CESifo WORKING PAPERS

7693
2019

June 2019

Welfare Losses of Road Congestion

Antonio Russo, Martin W. Adler, Federica Liberini, Jos N. van Ommeren

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7693*
Category 9: Resource and Environment Economics

Welfare Losses of Road Congestion

Abstract

We estimate the marginal external congestion cost of motor-vehicle travel for Rome, Italy, using a methodology that accounts for hypercongestion (a situation where congestion decreases a road's throughput). We show that the external cost – even when roads are not hypercongested – is substantial, equaling about two thirds of the private (time) cost of travel. About one third of this cost is borne by public transport users. Most roads are never hypercongested, but some are hypercongested for more than one hour per day. Hypercongestion accounts for about 40 percent of congestion-related welfare losses. Welfare losses incurred on roads that are hypercongested are substantial, predominantly because of a reduction in speed rather than throughput. Our results suggest policies that reduce congestion can result in important welfare gains.

JEL-Codes: D620, R410, H230.

Keywords: marginal external congestion cost, deadweight loss, hypercongestion, public transport.

Antonio Russo
ETH Zurich / Switzerland
russo@kof.ethz.ch

Martin W. Adler
VU University Amsterdam / The Netherlands
economics@atadleradvisory.com

Federica Liberini
ETH Zurich / Switzerland
liberini@kof.ethz.ch

Jos N. van Ommereen
VU University Amsterdam / The Netherlands
jos.van.ommeren@vu.nl

This version: May 2019
We thank Rome's Mobility Agency (Agenzia per la Mobilita) and the Italian regulator for public sector strikes (Commissione di Garanzia per gli Scioperi) for kindly providing data. We would like to thank Alex Anas, Richard Arnott, Pierre-Philippe Combes, Gilles Duranton, Dan Jaqua, Ismir Mulalic, Luis Rizzi, Ken Small and Erik Verhoef for insightful comments. We are grateful to audiences at UC Irvine, University of Toronto, Newcastle University, London School of Economics, VU Amsterdam, Brno University of Technology, Vienna University of Economics and Business, Institut d'Economia de Barcelona, Universitat Autonoma de Barcelona, Danish Technical University, GATE Lyon, PUC Santiago, the UEA meeting in Minneapolis, the IIPF conference in Lake Tahoe, the International Trade and Urban Economics workshop in St Petersburg, the Verkehrsökonomik und –politik Conference in Berlin and the meeting of the Italian Society for Transport Economics. All errors are our responsibility. We acknowledge financial contribution by the European Research Council-OPTION program.

1. Introduction

This paper measures the welfare losses due to travel delays caused by road congestion, focusing on private motor-vehicle (i.e. cars, motorbikes) and public bus traffic, for a wide set of road segments in the city of Rome, Italy. We explicitly account for a situation where congestion decreases a road's throughput, originally labelled by William Vickrey (1973) as *hypercongestion*.

Road congestion is a major issue in urban areas throughout the world. The average road user in France, Germany, the UK and the US spent 36 hours in gridlock in 2013 (CEBR, 2014). To deal with congestion, policymakers have several options, including road tolls, quantity restrictions (e.g. license-plate rationing), subsidized public transit supply and infrastructure expansion. None of these options comes at a low price, however. Tolls and quantity restrictions are politically controversial (De Borger and Proost, 2012), transit subsidies can be expensive (Parry and Small, 2009) and road infrastructure expansions produce additional traffic (Duranton and Turner, 2011). Given that policies that reduce congestion are not cheap, quantifying the costs imposed by congestion is important. However, we still know surprisingly little about these costs in large urban areas.

Since Pigou (1920), the most common way of characterizing road congestion externalities has been to focus on a (small) road segment and postulate a positive monotone relationship between the time to travel this segment and traffic flow, i.e. the number of vehicles that pass the segment per unit of time, also called throughput.[1] However, this characterization is inaccurate, because congestion relates to a more complex externality: as more cars take the road, higher *density* of vehicles – not flow – forces drivers to slow down, because it is the distance between cars that affects their speed (Greenshields, 1935). Importantly, when density is large enough, increasing it further slows traffic down to such an extent that flow decreases. Hence, the relation between traffic throughput and travel time, i.e. the road supply curve, is not monotonically upward sloping, but rather backward bending. This throughput-reducing phenomenon is called hypercongestion, and has been observed on several types of roads, e.g. highways and dense city road networks (Keeler and Small, 1977; Geroliminis and Daganzo, 2008).[2]

---

[1] This includes academic publications (e.g., Mayeres et al., 1996), authoritative reports by the US Federal Highway Administration (e.g., FHWA, 1997) and in much-cited handbooks (e.g., Maibach et al., 2008).

[2] The situation where congestion decreases a road's throughput is originally labelled by William Vickrey (1973) as hypercongestion. There are many other situations where congestion reduces throughput (Small and Chu, 2003). For example, overloads of switching equipment cause breakdowns in telephone networks. Storms drain clogs when high water flow carries extra debris. Service desk clerks take longer to handle enquiries when facing high

Hypercongestion has been extensively analyzed by the theoretical literature. Some authors have modeled this phenomenon as the consequence of downstream bottlenecks (Verhoef, 2003, 2005; Hall, 2018). Others have analyzed hypercongestion in isotropic road conditions (Arnott and Inci, 2010). However, the *empirical* economic literature on congestion within cities tends to ignore hypercongestion.[3] We aim to introduce a straightforward approach to measure the presence and welfare losses due to hypercongestion, and to calculate the marginal external congestion cost of travel.

The city of Rome provides an interesting setting for our study. First, congestion is heavy compared to other cities of similar size. This is the result of a high modal share of cars and motorbikes, combined with a limited supply of public transit infrastructure.[4] Second, Rome's public transport system relies primarily on buses, which share road space with private traffic.[5] This enables us to quantify the costs of road congestion through travel delays for bus travelers. Third, public transport strikes are extremely frequent in Rome and vary in intensity. We have information about *hourly* reductions in public transit supply due to strikes, which we use as an exogenous shock for identification purposes.

The first contribution of our paper is to propose a methodology to consistently estimate road supply curves for city roads, allowing for hypercongestion. In an early contribution, Keeler and Small (1977) address the issue by estimating travel time as a function of flow (on highways) and then inverting the estimated function. We improve upon their methodology by following the transportation science literature, which estimates the effect of vehicle *density* (e.g. the number of vehicles divided by the length of the road segment) on travel time and then derive the travel time-flow relation by applying fundamental identities (Hall, 1996; Geroliminis and Daganzo, 2008).[6] However, contrary to the transportation literature, we account for potential endogeneity issues. Common unobservable shocks, e.g. road accidents, may affect density and travel time simultaneously, producing an omitted variable bias. More fundamentally, density is

---

paperwork flows. Hall (2018) provides a valuable survey of engineering studies indicating the presence of hypercongestion.

[3] Theory indicates that the deadweight loss of congestion is significantly higher in the presence of hypercongestion, because the excessive demand for car travel not only raises travel time or queueing (Verhoef, 2005).but also because the road's throughput is reduced (Arnott, 2013; Fosgerau and Small, 2013; Hall, 2018).

[4] Traffic congestion indexes rank Rome among the world's most congested cities, similarly to Mexico City, Jakarta and Bangkok. Note that these cities are much bigger than Rome. The TomTom Traffic Index ranks Rome as the sixth most congested city during the morning peak. The Castrol Stop-Start Index places Rome just behind Mexico City.

[5] Until 2015, Rome had only two subway lines, recently augmented by a third short line, which is exceptionally low for a European city of comparable size (2.8 million inhabitants). Limited public resources and a high concentration of archeological sites are regularly cited as the main causes for the lack of infrastructure provision.

[6] In a dynamic model of congestion, Henderson (1974) also models travel time as a function of density, measured as the quantity of commuters on a road at a given time. See also Henderson (1981).

defined as the product of flow and travel time. Hence, any measurement error in travel time induces a positive correlation with density, implying an overestimate of its effect.

To overcome endogeneity issues, we instrument for density using two different demand-shifting instruments (which generate similar results). The first exploits weekly regularities in travel demand, captured by hour-of-the-week dummies. Specifically, we use the fact that, for example, traffic demand on Mondays at 9am is higher than at 12am. The second instrument exploits shocks in hourly public transit supply due to strikes, which induce motor vehicle travel demand shocks. Both instruments are arguably valid conditional on a range of controls for time (e.g., hour-of-the-day fixed effects) and weather. The validity of the strike instrument is enhanced by the fact that we observe many strikes in Rome.

Our second contribution is to employ the road supply estimates to quantify the marginal external cost of congestion and the associated deadweight losses, using data from a large set of roads in Rome. We find that the marginal external cost of congestion is high even when roads are not hypercongested. On average, we find that each motor-vehicle kilometer produces an external cost of 0.88 minutes per kilometer of additional travel delay. This value is equal to about 66 percent of the average motor vehicle travel time (1.33 min/km) in Rome. The external cost is much larger in the morning peak (1.4 min/km). Our results imply that, on average, the hourly deadweight loss per kilometer of road lane from congestion equals 2.21 vehicle-hours.

In our sample, hypercongestion is rather rare: it is present in only about 1.5 percent of the observations and limited to a subset of roads. Most roads are never hypercongested, but some are hypercongested for more than one hour per day. Nevertheless, hypercongestion accounts for a significant share (about 40 percent) of the overall deadweight loss. The deadweight loss due to congestion per unit of time is about 50 times higher when there is hypercongestion.[7] Welfare losses incurred on roads that are hypercongested are predominantly due to reductions in speed, rather than in throughput.

These results suggest that policy interventions that reduce congestion such as road pricing can bring significant welfare gains. However, if pricing is (for political reasons) unavailable, it is possible to achieve gains just by limiting hypercongestion, for example by adopting demand management measures such as adaptive traffic lights (Kouvelas et al., 2017).

Our third contribution is to estimate the external costs of congestion on *bus users*, which have so far been ignored in the empirical literature (see Small, 2004, for a simulation study). We show that in Rome, where buses travel on mixed traffic roads, about one third of the

[7] The low frequency of hypercongestion is consistent with recent findings of the traffic engineering literature. See, e.g., Loder et al. (2017).

marginal external congestion cost of motor vehicle travel is borne by bus travelers. These results have policy implications particularly for large cities in emerging and less developed economies, where buses are often the mainstay of the public transport system.

Our welfare calculations are based on a static framework, assuming an isotropic road segment in steady state (in line with Arnott and Inci, 2010). This framework fits with our traffic data, collected using loop detectors, which involves observations of traffic conditions for specific *discrete* periods (of one hour) and locations (roads). Applying this framework is straightforward but requires two important – potentially restrictive – assumptions. First, we ignore that individuals have preferences regarding the time of travel (e.g. arrival time at work) and rescheduling of the timing of trips might be costly. Hence, we do not capture the welfare losses of congestion through rescheduling of trips (to other time periods). The rescheduling cost are likely much smaller than the losses through increases in travel time.[8] Consequently, we underestimate the welfare losses from congestion.

In addition, road networks are typically not isotropic and may include bottlenecks that cause queueing. As the theoretical literature has pointed out, hypercongestion may also result from queuing before a bottleneck (Vickrey, 1969; Arnott et al., 1990; 1993, 1994; Van den Berg and Verhoef, 2011). Our data does not allow to distinguish between different causes of hypercongestion.[9] Therefore, we also employ an "approximated measure" of welfare losses, which assumes that hypercongestion is due to a downstream bottleneck operating at maximum capacity (Verhoef, 2003, Arnott et al., 1990). The latter allows for travel delays due to queueing before the bottleneck which results into welfare losses. We show that the welfare losses using this alternative measure are almost identical to those using our static framework. This is intuitive, because we will show that for hypercongested roads, the welfare loss due to a reduction in throughput is much smaller than that due to reductions in speed.

Our findings contribute to the literature measuring the costs of congestion (Small and Verhoef, 2007). We adopt a disaggregate framework that measures such costs at the level of single roads and is thus complementary to recent studies by Couture et al. (2018), who estimate

---

[8] In the extreme case that travelers have identical preferences regarding arrival times and all go through the same bottleneck, then the rescheduling losses are exactly equal to the travel time losses (Small and Verhoef, 2007). Because travelers are extremely heterogeneous regarding arrival times and many travelers do not reschedule, it is plausible that these rescheduling cost are an order of magnitude smaller than the travel time losses.

[9] Recently, Hall (2018) has shown that queuing may also cause reductions in a bottleneck's throughput and evaluates the implications for pricing and welfare.

aggregate travel supply relationships for a large sample of U.S. cities, and Akbar and Duranton (2016), who estimate this relationship at a citywide level for Bogotá.[10]

The paper proceeds as follows. Section 2 introduces the theory that underlies our identification strategy. Section 3 and 4 present the empirical approach and the data. Section 5 provides estimates of the marginal external cost of road congestion and welfare losses. Section 6 concludes.

2. Theoretical background

We develop a theoretical framework to guide the estimation of the relations between traffic flow (or throughput) and travel time, the marginal external cost of congestion and the ensuing welfare losses. There are two travel modes: private motor vehicles (cars and motorbikes) and public transit, which consists only of bus service. We consider an *isotropic* one-lane road segment of unit length (say, 1 km) in a *stationary steady-state*. Individuals choose whether to travel and which mode to use.

In our empirical application, we observe traffic conditions at the hourly level. Hence, our steady-state assumption implies that we ignore dynamic adjustments in traffic conditions *within* each hour. The assumption of an isotropic road segment implies we do not model the presence of downstream bottlenecks, i.e. reductions in road capacity that may cause queueing. In the welfare analysis, we shall show how our framework can take into account the presence of bottlenecks and their impact on welfare.

2.1 The demand for travel

There is a given number of heterogeneous individuals, N, who have some valuation for traveling, either by motor vehicle or public transit. Each individual takes at most one trip. Let N_{PT} be the demand for public transit trips, i.e. the number of individuals who travel by public transit. Similarly, let N_M be the demand for motor-vehicle trips.[11] The demand functions for public transit as well as motor-vehicle travel are negatively sloped with positive cross-price elasticities (i.e. the modes are substitutes). For public transit travelers, the generalized price of travel, p_{PT}, encompasses fares and time costs. For motor-vehicles, the price of travel only consists of travel time T. We have:

(1) $$N_{PT} = N_{PT}(p_{PT}, T),$$

[10] Our approach may therefore be less representative of travel costs at a wide area level, for example because it does not account for the possibility that drivers avoid heavily congested roads by taking detours. On the other hand, our approach provides a more fine-grained view of congestion costs at the street level.
[11] Hence there are N_N individuals who do not travel, where $N_N = N - (N_{PT} + N_M)$.

(2) $$N_M = N_M(T, p_{PT}).$$

In the welfare analysis, we assume that the demand function for motor-vehicle travel is linear:

(3) $$T = \mu + \theta p_{PT} - \varphi N_M,$$

where $\mu > 0$, $\varphi > 0$ and $\theta > 0$.[12] We shall assume that φ is a constant (we considering a wide range of values for this slope in the analysis), but we will estimate θ and μ separately for each hour and road. This is important as one expects the intercept of demand to change, for instance, between peak and off-peak hours. A key assumption for our empirical analysis is that shocks captured by changes in θ and μ, as well as public transit strikes, shift the demand for motor vehicle travel and can thus be exploited for identification purposes.

2.2 Cost functions for motor vehicle and bus travel

2.2.1 The road supply curve

Let *D* be the density of motor vehicles (e.g., 20 vehicles per kilometer), and *F* be the flow (or throughput) of vehicles passing the segment per unit of time (e.g. 10 vehicles per minute). The *road supply curve* is defined here as the relation between the time cost of motor vehicle travel *T* (e.g., two minutes per kilometer) and the flow, *F*. This is a supply relation, describing how the price of travel changes with the road's travel output. We denote this relation as $T(F)$ and derive it from fundamental physical relations, as follows. In line with the transport engineering literature (Helbing, 2001), we assume T is an increasing and convex function of *D*:

(4) $$T = h(\mathrm{D}),$$

where $\frac{\partial h}{\partial D} > 0$ and $\frac{\partial^2 h}{\partial D^2} > 0$. This assumption is intuitive: drivers choose their speed based on the distance to the car in front of them: a higher implies a shorter distance between cars and thus lower speed.[13] Furthermore, we use the following fundamental identity:

(5) $$D \equiv F \times T.$$

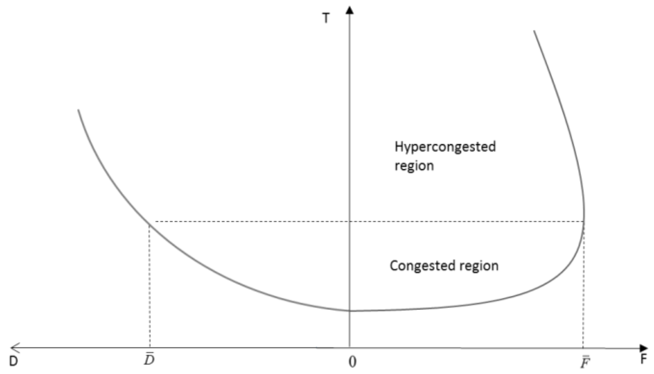
Combining (4) and (5), and applying the Implicit Function Theorem, we obtain the relation between travel time and flow:

(6) $$\frac{dT}{dF} = -\frac{-\frac{\partial h(FT)}{\partial F}}{1 - \frac{\partial h(FT)}{\partial T}} = \frac{\frac{\partial T}{\partial D} T}{1 - \frac{\partial T}{\partial D} F}.$$

[12] Demand for transit is also linear. We characterize it in Appendix A as it is not of importance here.

[13] For simplicity, we only focus on density of motor-vehicles, ignoring buses. Buses have a stronger effect on travel time delays than cars. However, in Rome, less than 1 percent of total traffic consists of buses. Hence, our results remain essentially unchanged even if one bus creates the same congestion as ten private cars.

Figure 1 – Fundamental diagram of traffic congestion.



To understand this relationship, note that (5) implies when D is zero, F is also zero. Now consider an increase in D. Expression (6) implies that when $\partial T/\partial D < 1/F$, a higher D raises travel time and F. [14] As D increases, a critical level $\bar{D}$ is reached, where the denominator of (6) becomes zero and dT/dF approaches infinity. Above $\bar{D}$, $\partial T/\partial D > 1/F$ holds, so $dT/dF < 0$ and $dF/dD < 0$. Consequently, due to the fundamental identity (5), D has a positive mechanical effect on F, but also an indirect negative effect because vehicles travel at lower speed. When the latter dominates, the supply relationship bends backwards, and there is *hypercongestion*.[15,16] Figure 1 provides an illustration: the left-hand side of the figure denotes the postulated effect of D on T, whereas the right-hand side shows the implied relationship between T and F.

The above discussion implies that there is a maximum flow on the road segment, $\bar{F}$, and a corresponding critical level of density $\bar{D}$, such that:

$$1 - \bar{F}\,\frac{\partial T}{\partial D}\Big|_{D=\bar{D}} = 0. \tag{7}$$

Hypercongestion occurs when $D > \bar{D}$. Henceforth, we specify (4) as follows:

$$T = \beta e^{\alpha D}, \tag{8}$$

where $\alpha, \beta > 0$, as proposed by Underwood (1961). This relation provides an accurate description of the travel time-density relation for roads in Rome. Given (8), we have:

$$\frac{dT}{dF} = \frac{\alpha T^2}{1 - \alpha D}. \tag{9}$$

[14] It also follows that dF/dD = (1-F ∂T/∂D)/T. Note that dT/dF and dF/dD have the same sign.

[15] Convexity of *h(.)* is crucial for this argument: if the function is linear, hypercongestion cannot occur.

[16] There is a debate in the transport literature, whether hypercongestion is a stable phenomenon in a static framework. Arnott and Inci (2010) show that hypercongestion is a stable phenomenon given similar steady-state assumptions as in the current paper. Gonzales (2015) shows that hypercongested traffic can be part of a stable equilibrium state given the presence of public transit.

Consequently, the critical density is $\overline{D} = 1 - \alpha\overline{D} = 0$, so $\overline{D} = 1/\alpha$.[17] Hence, by estimating (8), one can calculate whether hypercongestion is present (i.e. whether $D > \overline{D}$).

2.2.2 The generalized price of public transit travel

The generalized price of public transit travel, p_{PT}, increases with the in-vehicle bus travel time, T_{PT}, and the fare, f, but decreases with the supply of public transit, S.[18] We assume p_{PT} is an additive function of its arguments (note that we normalize the monetary value of time to one for simplicity):

$$p_{PT} = T_{PT}(D) - \vartheta(S) + f, \quad (10)$$

where $\vartheta(.)$ is a positive and increasing function, and where T_{PT} increases with motor-vehicle density, because, like other vehicles, buses drive slower in congestion. We aim to estimate the effect of motor-vehicle density on travel time of bus travelers, T_{PT} (and, hence, on p_{PT}). Accordingly, it makes sense to assume the same functional form as (8), but allowing for different parameters:

$$T_{PT} = \gamma e^{\sigma D}. \quad (11)$$

Observe that in-vehicle bus travel time, T_{PT}, consists of two components: time *between* stops and idle time *at* stops. The latter depends on traffic density as well as the number of boarding/alighting passengers at each stop, which we do not have information on.[19] Hence, in the empirical analysis we will ignore time at stops, although we take it into account in the welfare analysis. Note also that buses may either share the road with other vehicles, which we label as *mixed traffic lanes*, or use roads that enclose a *dedicated bus lane*, where bus traffic is largely – but possibly not fully – separated from other vehicles.[20] In the empirical analysis, we allow γ and σ to differ between mixed traffic and dedicated bus lanes.

We are also interested in the relationship between traffic flow and bus travel time. Combining (11) with (5) and (6), we write the marginal effect of an increase in private motor-vehicle flow on bus travel time as:

---

[17] We will see that α is around 0.02 when estimating density (expressed in veh/km-lane), which implies that $\overline{D}$ is about 50 vehicles per kilometer of road lane.

[18] For instance, higher supply of public transit, *S*, results in higher service frequency, which reduces waiting time at bus stops (Mohring, 1972).

[19] Congestion may affect time at stops, for example, because dense traffic makes it harder for buses to maneuver in and out of stops. In the welfare analysis, we assume that time at stops is not affected by congestion, hence we likely underestimate the negative welfare effect of congestion.

[20] In Rome (as in other large cities) dedicated lanes are shared with taxis, ambulances, police and public official vehicles.

$$(12) \qquad \frac{dT_{PT}}{dF} = \gamma e^{\sigma D} \sigma \frac{dD}{dF} = T_{PT}\sigma \left(T + F\frac{dT}{dF}\right) = T_{PT}\sigma \frac{T}{1-\frac{\partial T}{\partial D}F}.$$

We use this expression to derive the marginal external cost of private motor-vehicle travel on bus travelers.[21]

Our empirical identification strategy exploits public transit strikes as an exogenous shock to transit supply, S. To motivate this choice, let us focus on the theoretical effect of strikes. We assume that, in the absence of strikes, S does not vary over time and is normalized to one.[22] If a public transit strike takes place, however, S drops below one, so $S \in [0,1]$, causing an increase in the generalized price of public transit, p_{PT}. Note also that, in addition to this direct effect (that is, given the level of congestion), the reduction in transit supply also generates an increase in demand by motor-vehicle travel and, hence, in D. Therefore, there is also an indirect effect on the price of transit travel, because T_{PT} goes up due to the increased congestion (see (11)).

2.3 Welfare Analysis

We now focus on the steady-state equilibrium, where motor-vehicle travel demand, N_M, equals supply, F. Furthermore, demand for public transit equals supply.[23]

2.3.1 Social and external costs of motor vehicle travel

To understand the external congestion cost of travel, we focus on the time costs of travel (we ignore other external costs, such as fuel consumption, pollution and noise). The aggregate time cost is then the sum of aggregate time cost of travel for motor vehicle users, $F \times T$, plus the aggregate time cost for public transit travelers, $N_{PT} \times T_{PT}$.

The marginal *external* cost of motor-vehicle travel, which we denote as MEC, is the difference between the marginal social cost of a trip, *MSC* (the increase in the aggregate time cost due to the marginal motor-vehicle kilometer) and the user cost of this kilometer, T. The external cost can be divided in two components: the external cost to motor vehicle users,

[21] Note that the last term in the equality is positive in the absence of hypercongestion.
[22] This assumption literally holds for Rome between 8 a.m. and 17 p.m. It does not hold outside these hours, hence in our empirical analysis we will control for hour of the day.
[23] We use hourly observations for many road segments in the empirical analysis. Hence, for each road segment, we assume a steady-state equilibrium for a period of one hour. We ignore any variation *within* the hour, which might lead to underestimates of the welfare losses of congestion (and the pervasiveness of hypercongestion). This can be shown by noting that travel time is a convex function of density, and therefore that travel time is a convex function of flow, when density is in the hypercongested range.

denoted MEC_M, and that to public transit users, denoted MEC_{PT}. Hence, $MSC = MEC_M + MEC_{PT} + T$. We measure these external costs in the empirical analysis.

Let us focus first on the marginal external cost on *motor-vehicle users*, MEC_M. Differentiating $F \times T$ with respect to F using (6) and subtracting the user cost, T, shows that:

$$MEC_M = \frac{d[FT]}{dF} - T = \frac{dT}{dF}F = \frac{\frac{\partial T}{\partial D}D}{1 - \frac{\partial T}{\partial D}F}. \tag{13}$$

Given (8), MEC_M is specified as:

$$MEC_M = \frac{\alpha DT}{1 - \alpha D}. \tag{14}$$

Let us now focus on the external cost of motor-vehicle travel on bus users (i.e. the effect of an increase in motor-vehicle trips, *F*, on travel time of bus travelers through an increase in T_{PT}). We now differentiate $N_{PT} \times T_{PT}$ with respect to *F*. Given (12) and (8), the marginal external cost of motor vehicle travel on public transit users can be written as: [24]

$$MEC_{PT} = \frac{dT_{PT}}{dF}N_{PT} = T_{PT}\sigma N_{PT}\left(\frac{T}{1 - \frac{\partial T}{\partial D}F}\right) = T_{PT}\sigma N_{PT}\left(\frac{T}{1 - \alpha D}\right). \tag{15}$$

The marginal social cost of motor vehicle travel is then:

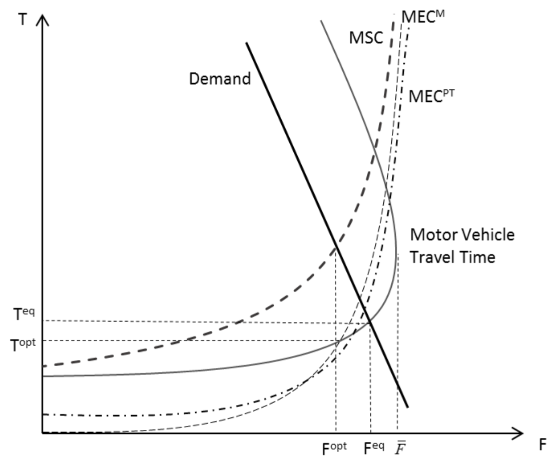
$$MSC = (\alpha D + T_{PT}\sigma N_{PT})\frac{T}{1 - \alpha D}. \tag{16}$$

Consider an equilibrium which is *not* hypercongested, i.e. $1 - \alpha D > 0$, so that an increase in D causes an increase in F. It follows from the above expressions that MEC_M, MEC_{PT} and MSC are positive and increase with F, with a slope that tends to infinity as F reaches the critical level $\bar{F}$. See Figure 2, which shows a backward-bending supply curve plus MEC_M, MEC_{PT} and MSC.

The socially-optimal traffic flow and travel time, and therefore density, are defined where the inverse demand for motor-vehicle travel (expression (3)) equals the marginal social cost of travel. See Figure 2, where the superscript *eq* denotes the equilibrium and *opt* the optimum.

[24] We ignore the welfare effect of the marginal user of public transport (as an increase in motor vehicle flow goes along with a smaller decrease in public transit use). If public transport users pay the marginal social cost of transit trips, the latter effect is zero.

Figure 2 – Marginal external and social costs of motor vehicle travel.



The deadweight loss of traffic congestion, DWL, in a given equilibrium is the increase in welfare when moving from this equilibrium to the optimal one. DWL is straightforward to calculate in our setup and equal to the integral of the difference between MSC and the inverse demand function for motor vehicle travel, computed between the optimal flow and the equilibrium one. See the top-right quadrant of Figure 3. A portion of the deadweight loss is due to the increase in travel time for transit users, denoted by DWL_{PT}. This portion is shown in the bottom-right quadrant of the same figure. A welfare-maximizing policy intervention is to set a road toll equal to $MEC_M + MEC_{PT}$ (evaluated at the optimum). In conclusion, when the equilibrium is not hypercongested, our theoretical framework is able to calculate *the marginal external cost of travel* as well as the *welfare losses of congestion*.

Consider now an equilibrium where the road is hypercongested, i.e. such that $1 - \alpha D < 0$ holds. Therefore, an increase in D causes a *reduction* in the road's throughput, *F*. The above expressions imply then that MEC_M, MEC_{PT} and MSC are negative (and not shown in Figure 2). Intuitively, density is so large that a reduction in density results in lower travel time but higher flow, i.e. lower costs for society as well as higher travel output. Therefore, an equilibrium with hypercongestion can never be optimal. Hypercongestion is an inefficient (i.e. exceedingly slow) way of 'producing' travel, not only because of the exceedingly high travel time, but also because the road's throughput (flow) is lower than optimal. Accordingly, the associated deadweight loss tends to be large. Figure 4 provides two illustrations of this type of equilibria and the associated deadweight loss on the motor-vehicle market (given by the extra travel time as well as the possible loss in throughput). In one illustration (left panel), the demand for motor vehicle travel is perfectly elastic, whereas in the other illustration (right panel), it is quite inelastic. Contrary to the non-hypercongested case, the deadweight loss can be large even if demand is completely inelastic.

Figure 3 – Deadweight loss of motor-vehicle travel in congested equilibria

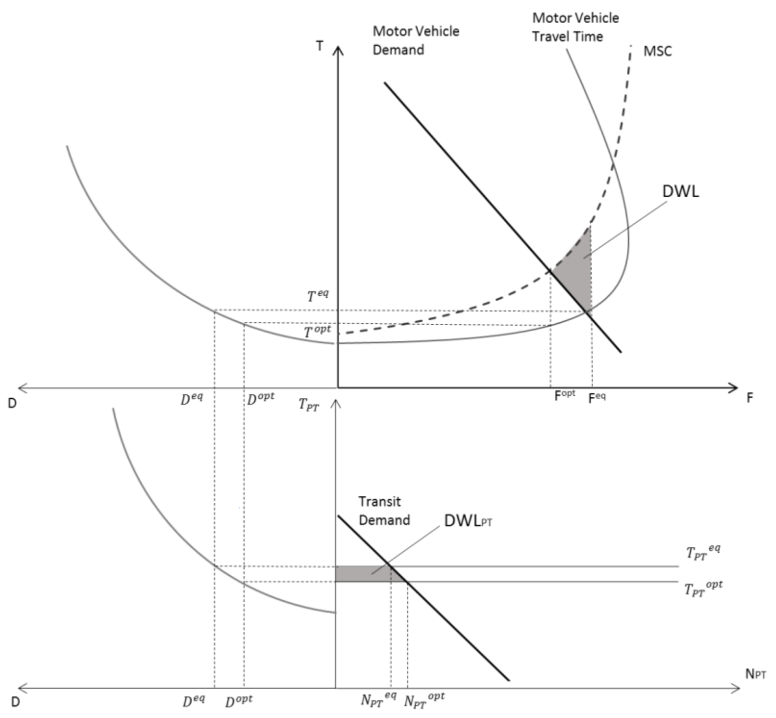
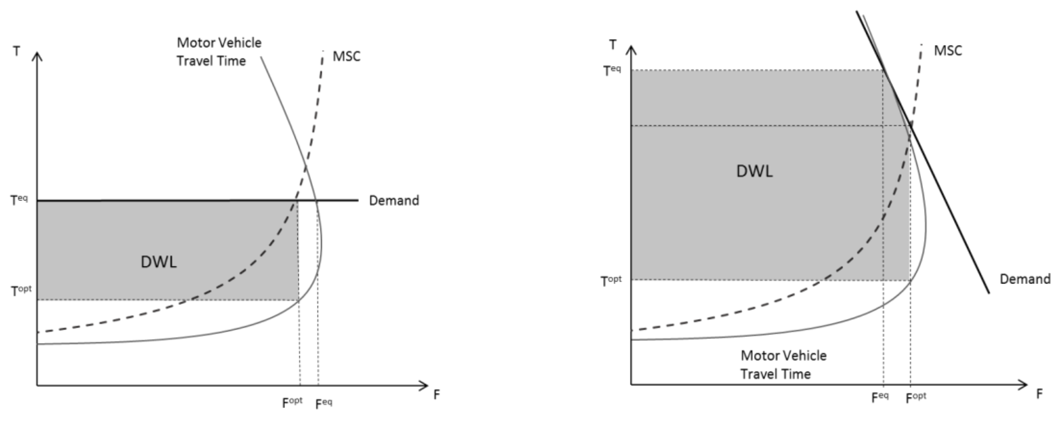


Figure 4 – Deadweight loss of motor-vehicle travel in hypercongested equilibria



2.3.2 Hypercongestion, bottlenecks and approximate welfare losses

We take a static approach to describing the road supply conditions. Specifically, we ignore the possibility that an equilibrium with hypercongestion may result from traffic queuing before a bottleneck (e.g. the junction between two roads or a reduction in the number of available lanes).

In the presence of bottlenecks, queueing is the main source of inefficiency (Arnott, 2013; Fosgerau and Small, 2013; Hall, 2018).

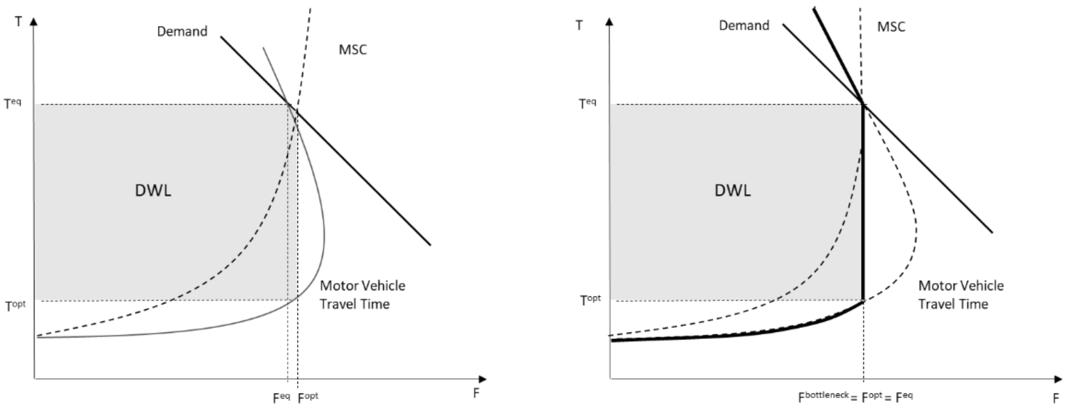
Given the nature of our data (based on vehicle counts and average speed measurements at specific points on a road, see below), we cannot identify whether hypercongestion results from queues before a bottleneck. Nevertheless, it is important to take the possibility of bottlenecks into account for the welfare analysis. Accordingly, we now adapt the conceptual framework to allow for bottlenecks and discuss the main implications for the way we characterize the optima and the deadweight losses from hypercongestion.

Suppose there is indeed such a bottleneck and that we observe traffic conditions at a point before it. The key difference with respect to the framework considered so far is that, by definition, flow on this road cannot exceed the maximum defined by the bottleneck's capacity. Hence, assuming the bottleneck is used to capacity in the equilibrium we observe, the optimum cannot entail a higher flow, but only lower travel time due to removing queues. Accordingly, in this alternative interpretation, we will estimate approximate welfare losses by making the conservative assumption that the observed (hypercongested) flow cannot be exceeded in the optimum.[25] The approximate welfare loss is then obtained by multiplying the observed flow with the difference between the observed travel time and the counterfactual travel time obtained keeping flow constant, but without queueing (and hypercongestion). See Figure 5 for an illustration. Note that this approach potentially underestimates actual welfare losses because it ignores potential losses in throughput due to hypercongestion.[26] However, we will show that both measures give essentially identical results. The reason is that the welfare loss due to the reduction in throughput when roads are hypercongested is small compared to the welfare loss due to reduced travel speed. In other words, the backward-bending part of the supply curve is close to being vertical in the range of values observed in our data.

[25] An alternative interpretation of this assumption is that the road supply curve is effectively vertical, which is in line with Verhoef (1999; 2003). We emphasize that this assumption does *not* imply that we assume away the welfare losses of congestion when roads are hypercongested. Quite the opposite, we will see that substantial welfare gains can be obtained by reducing demand which reduces travel times on the road.

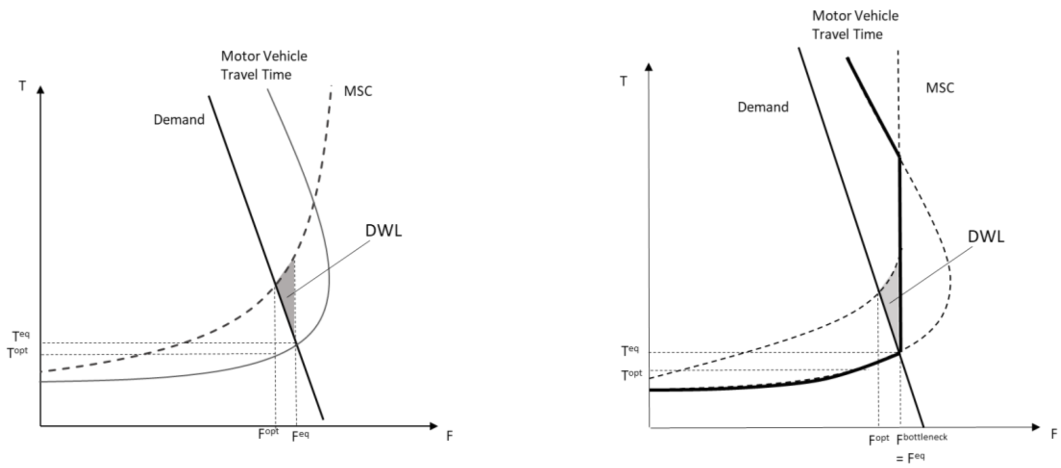
[26] See Hall (2018) for an analysis allowing for reduced throughput due to queueing at a dynamic bottleneck.

Figure 5 – Hypercongested equilibria without (left panel) and with bottleneck (right panel)



Finally, note that, as Figure 6 shows, the alternative interpretation allowing for bottlenecks has no practical consequence for the way we characterize optima and deadweight losses if the observed equilibrium is not hypercongested. Indeed, without hypercongestion the optimal flow is always smaller than the equilibrium one.

Figure 6 – Congested equilibria without (left panel) and with bottleneck (right panel)



3. Empirical Approach

3.1 Estimation of the road supply curve

We are interested in estimating the marginal external cost of congestion through increased travel time of motor-vehicle and public transit travelers. We first focus on the marginal external cost through increased travel time of motor-vehicle travelers. To estimate this cost, one needs information about the relationship between motor-vehicle travel time and flow, i.e. the road supply curve.

We cannot apply standard econometric techniques to estimate the road supply relation, because, as shown above, due to hypercongestion the relationship between T and F is a correspondence and *not* an injective function. We therefore proceed as follows: we first estimate the effect of motor-vehicle density on motor-vehicle travel time by making parametric assumptions on the functional form of $h(D)$, see (4). We then combine this estimate with (6) to derive dT/dF. Given estimates of $h(\mathrm{D})$, denoted $\hat{h}(\mathrm{D})$, for each observation of D, we calculate the predicted travel time $\hat{T} = \hat{h}(\mathrm{D})$, as well as the predicted flow $\hat{F}(\mathrm{D}) = \mathrm{D}/\hat{T}$. We show below that the travel-time flow relationship obtained in this way accurately predicts the observed one.[27]

In our empirical application, we assume for $h(D)$ the functional form in (8), which implies that the logarithm of travel time is a linear function of density. We use observations which vary by road and hour during the week, estimating separate models for each road. To be specific, we assume that $logT_{i,t}$, at road *i* and hour *t* is a linear function of density $D_{i,t}$, given several controls X_t and an error term $u_{i,t}$, so that:

$$logT_{i,t} = \tau_i + \alpha_i D_{i,t} + \kappa_i X_t + u_{i,t}. \quad (17)$$

The controls X_t include weather (e.g. temperature using a third-order polynomial, precipitation) and *three types of time controls*: hour-of-the-day, day-of-the-week fixed effects and week-of-year fixed effects. These time controls aim to capture unobserved changes in supply conditions (e.g. due to road works which only occur during certain periods or due to changes in the amount of sunlight). We emphasize that the estimates without these controls are almost identical. We cluster standard errors by hour, so we allow $u_{i,t}$ and $u_{j,t}$ to be correlated.[28]

We first estimate (17) using OLS. This approach assumes that $u_{i,t}$ is not correlated to density, conditional on controls. However, density may be endogenous, because (5) implies that density is equal to the flow multiplied with travel time – which is the dependent variable of interest. This aspect is especially problematic, because in many studies – including the current one – density is derived from observations of flow and travel time, rather than being explicitly observed. Therefore, any measurement error in travel time causes a positive correlation between

[27] Keeler and Small (1977) estimate flow *directly* as a quadratic (and therefore non-monotonic) function of travel time and then invert the estimated function. There are two disadvantages of this approach. First, it generally does not provide the causal effect of flow on travel time. Second, we have compared their approach for Rome with our approach and it appears that the fit of the estimated relation between flow and travel time is worse, despite relying on more parameters. The intuition for the lower fit is that the relationship between the logarithm of travel time and density is approximately linear, and therefore straightforward to estimate, whereas the relationship between flow and travel time is non-monotonic, and therefore difficult to estimate.

[28] Hence, each cluster contains a number of observations equal to the number of road segments observed.

travel time and density, resulting in overestimation of the effect of density.[29] Besides, measurement error is not the only source of endogeneity. For example, many unobserved supply shocks, such as road closures, accidents or bad weather, may simultaneously affect density and travel time. In sum, it is possible that OLS estimates of α_i are inconsistent. Formally, the endogeneity problem with the estimation of (17) is that the key requirement that $E(D_{i,t}u_{i,t}|X_t) = 0$ may fail.

To deal with endogeneity, we note that (17) essentially describes a (technological) supply relationship. To identify this relation, therefore, we use an instrumental variable approach exploiting variation in demand. Our approach is based on two demand-shifting instruments. The first instrument exploits regularities in travel demand over the hours of the week, as in Adler et al. (2019). Specifically, we use *hour-of-the-week* dummies, z_t, as instruments (e.g. a dummy for Monday morning between 9 and 10 AM is one instrument). Hence, we assume that $E(z_t u_{i,t}|X_t) = 0$. Importantly, X_t includes three other types of time fixed effects – hour-of-the-day, day-of-the-week and week-of-the-year dummies – as controls. Thus, the variation we exploit is that demand is higher during a certain hour of the week, but we control for the hour of the day (i.e., we control for daily variation in sunlight or policies that apply only on certain hours of the day, e.g. traffic light changes), day of the week and week of the year (i.e., we control for roadworks that tend to occur only on certain days or that are specific to a certain period of the year). So, for example, we exploit the fact that demand is lower at 7am in on Mondays compared to 8am on the same day, and we control for that at 7am there might be less light, which potentially influences driving speed for given levels of traffic density.

Our argument for why $E(z_t u_{i,t}|X_t) = 0$ must hold is that these *hour-of-the-week* dummies capture shifts in demand, conditional on other time fixed effects that control for any possible shifts in supply (e.g., for a given density, drivers may reduce speed in the evening because it gets darker).[30] Note that a hour-of-the-week dummy essentially measures the demand for a certain hour of the week averaged over the whole period. Hence the exclusion restriction is that, conditional on other time fixed effects, variation in *average* density over hour of the

[29] This problem is standard in many applications. One example are labor supply models where the number of hours worked per week is regressed on the hourly wage rate which is calculated as the weekly wage divided by the number of hours worked. See, e.g., Borjas (1980). We ran simulations – available upon request – indicating that measurement error in travel time is important: when its standard deviation is only 10 percent of the standard deviation of travel time, then the upward bias in the estimate of α is about 30 percent. Note also that, in presence of measurement error in flow, one would expect some attenuation bias (Wooldridge, 2002, p.75). However, our simulations indicate that measurement error in flow produces an almost negligible downward bias.

[30] Note that the travel demand function is usually expressed as a relationship between travel time and flow. Because density is the product of travel time and flow, it means that a shift of the demand function results in a shift in the in the demand relationship between travel time and flow.

week, where we average over the full period of observation, is entirely due to changes in demand. Consequently, the instrument is valid given the mild – and realistic in the context of Rome – assumption that that there are no supply shocks – including policies – that change the speed of vehicles systematically at a certain hour for a specific day of the week. Thus, this IV approach allows for policies that adapt road supply with a fixed pattern over the time of the day across different days of the week. For example, it allows for roadworks which only take place in the evening, or only on Fridays. Our controls also take care of environmental conditions that affect driving speed for given density at certain hours of the day, as well as weather conditions (rain and temperature).

Let us suppose now that the above instrument is invalid, which would be the case if there are road supply shocks that are specific to certain hours of the week (e.g., on Monday morning, traffic lights do not function properly). In this scenario, one wishes to control for systematic hour-of-the-week variation. Hence, we introduce an alternative instrument where – in addition to the other time controls mentioned – we also control for these hour-of-the-week fixed effects. The second instrument uses variation in the supply of public transit, S_t, due to strikes, which cause positive shocks to the demand for motor vehicle travel as argued above. This instrument is arguably valid, i.e. random, conditional on our extensive set of time controls. Furthermore, this instrument is more likely to be valid *given a large number of strikes,* as we observe for Rome. This reduces the probability of relying on a set of strikes which accidentally occur on days where the congestion level differs from the norm.

When adopting this second instrument, the use of hour of the week controls in (17) is key for two reasons. First, the occurrence of strikes is not completely random with respect to hours of the week (for example, strikes are common on Friday outside peak hours). Second, hour-of-the-day fixed effects capture any variation in the supply of *scheduled* public transit (i.e., the schedule in the absence of strikes), which makes it plausible that changes in public transit supply are entirely due to strikes and therefore exogenous.[31]

A potential disadvantage of using strikes is that demand shocks due to strikes may be less representative for identifying road supply curves than demand shocks using travel demand regularities (for example, strikes may have a larger effect at times when roads are busy). Hence, the LATE interpretation of instrumental variable outcomes suggest that the obtained estimate using strikes has less external validity than using travel demand regularities. Another potential disadvantage, when using public transit strikes as an instrument, is that changes in public transit

[31] The results do not change when we do not control for hour-of-the-week fixed effects. Hence, in the context of Rome, it is a matter of taste whether one prefers to control for these fixed effects.

supply directly change the number of vehicles on the road, which may invalidate the assumption that bus strikes are valid instruments of motor-vehicle density. This is a minor issue however, because on average less than 1 percent of all vehicle flow in Rome refers to buses (specifically, only six buses pass a road per hour). Nevertheless, we have addressed this issue by estimating models where we explicitly acknowledge that an increase in public transit supply increases the number of vehicles on the road. We did not find major changes in the results.

Finally, it is not a priori guaranteed that the demand shifters we described are valid instruments for density to identify the supply relation. The LATE interpretation of instrumental variable outcomes allows for heterogeneous effects of the demand-shifting instrument, but requires that these demand-shifting instruments *monotonically* affect density (Imbens and Angrist, 1994; Angrist and Pischke, 2009). It is not obvious that this requirement holds in our setup.[32] In Appendix C, we show that the monotonicity assumption holds when the following condition is satisfied:

$$\alpha T < \frac{T\varphi}{\varphi D - T^2} \qquad \text{or} \qquad T^2 > \varphi D. \tag{18}$$

We will show that this condition generally holds in our data, see Section 5.1.

3.2. Estimating the MEC on motor vehicle travelers

Given estimates of α based on (17), we can derive MEC_M using (14). Intuition suggests that this approach does not generate precise estimates when density approaches the critical level $\bar{D}$, because the supply curve is vertical. This can be formally shown by assuming that $\partial T/\partial D$ is a random variable with a given variance, $var(\partial T/\partial D)$. Because the ratio of two random variables does not have a well-defined variance, we approximate the variance using a Taylor expansion. Using this expansion, the variance of MEC_M can be written as follows:

$$var(MEC_M) \approx \frac{var(\partial T/\partial D)\, D^2}{\left(1 - \frac{\partial T}{\partial D}F\right)^4} = \frac{var(\alpha)(TD)^2}{(1 - \alpha D)^4}. \tag{19}$$

The denominator of this expression contains a power of *four*. This implies that the estimate of MEC_M using (14) divided by its standard error, defined by the square root of (19), is equal to $\alpha(1 - \alpha D)$ and therefore goes to zero when density approaches the critical density. Thus, estimates of MEC_M for levels of flow close to its maximum are extremely unreliable, because its standard error is large relative to the estimate. Although there are only few observations of

[32] A positive shock to μ (or p_{PT}) in (3) implies that the demand for motor-vehicle travel shifts outwards in travel time – flow space. Given hypercongestion, this shock may cause either an increase or a decrease in density in equilibrium.

flow close to the maximum in our data, we will exclude these observations (our estimate of the total welfare loss of congestion remains unaffected by this issue).

Another issue is that MEC_M is a highly non-linear function of the effect of density on travel time, α. This raises the question of how a bias in the estimate of α affects the estimate of MEC_M. Given (14), we note that MEC_M is increasing in α:

$$\frac{\partial MEC_M}{\partial \alpha} = \frac{DT + DMEC_M}{1 - \alpha D} > 0. \tag{20}$$

Hence, any overestimate of α results in an overestimate of MEC_M.[33] The elasticity of MEC_M with respect to α can be shown to exceed one and is increasing in density, because:

$$\frac{\partial MEC_M}{\partial \alpha}\frac{\alpha}{MEC_M} = 1 + \frac{\alpha D}{1 - \alpha D} > 1. \tag{21}$$

We note that in our application, on average, αD is equal to 0.2, so, the elasticity is 1.20 and only slightly above one. Therefore, for the average estimate, the bias in MEC_M is roughly proportional to the bias in α.[34]

3.3 Estimating the MEC on bus users

We also aim to estimate the marginal external cost of congestion on bus travelers, exploiting hourly data on travel time on road segment *i* at time *t*, $T_{PTi,t}$. Combined with information about density of motor vehicles that use the same road, this allows us to estimate parameter σ in (11). Given α and σ, we calculate the MEC_{PT} as described in (15).

To estimate σ, we use a similar approach as to estimate α as described in 3.1. We estimate separate models for each road, using a log-linear specification including time controls (hour-of-the-day, day-of-the-week and week-of-the-year fixed effects). These time controls aim to capture unobserved supply shocks for bus travel (e.g., roadworks). Furthermore, we include weather controls and bus stop fixed effects. Hence, the equation we estimate is:

$$logT_{PTi,t} = \pi_i + \sigma_i D_{i,t} + \vartheta_i X_t + v_{i,t}, \tag{22}$$

where $v_{i,t}$ is the error term.

Arguably, endogeneity of traffic density due to reverse causality is less of a problem here than when analyzing the effect of density on motor-vehicle travel time, because bus travel time does *not* enter the expression for density, $D \equiv FT$. Nevertheless, there are still reasons to

[33] Note that the numerator and denominator of (20) are both positive when hypercongestion is absent, and they are negative when hypercongestion is present.

[34] For higher levels of density, this issue is more serious. For example, where density is about 2.5 times the average density, the relative overestimate equals two. Hence, any overestimate of α will result in a disproportional overestimate of on MEC_M.

suspect that OLS estimates are biased, particularly on mixed traffic roads (e.g., accidents may affect both car density and travel speed of buses). To account for endogeneity, we use hour-of-the-week fixed effects as an instrument, based on the same arguments as when estimating α in section 3.1. Specifically, these fixed effects capture regular changes in travel demand. However, we do not use public transit strikes as an instrument when estimating this model. The reason is that during the two-month period for which we observe bus travel times, we observe relatively few strikes. Although the strikes instrument turns out to be strong for most roads in our sample, using few strikes reduces the strength of our identification strategy, increasing the likelihood of spurious results.[35]

4. Data

The city of Rome is heavily dependent on motorized travel: 66% of trips are made by motor vehicles (50% by car and 16% by motorbike/scooter). Roughly, 28% of trips take place by public transport (ATAC SpA, 2013).[36] The main share of public transit supply is through buses (about 70% in terms of vehicle-kms as well as passenger-kms), see Table B1 in Appendix.

4.1 Travel data

Our data on motor vehicle traffic is derived from loop detectors. We employ 422,691 hourly observations (of which about 5 percent is during strikes) on flow and travel time for 33 measurement points of different roads between 5am and midnight for 769 working days, during a period from the 2nd of January 2012 to the 22nd of May 2015.[37] Motor vehicles refers to cars, commercial trucks and motorbikes, whereas flow refers to the number of motor vehicles passing a road per minute per road lane. Travel time is measured in minutes per kilometer.[38] Density is

[35] One may also suspect that, due to reduced frequency, bus occupancy tends to increase during strikes, thereby also raising idle time at bus stops. This effect could potentially invalidate strikes as instruments for density. However, recall that our data on bus travel time only refers to travel between stops, excluding time at stops. Hence, it is reasonable to expect no direct effect of strikes on bus travel time (given traffic density).

[36] The rate of motorization is high with 67 cars and 15 motorcycles per 100 inhabitants. There are about 1.6 cars per household. The high car ownership rate combined with substantial public transit use indicates that many regular transit users have access to a private vehicle, and are able to switch mode in the event of a transit strike.

[37] Measurement locations include twelve one-lane (per direction) roads – all located in the city center and with a speed limit of 50km/h (1.2 min/km). The other 21 roads have two lanes. These include seven large arterial roads with a speed limit of 100 km/h (0.6 min/km), eight with speed limits between 60 and 100 km/h and six with a speed limit of 50 km/h.

[38] We observe the average speed of vehicles at an hourly level (we invert speed to obtain average hourly travel time). We also observe counts (i.e., the number of vehicles passing a detector) per hour and convert it in flow per minute, ignoring within-hour variation.

calculated as the product of flow and travel time and measures the number of motor vehicles per kilometer of road lane. See Table 1.[39]

Table 1 – Motor vehicle travel

	Travel time [min/km]	Density [veh/km-lane]	Flow [veh/min-lane]	Obs.
Strike	1.365	14.6	11.1	23,018
No strike	1.327	13.4	10.5	399,673
Total	1.33	13.5	10.6	422,691

On average, travel time of private motor vehicles is 1.33 min/km, which corresponds to an average (instantaneous) speed of 46 km/h. This speed is far above the average speed of an entire trip, e.g. because we exclude waiting time at traffic lights. Flow per lane is above 11 vehicle-kms per minute and density is about 13.5 motor vehicles per kilometer. The above figures provide information for average traffic conditions, and thus mask substantial differences in congestion levels over time and between roads. We define a road as heavily congested during a certain hour when the speed on that road is less than 60 percent of free-flow speed (defined by the 95 percent percentile of the speed distribution observed on that road). Using this definition, on average roads are heavily congested about one hour per day, or 5 percent of the time. However, there is extreme variation between roads. Figure 7 shows the share of hours per day that a road is heavily congested. We single out ten ’heavily congested roads’, which are heavily congested at least one hour per day, with an average of about three hours per day, whereas the other 23 roads are heavily congested less than one hour per day (Figure B10 in Appendix shows the frequency of heavy congestion on roads in our sample).

A road is defined as hypercongested when, for given flow, the travel time lies on the backward bending portion of the supply curve. Visual inspection suggests that the 10 heavily congested roads (as defined above) are hypercongested for some time of the day, whereas the other 23 are not.[40] In Figures 8 and 9, we provide a scatterplot of the relationship between travel time and flow for two roads: one that clearly shows signs of hypercongestion and one where hypercongestion is absent. Clearly, hypercongestion is an empirically relevant occurrence. In

[39] We drop a few observations when travel time either exceeds 5 min/km or is below 0.4 min/km, when flow is zero or exceeds 2,100 vehicles per hour. The results are robust to the inclusion of these outliers. Information from the measurement locations is sometimes missing (e.g., meters are malfunctioning). Information on the whole month of August 2012 is missing, because the data collection agency moved to another office in this month. A few other days are missing for unknown reasons.

[40] Note that our definition of ’heavily congested road’ does not imply that a road is hypercongested. Traffic on a road may be slow on a given hour for reasons not directly related to density (e.g., because cars cruise for parking). On the other hand, all roads that we identify as hypercongested also turn out to be heavily congested.

both figures, we have also drawn the predicted supply curve, derived from one of our travel time-density estimates discussed later on in Section 5.1.

Figure 7 – Daily share heavily congested

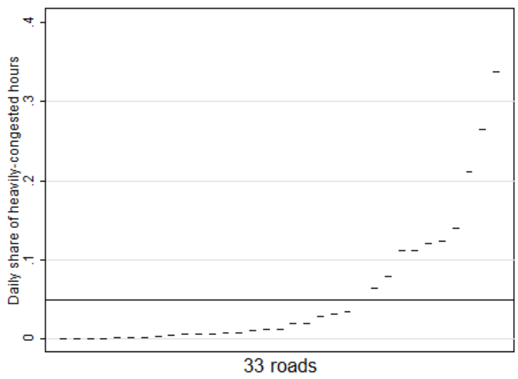
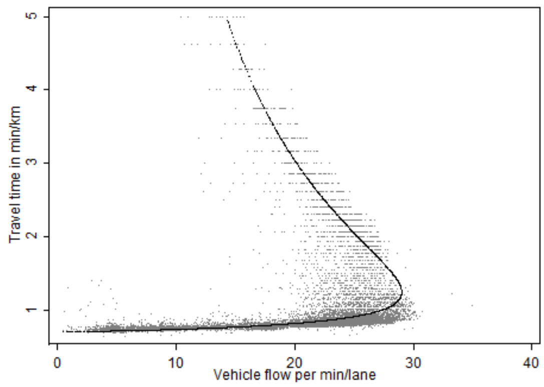


Figure 8 – Hypercongested road



A priori, it is sometimes ambiguous whether a road is hypercongested during a certain hour. To illustrate, consider the road in Figure 8 – which exhibits hypercongestion – and focus on observations with a flow around 25 motor vehicles per minute per lane, and where travel times are in between the (to-be-estimated) backward-bending supply curve. Without some theoretical basis, it is unclear whether these observations refer to hours where the road is hypercongested. We address this ambiguity by defining a road as hypercongested during an hour if and only if traffic density during that hour exceeds the critical level $\overline{D}$. Note that this definition implies that if a road is hypercongested for only a couple of minutes during a certain hour, we do not consider it as hypercongested during that hour. Hence, we most likely underestimate the pervasiveness of hypercongestion. Finally, recall that we estimate travel-time density relationships that assume that the logarithm of travel time is a linear function of density. In Figure 10, we show a scatterplot of this relationship for the hypercongested road depicted in Figure 8, which indicates that assuming this functional-form is reasonable.[41] A similar conclusion applies to other roads in our sample.

Furthermore, we have a sample of 27 roads used by the city's bus network, of which four have dedicated bus lanes (see Table 2). We distinguish between 58 bus line sections (i.e., the part of the ride between two successive stops), 14 of which are on dedicated lanes. We

[41] This figure suggests that for density levels below eight, which occur mainly outside peak hours, the marginal effect of density on log travel time is smaller. Excluding these observations generates almost identical, but somewhat more pronounced, results in our regression analysis, because we will estimate models with weights proportional to the hourly flow, and flow levels are low outside the peak, so low density observations will receive little weight.

employ data on the hourly bus travel time for each bus line section for the months of March in 2014 and 2015. We have hourly information on 71,645 sections for mixed traffic roads and on 31,024 sections for dedicated bus lanes. On mixed traffic roads, average travel time in the bus is about 1.56 minutes per kilometer (about 30 km/h), whereas on dedicated bus lanes it is substantially lower and equal to 1.08 minutes per kilometer (about 20 km/h).[42]

Figure 9 – Congested road

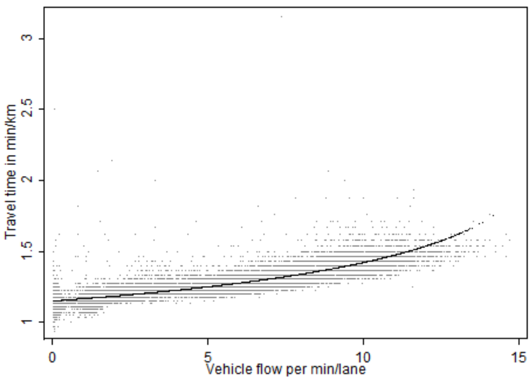


Figure 10 – Travel time-density

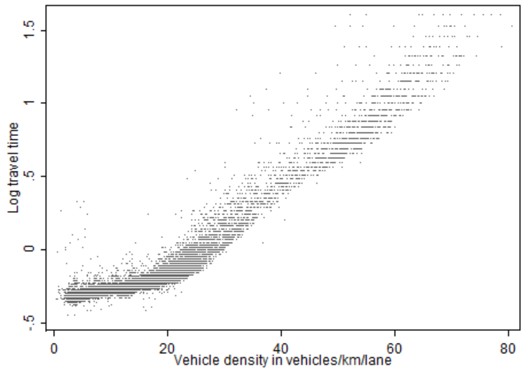


Table 2 – Bus travel

	Mixed Traffic	Dedicated Bus Lanes		Mixed Traffic	Dedicated Bus Lanes
Bus travel time between stops [min/km]	1.56	1.08	Bus users per section [pass-km/min]	5.16	9.96
Bus time at stops [min/stop]	0.69	0.78	Travel time motor veh. [min/km]	1.41	1.20
Bus travel time (incl. at stops) [min/km]	3.02	1.99	Density motor veh. [veh/lane-km]	14.8	13.5
Line section length [km]	0.47	0.85	Number of roads	23	4
Bus flow per lane [veh/min]	0.08	0.24	Number of bus lines	15	2
Bus flow per road [veh/min]	0.12	0.24	Number of bus line sections	44	14

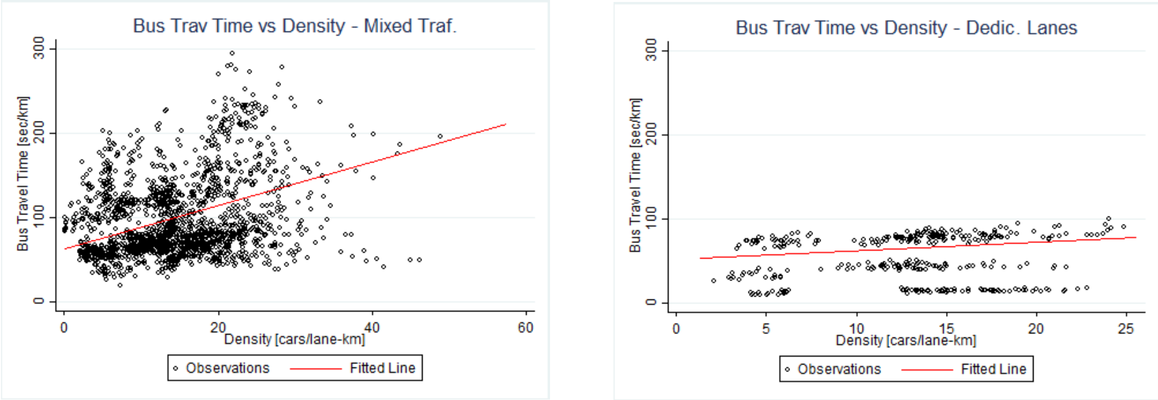
Note: 71,645 observations for mixed traffic roads and 31,024 observations for dedicated bus lanes

Motor-vehicle traffic conditions are quite similar for both types of roads: roads with dedicated bus lanes have slightly lower motor-vehicle travel times and densities than mixed traffic roads. Finally, note that one can expect the congestion-relief benefit for bus travelers to be substantial only if buses are strongly affected by road congestion. This seems to be the case for Rome. Figure 11 indicates that bus travel time (between stops) strongly increases with the

[42] Bus travel time is derived from micro data on the time of arrival and departure at each stop of every bus running on the city's bus network. This data is provided by the Mobility Agency. For most road traffic measurement locations, we are able to precisely identify the bus line section that encompasses the location. For some locations, however, we do not have exact coordinates. In those cases, we use two or three successive bus line sections (per road direction), which surely encompass the measurement location. We consider at least two bus line sections per location (one for each traffic direction).

density of motor-vehicles on mixed traffic roads. By contrast, for roads with dedicated lanes, bus travel times are hardly affected by density.

Figure 11 – Bus travel time and motor-vehicle density



Note: Bus travel time includes only time between stops.

4.4 Transit strikes in Rome

Information on strikes is provided by the Italian strike regulator (Commissione di Garanzia per gli Scioperi). During the 769 working days we observe, there are 43 with a transit strike, of which 27 took place only in Rome, whereas the other 16 were national. Strikes were announced to the public several days in advance (seven were partially cancelled).

We use information about hourly strike intensity at the city level. Specifically, Rome's Mobility Agency provided us with the share of provided service (relative to the regular schedule, i.e. without strikes) during strike hours (in terms of veh/kms). Consequently we are able to exploit *hourly* variation in the share of available public transit for identification purposes. The average share of transit service available during strikes is 0.56. That is, during the average strike hour, slightly more than half of public transport is still supplied. We plot the distribution of this variable in Appendix B (Figure B6). There is substantial variation in the share during strikes: it varies mainly between 0.30 and 0.83 (with few observations below 0.3). The share of transit available is highest during the 8 a.m. morning peak (the median is about 0.75) and the 7 p.m. evening peak hour (the median is about about 0.65). During these hours, the variation in the share is also small. From 9 a.m. to 3 p.m., the share is substantially lower, but the range is also much wider (see Figure B7).[43] The reason is that Italian law does not allow full transit service shutdowns during strikes, mandating a minimum service level during peak

[43] We also have information on the non-strike *scheduled* service level, i.e., the usual number of buses operating (see Figure B5). The number of scheduled buses in Rome hardly varies between 8am and 5pm except on strike days. These observations support the use of strikes as a way of identifying the effects of public transit supply.

hours. Moreover, regulation forbids strikes during weekends and holiday months, mainly in August and September. Excluding these months, the distribution of strike activity is quite even over the year, with somewhat higher concentration in the spring period. Most strikes take place on Mondays and, in particular, on Fridays (see Figures B8 and B9).

5. Empirical Results

To quantify the welfare losses caused by congestion, our first step is to estimate the effect of traffic density on travel time of motor-vehicles and buses, using (17) and (22). We use OLS and the two IV approaches described above. We report the full results for each road separately in the Appendix D, see Tables D1 and D2. In Table 3, we provide the average and the standard deviation of the marginal effect of density, α, on log motor-vehicle travel time (see equation (8)). Table 4 reports similar information regarding the effect of density, σ, on log bus travel time (see equation (11)).

The first column of Table 3 shows that when using OLS, a marginal increase in density increases log travel time by 0.021, on average.[44] Hence, increasing density by one vehicle per km-lane increases travel time by about 2 percent. The standard deviation of this effect is about 0.01, indicating that the marginal effect does not differ much between roads. The second column reports the IV results using the hour-of-the-week instrument. This instrument is strong for all roads. It appears that the effect of density is about 0.019, on average, suggesting that the OLS estimates are somewhat biased upwards, by about 10 percent. This upward bias is statistically significant for the majority of roads at the 5 percent level.[45] We will use the estimates of column (2) for the welfare analysis (while providing results based on other estimates in a sensitivity analysis).

[44] These results are largely consistent with the transport engineering literature, see for example Greenberg (1959). We have also estimated models where we explicitly acknowledge that a strike through a decrease in public transit supply directly decreases the number of vehicles on the road, which invalidates using strikes as an instrument by making assumptions on the effect of removing a bus compared to the effect of a standard motor-vehicle. Even when we assume that one single bus causes the same travel delays as 10 motor vehicles, we get identical results.

[45] For 20 of the 33 roads, the Hausman t-test exceeds two (in absolute value). See, Wooldridge (2002, p. 99) for details about this test.

Table 3 – Log travel time

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) IV
Density (average)	0.0214	0.0193	0.0233	0.0211	0.0172
Standard deviation	[0.0095]	[0.0087]	[0.0091]	[0.0084]	[0.0101]
Instrument		Hour-of-week		Hour-of-week	Public transit
Controls	Yes	Yes	Yes	Yes	Yes
Number of roads	33	33	30	30	30
Number of obs.	422,691	422,691	371,687	371,687	371,687

Note: standard deviation of density effect between square brackets. The dependent variable is the logarithm of travel time (min/km). We estimate the marginal effect of density for each road separately and then report the average as well as the standard deviation of the effect. The controls include temperature, rain, hour-of-day, day-of-week and week-of-the-year fixed effects. In the last column, we also control for hour-of-the-week fixed effects; controlling for this variable is however immaterial for the results.

It appears that the alternative instrument, public transit, is strong for 30 out of 33 roads, as the F-test exceeds the recommended value of 10 for three roads (Wooldridge, 2002, p. 105). For these 30 roads, we provide estimates given the public transit instrument, as well as the two previous approaches for comparison. Also for this subsample we find that the IV estimates when using the hour-of-the-week instrument are somewhat smaller than the OLS ones. Using public transit as an instrument reduces the estimates further, suggesting a larger upward bias of OLS. Using these estimates for the welfare analysis generates very similar results.

To check the consistency of our instrumental variable approach, we test whether the monotonicity condition (18) holds. This condition is satisfied for all our observations as long as demand is sufficiently elastic, i.e. when $\varphi \leq 0.5$ which corresponds to a demand elasticity of approximately -0.28. Furthermore, it holds for more than 90% of observations even when $\varphi = 2$ (which corresponds to an average demand elasticity of - 0.04, i.e. an almost vertical demand for motor vehicle travel, which is rather implausible). Therefore, the condition for the validity of our instruments is practically always satisfied.

Table 4 – Log bus travel time

	Mixed traffic OLS	Mixed traffic IV	Dedicated lanes OLS	Dedicated lanes IV
Density (average)	0.0160	0.0195	0.0047	0.0088
Standard deviation	[0.0239]	[0.0257]	[0.0076]	[0.0093]
Instrument		Hour-of-week		Hour-of-week
Controls	Yes	Yes	Yes	Yes
Number of roads	23	23	4	4
Number of obs.	71,645	71,645	31,024	31,024

Note: standard deviation of density effect between square brackets. The dependent variable is the logarithm of bus travel time (min/km) in between bus station stops. We estimate the marginal effect of density for each road separately and then report the average as well as the standard deviation of the effect. The controls include temperature, rain, hour-of-day, day-of-week, week and bus line section fixed effects.

Table 4 reports the results on the effect of density on the log of bus travel time, σ. When applying the IV approach, it appears that the instrument is strong for all roads included. Using OLS, the estimated effect of density for all mixed traffic roads is positive and statistically significant at the 5 percent level. Using IV, the effect is also always positive but statistically insignificant for some roads. On average, the results suggest that the OLS estimates are downward biased by about 20 percent (see Table 6). The IV specification implies that a unit increase in traffic density increases bus travel time on mixed traffic roads by roughly 2 percent, which is almost identical to our estimate for the effect on motor vehicle travel time (see Table 3).[46] There is, however, a much smaller effect on dedicated lanes, and this effect is only statistically significant on one road.

Using the estimates described above, we predict each road's supply curve – i.e. the travel-time flow relationship – as explained at the beginning of Section 3.1.1. Figure 9 provides an example of such prediction for the hypercongested road discussed earlier (black line). The predicted travel-time flow relationship is backward-bending, in line with traffic engineering studies (Helbing, 2001; Geroliminis and Daganzo, 2008).

We now turn to the analysis of external costs and welfare losses from congestion. In Table 5, we describe the results for the observed as well as for the optimal equilibria, averaged for all roads and hours in our sample. The first column refers to the observed equilibrium. In the first five rows of this table, we report the average of density, flow and travel time of motor vehicles as well as the bus travel time (between stops) and number of bus users. Given the estimates of density on motor-vehicle time, α, combined with information about the observed density, we calculate how often hypercongestion occurs (i.e. when $D > 1/\alpha$). As we report in Table 5, it appears that hypercongestion occurs on average only about 17 minutes per day (out of 19 hours), i.e. for about 1.5 percent of the time. Such a low estimate is in line with our descriptive statistics, which indicated that the majority of roads are heavily congested less than one hour per day.[47] We have also examined when hypercongestion occurs most during the day. As one may expect, hypercongestion occurs almost exclusively during the peak hours, and particularly during the morning peak.

[46] The large majority of roads has a positive coefficient, although not always statistically significant because of larger standard errors. A couple of roads have a negative sign but not significant. These tend to be the roads with lower OLS estimates. See Table D3 in Appendix D.

[47] We weigh this measure by hourly flow on a given road, although the unweighted measure is almost identical. Recall that the roads in our sample are heavily congested on average about one hour per day, hence roads that are heavily congested are also hypercongested less than 20 percent of the time. Another possible explanation for the low frequency of hypercongestion is that we use data that is averaged at the hourly level and, therefore, possibly miss instances where roads are hypercongested for only some minutes during an hour.

These results focus on averages per hour and road, thereby masking large differences between roads: for many roads there is either never hypercongestion (18 roads), or hardly any hypercongestion (9 roads, less than 25 minutes per day on average), but *three roads are hypercongested for at least one hour per day* (see Appendix D, Tables D1- D3).

Table 5 – Observed and optimal equilibria – Full Sample

	Equilibrium	Optimum	Approx. optimum
Density (veh/km-lane)	13.55	9.46	9.32
Flow (veh/min-lane)	10.6	8.40	8.27
Travel time, motor veh. (min/km)	1.33	1.21	1.20
Travel time, bus (min/km)	1.44	1.24	1.22
Bus users (pass-km/min)	5.87	6.04	6.07
Hypercongestion (min/day)	17.43	0.00	0.00
MEC (min/km)	0.88	0.40	0.40
MEC_M, motor veh. (min/km)	0.6	0.21	0.21
MEC_{PT}, buses (min/km)	0.28	0.21	0.21
DWL (min/km-lane)		2.21	2.17
only hypercongestion		62.87	52.81
no hypercongestion		1.26	1.26
Number of roads		33	

In Table 5, first column, we report the MEC for our full sample of roads, again averaging over roads and hours (note that we exclude observations when roads are hypercongested). The MEC produced by a motor-vehicle travelling one km is about 0.88 minutes on average.[48] This cost is substantial when compared to the average travel time per km (1.33 minutes). About two thirds of this cost is on motor-vehicle travelers, whereas one third is on bus travelers. Assuming a value of time equal to 15.59€/h for car users and 9.54€/h for bus users,[49] the monetary MEC per vehicle-km is €0.2 (0.6×15.59€/60+0.28×9.54€/60), one fourth of which is on bus travelers.

We have performed a range of sensitivity analyses to examine these results. In particular we have examined to what extent our MEC_M and MEC_{PT} estimates are robust to alternative specifications. To start with, we have examined to what extent MEC_M is sensitive to individual road estimates, given that MEC_M is a highly nonlinear function of the estimates of α. We have

[48] In Table 5, we report the weighted average of the marginal external time cost for a road, using the flow per road as weight. This masks uncertainty about the estimates of the marginal external cost for individual roads. The t-value of MEC_M is equal to $1 - \alpha D$ multiplied with the t-value of α. In our data, $1 - \alpha D$ is on average 0.6. Hence, because α is precisely estimated for most roads with a high t-value, MEC_M is also precisely estimated for these roads. This is not true for MEC on buses, because σ is not precisely estimated at the individual road level, due to the much smaller sample size (we have only two months of bus data).

[49] These are the median values for Milan, the second-largest Italian city, reported by Rotaris et al. (2010). We did not find this information for Rome.

therefore estimated road supply curve models imposing that all roads have the same functional form (i.e., α is identical for all roads). Results hardly changed, showing they are not due to extreme estimates for specific roads, as suggested by (21). We have then calculated a 95 percent confidence interval for MEC_M (see last columns of Table D2). The average interval varies between 0.52 and 0.69. Finally, we have estimated all bus travel time models using OLS, instead of the instrumental variable approach, and find that the *average* estimate of MEC_{PT} is robust across alternative methodologies (see Table D3).[50]

Table 5 also describes the optimal equilibria – in terms of density, flow and travel time. These equilibria are characterized for each hour t and road i, based on our estimated supply relations for motor vehicle and bus travel and assuming linear demands, as specified in Section 2.1. Specifically, we characterize the demand for motor vehicle travel (3) by estimating the parameters $\mu_{i,t}$ and $\theta_{i,t}$ assuming that each hourly observation of motor-vehicle and bus travel (time, flow and density) describe an equilibrium and that the slope, φ, is constant. We also compute the number of bus users in the optimum. We refer the reader to Appendix A for a detailed description of this procedure. We consider the case where $\varphi = 0$, 0.1, 0.3 and 1. The implied corresponding average demand elasticities are, respectively, minus infinity, -1.5, - 0.5 and - 0.14.[51] Hence, we consider a rather broad spectrum of demands, spanning from perfectly elastic to almost perfectly inelastic. For reasons of space, however, in Table 5 we describe the optimal equilibria only for the case $\varphi = 0.1$. Qualitatively, the results for the other values of φ are similar (Appendix E).

As shown in Table 5, density decreases when moving from the observed to the optimal equilibria. The average reduction is substantial: from 13.55 to 9.46 vehicles per km per road lane, i.e. by about 30 percent. Average travel time for motor vehicles falls from 1.33 to 1.21 min/km, i.e. about 5 percent. This reduction may seem small, but the drop is larger on more congested roads. For example, for the road depicted in Figure 9, average travel time falls from 0.96 to 0.81 minutes/km, i.e. about 15 percent. In addition, average flow decreases by about 15 percent. Furthermore, the reduction in bus travel time is more pronounced: on average, travel time falls from 1.44 min/km to 1.24 min/km, i.e. about 15 percent.

---

[50] Given the presence of hypercongestion, one expects that an approach where one regresses (log) travel time of motor-vehicles on *flow* generates a bias in MEC_M. We have applied such an approach using OLS, which imply a MEC_M of about 0.20, which is less than one third of the value reported in Table 5. Instrumenting flow with hour-of-week fixed effects does not solve the issue. We obtain then an average MEC_M of about 0.34, i.e. 50 percent of the estimate in Table 7. When we use public transit strikes as an instrument, the estimate is 0.72 which exceeds in Table 5. Hence, in contrast to our proposed methodology, this "travel time – flow approach" is extremely sensitive to the estimation method used.

[51] We have also examined non-linear specifications for demand. To be more specific, we have assumed log-linear demand specifications assuming the given elasticities. Again the results do not change much.

The average MEC computed in the optimum is equal to 0.40 min/km, roughly half than the average MEC in the observed equilibria. In monetary terms, the MEC in the optimum is equal to €0.086 per vehicle-km (0.21×15.59€/60+0.21×9.54€/60). This figure is indicative of the size of the optimal road toll in Rome. Assuming an average trip length of 13 kilometers, as reported by the Mobility Agency (PGTU, 2014), the optimal toll is about 1.12 Euros per trip.

Furthermore, we calculate the welfare change of inducing a shift from the observed equilibrium to the optimum, i.e. the deadweight loss (DWL), expressed in minutes of travel time per kilometer of road lane. We find that the average DWL is 2.24 vehicle-minutes for every minute of the day (or about 134 minutes each hour) per kilometer of road lane. To provide a sense of the relevance of hypercongestion, we also report the average DWL for hypercongested equilibria and non-hypercongested ones separately. The penultimate row in Table 5, indicates that the DWL is 1.28 minutes per lane-km on average on non-hypercongested roads, i.e. about 55% of the total. Therefore, despite the roads in our sample are seldom hypercongested, about half of the overall deadweight loss comes from hypercongestion. Furthermore, the average DWL on a hypercongested road is very large: it is about 62 minutes per lane-km on an average minute of the day. Using the same monetary values of time as above, we find that this is equivalent to roughly €15 every minute. Hence, the loss per vehicle travelling one km is about €1.78 (€15/8.4). To put this in perspective, the *hourly* deadweight loss is about €1,800 for a hypercongested two-lane road segment of one kilometer length (€15×60×2).

Table 5 also characterizes the "approximate optima" under the assumption that the road includes a downstream bottleneck. Under this assumption, the optima that corresponds to equilibria with hypercongestion differ from the case without bottlenecks, because the optimal flow cannot exceed the observed one (see Section 2.3.2, Figures 5 and 6), while the welfare loss is entirely due to the extra travel time caused by queuing. Note that, as explained in Section 2.3.2, this assumption only affects the equilibria with hypercongestion (there is no change in non-hypercongested equilibria by construction). Thus, given the small percentage of observations with hypercongestion, allowing for bottlenecks brings to very little change in the results. The most notable change is that the DWL in hypercongested equilibria is about 20 percent lower when one does not allow for throughput reductions. The intuition for this finding is that the welfare loss due to the reduction in throughput when roads are hypercongested is small compared to the welfare loss due to reduced travel speed. In other words, the backward-bending part of the supply curve is close to being vertical.

To illustrate how the external costs of congestion vary during the day, we show the marginal external cost for the observed equilibria (excluding those with hypercongestion) in

Figure 12, as well as the deadweight loss per hour of the day (when $\varphi = 0.1$) in Figure 13. Unsurprisingly, both quantities fluctuate over the day and are much larger during peak hours.

Figure 12 – Marginal external cost

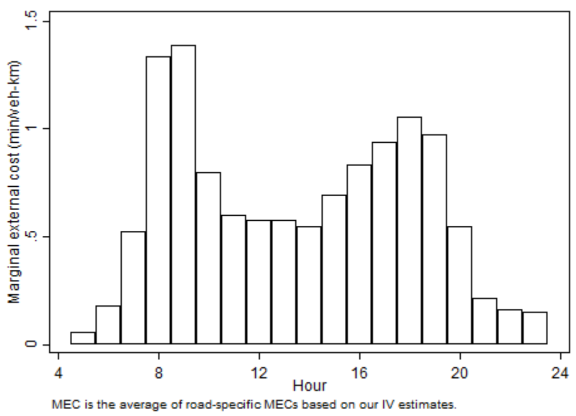
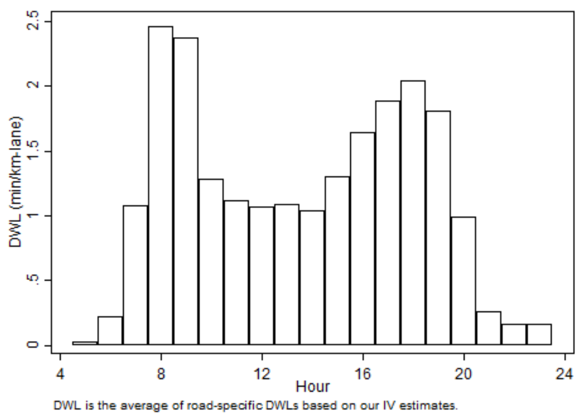


Figure 13 – Deadweight Loss



Finally, in Table 6 we break down the results for two subsamples of roads: those where bus and motor vehicles share the same lanes (mixed traffic) and those where buses travel on dedicated lanes. The main aspect to notice is that the MEC tends to be higher on mixed traffic roads, particularly due to the external cost of motor vehicle travel on bus users. By contrast, the MEC on bus users is almost non-existent when buses travel on dedicated lanes, as one would expect.

Table 6 – Observed and optimal equilibria – Mixed Traffic and Bus Lane roads

	Mixed Traffic			Dedicated Lanes		
	Equilibrium	Optimum	Approx. Optimum	Equilibrium	Optimum	Approx. Optimum
Density (veh/km-lane)	13.83	9.59	9.45	11.57	8.51	8.39
Flow (veh/min-lane)	10.72	8.56	8.43	9.07	7.28	7.18
Travel time, motor veh. (min/km)	1.35	1.22	1.20	1.25	1.16	1.15
Travel time, bus (min/km)	1.56	1.28	1.26	1.08	1.04	1.03
Bus users (pass-km/min)	5.16	5.37	5.40	9.96	9.87	9.92
Hypercongestion (min/day)	19.22	0.00	0.00	2.70	0.00	0.00
MEC (min/km)	0.94	0.42	0.42	0.59	0.27	0.27
MEC_M, motor veh. (min/km)	0.63	0.21	0.21	0.50	0.25	0.25
MEC_{PT}, buses (min/km)	0.31	0.22	0.22	0.09	0.08	0.08
DWL (min/km-lane)		2.28	2.25		0.97	0.95
only hypercongestion		68.98	58.63		16.23	13.79
no hypercongestion		1.29	1.29		0.79	0.79
Number of roads	29			4		

Taken together, the results of this section indicate that the welfare losses due to road congestion in Rome are substantial. However, there are some caveats. First, although we observe traffic data from many measurement locations that are quite evenly spread across the city, our sample may not be entirely representative of the road network in Rome. Second, we estimate road supply curves at the individual road level, and not at an area- or network-wide level. Hence, our estimates of the external costs do not account for the possibility of avoiding heavily-congested roads by using alternative routes or by traveling at other times.[52] If, for example, drivers may choose alternative uncongested routes, our analysis may overestimate the aggregate congestion costs. However, in Rome, the extent to which drivers can avoid congested arteries without taking substantial detours on secondary roads (which may also easily become congested), is unclear. If indeed they cannot, the extra-vehicle kilometers may increase the aggregate travel time losses, implying that we are somewhat underestimating these losses.

6. *Conclusion*

We estimate the marginal external cost of road congestion, and the associated deadweight losses, allowing for hypercongestion and considering the travel times of cars as well as public transit (bus) users. We exploit variation in public transit strikes and hourly variation in demand over the day to account for endogeneity issues. We demonstrate that, for the city of Rome, the marginal external cost is substantial: it is, on average, equal to about 66 percent of the private time travel cost, while reaching considerably higher levels during peak hours. When roads are not hypercongested, the marginal external cost of motor vehicle travel is €0.2 per kilometer on average, but almost double during peak hours. About one third of the marginal external cost of road congestion in Rome is borne by bus travelers. We find that most roads are never hypercongested, but some roads are hypercongested for more than one hour per day, on average. The welfare losses produced by congestion can be up to 50 times larger for hypercongested than for normally congested roads. As a result, hypercongestion accounts for about 40 percent of congestion-related welfare losses. We also demonstrate that welfare losses incurred on roads that are hypercongested are substantial, predominantly because of reductions in speed rather than in throughput.

Our findings suggest that policies designed to reduce congestion can bring substantial welfare gains. For example, the high deadweight losses of hypercongestion suggests that,

[52] Akbar and Duranton (2016) provide citywide estimates of supply and demand functions for Bogota', using information from travel surveys and Google Maps. Couture et al. (2018) also provide estimates at the aggregate level from a sample of US cities.

particularly if road pricing is unavailable, quantitative measures to curb traffic on heavily congested roads (e.g., through adaptive traffic lights) may be warranted (Fosgerau and Small, 2013). These measures can prevent hypercongestion, particularly in presence of unexpected demand or supply shocks such as accidents (Kouvelas et al., 2017).

References

Adler, M.W., Liberini, F., Russo, A. and Van Ommeren, J.N. (2019). The congestion relief benefit of public transit: evidence from Rome. CESifo Working Paper.

Akbar, P.A. and Duranton, G. (2016). Measuring congestion in a highly congested city: Bogota'. Mimeo.

Angrist, J. D. and Pischke, J. S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.

Arnott, R., De Palma, A. and Lindsey, R. (1990). Economics of a bottleneck. *Journal of Urban Economics*, 27(1), 111-130.

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

Arnott, R. and Inci, E. (2010). The stability of downtown parking and congestion, *Journal of Urban Economics*, 68, 3, 260-276.

Arnott, R. (2013). A bathtub model of downtown traffic congestion. *Journal of Urban Economics*, 76, 110-121.

ATAC SpA (2013). Carta Generale dei Servizi. Rome.

Borjas, G.J. (1980). The relationship between wages and weekly hours of work: the role of division bias. *Journal of Human Resources* 15 (3), 409–23.

CEBR (2014). Economic and Environmental Costs of Gridlock: an assessment of the direct and indirect economic and environmental costs of idling during heavy road traffic congestion to households in the UK, France and Germany. London, UK.

Couture, V., Duranton, G. and Turner, M. A. (2018). Speed. *Review of Economics and Statistics*, 100(4), 725-739.

De Borger, B. and S. Proost (2012). A political economy model of road pricing. *Journal of Urban Economics* 71, 79-92

Duranton, G. and Turner, M. A. (2011). The fundamental law of road congestion: Evidence from US cities. *American Economic Review*, 2616-2652.

FHWA (1997). *1997 Federal highway cost allocation study final report.* US Federal Highway Administration, Department of Transportation, Washington, D.C.

Fosgerau M. and Small, K. (2013). Hypercongestion in downtown metropolis. *Journal of Urban Economics* 76, 122-134.

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

Gonzales, E. (2015). Coordinated pricing for cars in transit in cities with hypercongestion. *Economics of Transportation,* 4, 64-81

Greenberg, H. (1959). An analysis of traffic flow. *Operations research*, 7(1), 79-85.

Greenshields, B. D. (1935) A Study of Traffic Capacity. *Highway Research Board Proceedings*, Vol. 14. pp. 448–477.

Hall, F. L. (1996). Traffic stream characteristics. *Traffic Flow Theory*. US Federal Highway Administration.

Hall, J.D. (2018). Pareto Improvements from Lexus Lanes: The Effects of Pricing a Portion of the Lanes on Congested Highways, *Journal of Public Economics* 159, 113-125

Helbing, D. (2001). Traffic and related self-driven many-particle systems. *Reviews of Modern Physics*, *73*(4), 1067.

Henderson, J.V. (1974). Road congestion: a reconsideration of pricing theory. *Journal of Urban Economics* 3, 346-365.

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

Imbens, G. and Angrist, J. (1994). Identification and estimation of local average treatment effect, *Econometrica*, 62, 467-476.

Loder A., Ambuehl L., Menendez M. and Axhausen K. (2017). Empirics of multi-modal traffic networks – Using the 3D macroscopic fundamental diagram. *Transportation Research C* 82, pp. 88-101

Keeler, T. E. and Small, K. A. (1977). Optimal peak-load pricing, investment, and service levels on urban expressways. *Journal of Political Economy*, *85*(1), 1-25.

Kouvelas, A., Saeedmanesh, M. and Geroliminis, N. (2017). Enhancing model-based feedback perimeter control with data-driven online adaptive optimization. *Transportation Research B*, 96, 26-45.

Litman, T. (2015). *Transit price elasticities and cross-elasticities*. www.nctr.usf.edu/jpt/pdf/JPT 7-2 Litman.pdf). Retrieved: 23/09/2015.

Maibach, M., Schreyer, C., Sutter, D., Van Essen, H. P., Boon, B. H., Smokers, R. and Bak, M. (2008). *Handbook on estimation of external costs in the transport sector*. CE Delft.

Mayeres, I., Ochelen, S. and Proost, S. (1996). The marginal external costs of urban transport. *Transportation Research D*, 1(2), 111-130.

Mohring, H. (1972). Optimization and scale economies in urban bus transportation. *American Economic Review*. 1972, 591–604.

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

Parry, I. W. and Small, K. A. (2009). Should urban transit subsidies be reduced? *American Economic Review*, 99(3), 700-724.

PGTU. Piano Generale del Traffico Urbano (2014). City of Rome, 2014.

Pigou, A. C. (1920), *The Economics of Welfare*,, London, Macmillan.

Rotaris, L., Danielis, R., Marcucci, E. and J. Massiani (2010). The urban road pricing scheme to curb pollution in Milan, Italy: Description, impacts and preliminary cost–benefit analysis assessment. *Transportation Research A* 44, 359–375.

Small, K. (2004). Road pricing and public transport. *Research in Transportation Economics*, 9(1). 133-158.

Small, K. A. and Chu, X. (2003). Hypercongestion. *Journal of Transport Economics and Policy*, 37(3), 319-352.

Small, K. A. and Verhoef, E. T. (2007). *The Economics of Urban Transportation*. London, Routledge.

Underwood, R.T. (1961). *Speed, volume and density relationship*. Quality and theory of traffic flow, Yale Bu. Highway traffic, 141-188.

Van den Berg V. and E. Verhoef (2011). Winning or losing from dynamic bottleneck congestion pricing? *Journal of Public Economics*, 95(7), 983-992

Verhoef, E.T. (2003). Inside the queue: hypercongestion and road pricing in a continuous time - continuous place model of traffic congestion. *Journal of Urban Economics* 54, 523-535.

Verhoef, E. T. (2005). Speed-flow relations and cost functions for congested traffic: Theory and empirical analysis. *Transportation Research Part A*, 39(7-9), 792-812

Vickrey, W. (1973). The Economics of Congestion Control in Urban Transportation. In Emerson F.C., editor: *The Economics of environmental problems; five lectures on the economic aspects of public policies to control the environment*. University of Michigan.

Wooldridge, J. (2002), *Econometric Analysis of Cross Section and Panel Data.* Cambridge, MIT press.

Appendix A: Computing the optima and the DWL

We characterize the optima in Table 5 and the deadweight loss of congestion. We assume that the market is in equilibrium, so $F_{i,t} = N_{M,i,t}$. The first step is to characterize the demand for motor vehicle travel *by hour and road* using (3). To characterize this demand, we compute for each hour (indexed by t) and road (indexed by i) the values of μ and θ. Inverting (3), we get:

(A1) $$F_{i,t} = \frac{\mu_{i,t}}{\varphi} + \frac{\theta_{i,t}}{\varphi} p_{PT\,i,t} - \frac{T_{i,t}}{\varphi},$$

which implies that the cross-price elasticity of demand for motor vehicle travel with respect to the price of public transit is:

(A2) $$\varepsilon_{F,PT\,i,t} \equiv \frac{dF_{i,t}}{dp_{PT}} \frac{p_{PT\,i,t}}{F_{i,t}} = \frac{\theta_{i,t}}{\varphi} \frac{p_{PT\,i,t}}{F_{i,t}}.$$

In a companion paper, using data on monetary price changes in public transport on travel demand for Rome, see Adler et al. (2019), we find that $\varepsilon_{F,PT}$ is about 0.1. We assume the same elasticity applies for travel time changes in public transport. Given the assumed φ, and given hourly observations of $F_{i,t}$ and $T_{PT\,i,t}$ for each road, we compute the value of $\theta_{i,t}$ for the given road-hour pair as follows:

(A3) $$\theta_{i,t} = \frac{0.1\varphi F_{i,t}}{T_{PT\,i,t}}.$$

The value of intercepts $\mu_{i,t}$ can be calculated given the assumption that, on a given road-hour pair, the market is in equilibrium. Given φ, $\theta_{i,t}$ and information on $T_{i,t}$, $T_{PT\,i,t}$ and $F_{i,t}$, one calculates $\mu_{i,t}$ using (3).

Concerning public transit travel, we assume a linear (inverse) demand, with the following form:

(A4) $$T_{PT\,i,t} = \varsigma_{i,t} + \varrho_{i,t} \times T_{i,t} - \gamma_{i,t} \times N_{PT\,i,t}.$$

Where $\varsigma_{i,t}$, $\varrho_{i,t}$ and $\gamma_{i,t}$ are positive parameters. To characterize this function, we first need to characterize these parameters. Inverting (A4), we get:

(A5) $$N_{PT\,i,t} = \frac{\varsigma_{i,t}}{\gamma_{i,t}} + \frac{\varrho_{i,t}}{\gamma_{i,t}} T_{i,t} - \frac{T_{PT\,i,t}}{\gamma_{i,t}}.$$

To determine $\gamma_{i,t}$, we assume the price elasticity of bus travel in Rome is -2.2 (this is the value that Parry and Small (2009) assume for peak-hour travel in London). This elasticity writes:

(A6) $$\varepsilon_{PT} \equiv \frac{dN_{PT}}{dT_{PT}} \frac{T_{PT\,i,t}}{N_{PT\,i,t}} = -\frac{1}{\gamma_{i,t}} \frac{T_{PT\,i,t}}{N_{PT\,i,t}}.$$

Using this expression and our observations of $N_{PT\,i,t}$ and $T_{PT\,i,t}$ we can calculate C for the given hour and road as:

(A7) $$\gamma_{i,t} = \frac{2.2 \times N_{PT\,i,t}}{T_{PT\,i,t}}.$$

We can then calculate $\varrho_{i,t}$ as follows. We assume the cross-price elasticity of public transport (bus) with respect to motor vehicles is 0.14, as reported in Litman (2017). We then get:

(A8) $$\varrho_{i,t} = \frac{0.14 \times \gamma_{i,t} \times N_{PT\,i,t}}{T_{PT\,i,t}}.$$

Finally, we determine the intercept $\varsigma_{i,t}$ using (A4) and information on $T_{PT\,i,t}$, $T_{i,t}$, $N_{PT\,i,t}$ and the parameters determined previously.

The next step is to characterize the *optimal equilibrium* – in terms of density, flow, number of bus users and travel-time of motor vehicles and buses – corresponding to each observed equilibrium (per hour and road). To do so, we combine the information on demand with an estimate of the road-specific road supply curve using our IV estimates in Tables 5 and 6. Optimality requires that marginal benefit equals marginal social cost. Hence, in the optimal equilibrium, $\mu + \theta T_{PT} - \varphi F = MEC_M + MEC_{PT} + T$ must hold. Given (8), (11), (14) and (15), the optimal density is found by numerically solving the following equation (we omit the *i* and *t* indexes for ease of notation):

(A9) $$\mu + \theta\gamma e^{\sigma D} - \varphi\,(\mathrm{D}/\,\beta e^{\alpha \mathrm{D}}) = \beta e^{\alpha \mathrm{D}} + \alpha\mathrm{D}\,\beta e^{\alpha \mathrm{D}}/(1 - \alpha\mathrm{D}) + \gamma e^{\sigma D}\sigma N_{PT}\left(\frac{\beta e^{\alpha \mathrm{D}}}{1 - \alpha D}\right),$$

where the parameters in the equation are estimated empirically. Given the optimal density, we calculate the corresponding optimal travel time and flow as well as the optimal number of bus users (using (A5)) and bus travel time. One can then evaluate the MEC in the optimum. Also, we can find the corresponding DWL by comparing the optimum to the observed equilibrium for the given road-hour pair, using the estimated supply functions and the demand functions in (3) and (A5).

Appendix B: Figures

Figure B1 – Map of Rome and location of traffic measurement points

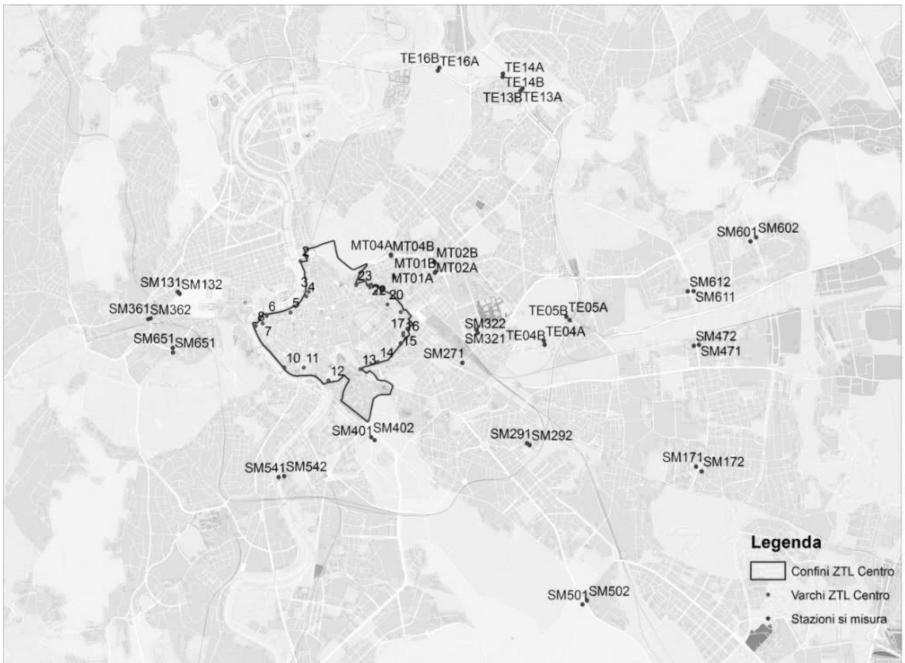


Figure B2 – Vehicle density histogram

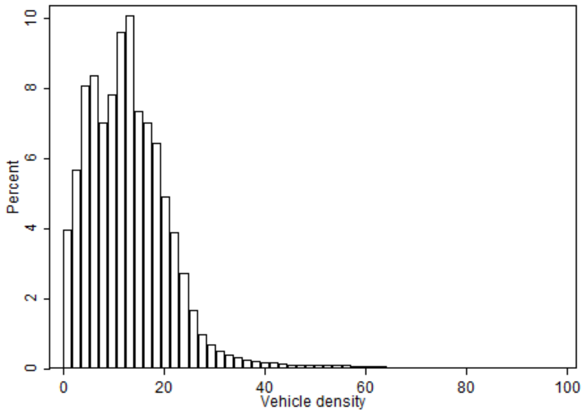


Figure B3 – Vehicle flow histogram

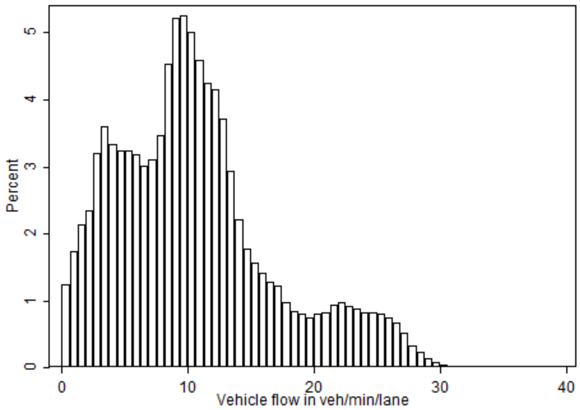


Figure B4 – Heavy congestion by hour

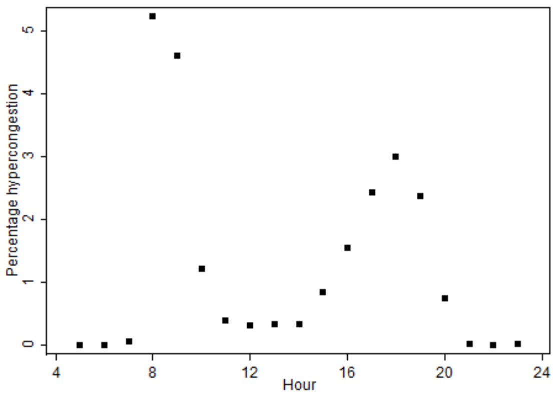


Figure B5 – Public transit on non-strike day

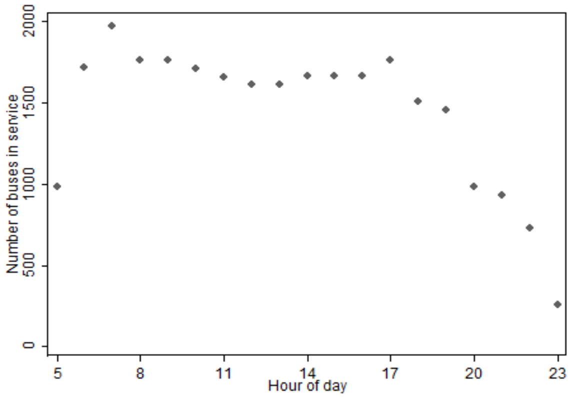


Figure B6 – Public transit share for strikes

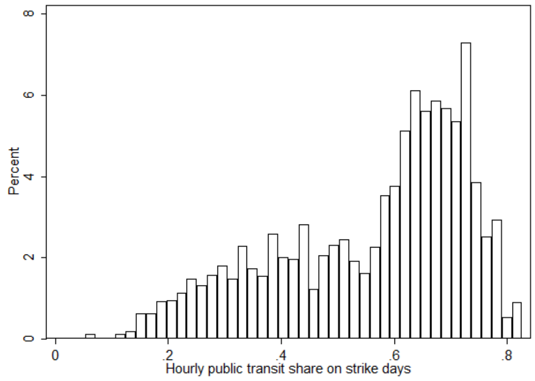


Figure B7 – Public transit share per strike hour

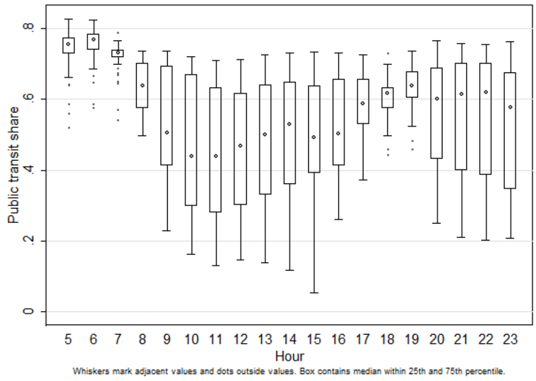


Figure B8 – Strikes by month

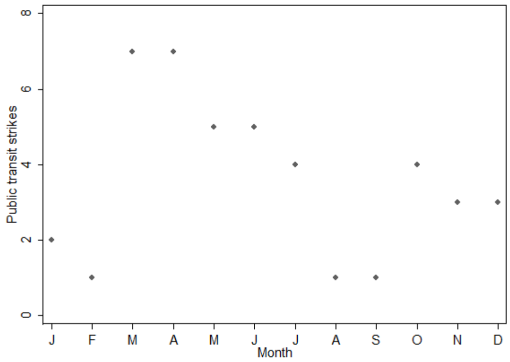


Figure B9 – Strikes by day

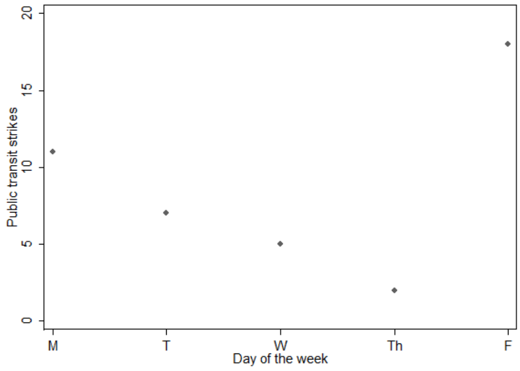


Figure B10 – Frequency heavy congestion

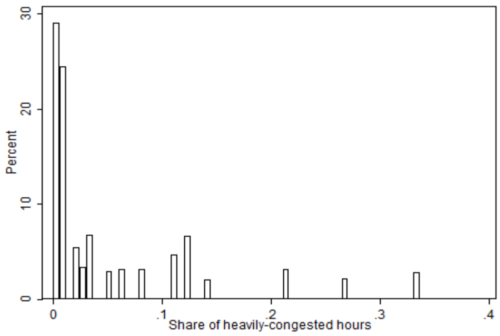


Table B1 – Travel in Rome's metropolitan area

	Car		Bus		Rail	
	Peak	Off-Peak	Peak	Off-Peak	Peak	Off-Peak
Annual veh-kms, millions	6,116	8,445	66.7	67.7	10.24	7.2
Annual passenger kms, millions	8,623	12,837	3,403	2,304	1,639	628
Vehicle occupancy (pass-km/veh-km)	1.4	1.51	51	34	160	87

Source: Own calculations based on information from Rome's General Traffic Plan (PGTU, 2014). The data refer to the year 2013.

Appendix C: sufficient conditions for monotonic effect of instruments on density

We derive (18) and show that this condition is sufficient for demand-shifting instruments to affect density monotonically. We start from (3) and rewrite this relation in the time-density space. Noting that N_M captures traffic flow and using (5) we can write (3) as:

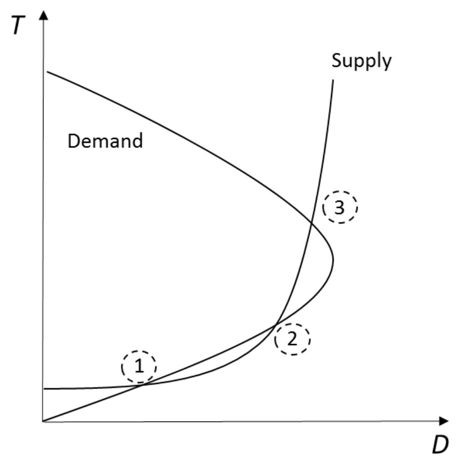
(C1) $$T = \mu + \theta p_{PT} - \varphi \frac{D}{T}.$$

We have assumed that φ is constant. Hence, an increase in demand implies an increase in $\mu + \theta p_{PT}$. Furthermore, this equation implies that the demand curve is such that:

(C2) $$\frac{dT}{dD} = -\frac{T\varphi}{T^2 - \varphi D}.$$

Hence, the demand relationship is backward bending (in the time-density space): it is upward-sloping if and only if $T^2 - \varphi D < 0$. Demand crosses the vertical axis at the origin ($D = 0$ and $T = 0$) and where D = 0 and $T = \mu + \theta p_{PT}$. By contrast, the supply function, $T = h$(D), is upward sloping with a positive intercept. See the specification of this function in expression (8). Hence, there can be *at most* three equilibria, two on the upward sloping part of the demand relationship and one on the downward sloping part. See Figure C1 for an illustration. We now provide two sufficient conditions for the equilibrium to lie on a part of the demand curve such that an increase in $\mu + \theta p_{PT}$ results in an increase in the equilibrium level of density.

Figure C1



Consider the equilibrium marked by 1 in Figure C1. This equilibrium lies on the upward sloping part of demand. Furthermore, the supply function intersects the demand function from above. Hence, the following condition is satisfied:

(C3) $$\frac{dh}{dD} = \alpha T < -\frac{T\varphi}{T^2 - \varphi D}.$$

An increase in $\mu + \theta p_{PT}$ makes the demand relation rotate clockwise around the origin. To see this, rewrite (C1) as $D = \frac{(\mu+\theta p_{PT})T - T^2}{\varphi}$. This implies that $\frac{dD}{d(\mu+\theta p_{PT})} = \frac{T}{\varphi} > 0$ and $\frac{d^2 D}{d(\mu+\theta p_{PT})dT} = \frac{1}{\varphi} > 0$. Hence, the equilibrium marked by 1 is such that density increases when $\mu + \theta p_{PT}$ increases.

Consider now the equilibrium marked by 3 in Figure C1, which lies on the downward-sloping part of demand. Hence, the following condition must hold:

(C4) $$-\frac{T\varphi}{T^2 - \varphi D} < 0 \Rightarrow T^2 > \varphi D.$$

An increase in $\mu + \theta p_{PT}$ induces demand to rotate clockwise around the origin. Hence, this equilibrium is also such that density increases when $\mu + \theta p_{PT}$ increases.

Finally, notice that we cannot be in the equilibrium marked by 2 when either (C3) or (C4) hold, because this equilibrium is such that the supply function intersects the demand function from below and the demand function is positively sloped. Consequently, when either (C3) or (C4) hold, a positive shock to the in intercept in the demand for motor-vehicle travel demand causes a *monotonic* increase in *D*.

Appendix D: road-level effects of density on travel time

Table D1 – Log travel time, OLS

Road	OLS	Se(OLS)	Critical value ($\bar{D}$)	Hyper-conges. (min/day)	MEC_M	MEC_M (IV) 95% CI	MEC_M (IV) 95% CI
1	.011595	.000097	86.2407	0	0.5837	0.5184	0.6826
2	.007497	.000102	133.391	0	0.3427	0.2605	0.4485
3	.031770	.000116	31.4759	33.768	0.9595	0.8368	1.1028
4	.028459	.000078	35.1381	55.553	0.6011	0.5272	0.6795
5	.031108	.000169	32.1460	13.391	1.6703	1.3957	1.9961
6	.019860	.000298	50.3535	0	1.1489	1.0633	1.2884
7	.006302	.000191	158.668	0	0.0894	0.0778	0.1166
8	.018264	.000125	54.7532	1.7472	0.5778	0.5075	0.6719
9	.021675	.000113	46.1357	9.8928	0.5318	0.4609	0.6165
10	.034469	.000111	29.0111	116.03	1.7620	1.3112	2.2276
11	.015970	.000163	62.6166	0	0.1716	0.1301	0.2227
12	.007382	.000104	135.467	0	1.3554	1.1889	1.6149
13	.019441	.000153	51.4363	0.1452	0.6181	0.4618	0.8015
14	.038659	.000478	25.8669	0.0876	0.1443	0.0482	0.2695
15	.034491	.000193	28.9929	2.598	0.4484	0.4316	0.4734
16	.018867	.000153	53.0030	0	0.8342	0.7171	0.9940
17	.016399	.000141	60.9780	0	0.7811	0.6943	0.9039
18	.021117	.000518	47.3545	0	0.0802	0.0690	0.1090
19	.005412	.000127	184.782	0	0.1483	0.1223	0.1914
20	.014548	.000154	68.7357	0	0.1865	0.1526	0.2318
21	.025971	.000070	38.5043	83.404	0.8200	0.7000	0.9473
22	.022089	.000115	45.2712	1.002	0.6353	0.5775	0.7133
23	.027194	.000156	36.7721	5.1756	1.1688	1.1004	1.2754
24	.021836	.000128	45.7954	3.4044	0.5999	0.5392	0.6864
25	.02178	.000267	45.9137	8.3244	0.9546	0.8324	1.1838
26	.023400	.000131	42.7347	3.1524	0.3422	0.3139	0.3820
27	.029340	.000067	34.0826	113.00	0.9719	0.8509	1.1050
28	.027826	.000197	35.9377	47.217	0.6926	0.6592	0.7364
29	.029670	.000073	33.7047	128.32	1.0878	0.9478	1.2423
30	.035570	.000283	28.1135	20.006	0.6273	0.6042	0.6723
31	.027515	.000233	36.3412	1.794	0.3452	0.3297	0.3750
32	.005390	.000147	185.536	0	0.0602	0.0502	0.0788
33	.008079	.000232	123.778	0	0.1112	0.1107	0.1238
Avg.	0.02148			19.644	0.6501	0.5715	0.7543

Note: Road segment specific estimations. Dependent variable is log of travel time. We also list the critical value, the extent of hypercongestion and MEC_M. In the last two columns, we provide the 95 percent confidence interval estimates for MEC_M.

Table D2 – Log travel time, IV - instrument public transit share

Road	IV	Se(IV)	Critical value ($\bar{D}$)	Hyper-conges. (min/day)	MEC_M	MEC_M (IV) 95% CI	MEC_M (IV) 95% CI
1	0.0155	.002283	64.7139	0.5856	0.9129	0.7288	1.1916
2	0.0054	.001275	186.415	0	0.1929	0.0349	0.3962
3	0.0247	.002269	40.4601	21.185	0.6870	0.3810	1.0443
4	0.0261	.001403	38.3722	49.988	0.5641	0.3211	0.8221
5	0.0237	.002678	42.2182	6.1596	0.8553	0.5140	1.2602
6	0.0173	.002781	57.6817	0	0.4057	0.0484	0.9880
7	0.0054	.005816	185.087	0	0.1012	0.0838	0.1420
8	0.0163	.001306	61.4788	0.4992	0.5589	0.3879	0.7877
9	0.0182	.001272	55.0537	6.0684	0.4600	0.1900	0.7825
10	0.0358	.001020	27.9431	129.31	1.8403	1.0694	2.6364
11	0.0144	.001337	69.6603	0	0.1761	0.0931	0.2782
12	0.0131	.001400	76.3813	16.392	2.0451	1.1794	3.3939
13	0.0133	.000800	75.2664	0	0.3561	-0.0443	0.8259
14	0.0405	.003173	24.6870	0.1752	0.4073	0.1383	0.7576
15	0.0328	.003182	30.5338	2.4312	0.3778	0.2813	0.5217
20	0.0028	.005114	352.641	0	0.0325	-0.0113	0.0911
21	0.0084	.008911	119.616	0	0.4085	0.3660	0.4536
22	0.0066	.004523	150.767	0	0.1247	0.0347	0.2462
23	0.0039	.005990	257.685	0	0.0860	-0.0455	0.2909
24	0.0192	.001826	52.2163	2.724	0.5823	0.3330	0.9379
25	0.0169	.002092	59.1833	1.494	0.7611	0.2947	1.6359
26	0.0169	.002457	59.2331	2.1324	0.2496	0.1294	0.4190
27	0.0271	.001262	36.8364	102.33	0.9280	0.5558	1.3375
28	0.0000	.042489	25613.7	0	0.0007	-0.0229	0.0316
29	0.0265	.001261	37.7723	111.46	1.0009	0.5737	1.4725
30	0.0103	.016611	97.4756	0	0.1555	0.1042	0.2553
31	0.0065	.003866	154.782	0	0.0755	0.0368	0.1498
Avg.	0.0166	0.00476		16.775	0.531	0.287	0.857

Note: Road segment specific estimations. Dependent variable is log of travel time. For the IV estimation, we use public transit share as an instrument. We do not report IV estimates for roads where public transit is a weak instrument. We also list the critical value, the extent of hypercongestion and MEC_M. In the last two columns, we provide the 95 percent confidence interval estimates for MEC_M.

Table D3 – Log travel time, IV - instrument hour of week

Road	IV	Se(IV)	Critical value ($\bar{D}$)	Hyper-congestion (min/day)	MEC_M (IV)	MEC_M (IV) 95% CI	MEC_M (IV) 95% CI
1	0.0126	0.00079	79.10509	0.252	0.6508	0.5318	0.8054
2	0.0108	0.0005	92.70183	0	0.4586	0.4029	0.5185
3	0.0312	0.00081	32.1069	32.232	0.9257	0.8571	1.0008
4	0.0275	0.0006	36.41183	53.4	0.5847	0.5681	0.6034
5	0.0299	0.00075	33.41542	11.388	1.3144	1.2185	1.4458
6	0.0248	0.00186	40.27681	0.42	0.6948	0.5341	0.8708
7	0.0068	0.00129	147.8345	0	0.1291	0.0783	0.1838
8	0.0129	0.00063	77.25784	0	0.4125	0.3545	0.4744
9	0.0199	0.00074	50.37664	7.896	0.5093	0.4641	0.5544
10	0.0340	0.00044	29.45576	110.892	1.8097	1.7870	1.8455
11	0.0151	0.00066	66.21012	0	0.1872	0.1681	0.2070
12	0.0119	0.00087	84.02668	10.956	1.7723	1.3931	2.1704
13	0.0200	0.00056	50.05761	0.144	0.6256	0.5720	0.6712
14	0.0450	0.00206	22.20009	0.348	0.4747	0.4142	0.5395
15	0.0195	0.00188	51.24009	0.588	0.1992	0.1633	0.2436
16	0.0181	0.00092	55.27551	0	0.8247	0.7042	0.9606
17	0.0143	0.0009	69.95548	0	0.6630	0.5539	0.7841
18	0.0163	0.00317	61.41909	0	0.1188	0.0688	0.1763
19	0.0117	0.00116	85.78853	0	0.3123	0.2375	0.3935
20	0.0180	0.00101	55.52052	0	0.2581	0.2220	0.2969
21	0.0260	0.00047	38.42308	83.568	0.8228	0.8056	0.8547
22	0.0194	0.0009	51.66095	0.252	0.5292	0.4524	0.6108
23	0.0206	0.00134	48.59289	1.164	0.7388	0.5878	0.9159
24	0.0187	0.00098	53.60258	2.64	0.5556	0.4646	0.6625
25	0.0188	0.00146	53.17796	1.92	0.9850	0.6683	1.2534
26	0.0193	0.00117	51.79839	2.688	0.3024	0.2524	0.3558
27	0.0279	0.00051	35.79173	106.344	0.9668	0.9211	1.0132
28	0.0192	0.00243	52.2117	26.82	0.4530	0.4126	0.5397
29	0.0279	0.00053	35.86225	116.964	1.0731	1.0164	1.1220
30	0.0214	0.00323	46.76476	5.556	0.3737	0.2639	0.5133
31	0.0112	0.0015	89.18195	0	0.1434	0.0997	0.1916
32	0.0051	0.00072	198.1779	0	0.0695	0.0491	0.0910
33	0.0022	0.00104	451.7093	0	0.0286	0.0023	0.0566
Average	0.0193	0.00115		17.432	0.6051	0.5239	0.6947

Note: Road segment specific estimations. Dependent variable is log of travel time. For the IV estimation, we use hour-of-week dummies as an instrument. We also list the critical value, the extent of hypercongestion and MEC_M. In the last two columns, we provide the 95 percent confidence interval estimates for MEC_M.

Table D4 – Log bus travel time, instrument hour of week

Road	OLS	Se(OLS)	IV	se(IV)	Bus users	Ded. Lane	MEC_{PT} (OLS)	MEC_{PT} (IV)
1	0.020	0.003	0.007	0.010	6.864	No	0.9211	0.270
2	0.007	0.002	0.001	0.004	8.274	No	0.3068	0.069
3	0.034	0.001	0.028	0.004	5.722	No	0.4013	0.230
4	0.038	0.001	0.036	0.003	5.775	No	0.6004	0.524
5	0.004	0.002	0.009	0.006	10.132	Yes	0.0213	0.052
6	0.003	0.003	-0.006	0.009	10.080	Yes	0.0539	-0.078
8	0.018	0.001	0.024	0.003	2.002	No	0.1538	0.165
9	0.021	0.002	0.027	0.006	2.094	No	0.1288	0.160
10	0.023	0.002	0.021	0.004	1.924	No	0.3041	0.275
11	0.012	0.004	0.005	0.010	1.692	No	0.0556	0.025
12	0.014	0.003	0.031	0.007	2.346	No	0.3324	1.609
13	0.017	0.004	0.020	0.007	4.192	No	0.1850	0.261
14	0.125	0.009	0.133	0.021	3.235	No	1.9528	2.651
15	0.001	0.005	-0.002	0.016	11.709	No	0.0091	-0.015
16	0.011	0.005	0.004	0.010	5.360	No	0.1922	0.050
17	0.003	0.007	0.017	0.015	5.261	No	0.0503	0.343
18	0.014	0.009	0.047	0.027	2.711	No	0.0205	0.117
19	0.002	0.002	0.003	0.005	2.834	No	0.0033	0.004
21	0.011	0.001	0.010	0.003	11.969	No	0.3830	0.328
22	0.009	0.004	0.008	0.012	5.284	No	0.1021	0.074
23	-0.013	0.003	0.012	0.008	11.080	No	-0.3696	0.375
27	0.000	0.002	0.000	0.003	5.084	No	0.0027	0.002
29	0.009	0.001	0.011	0.002	3.451	No	0.1306	0.130
30	0.019	0.005	0.029	0.007	3.684	No	0.1912	0.131
31	0.013	0.005	0.027	0.015	6.099	No	0.2945	0.467
32	0.018	0.003	0.019	0.005	9.165	Yes	0.2308	0.269
33	-0.002	0.004	0.004	0.006	10.466	Yes	-0.0545	0.107
Average	0.016	0.003	0.019	0.008	5.870	/	0.244	0.286

Note: Road segment specific estimations for all roads. Dependent variable is log of bus travel time. For the IV estimation, we use hour-of-week dummies as instruments. We also list the MEC_{PT} based on OLS and IV estimates. Roads 7,20,24,25,26 and 28 are omitted because we do not have traffic data for the months of March 2014 and 2015, hence we cannot estimate the effect of traffic density on bus travel time.

Appendix E: Sensitivity of results

Table E1 - Full Sample

	Equilibrium	Optimum				Approximate optimum			
FULL SAMPLE		φ=0	φ=0.1	φ=0.3	φ=1	φ=0	φ=0.1	φ=0.3	φ=1
Density (veh/km-lane)	13.55	5.53	9.46	10.87	11.94	5.45	9.32	10.70	11.75
Flow (veh-km/min-lane)	10.6	4.99	8.40	9.34	9.96	4.92	8.27	9.20	9.81
Travel time, private veh. (min/km)	1.33	1.18	1.21	1.24	1.27	1.17	1.20	1.23	1.25
Travel time, bus (min/km)	1.44	1.12	1.24	1.28	1.33	1.10	1.22	1.26	1.30
Bus users (pass/min-lane)	5.87	5.78	6.04	6.04	6.05	5.81	6.07	6.07	6.08
Hypercongestion (min/day)	17.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MEC (min/km)	0.88	0.30	0.40	0.48	0.57	0.30	0.40	0.48	0.57
MEC, motor veh. (min/km)	0.6	0.14	0.21	0.28	0.36	0.14	0.21	0.28	0.36
MEC, buses (min/km)	0.28	0.20	0.21	0.26	0.29	0.20	0.21	0.26	0.29
DWL (veh-min/km-lane)		2.24	2.21	2.04	1.63	2.21	2.17	2.00	1.60
DWL in hypercongested eq.		64.99	62.87	55.31	51.09	53.29	52.81	47.02	45.47
DWL w/o hypercongestion		1.28	1.26	0.98	0.80	1.28	1.26	0.98	0.80

Table E2 - Mixed Traffic Roads

MIXED TRAFFIC	Equilibrium	Optimum				Approximate optimum			
		φ=0	φ=0.1	φ=0.3	φ=1	φ=0	φ=0.1	φ=0.3	φ=1
Density (veh/km-lane)	13.83	5.52	9.59	11.01	12.20	5.44	9.45	10.84	12.01
Flow (veh-km/min-lane)	10.72	5.00	8.56	9.49	10.21	4.93	8.43	9.35	10.05
Travel time, private veh. (min/km)	1.35	1.19	1.22	1.25	1.28	1.17	1.20	1.23	1.26
Travel time, bus (min/km)	1.51	1.17	1.28	1.32	1.37	1.15	1.26	1.30	1.35
Bus users (pass/min-lane)	5.16	5.57	5.37	5.38	5.38	5.59	5.40	5.41	5.41
Hypercongestion (min/day)	19.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MEC (min/km)	0.94	0.30	0.42	0.50	0.59	0.30	0.42	0.50	0.59
MEC, motor veh. (min/km)	0.63	0.13	0.21	0.29	0.37	0.13	0.21	0.29	0.37
MEC, buses (min/km)	0.31	0.22	0.22	0.29	0.32	0.22	0.22	0.29	0.32
DWL (veh-min/km-lane)		2.44	2.28	2.24	1.79	2.41	2.25	2.21	1.77
DWL in hypercongested eq.		71.95	68.98	65.30	60.21	59.00	58.63	54.86	53.58
DWL w/o hypercongestion		1.31	1.29	1.04	0.83	1.31	1.29	1.04	0.83

Table E3 - Roads with Dedicated Lanes

DEDICATED LANES	Equilibrium	Optimum				Approximate optimum			
		φ=0	φ=0.1	φ=0.3	φ=1	φ=0	φ=0.1	φ=0.3	φ=1
Density (veh/km-lane)	11.57	5.58	8.51	9.90	10.10	5.50	8.39	9.75	9.94
Flow (veh-km/min-lane)	9.07	4.80	7.28	8.25	8.27	4.73	7.18	8.13	8.14
Travel time, private veh. (min/km)	1.25	1.11	1.16	1.19	1.21	1.09	1.15	1.18	1.19
Travel time, bus (min/km)	1.06	1.02	1.04	1.05	1.05	1.01	1.03	1.04	1.04
Bus users (pass/min-lane)	9.96	9.89	9.87	9.87	9.83	9.93	9.92	9.92	9.88
Hypercongestion (min)	2.70	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
MEC (min/km)	0.59	0.26	0.27	0.33	0.39	0.26	0.27	0.33	0.39
MEC, motor veh. (min/km)	0.50	0.18	0.25	0.25	0.30	0.18	0.25	0.25	0.30
MEC, buses (min/km)	0.09	0.07	0.08	0.08	0.08	0.07	0.08	0.08	0.08
DWL (veh-min/km-lane)		1.36	0.97	0.57	0.44	1.34	0.95	0.57	0.43
DWL in hypercongested eq.		23.79	16.23	10.36	10.08	19.51	13.79	8.97	8.70
DWL w/o hypercongestion		1.24	0.79	0.54	0.52	1.24	0.79	0.54	0.52