief.[21] In other words, they show the (average) optimal tax or subsidy the social planner would charge to discourage or encourage HOV lane use. With no induced demand, $\alpha = 0$, the planner would always encourage HOV lane use. We see this from the negative values of marginal net external costs across routes, between -\$0.56 and -\$2.68 per car-mile. On the other hand with full induced demand, $\alpha= 1$, the planner would always discourage (tax) HOV lane use. In this case, marginal net external costs range from \$0.36 to \$1.39 per car-mile. For intermediate levels of induced demand, the policy prescriptions are more complicated. For modest levels of induced demand, $\alpha = 0.2$ and $\alpha = 0.4$, the planner would, on average still encourage HOV lane use. The one exception is I-210E where the planner would on average tax HOV lane use for $\alpha = 0.4$. For higher levels of induced demand, the planner is more likely to discourage HOV lane use. If $\alpha = 0.6$, the mean values suggest encouraging HOV lane use on half the routes and discouraging HOV lane use on the other half. If $\alpha = 0.8$, the planner would only encourage HOV lane use on I-10W.

Importantly, these results are robust to whether our travel time density data include observations with traffic incidents. Appendix Table 1 presents results dropping *all* observations that occur on a day with a recorded incident. We reestimate the travel time density relationships for this restricted sample and calculate the corresponding marginal net external costs for various levels of induced demand. The results are very similar to those presented in Table 3, despite dropping approximately 95 percent of the observations from our sample. As such, we use the full sample for the remainder of our analysis.

While the mean peak period values reported in Table 3 are useful for determining the overall trends in the data, these figures hide substantial variation across hours of the day. To get a sense of how marginal net external costs vary across hours of the day, Figure 5 plots average values for four routes, I-10E, I-10W, I-210E and I-210W for the full range of induced demand from $\alpha = 0$ to

[21] We assume each carpool consists of two commuters (three commuters for I-10E and I-10W) and ignore additional use-externalities to enable the reader to easily compare between the righthand and lefthand columns.

$\alpha = 1$. For low levels of induced demand, HOV lane marginal net external costs are negative during the peak hours, *i.e.* the morning peak on the west-bound routes and the afternoon peak on the east-bound routes. The I-10 routes are characterized by high access costs (τ) owing to the +3 HOV lane restriction on these routes. As a result, the HOV lanes are more likely under-consumed and the planner prefers very large subsidies if induced demand is small. Similarly, the planner would encourage HOV lane consumption on I-210E and I-210W for low-levels of induced demand, though the implied subsidies are much smaller in magnitude.

For higher levels of induced demand marginal net external costs increase. For I-210, the implied subsidies switch to taxes for modest levels of induced demand, $\alpha < 0.5$. On routes I-10E and I-10W, costs remain negative during the peak periods for much higher levels of induced demand. The planner would only discourage HOV lane consumption for values greater than 0.7 to 0.8. Again, these effects are consistent with higher access costs on I-10. Combined, these results underscore the critical need to understand the potential magnitude of induced demand effects. Optimal use of HOV lanes during peak periods could require either large subsidies or large taxes, greater than one dollar per car-mile, depending on the magnitude of induced demand.

Further, there is substantial variation in marginal net external costs even within the peak hour for each route. Table 4 summarizes points on the distributions of marginal net external costs for I-10E, I-10W, I-210E and I-210W during the peak morning or afternoon commute hour. We present statistics for three levels of induced demand, $\alpha = 0$, $\alpha = 0.5$ and $\alpha = 1$. For instance, consider costs on I-10E. With no induced demand, marginal net external costs are negative, \$3.49 per car-mile at the 25th percentile and \$1.73 at the 75th percentile. For higher levels of induced demand, costs become more positive but the range in values is still quite large. With full induced demand, marginal net external costs are positive, \$0.26 per car-mile at the 25th percentile and \$1.03 at the 75th percentile. To put these effects in perspective, a ten mile commute at the 75th percentile would be tolled at just over \$10 per vehicle. This pattern is similar across the other three routes shown in Table 4.

Moreover, the policy prescription can even change sign. For instance, with intermediate induced demand on I-210, policy makers would encourage HOV lane consumption at the 10th percentile but discourage HOV lane use at the 90th percentile. This means that policy makers cannot simply rely on mean levels or heuristics in setting policies towards HOV lanes.

Finally, note that marginal net external costs are essentially zero for all routes during the early morning and late evening hours on all routes. On the I-210 routes, costs are also near zero during the mid-day. Since congestion is low during these periods there is no need to encourage or discourage HOV lane use.

7 Discussion

As our results above show, whether encouraging HOV lane consumption lowers total congestion critically depends on the relative marginal external costs in the mainline and HOV lanes as well as induced demand for travel. We have shown congestion costs are easily estimated from observed traffic data. However, induced demand is somewhat more difficult to measure in a manner consistent with our framework.

Recent empirical studies estimate the relationship between increases in highway capacity and vehicle miles (kilometers) traveled. Examples of work in this area include Noland (2001), and Cervero and Hansen (2002). Hymel, Small, and Van Dender (2010) provide a thorough review of this literature. In the short-run, estimates of induced demand range from 0.10 to 0.6. In the longer run, estimates range from 0.7 to 1.0 (Noland, 2001; Cervero and Hansen, 2002; Hymel, Small, and Van Dender, 2010). Duranton and Turner (2011), provide the most recent estimates of induced demand in US metropolitan areas from 1983 to 2003. They estimate a long-run elasticity of vehicle kilometers traveled with respect to highway capacity of approximately 1.0. These estimates suggest encouraging HOV lane likely reduces total congestion costs in the short-run but increases congestion in the longer run. At the extremes, when induced demand is zero, policy makers would always choose to encourage HOV lane use. With full induced demand, encouraging HOV lane use always increases total congestion and other vehicle-related use externalities such as emissions.

However, between these extreme cases our ability to apply highway capacity based induced demand estimates is somewhat limited. Capacity expansions shift out the travel cost function by lowering congestion costs resulting in a greater number of commuters choosing to drive in equilibrium.[22] The movement of commuters from the mainline to the HOV lane is analogous to an expansion of mainline highway capacity since the net result is a reduction in mainline congestion. However, estimates based on capacity expansion do not map directly to our induced demand parameter α, which is a non-linear function of the outside option and mainline travel time density relationships. Nonetheless, these estimates are informative as low estimated induced demand corresponds to lower values of α in our model and vice versa.

All in all there is still much to learn when the true level of induced demand is unknown.[23] Whether encouraging HOV lane use lowers congestion varies substantially across route. When the transaction costs of carpool formation are higher, increasing HOV lane use is more likely to lower total congestion costs. In particular I-10, with its 3+ HOV lanes, stands out from the other routes as being more likely to benefit from increased carpooling. Higher access costs mean a higher level of induced demand is necessary to outweigh the congestion relief benefits of increasing HOV lane use. While our numerical results support this prediction, whether switching from a 3+ to 2+ restriction would be beneficial is an unresolved empirical question.[24] Note that while encouraging carpooling will lower congestion costs for moderate levels of induced

[22]The mechanisms by which commuters respond to increased highway capacity are numerous. For example, switching between driving routes, reduced use of public transit or telecommuting, or an increase in the overall level of travel.

[23]Policy experiments could be used to estimate induced demand by randomly assigning enticements to use the HOV lane and measuring changes in HOV and mainline traffic densities.

[24]Interestingly, Los Angeles experimented with a reducing the occupancy requirement on I-10 from 3+ to 2+ beginning on January 1, 2000. However, on July 24, 2000 the requirement was subsequently raised back to 3+ during peak periods due to increased HOV lane congestion and slower speeds (Turnbull, 2002).

demand, it is also more likely to raise costs during the highest peak demand times.

References

Bento, Antonio, Daniel Kaffine, Kevin Roth, and Matthew Zaragoza-Watkins. 2014. "The Effects of Regulation in the Presence of Multiple Unpriced Externalities: Evidence from the Transportation Sector." *American Economic Journal: Economic Policy* 6 (3):1–29.

Cervero, Robert and Mark Hansen. 2002. "Induced Travel Demand and Induced Road Investment: A Simultaneous Equation Analysis." *Journal of Transport Economics and Policy* 36 (3):pp. 469–490.

De Palma, André, Moez Kilani, and Robin Lindsey. 2005. "Comparison of second-best and third-best tolling schemes on a road network." *Transportation Research Record: Journal of the Transportation Research Board* (1932):89–96.

Duranton, Gilles and Matthew A. Turner. 2011. "The Fundamental Law of Road Congestion: Evidence from US Cities." *American Economic Review* 101 (6):261652.

Federal Highway Administration. 2016. "Federal-Aid Highway Program Guidance on High Occupancy Vehicle (HOV) Lanes." Tech. rep., U.S. Department of Transportation, Washington, DC.

Hughes, Jonathan E. and Daniel Kaffine. 2017. "When is increasing consumption of common property optimal? Sorting, congestion and entry in the commons." *Journal of Environmental Economics and Management* 81:227 – 242.

Hymel, Kent M., Kenneth A. Small, and Kurt Van Dender. 2010. "Induced demand and rebound effects in road transport." *Transportation Research Part B: Methodological* 44 (10):1220–1241.

Kockelman, Kara Maria. 1998. "Changes in flow-density relationship due to environmental, vehicle, and driver characteristics." *Transportation Research Record: Journal of the Transportation Research Board* 1644 (1):47–56.

Lévy-Lambert, HUBERT. 1968. "Tarification des Services à Qualité Variable–Application aux Péages de Circulation." *Econometrica: Journal of the Econometric Society* :564–574.

Liu, Louie Nan and John F. McDonald. 1998. "Efficient Congestion Tolls in the Presence of Unpriced Congestion: A Peak and Off-Peak Simulation Model." *Journal of Urban Economics* 44 (3):352 – 366.

Marchand, Maurice. 1968. "A note on optimal tolls in an imperfect environment." *Econometrica: Journal of the Econometric Society* :575–581.

Noland, Robert B. 2001. "Relationships between highway capacity and induced vehicle travel." *Transportation Research Part A: Policy and Practice* 35 (1):47 – 72.

Parry, I.W.H. and K.A. Small. 2005. "Does Britain or the United States have the right gasoline tax?" *The American Economic Review* 95 (4):1276–1289.

Parry, I.W.H., M. Walls, and W. Harrington. 2007. "Automobile Externalities and Policies." *Journal of Economic Literature* 45 (2):373–399.

Schrank, David, Tim Lomax, and Bill Eisele. 2012. "Urban Mobility Report." Tech. rep., Texas Transportation Institute, College Station, TX.

Small, Kenneth A., Clifford Winston, and Jia Yan. 2005. "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability." *Econometrica* 73 (4):1367–1382.

Turnbull, Katherine F. 2002. "Affects of Changing HOV Lane Occupancy Requirements: El Monte Busway Case Study." Tech. rep., Texas Transportation Institute.

Verhoef, Erik, Peter Nijkamp, and Piet Rietveld. 1996. "Second-Best Congestion Pricing: The Case of an Untolled Alternative." *Journal of Urban Economics* 40 (3):279 – 302.

Verhoef, Erik T. 1999. "Time, speeds, flows and densities in static models of road traffic congestion and congestion pricing." *Regional Science and Urban Economics* 29 (3):341–369.

Yang, Hai and Hai-Jun Huang. 1999. "Carpooling and congestion pricing in a multilane highway with high-occupancy-vehicle lanes." *Transportation Research Part A: Policy and Practice* 33 (2):139 – 155.

8 Figures

Figure 1: Mainline and HOV lane routes in the study area.

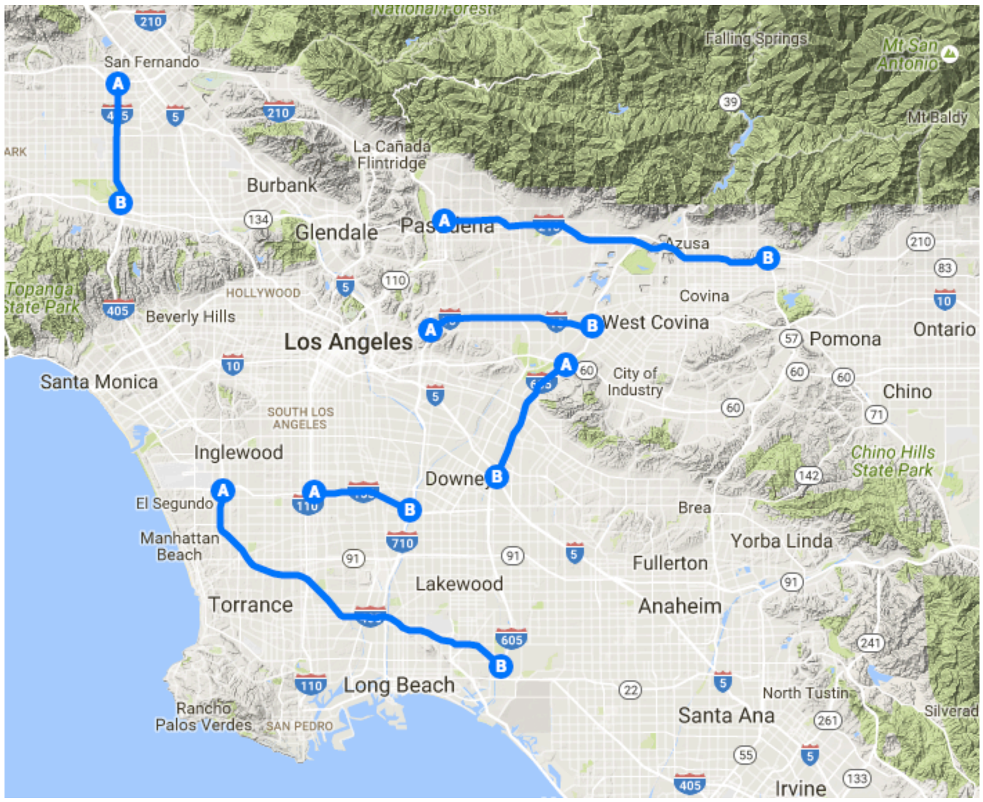
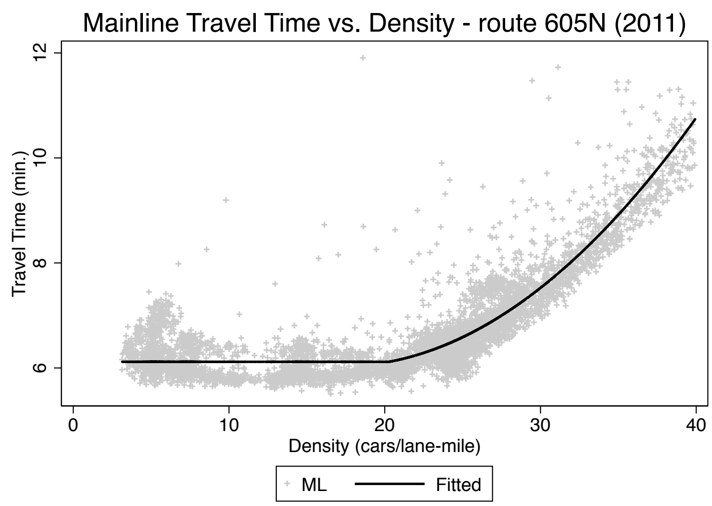


Figure 2: Representative travel time density relationships for mainline and HOV lanes.

(a) Route I-605N mainline lanes



(b) Route I-605N HOV lane

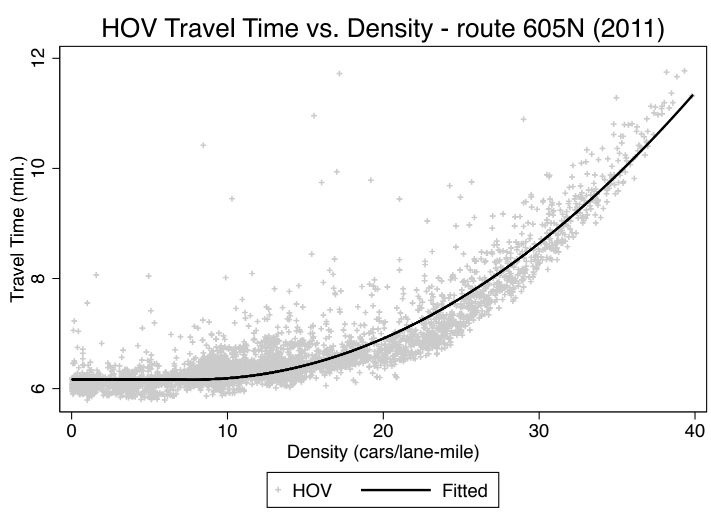
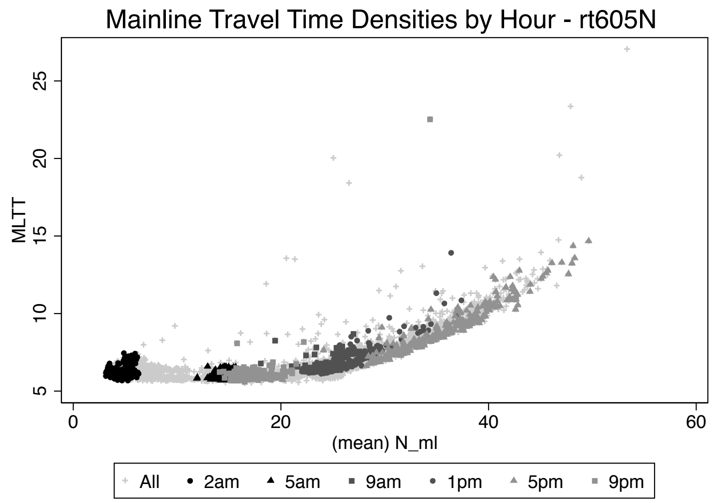


Figure 3: Representative travel time density relationships for mainline and HOV lanes by hour.

(a) Route I-605N mainline lanes



(b) Route I-605N HOV lane

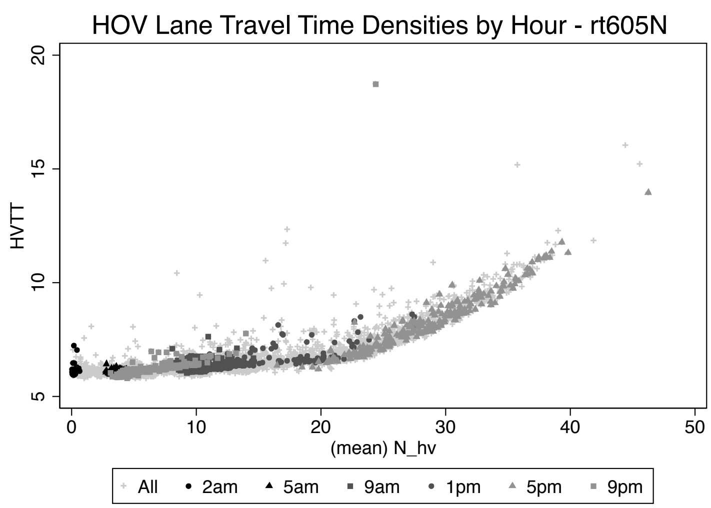
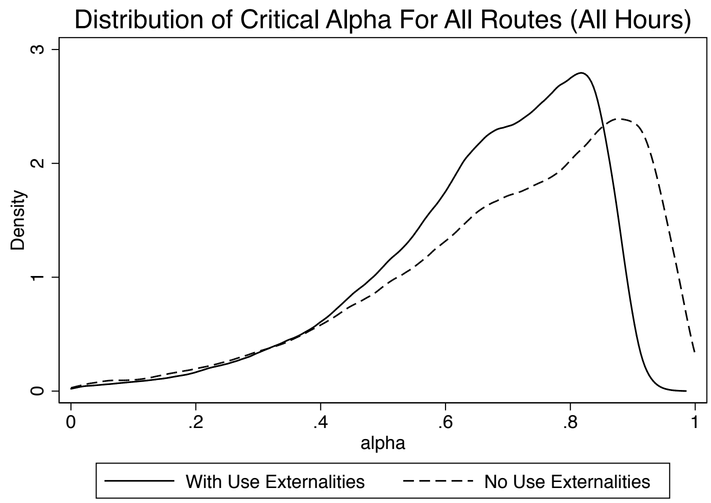


Figure 4: Distribution of critical alphas across twelve routes in Los Angeles with and without additional use-externalities.

(a) All hours



(b) Peak hour (8am/5pm)

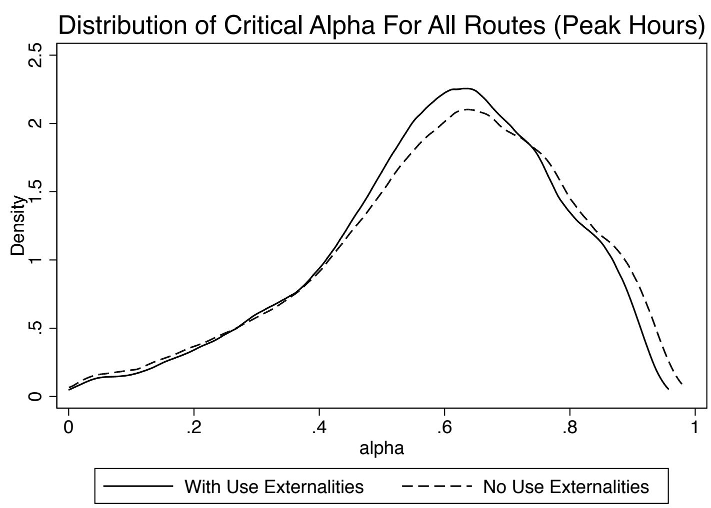
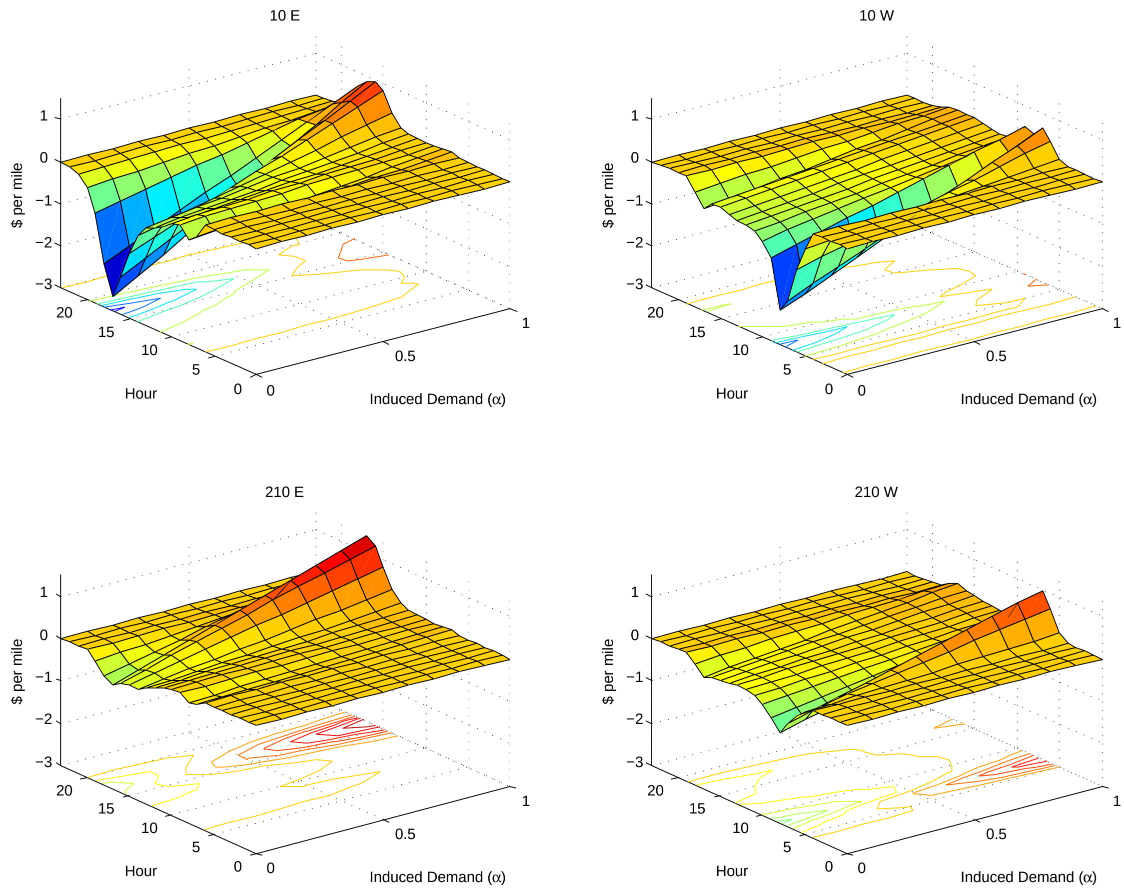


Figure 5: Optimal HOV lane tolls (subsidies) for I-10 and I-210.



9 Tables

Table 1: Descriptive statistics for Los Angeles freeway routes.

		Off-Peak Hour (2am)				Peak Hour (8am/5pm)			
		ML		HOV		ML		HOV	
Route	miles	mph	cars/lane-mi.	mph	cars/lane-mi.	mph	cars/lane-mi.	mph	cars/lane-mi.
105 E (PM)	5.0	65.3	4.3	64.7	0.4	34.6	43.2	44.1	30.0
105 W (AM)	5.5	65.6	3.5	65.0	0.2	42.8	36.6	48.0	26.6
10 E (PM)	8.7	65.8	3.5	63.2	0.2	29.3	47.3	47.6	22.1
10 W (AM)	9.0	65.2	3.3	63.6	0.3	35.2	43.0	54.7	15.7
210 E (PM)	18.2	62.1	2.8	64.8	0.3	32.9	43.2	34.6	38.5
210 W (AM)	18.3	61.2	2.8	64.2	0.4	35.8	41.3	45.1	26.1
405 N North Rt. (PM)	4.9	66.0	3.4	65.5	0.1	35.2	42.9	45.5	27.3
405 N South Rt. (AM)	19.3	65.2	2.4	64.7	0.4	43.7	37.9	51.3	23.6
405 S North Rt. (AM)	12.5	65.4	2.5	64.5	0.2	30.6	45.0	42.8	23.7
405 S South Rt. (PM)	19.5	66.4	2.7	65.4	0.2	34.1	43.9	40.0	35.1
605 N (PM)	6.7	62.2	4.5	65.7	0.2	38.1	41.3	42.5	33.7
605 S (AM)	6.7	60.6	4.4	64.8	0.2	38.8	38.8	52.8	20.5

Table 2: Points on the distribution of critical alphas for Los Angeles freeway routes.

(a) Critical induced demand ($\alpha^{\star}$) ignoring use-externalities.

Route	n	mean	min.	p5	p10	p50	p90	p95	max.
105 E (PM Peak)	2,401	0.51	-0.47	0.15	0.27	0.56	0.70	0.72	0.87
105 W (AM Peak)	2,401	0.51	-1.09	0.18	0.27	0.54	0.70	0.74	0.88
10 E (PM Peak)	2,401	0.74	-0.46	0.40	0.50	0.82	0.90	0.92	0.95
10 W (AM Peak)	2,401	0.81	0.06	0.54	0.66	0.84	0.93	0.95	0.98
210 E (PM Peak)	2,401	0.29	-1.36	-0.11	0.00	0.33	0.54	0.60	0.78
210 W (AM Peak)	2,401	0.60	0.04	0.39	0.44	0.61	0.74	0.76	0.84
405 N North Rt. (PM Peak)	2,401	0.64	0.17	0.45	0.50	0.65	0.77	0.80	0.90
405 N South Rt. (AM Peak)	2,401	0.63	0.15	0.43	0.47	0.64	0.79	0.81	0.88
405 S North Rt. (AM Peak)	2,401	0.68	-0.22	0.39	0.46	0.69	0.90	0.91	1.00
405 S South Rt. (PM Peak)	2,401	0.43	-1.22	0.08	0.16	0.46	0.67	0.70	0.86
605 N (PM Peak)	2,401	0.43	-1.23	0.07	0.18	0.46	0.65	0.68	0.81
605 S (AM Peak)	2,401	0.74	0.27	0.59	0.64	0.76	0.83	0.85	0.92

(b) Critical induced demand ($\alpha^{\star}$) including use-externalities.

Route	n	mean	min.	p5	p10	p50	p90	p95	max.
105 E (PM Peak)	2,401	0.51	-2.62	0.18	0.29	0.55	0.68	0.71	0.85
105 W (AM Peak)	2,401	0.51	-1.76	0.20	0.29	0.54	0.68	0.72	0.83
10 E (PM Peak)	2,401	0.73	-0.26	0.41	0.50	0.80	0.89	0.90	0.94
10 W (AM Peak)	2,401	0.79	-0.07	0.53	0.61	0.82	0.90	0.92	0.96
210 E (PM Peak)	2,401	0.30	-2.86	-0.07	0.04	0.34	0.54	0.59	0.77
210 W (AM Peak)	2,401	0.59	-0.88	0.39	0.44	0.60	0.72	0.74	0.82
405 N North Rt. (PM Peak)	2,401	0.62	-1.99	0.45	0.50	0.64	0.75	0.78	0.87
405 N South Rt. (AM Peak)	2,401	0.62	-0.76	0.44	0.47	0.62	0.76	0.79	0.86
405 S North Rt. (AM Peak)	2,401	0.67	-1.25	0.40	0.46	0.68	0.87	0.89	0.96
405 S South Rt. (PM Peak)	2,401	0.43	-1.32	0.11	0.18	0.46	0.66	0.69	0.84
605 N (PM Peak)	2,401	0.43	-2.43	0.10	0.21	0.46	0.64	0.67	0.79
605 S (AM Peak)	2,401	0.72	-0.67	0.58	0.62	0.73	0.81	0.82	0.88

Table 3: Mean calculated marginal congestion costs to mainline and HOV lane commuters in \$ per car-mile. Mean marginal net external cost for HOV lane commuters at various levels of induced demand α in \$ per car-mile.

				HOV Lane Mean Net MEC ($/car-mile)					
Route	**n**	MEC_l	MEC_h	$\alpha = 0$	$\alpha = 0.2$	$\alpha = 0.4$	$\alpha = 0.6$	$\alpha = 0.8$	$\alpha = 1$
105 E (PM)	2,401	0.79	0.76	-0.83	-0.51	-0.19	0.12	0.44	0.76
105 W (AM)	2,401	0.59	0.60	-0.59	-0.35	-0.12	0.12	0.36	0.60
10 E (PM)	2,401	1.18	0.85	-2.68	-1.97	-1.27	-0.56	0.15	0.85
10 W (AM)	2,401	0.87	0.43	-2.19	-1.67	-1.14	-0.62	-0.10	0.43
210 E (PM)	2,401	0.97	1.39	-0.56	-0.17	0.22	0.61	1.00	1.39
210 W (AM)	2,401	0.74	0.60	-0.87	-0.58	-0.29	0.01	0.30	0.60
405 N North Rt. (PM)	2,401	0.84	0.58	-1.10	-0.77	-0.43	-0.09	0.24	0.58
405 N South Rt. (AM)	2,401	0.60	0.43	-0.78	-0.54	-0.30	-0.06	0.19	0.43
405 S North Rt. (AM)	2,401	1.04	0.67	-1.41	-1.00	-0.58	-0.16	0.25	0.67
405 S South Rt. (PM)	2,401	0.95	1.08	-0.82	-0.44	-0.06	0.32	0.70	1.08
605 N (PM)	2,401	0.82	0.91	-0.73	-0.40	-0.07	0.26	0.58	0.91
605 S (AM)	2,401	0.72	0.36	-1.07	-0.79	-0.50	-0.21	0.07	0.36

Notes: Mean mainline and HOV lane marginal external costs, MEC_l *and* MEC_h calculated as $n_l T'_l V$ and $n_h T'_h V$ assuming two riders per HOV lane car on all routes except 10E and 10W where three riders is assumed. The constant value of time, V is \$21.46. We ignore additional vehicle-related use externalities.

Table 4: Variation in peak hour marginal net external cost for HOV lane commuters at various levels of induced demand α in $ per car-mile.

	I -10 E			I -10 W		
	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$
Mean	-2.68	-0.91	0.85	-2.19	-0.88	0.43
10th Percentile	-4.66	-1.95	0.26	-3.61	-1.56	0.12
25th Percentile	-3.49	-1.31	0.40	-3.03	-1.25	0.20
Median	-2.42	-0.84	0.61	-2.27	-0.88	0.36
75th Percentile	-1.73	-0.45	1.03	-1.38	-0.51	0.54
90th Percentile	-1.09	0.00	1.79	-0.29	-0.09	0.78

	I -210 E			I -210 W		
	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 0$	$\alpha = 0.5$	$\alpha = 1$
Mean	-0.56	0.41	1.39	-0.87	-0.14	0.60
10th Percentile	-1.11	-0.06	0.61	-1.37	-0.39	0.27
25th Percentile	-0.82	0.08	0.88	-1.10	-0.25	0.37
Median	-0.57	0.31	1.29	-0.82	-0.13	0.53
75th Percentile	-0.33	0.65	1.76	-0.60	-0.03	0.77
90th Percentile	-0.01	1.06	2.30	-0.44	0.10	1.00

Notes: MEC_l *and* MEC_h calculated as $n_l T'_l V$ and $n_h T'_h V$ assuming two riders per HO lane car on all routes except 10E and 10W where three riders is assumed. The constant value of time, V is $21.46. We ignore additional vehicle-related use externalities.

Appendix

A Substitution across K congestible routes

In the conceptual model and empirical exploration of HOV lane use we focus on two lane types, an HOV lane and the adjacent mainline lanes. As such, our calculations assume that increasing HOV lane use draws drivers from the adjacent mainline lane. However, two relevant empirical concerns emerge. First, new HOV lane users may come from not only the mainline lanes, but also from alternate congestible highways and surface streets. Second, congestion relief in the mainline may be offset not only by entry of new users from the uncongestible alternative option but also from congestible alternatives. These alternative congestible options can be thought of as belonging to the set of K congestible options in the analysis above. Here, we briefly argue that if some users are moving from other congested options, the congestion relief benefits calculated in the main text upper-bound the benefits of increasing HOV lane use. In this case, our main results are conservative in that they bias us towards concluding that HOV lanes are under-consumed.

Let τ_l represent the access cost for the mainline highway and τ_s represent the access cost for a parallel surface street (s). Because mainline highways are a faster travel option in the absence of congestion, then $\tau_l < \tau_s$.[25] Per our theory above (see also Appendix Figure 1), larger access cost on s implies that $n_l > n_s$ in the decentralized equilibrium. If the congestion functions are identical functions of the number of users, convexity of the congestion function and $n_l > n_s$ implies that $T_l'(n_l) > T_s'(n_s)$. As such, the congestion relief benefit from a driver leaving the surface street for the HOV lane ($n_s T_s'(n_s)$) is smaller than the congestion relief benefit for a driver leaving the mainline ($n_l T_l'(n_l)$). Similarly, any driver leaving a less congested option for the mainline will generate greater congestion cost in the mainline ($n_l T_l'(n_l)$) than congestion relief on their original route ($n_s T_s'(n_s)$).

[25] For example, because of higher speed limits, no traffic signals and more direct routes than surface streets.

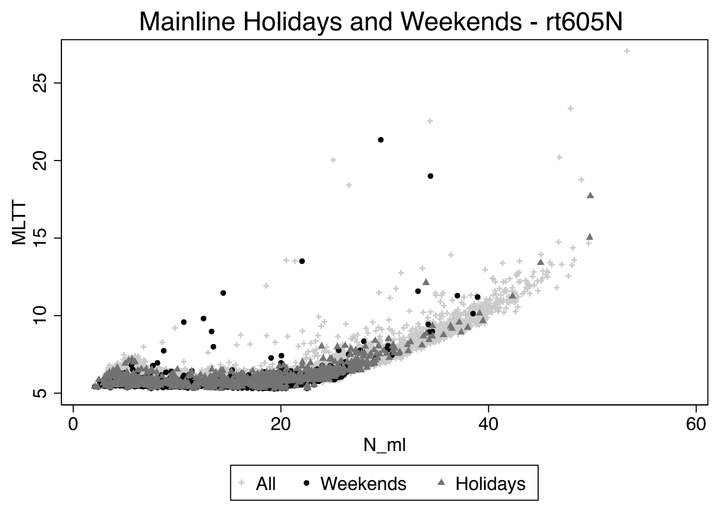
Relaxing the assumption that the congestion functions are identical raises the possibility that congestion relief on surface streets could be larger in magnitude than congestion relief in the mainline, i.e. if $T_s'(n_s) > \frac{n_l}{n_s} T_l'(n_l)$. For this to be true, the marginal change in congestion on the surface street ($T_s'(n_s)$) would have to be large enough relative to the marginal change in congestion in the mainline ($T_l'(n_l)$) to outweigh the fact that $n_l > n_s$. This seems unlikely to be true.[26] Therefore, we view the congestion benefits we calculate in the main text as an upper-bound on the potential congestion relief under the circumstances described above.

[26] For example, if there were twice as many users on the mainline highway as on a parallel surface street, then the marginal congestion function for the surface street would need to be twice as steep (despite having half as many users) for our calculation to not represent an upper-bound.

B Appendix figures

Figure 1: Representative travel time density relationships for mainline and HOV lanes on holidays and weekends.

(a) Route I-605N mainline lanes



(b) Route I-605N HOV lane

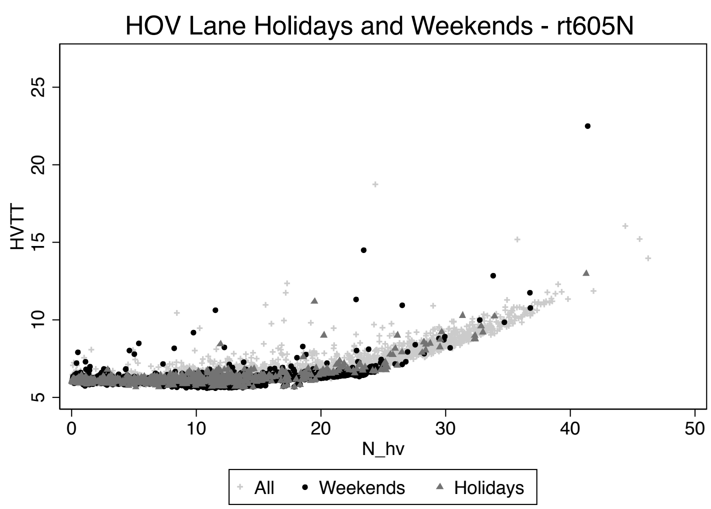
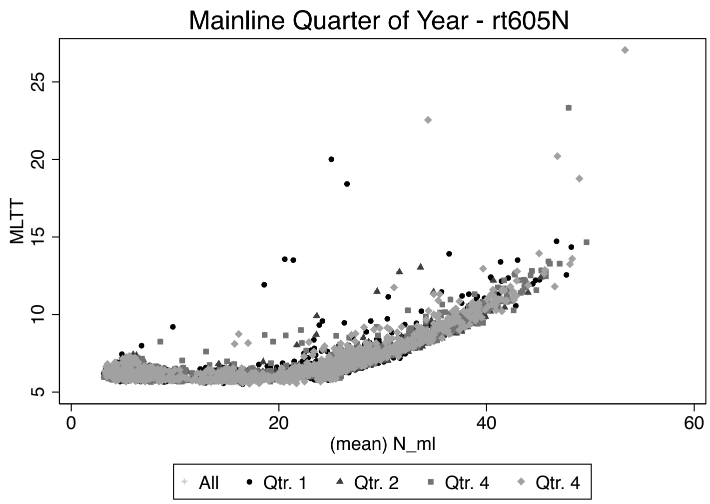


Figure 2: Representative travel time density relationships for mainline and HOV lanes by quarter.

(a) Route I-605N mainline lanes



(b) Route I-605N HOV lane

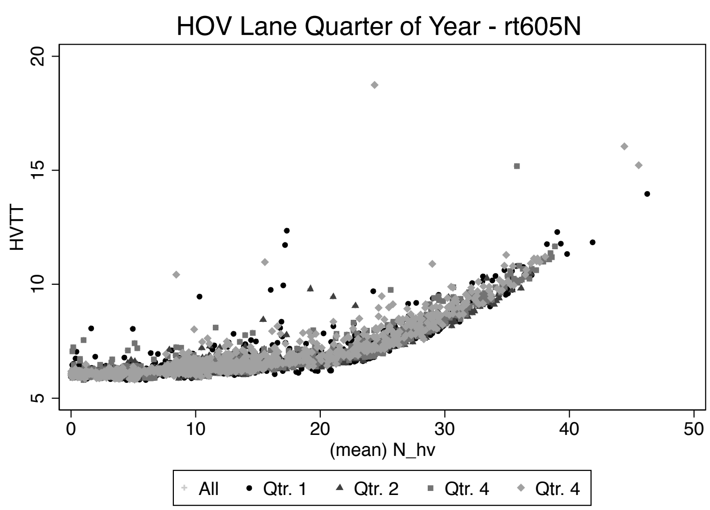
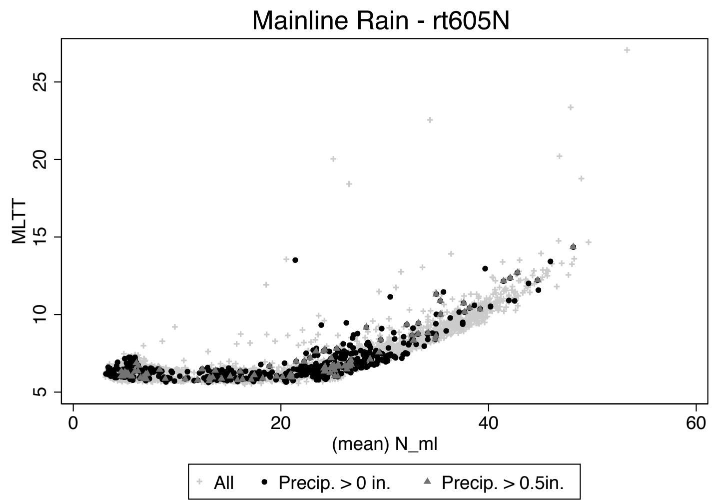


Figure 3: Representative travel time density relationships for mainline and HOV lanes by rainfall.

(a) Route I-605N mainline lanes



(b) Route I-605N HOV lane

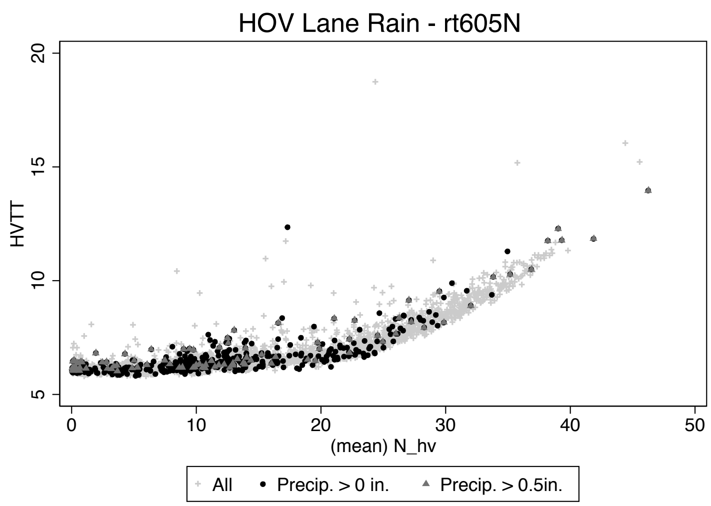
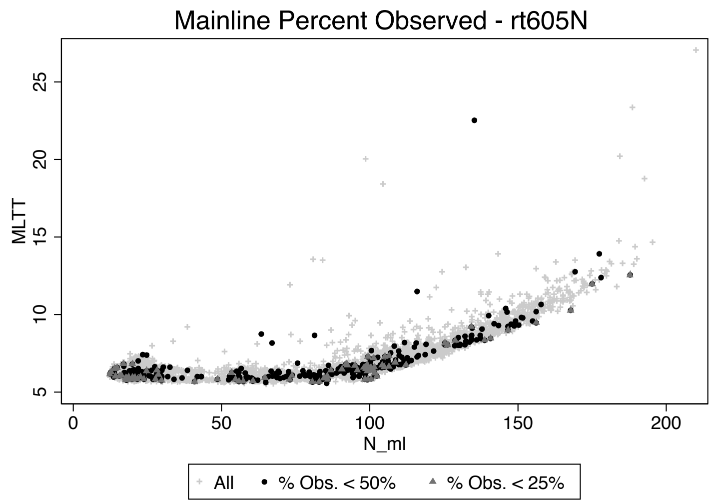


Figure 4: Representative travel time density relationships for mainline and HOV lanes by level of imputation.

(a) Route I-605N mainline lanes



(b) Route I-605N HOV lane

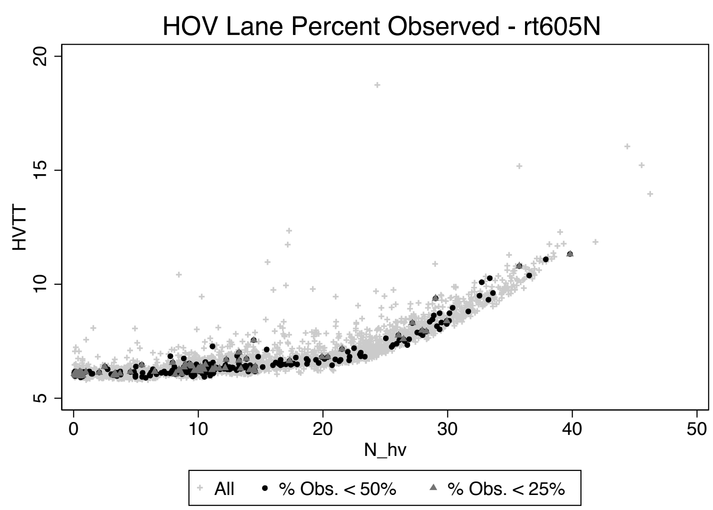
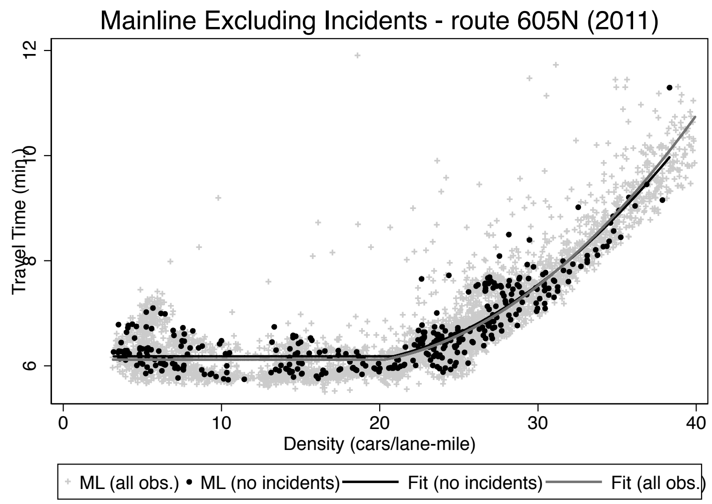
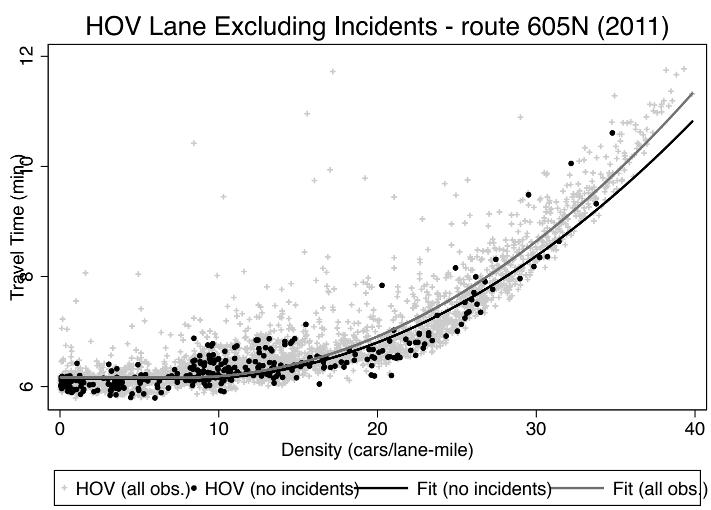


Figure 5: Representative travel time density relationships for mainline and HOV lanes excluding incidents.

(a) Route I-605N mainline lanes



(b) Route I-605N HOV lane



C Appendix tables

Table 1: Mean calculated marginal congestion costs to mainline and HOV lane commuters in $ per car-mile excluding all observations with a traffic "incident." Mean marginal net external cost for HOV lane commuters at various levels of induced demand α in $ per car-mile, excluding incidents.

				HOV Lane Mean Net MEC ($/car-mile)					
Route	**n**	***MEC_l***	***MEC_h***	**$\alpha = 0$**	**$\alpha = 0.2$**	**$\alpha = 0.4$**	**$\alpha = 0.6$**	**$\alpha = 0.8$**	**$\alpha = 1$**
105 E (PM)	215	0.78	0.77	-0.78	-0.47	-0.16	0.15	0.46	0.77
105 W (AM)	215	0.57	0.60	-0.54	-0.32	-0.09	0.14	0.37	0.60
10 E (PM)	89	1.19	0.87	-2.69	-1.98	-1.27	-0.55	0.16	0.87
10 W (AM)	93	0.97	0.51	-2.41	-1.83	-1.25	-0.66	-0.08	0.51
210 E (PM)	88	1.03	1.31	-0.74	-0.33	0.08	0.49	0.90	1.31
210 W (AM)	86	0.75	0.55	-0.94	-0.64	-0.34	-0.04	0.25	0.55
405 N North Rt. (PM)	88	0.85	0.54	-1.17	-0.83	-0.49	-0.15	0.20	0.54
405 N South Rt. (AM)	88	0.62	0.36	-0.88	-0.63	-0.38	-0.14	0.11	0.36
405 S North Rt. (AM)	86	1.22	0.90	-1.54	-1.05	-0.56	-0.08	0.41	0.90
405 S South Rt. (PM)	86	1.06	1.05	-1.07	-0.65	-0.22	0.20	0.63	1.05
605 N (PM)	168	0.75	0.84	-0.66	-0.36	-0.06	0.24	0.54	0.84
605 S (AM)	117	0.68	0.34	-1.02	-0.75	-0.48	-0.21	0.06	0.34

Notes: Mean mainline and HOV lane marginal external costs, MEC_l *and* MEC_h calculated as $n_l T'_l V$ and $n_h T'_h V$ assuming two riders per HOV lane car on all routes except 10E and 10W where three riders is assumed. The constant value of time, V is $21.46. We ignore additional vehicle-related use externalities.