Department of Economics- FEA/USP

The Political Economy of an Optimal Congestion Tax: An Empirical Investigation

CLAUDIO LUCINDA
RODRIGO MOITA

WORKING PAPER SERIES Nº 2019-10

DEPARTMENT OF ECONOMICS, FEA-USP
WORKING PAPER Nº 2019-10

The Political Economy of an Optimal Congestion Tax: An Empirical Investigation

Claudio Lucinda (claudiolucinda@usp.br)

Rodrigo Moita (rodrigomsm@insper.edu.br)

Abstract:

Despite widely prescribed by economists, a congestion tax is seldom used in practice. Why? This paper combines a structural econometric model with a simulation algorithm to estimate an optimal congestion tax and investigate its political acceptance. Results for Sao Paulo show the tax to be 2 USD per trip. Policy simulations indicate that (i) commuters that switch to the public transportation bears the largest share of the tax burden and (ii) revenue recycling is essential for the policy to be accepted. Skepticism about the use of tax revenues is the likely cause for the low use of the congestion tax.

Keywords: pigouvian tax; urban transportation; congestion; urban toll

JEL Codes: L92; R41; L91

The Political Economy of an Optimal Congestion Tax: An Empirical Investigation*

Claudio Lucinda† Rodrigo Moita‡

February 18, 2019

Abstract

Despite widely prescribed by economists, a congestion tax is seldom used in practice. Why? This paper combines a structural econometric model with a simulation algorithm to estimate an optimal congestion tax and investigate its political acceptance. Results for Sao Paulo show the tax to be 2 USD per trip. Policy simulations indicate that (i) commuters that switch to the public transportation bears the largest share of the tax burden and (ii) revenue recycling is essential for the policy to be accepted. Skepticism about the use of tax revenues is the likely cause for the low use of the congestion tax.

Keywords: pigouvian tax, urban transportation, congestion, urban toll

1 Introduction

People in large cities talk and think about traffic congestion; they change schedules, search information about traffic conditions, listen to radio stations broadcasting road conditions, use apps such as Waze and Google Maps among many other things. As a response, cities invest large sums in urban transportation infrastructure: a larger subway network, new wider roads, bridges, tunnels and so on. Nonetheless, traffic is getting worse in most cities.

A well known fact in economics of transportation is that a congestion tax is the main prescription for addressing this problem. Congestion pricing is far from being a novel idea; indeed, as Lindsey [2006] points out, even Adam Smith has written several pages on toll pricing. However, the discussion took its current form only from the second half of the 20th century, facing the question of what would be the optimal charge for a monopolist road operator. Authors such as Vickrey [1948], Walters [1961] and Downs [1962], discussed what would be the relevant toll prices to reduce congestion externalities.

*We are thankful to comments from Ciro Biderman, Danilo Igliori, Leonardo Basso, Juan Pablo Monteiro, and several participants on EARIE Annual Conference, ITEA Conference, Fundacao Getulio Vargas, Insper, State University of Rio de Janeiro and University of Sao Paulo.

†University of Sao Paulo. Email: claudiolucinda@usp.br

‡Insper. Email: rodrigomsm@insper.edu.br.

A second and somewhat surprising fact is that urban tolls are mostly used in developed countries, such as England, Sweden or Singapore, but not at all in poor or developing countries, where traffic externalities might be even more pressing. Both facts naturally lead to the question: why isn't a congestion toll policy more widely used? Instead, several cities adopted quantity restriction mechanisms, such as license plate restrictions, which are clearly sub-optimal.[1] We believe that the answer to this question is a political economy mechanism which hinders developing countries' ability of taking advantage of this tool, besides all operational aspects of selecting the optimal congestion tax.

This puzzle can be addressed by taking an empirical approach to the problem; but to the best of our knowledge no paper has already done that. We aim to fill this gap by proposing a structural econometric model of driving decisions including driving generated externalities to estimate the relevant demand parameters, also taking into account the dynamic aspects of mode choice in a day with multiple trips. Then we compute a pigouvian congestion charge for Sao Paulo, Brazil, taking into account both heterogeneity in the value of time and commuters income, as well as differences on congestion levels on distinct parts of the city. From this value, we simulate alternative policies differing about how the revenues from the congestion tax are recycled, whether by a lump sum transfer or a bus subsidy, and do a welfare analysis to understand the impact of such policies, and their political acceptability.

Main Findings. Our econometric results imply own price elasticities that nearly match previous studies such as Batarce and Ivaldi [2014]. They also imply an optimal linear congestion tax of 6.25 BRL, or 2 USD, per car trip in downtown Sao Paulo.

The tax's welfare impact shows that the group which switches to public transportation is the one that carries most of the tax's burden. This finding differs from the theoretical results of De Borger and Proost [2012], which says remaining drivers are the most affected group. It is also in line with the stylized fact that the middle class would be the most affected by this type of policy.

The reason is commuters' heterogeneity in value of money and value of time. Remaining drivers tend to be the wealthiest part of the population, which have the lowest marginal utility of money and therefore are less affected by the tax. They also have the highest value of time, benefiting more from the decrease in travel time.

The political economy analysis also shows revenue recycling as a major feature when implementing a congestion tax. Simulations show that it increases the political acceptability of the tax, specially if it is targeted to the citizens most harmed by the policy (mostly new bus users). Also, there is a substantial difference in terms of traffic reduction between the lump sum rebate and the bus subsidy: the bus subsidy reduces car usage substantially more than a lump sum transfer would. Such increases could even reach a level in which the expected shifts in public transportation demand would require substantial capacity investments.

The scenario without revenue recycling mimics the situation in which the tax structure as well the governance of public transportation is as such that revenues from the congestion tax cannot be funneled to individuals, either as a lump sum transfer or as a reduction in bus fares. The majority rejection of the urban toll in the situation with no revenue recycling helps explain why this is not adopted in developing countries, where the cost of public funds

[1] Sao Paulo, Beijing, Santiago, Mexico City, and Bogota, among many others, all adopted similar variations of license plate restrictions.

is higher than in developed countries.[2]

Contribution and Related Literature. This paper relates to different branches of the congestion charge literature. Closest to this paper is the literature that analyze political economy aspects of congestion pricing. De Borger and Proost [2012] set up a theoretical model to analyze the acceptance of a congestion pricing. They show who the policy winners and losers are under different scenarios, and the important role of revenue recycling. Armelius and Hultkrantz [2006] simulate the welfare effect of an urban toll on Stockholm city and show the importance of improving public transportation to increase the policy's acceptance. Eliasson and Mattsson [2006] analyze the claim the resistance to implement congestion pricing in Stockholm is due to the policy regressive nature. The authors show that, on the contrary, the policy is in fact progressive, but it depends on the redistribution of the revenue collected.

More generally, this paper relates to the broad literature which analyzes the question of why some optimal policies are not adopted in practice. An important paper is Grossman and Helpman [1994], which analyzes why free trade is not widely adopted, despite being prescribed by the economic theory as welfare maximizing.

A related issue is the fact that congestion is endogenous to the policy implemented; that is, policy changes will feed back into driving decisions and they will lead to different congestion levels. Batarce and Ivaldi [2014] estimate a structural demand model for travel decisions including the network congestion. In the present paper, we also take into account the effects of endogenous congestion by simulating traffic across routes. They simulate an optimal bus fare (welfare maximizing), while we estimate in our paper an optimal - also welfare maximizing - congestion charge.

Parry and Bento [2001] calibrate a model to compute the optimal tax. But to the best of our knowledge, our paper is the first to estimate an optimal congestion tax using microdata. As the results show, including heterogeneity in travellers preferences through microdata is what enable us to identify winners and losers of the policy.

In an earlier work, Lucinda et al. [2017], compare the welfare impact of a license plate restriction with that of a congestion tax. However, the congestion tax was not optimal. Instead, it was set to induce the same level of congestion as the license plate restriction.

Following this line of argument, papers such as Small [1983], Layard [1977] and Safirova et al. [2004] investigate the distributional effects of toll prices and whether they are regressive. If so, this effect could be a reason for the low acceptance of congestion charges and toll prices worldwide.

Our paper analyzes a situation where the revenue from the congestion tax is used to subsidize bus fare. Basso and Silva [2014] consider the situation where both policies coexist. Their results show a substitutability between the tax on cars and subsidy on transit in terms of welfare. Our results show that int terms of political acceptability both policies are complimentary.

This paper also relates to the econometrics of discrete choice literature that incorporate dynamics on the decisions. Train [2003] provides a clear explanation of a logit with dependent decisions over time. Adamowicz [1994] and Erdem [1996] also pose some models in which there is some persistence in individual decisions. Here, we take

[2] In this context, for cost of public funds we mean that for each monetary unit raised by the tax, only a fraction of it returns to the population.

into account the fact a prior driving trip for an individual could influence the probability of driving in subsequent decisions.

2 Optimal Congestion Charge – Conceptual Framework

Our conceptual framework to define the optimal congestion charge is composed of four parts: a travel mode demand system in which individuals choose their transportation mode, a model of traffic that aggregates individual trip choices into route traffic, an externality equation relating traffic to travel time and an algorithm that uses these three pieces just to compute the optimal tax.[3] This section develops each of these models and the algorithm.

2.1 Travel Mode Choice Model

We use the mixed logit model (see Train [2009] for a full discussion of this class of models). The mixed logit approach is used here because it can provide more flexible substitution patterns than both the multinomial logit and the nested logit – for instance, used by Molnar and Mesheim [2010]. The starting point is a population of I potentially heterogeneous individuals, with an individual denoted i. Each individual has to decide which transportation mode to take to work. From the choice set defined above, it is assumed the respondents must choose a single alternative from a subset of the discrete choice set $J = \{\text{Bus, Rail, Driving, Motorcycle, Taxi, Other}\}$. Not all alternatives might be available to all individuals, though. Depending on infrastructure availability, some of them might not be available.

For each $j \in J$ alternative, individual i derives utility U_{ij}. Assuming that individual's behavior is utility-maximizing, choices can be represented by a binary variable defined as:

$$y_i = \begin{cases} 1 & \text{if } \ U_{ij} > U_{ij'} \forall j' \neq j \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Individual choices depend on many factors affecting their utility from each alternative. Assuming utility can be decomposed into two parts, one can write U_{ij} as:

$$U_{ij} = \beta_i z_{ij} + \varepsilon_{ij} \tag{2}$$

In which z_{ij} is a set of observed variables relating to alternative j for individual i which might depend on individual specific variables x_i as well as alternative specific variables w_j such that it can be expressed as $z_{ij} = z(x_i, w_j)$ for some function z. It is assumed here the function z is the identity, which makes $z_{ij} = [x_i : w_j]$. β_{ij} is a corresponding vector of coefficients for the observed variables, and ε_{ij} captures the impact of all unobserved factors that affect the individual's choice. When ε_{ij} follows an extreme-value distribution, we say that it belongs to the Generalized Extreme Value (GEV) family of models [Fisher and Tippett, 1928, Gnedenko, 1943].

[3] We are assuming here the only externality is the congestion externality; thus, we are abstracting from pollution, traffic accidents and other externalities from car usage.

We assume the coefficients have some sort of heterogeneity – in our case related unobserved characteristics - and are given by $\beta_i = \beta + \sigma v_i$. In which σ is a vector capturing heterogeneity in (possibly some of) the β parameters, and v_i is an additional random perturbation.

Adapting the previous notation, the utility for consumer n of choosing alternative i as follows:

$$U_{ij} = \beta z_{ij} + \sigma v_i z_{ij} + \varepsilon_{ij}$$

Still assuming a GEV distribution for the ε_{ij} term, the choice probability for alternative j could be derived for individual i:

$$P_{ij} = \int \frac{\exp[\beta z_{ni} + \sigma v_i z_{ij}]}{\sum_{l=1}^{J} \exp[\beta z_{il} + \sigma v_i z_{il}]} dP(v) \tag{3}$$

This formula is somewhat different from its MNL analogue because of the integral required to account for the random nature of the v_n terms. Depending on the assumed distribution, which is denoted by $P(v)$ in the previous formula, the integral for the choice probability could be computed numerically (by some sort of quadrature method), or by simulation. In the present paper the relevant integrals – assuming a standard normal distribution for the v terms – are estimated by simulation, using 50 Halton draws. These simulated choice probabilities are used to estimate the β_{ni} and σ parameters by Maximum Likelihood.

Closed forms can also be derived for the Equivalent Variation, which will be used to compute the effects of counter-factual measures, such as the urban road tax discussed here. The consumer surplus is approximated by the so-called logsum measure [Small and Rosen, 1981]:

$$E(CS) = \int \frac{\ln\left[\sum_{l=1}^{J} \exp[\beta_{il} z_{il} + \sigma v_l z_{il}]\right]}{\alpha} dP(v) + C \tag{4}$$

In which α represents the marginal effect of the cost variable on the non-random part of utility, and C is an integration constant which will be ignored throughout the analysis. The next stop is aggregate individual choices into traffic flows in a given route. As usual in discrete choice models, an important economic quantity is the ratio of marginal (dis)utility of time to the marginal (dis)utility of trip costs, which will be denoted as Value of Time (VOT_i):

$$VOT_{ij} = \frac{\partial U_{ij}/\partial t_j}{\partial U_{ij}/\partial C_j} \tag{5}$$

In which t_j represents the trip time for mode j and C_j trip costs for mode j

2.2 Route Traffic Model

There are I_k travelers who travel on route $k \in K$, from origin O_k to destination D_k. As in the previous part of the conceptual framework, the individual has a choice set composed of J mutually exclusive alternatives. The set of travelers choosing mode j on route k is the sum of y_{ij} defined as in equation (1), but only for those individuals who choose route k:

$$D_{jk} = \sum_{i \in I_k} y_{ij} \tag{6}$$

and total citywide demand for mode j is

$$D_j = \sum_k D_{jk}. \tag{7}$$

The next step is to embed these results into a model of citywide traffic, which is a problem similar to the "assignment problem" in the transportation literature. Our goal is to have an equilibrium model where individual mode choice affects travel time. In equilibrium, no citizen wants to change her transportation mode. Perhaps the trickiest part of this model is to link individual choices to aggregated traffic and travel time. There are many ways to do it. The key distinction among them is the level of network detail used. See de Dios Ortuzar et al. [1994] for a full discussion of these methods.

At one extreme, you have no network at all, only one aggregated equation relating the total volume of cars to travel time, like in Anderson [2014]. He uses the BUREAU OF PUBLIC ROADS [1964] equation that relates flow in a route to travel time. At the other end, there is a full specification of the network coupled with a geo-referenced map. In this case, the number of cars assigned to a specific route generates traffic. In these models, drivers are assumed to be rational and to minimize trip (generalized) costs. In between those two extremes, there are more stylized models of the network. The most common type is a network specified by nodes and links. The nodes are usually the centroid of a region, and links are the connections of this region to the others. These models can vary in the details of the links it uses, such as capacity, speed and so on.

In this paper, we use a version of this stylized model of city traffic network. Following the Origin Destination geographic division of Sao Paulo Metro area, we build an adjacency matrix for city zones, with a 1 on the (a, b) entry if zones a and b have a connecting border to each other, and 0 otherwise. Then we apply Dijkstra [1959] algorithm to find the shortest path (that crosses the smallest number of zones) between any regions a and b. This will be the assumed path of each commuter, and the traffic in a given zone is the sum of people that passes through that zone. The route traffic is defined similarly, as the sum of traffic of all zones this route crosses. Let $\Delta_n \subset K$ be the set of all routes whose shortest path (according to Dijkstra algorithm) crosses region n. Total traffic on region n is

$$trff_n = \sum_{k \in \Delta_n} D_{jk} \tag{8}$$

and total traffic on route k is the sum of all $trff_n$ for all n in the shortest path between origin O_k and destination D_k, which will be denoted as N_k:

$$T_k = \sum_{n \in N_k} trff_n \tag{9}$$

Equations 6, 7, 8 and 9 define the mapping from individual choices to route traffic: $T_k = \Phi(\mathbf{y})$, where $\mathbf{y}$ is the vector of individual choices.

2.3 Externality Model

Traffic on a given route also has some external effects. The external effects considered here are related to congestion, reflected on increased trip times. Trip time for a traveler on route k using mode j, represented by t_{jk}, depends on traffic and distance.

$$t_{jk} = t_j\left(T_{\cdot k}, dist_k\right) \tag{10}$$

This equation captures the negative externality generated by car users. We assume that only bus and car trip times are affect by traffic, and we consider only the negative externality of car users for congestion pricing. Since traffic and travel times are dependent on trip costs, the imposition of a congestion charge must take into account these effects.

2.4 Optimal Congestion Charge Algorithm

Given the feedback mechanism implied by our conceptual framework, under which changes in mode choice imply changes in travel times through the traffic model and the externality equation, we propose an iterative algorithm for finding the effects of a change in costs for any travel mode. Given a tax (or any change in trip costs) p, a traffic equilibrium is a fixed point of the system composed by equations 6, 8, 9 and 10. Let the pre-tax travel time be denoted as t^0. The iteration, written in pseudocode, is as follows:

Algorithm 1 Equilibrium Mode Choices

Require: $p \neq 0$ and $\epsilon > 0$

$tol > \epsilon$
$i = 1$
$maxit = 100$
while $tol > \epsilon, i < maxit$ **do**
    $D^0_{jk} = f(p, t^0_{jk})$ using equations 1 and 6
    $trff^0_n = g(D^0_{jk})$ for car mode, using equation 8 for each zone
    $T^0_k = \sum_{n \in N_k} trff^0_n$ as in equation 9
    $t^1_{jk} = t_j(T^0_k, dist_k)$ as in equation 10
    $D^1_{jk} = f(p, t^1_{jk})$ using 1 and 6
    $tol \to \|D^1_{jk} - D^0_{jk}\|$
    $t^0_{jk} \to t^1_{jk}$
    $i \to i + 1$
end while

The resulting choices after the recursion comes to an end are the equilibrium mode choices. The next step is to use this iterative algorithm to find an optimal congestion charge. The algorithm employs an inner while loop similar to the one in Algorithm 1, with an addition returning the value of time – which will be the source for the marginal external effect of congestion (as in Parry [2009]). The linear optimal congestion charge is one that makes the marginal social cost of traveling equal the marginal social benefit, regardless of any differences in the externality levels created by any specific trip. It is also presented in pseudocode as Algorithm 2.

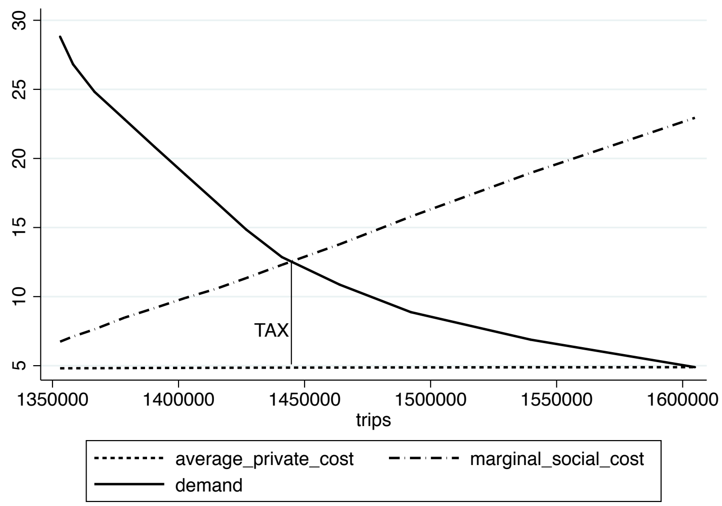
Algorithm 2 Optimal Linear Congestion Charge

Require: $\epsilon > 0, \epsilon_2 > 0$

p_0
$tol_2 > \epsilon_2$
$i_2 = 1$
$maxit_2 = 100$
while $tol_2 > \epsilon_2, i_2 < maxit_2$ **do**
    $tol > \epsilon$
    $n = 1$
    $maxit = 100$
    while $tol > \epsilon, n < maxit$ **do**
        $D^0_{jk} = f(p, t^0_{jk})$ using equations 1 and 6
        $trff^0_n = g(D^0_{jk})$ for car mode, using equation 8 for each zone
        $T^0_k = \sum_{n \in N_k} trff^0_n$ as in equation 9
        $t^1_{jk} = t_j(T^0_k, dist_k)$ as in equation 10
        $D^1_{jk} = f(p, t^1_{jk})$ using 1 and 6
        $tol \to \|D^1_{jk} - D^0_{jk}\|$
        $t^0_{jk} \to t^1_{jk}$
        $n \to n + 1$
    end while
    $VOT_{ij} = \frac{\partial U_{ij}/\partial t_j}{\partial U_{ij}/\partial C_j}$ as in equation 5
    $EXT_i = VOT_{ij} \times \frac{\partial t_j}{\partial T_k}$, with the last derivative from equation 10
    $EXT = \sum_i EXT_i$
    $p_1 = EXT - ((EXT/p_0) - 1)$
    $tol_2 = \|p_0 - p_1\|$
    $p_0 \to p_1$
    $i_2 \to i_2 + 1$
end while

The results of Algorithm 2 can also be presented graphically as in figure 1. In this figure, the simulation results of an optimal congestion charge for the morning peak are presented, as well as the marginal social benefit, marginal social cost and marginal private cost curves.

Figure 1: Equilibrium with Optimal Congestion Charge - morning peak



In the next section, the dataset used will be presented.

3 Data and Analysis

3.1 Dataset

The main source of transportation statistics to be used will be the Origin-Destination (OD) survey carried out by the Sao Paulo Subway Company. This survey has 169,625 observations, with information about trips made in 2007 in the region composed by 38 municipalities besides Sao Paulo city, depicted in Figure 2, an area known as Metropolitan region of Sao Paulo marked in yellow (38 municipalities) and in orange (Sao Paulo city).

In this survey, detailed information on the origin, destination, mode choice and attributes (both individual and chosen transportation mode) were recorded. The survey was carried out by a team of 370 researchers, visiting 54,700 households, with approximately 30,000 of those considered valid after the vetting process of the raw data. The Survey uses a stratified sampling technique, with error margins below 5%.

Figure 2: Sao Paulo Metro Area

Source: OD Survey, Sao Paulo (2007).

The survey has a wide range of individual information, such as income, car and home ownership, household size and other characteristics which could influence the decision of whether or not to take the trip and what transportation mode to use. The survey also has trip information as departure and arrival time, latitude and longitude for departure and arrival, mode of transportation, reason of the trip – leisure or work – and so on.

The survey classifies many different types of transportation modes. For simplicity, we reduced the choice set to the following set of alternatives:

- Bus: trips that respondents answered as their main mode choice either municipal, inter municipal or hired buses
- Rail: trips that respondents answered as main mode choice as subway or rail.
- Driving: trips that respondents answered as main mode choice as driving
- Motorcycle: trips that respondents answered as main mode choice as motorcycle
- Taxi: trips that respondents answered as main mode choice as taxi
- Other: other main mode choices - usually shared ride, walking or bicycle.

Another relevant issue for our dataset is respondents are followed during the day. That is, several trips in the dataset refer to trips made by the same individual in different moments of the day. From a raw total of 91,405 unique individuals, only 27,814 have only one trip recorded, reflecting the sometimes complex travel arrangements of individuals in Sao Paulo. Especially with respect to car usage, this will pose some challenges for the econometric analysis, since trip decisions in cases such as these are not to be independent, as to be seen below.

3.2 Travel time and cost matrices

Another important part of the database preparation is computing time and cost of options that were available but not chosen by the decision maker. Some alternatives had costs that could be determined the same way as costs for trips actually made – such as bus, subway and rail trips. For the other alternatives, the costs and duration for the other choices were estimated from a regression model using the observed choices for each mode. The dependent variable was the logarithm of trip cost or trip duration, and the independent variables were as such:

- Dummy Variables for Departure Hour
- Dummy Variables for Arrival Time
- Dummy Variables for Trip Motivation
- Distance in kilometers

The latter variable was estimated as the Euclidean Distance between the geographical coordinates of the OD zones (according to the survey's division in 460 zones). The estimated coefficients were used to estimate expected trip time and cost for alternatives not chosen[4][5], as in Lucinda et al. [2017].

Not all modes were available to all respondents, though. This problem was more pronounced for rail and subway modes, given the lack of infrastructure for all zones. This problem was addressed by restricting the availability of rail and subway choices for only those zones which had some respondents choosing these modes. The dataset which will be used in the following econometric analysis has the following descriptive statistics for the individual and household characteristics:

Table 1: Descriptive Statistics

Individual Specific Variables. N=34162 Individuals		
	Mean	Std. Deviation
Income (Monthly) in 1000BRL	2.7289	2.3942
Age	36.1454	12.7818
Not Student	0.8728	0.3332
Female	0.4282	0.4948
Household Variables. N=20820 Households		
	Mean	Std. Deviation
Number of HH Members	3.7743	1.8525
Number of cars	0.7117	0.7732

The results of table 1 indicate the average age of respondent is 36 years old, 87.3% of them are not currently studying and 42% female. The average household size is 3.74 members and the average number of cars is 0.71 per

[4]The maximum predicted time was trimmed at 300 minutes and the maximum predicted cost also in 300 BRL.
[5]Results available upon request.

household. As for the trip specific variables, the descriptive statistics across the transportation modes are in Table 2.

Table 2: Descriptive Statistics – Trip Specific Variables

	Observed Choices		Imputted Values	
	Mean	Std. Dev.	Mean	Std. Dev.
Trip Cost (in BRL)				
Bus	3.0349	1.2318	2.3000	0.0000
Rail	3.8121	2.3005	2.3000	0.0000
Driving	2.4370	3.8657	4.6111	7.6080
Motorcycle	0.6441	0.4337	0.9292	1.0808
Taxi	21.7254	20.3965	30.2214	29.9220
Shared Ride/Walking/Bike	0.0001	0.0275	1.5948	0.5945
Trip Time (Minutes)				
Bus	73.3150	36.1439	62.6896	47.3669
Rail	94.4100	39.0813	57.1711	23.7539
Driving	39.2598	35.1201	60.8492	55.2158
Motorcycle	29.3690	18.3345	40.4610	38.0708
Taxi	39.0545	43.7068	61.9800	63.7472
Shared Ride/Walking/Bike	21.9505	22.1204	58.1907	53.5581

OBS: Weighted Sample. N=66,693 trips observed and N=259,152 trips imputted.

The descriptive statistics are in line with what is to be expected, with both the cost for observed choices and the cost for the alternatives not chosen by the respondent are roughly in line with each other. The zero standard deviation for the imputed values for the trip costs for bus and rail is due to the fact these values were imputed according to the prevailing fares at the time.

In terms of number of trips using different modes and trip location, table 3 presents some interesting results:

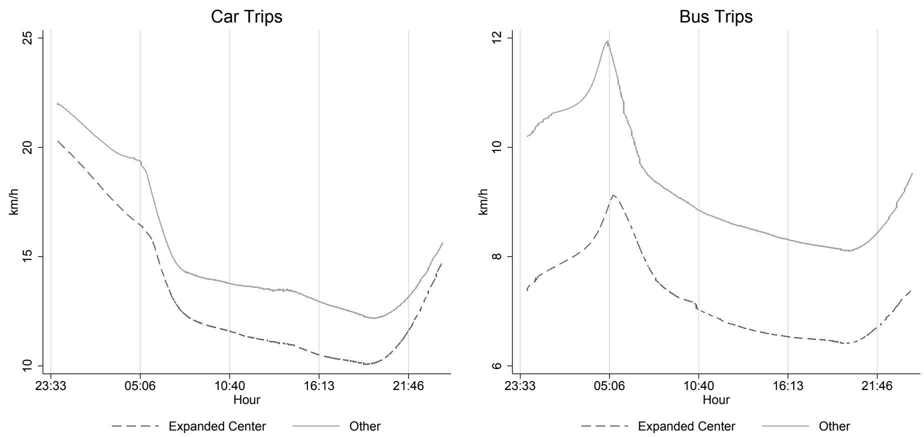
Table 3: Trip Distribution according to trip mode, motivation and region

	Trip Location		
	Other	Expanded Center	Total
	Work Related Trips		
Bus	5,674,147	273,843	5,947,989
Rail	2,024,716	182,456	2,207,172
Driving	3,788,459	660,198	4,448,657
Motorcycle	533,584	29,277	562,861
Taxi	18,664	16,728	35,392
Other	4,232,609	527,165	4,759,774
Total	16,272,178	1,689,667	17,961,845
	All Motives		
Bus	10,298,255	566,647	10,864,901
Rail	2,718,583	315,680	3,034,263
Driving	6,050,724	1,231,198	7,281,922
Motorcycle	671,668	38,212	709,881
Taxi	43,557	47,016	90,573
Other	14,537,558	1,463,566	16,001,124
Total	34,320,346	3,662,318	37,982,664

First of all, comparing the upper and lower panels of table 3, we can see that work related trips (that is, trips with were motivated by work reasons either in trip origin or destination) are more than half of all trips. This share is roughly similar across modes. The second issue is the number of trips to/from the expanded center is actually about 10% of all trips in Sao Paulo metropolitan area.

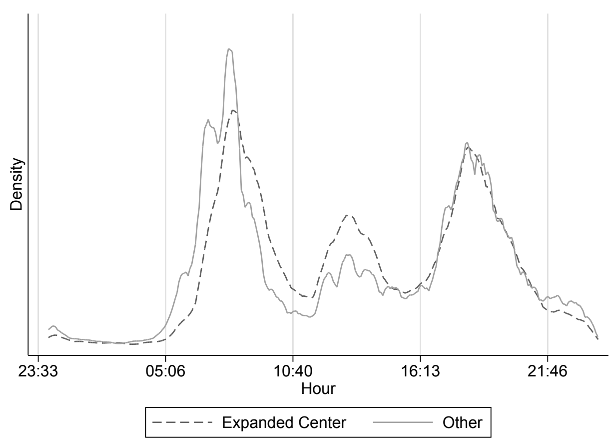
Although the share of trips to the expanded center is somewhat small, the issue here is not one of the absolute number of trips. The congestion problem is the effect of this traffic on a limited road infrastructure. Figure 3 shows a smoothed estimate for car and bus trip speed during the typical day. One can see in this figure that (i) car trips tend to be faster than bus trips, and (ii) trips to/from the expanded center tend to be slower than trips to/from areas off the expanded center.

Figure 3: Speed in km/h for trips in the Expanded Center or off the Expanded Center



The marked decrease in average speeds during the day can be related to the density of trips across the day, as we can see in the figure 4:

Figure 4: Smoothed Density of Trips during the day



In this figure, one can see the density of trips is higher during the period from 7am to approximately 4pm; after that both densities almost coincide. Both curves show a marked increase in the morning peak at about 7am. Given this dataset whose descriptive statistics are presented in tables 1 and 2, the next section will be focused on the modeling to be used.

4 Econometric Results

The econometric model follows the discrete choice approach discussed in equation 1 in section 2. Since the OD survey is a panel dataset, in which all individuals are followed for a period, we used an approach suggested by Train [2009, chap. 2] to model state dependence. We consider the respondent having to use her car when choosing the "Driving" alternative induces dynamics related to return home the car at the end of all trips. For example, a person that drives to work needs to return home driving later on. This is a substantial improvement over previous models, such as Lucinda et al. [2017], in which this dynamic was not considered and all trips were seen as independent.

The complete utility specification for an individual i from choosing mode alternative j is as follows:

$$U_{ij} = f(C_{ij}, t_{ij}) + SD_i \beta_{ij}^{SD} + \mathbf{Z_i} \beta_{\mathbf{ij}}^{\mathbf{Z}} + \varepsilon_{ij} \tag{11}$$

In which SD_i is a variable intended to capture the dynamics from driving, due to the fact that the trips in our dataset have a panel data structure (that is, several trips for the same individual are recorded). SD_i is a dummy variable which takes the value of zero if the person have not used the car, becomes 1 when the person uses the car and remain 1 until the car is returned home. It has value zero if the person take any non car trip after returning home the car. In summary, it is a variable that has value 1 if the car was used and not returned home. Since it is a variable which has a single value (zero or one) for all choices individual i faces in a given moment, it is a case specific variable. The $\mathbf{Z_i}$ is a vector of case specific variables as follows:

- Whether the individual is coming/going to the expanded center of Sao Paulo
- Whether the is respondent is female
- Whether the trip begins or ends at a zone with dedicated bus lanes
- Whether the trip begins or ends with a rail station
- Income level, denoted by I_i

And finally, the $f(C_{ij}, t_{ij})$ is a function of trip cost and trip time, intended both to capture unobserved alternative characteristics as well as dealing with the problem of independence of irrelevant alternatives. This function is specified as follows:

$$f(C_{ij}, t_{ij}) = \beta^C C_{ij} + \beta^t t_{ij} + \beta^{CI} \frac{C_{ij}}{I_i} + \beta^{tI} \frac{t_{ij}}{I_i} + \sigma^C v_{ij}^C C_{ij} + \sigma^t v_{ij}^t t_{ij}$$

In which v_{ij}^C and v_{ij}^t are Halton draws from a standard normal distribution and $\beta^C, \beta^t, \beta^{CI}, \beta^{tI}, \sigma^C$ and σ^t are coefficients to be estimated by Maximum Simulated Likelihood[6]. The following table shows the econometric results, using the sample of working trips and the survey sampling weights.

[6]The software used was Stata, version 14, and the code was the `mixlogit` by Arne Hole

Table 4: Coefficient Estimates for the Alternative Specific Variables

	Estimates	
Fixed		
Trip Time – t_{ij}	-0.0384	***
	(-18.0153)	
Trip Cost – C_{ij}	-0.2253	***
	(-13.1769)	
Trip Time/Income – $\frac{t_{ij}}{I_i}$	0.0187	***
	(8.2267)	
Trip Cost/Income – $\frac{C_{ij}}{I_i}$	-0.0562	***
	(-3.9220)	
Random		
Trip Time – σ^t	0.0984	***
	(32.7974)	
Trip Cost – σ^C	0.0982	***
	(16.4952)	
Observations	3.9e+05	
LR chi2	7.5e+03	
Log-Lik.	-1.6e+07	

T-Stats in Parentheses. ***-p-value<0.01,** p-value<0.05

Time spent on the trip has a negative effect on mode choice, and it increases as income increases. It corroborates the well known result (see Small et al. [2007]) that the value of time increases with the income level. On the other hand, trip cost also has a negative impact o mode choice, but it decreases as income increases. Which is in accordance with the fact that the marginal utility of income decreases as income increases.

The next table also present coefficients for the individual specific variables.

Table 5: Estimates for the Using Car Dummy – β_j^{SD}

	Estimates	
Using Car Dummy-Bus	-3.3354	***
	(-7.0868)	
Using Car Dummy-Rail	-3.3583	***
	(-11.3161)	
Using Car Dummy-Driving	3.6007	***
	(34.5392)	
Using Car Dummy-Motorcycle	-4.2680	***
	(-8.7837)	
Using Car Dummy - Taxi	-1.5985	***
	(-4.9055)	

T-Stats in Parentheses. ***-p-value<0.01,** p-value<0.05

As for the case specific variables, the most interesting result is about the State Dependence dummy - represented as "Using Car Dummy" in table 5, where it has a positive coefficient only for the "Driving" alternative, and negative for all others. From the estimated coefficients, the table 6 presents the price elasticities[7].

Table 6: Trip Mode Elasticity Matrix

		Responses in Choice Probabilities					
		Bus	Rail	Driving	Motorcycle	Taxi	Other
From changes in trip cost of	Bus	-0.3454	0.0694	0.0975	0.1102	0.0615	0.1073
	Rail	0.0632	-0.2803	0.0434	0.0571	0.017	0.0486
	Driving	0.1406	0.0733	-0.3623	0.1310	0.0621	0.1254
	Motorcycle	0.0062	0.0080	0.0039	-0.1783	0.0033	0.0105
	Taxi	0.0124	0.0063	0.0093	0.0174	-2.1883	0.0137
	Other	0.0488	0.0389	0.0499	0.1019	0.0286	-0.1637

Our own price demand elastcities for bus, rail and car trips are -0.34, -0.28 and -0.36, in line with the results of Batarce and Ivaldi [2014], which report values of -0.34, -0.43 and -0.39. These values are also in line with the results presented in the meta analysis of Oum and Walters II [2000], which report a range of own price elasticities for car trips between 0.00 and -0.52 and for bus trips between 0.01 and -0.96.

[7]The elasticities are arc elasticities, computed as the percentage change in choice probabilities of a 1% change in trip costs.

After estimating the demand parameters from the discrete choice model as in equation 11, we turn to the equation relating travel time to traffic. Table 7 shows the model estimates relating time to traffic.

It is usual in the economics of transportation literature to assume a pre-established relationship between travel time and traffic, as given by the BUREAU OF PUBLIC ROADS [1964] equation. We do not believe this is a good assumption, and opt to estimate this relationship from data. There are two reasons for that. First, the Bureau equation refers to traffic and time on a highway link, and that clearly is not the case here, where we look for traffic in a city. Second, given that it is a large city we want to have heterogeneity in delays across the city. We expect downtown delays in peak times to be much worse than delays on the suburbs.

Table 7: Traffic-time regressions

	Bus (log)		Car (linear)	
Total Traffic	0.2176	***		
	(20.9190)			
Distance	0.0182	***	0.8341	***
	(14.4756)		(23.5888)	
Origin of a Bus Lane	-0.0821	***	-2.5963	***
	(-4.6862)		(-5.2843)	
Destination of a Bus Lane	0.0004		3.5927	***
	(0.0228)		(7.6034)	
Car Traffic Volume			0.0003	***
			(30.8019)	
Constant	1.3717	***	16.3776	***
	(12.9548)		(38.9624)	
Observations	3.6e+03		6.5e+03	
R2	0.2607		0.3627	

T-Stats in Parentheses. ***-p-value<0.01,** p-value<0.05

The algorithm presented in 2 was employed to find the optimal linear congestion charge, which came at about 6.BRL. A visual depiction of this equilibrium is presented in figure 1, where the the numbers refer to the morning peak hours. The graph shows the computed demand function for car trips, the average private cost of and the social marginal cost (average cost plus the negative externality).

Economic theory asserts this is an improvement compared to the decentralized equilibrium, since it forces consumers to internalize the congestion externality from their decisions. However, it is not guaranteed a society will actually implement such a solution. In the next section, we use the model estimates and the algorithm to simulate the effects of an optimal congestion tax, and the political aspects of such a policy.

5 The Political Economy of a Congestion Tax

As for the effects of a congestion tax, De Borger and Proost [2012] theoretically show such a policy does have an heterogeneous effects on the population, with winners and losers. In this section, we use the conceptual framework previously developed to assess the theoretical claims these authors made regarding the political economy of a congestion tax under three scenarios. These scenarios all include a congestion tax, but differ with regard to the destination of the revenue raised by it. In the appendix, a fuller exposition of the De Borger and Proost model is presented and here only their main conclusions are reviewed. They start by positing individuals vote according to their welfare gains or losses. An individual who suffers a welfare gain is a supporter of a given measure and, conversely, an individual who has a decrease in his or her welfare opposes it.

Congestion Tax without Revenue Recycling Their baseline scenario is a congestion tax with no revenue recycling effect; that is, all revenues from the congestion tax will be spent in activities without any effect on consumers' welfare. They conclude in this case drivers are worse off, since they continue driving but have to pay the congestion tax. People who change their mode choice away from driving (for instance, towards public transportation) also lose, since they are moving to a (previously) worse mode choice. On the other hand, individuals who originally used other transportation modes (such as public transportation) do benefit from the congestion charge, since they reap the gains from reduced travel times with less congestion. With majority voting and a not too high number of drivers before the measure, a congestion charge can be implemented

Congestion Tax with Revenue Recycling – Lump-Sum transfer This scenario assumes all revenue raised from the congestion tax is returned to all individuals as a lump sum transfer. Assuming individuals who keep driving even after paying the congestion tax and receiving the lump sum transfer still experience a welfare loss, the main change with regard to the previous scenario is about the individuals who switch away from driving; now some of them might not lose from the measure, because the welfare losses from changing to a worse alternative are compensated by gains from the lump sum transfer. The individuals who originally used other transportation means are doubly benefited by this policy, since they experience reduced travel times and gain the lump sum transfer. As in the previous scenario, with majority voting and a not too high number of drivers before the measure, a congestion charge can be implemented. However, the number (before the congestion tax) of drivers required to block a congestion tax is now higher.

Congestion Tax with Revenue Recycling – Public Transportation Subsidies In the last scenario, all revenue is recycled as a reduction in public transportation fares. Users of public transportation now have a larger gain than before (since the expenditure is focused on public transportation users), whereas individuals who continue driving lose just as in the previous scenario. However, this scenario provider better gains for individuals who switch away from driving towards public transportation, which restricts even more the initial mass of drivers required to block the imposition of a congestion tax – that is, this is the scenario in which such a measure is most likely to be implemented.

An important thing to notice is we are not taking into account the effects such a policy might have on other distortions in the economy, which might influence welfare gains for all individuals. For instance, Parry and Bento [Parry and Bento, 2001]show the welfare effects on the labor market of a congestion can easily exceed the welfare effects from transport mode choices. Another distortion which we do not take into account in the present paper is the role of other distortions in the economy. Parry [Parry, 2002] points out that when public transportation fares are set below their marginal cost, a congestion tax could have an additional welfare effect from the additional commuters whose cost must be partially covered by taxes. In the next section, we will use the results of Algorithm 1 to estimate the welfare effects of a Congestion Tax in each scenario.

5.1 Scenario Results

Scenario I: congestion tax with no revenue recycling We start by looking at the simplest policy, simply implementing an optimal congestion tax on downtown Sao Paulo, and revenues are used with no positive welfare effects to the population. Table 8 shows that even an optimal tax would face resistance by a majority of the population. One individual favors the policy if its average logsum (see equation 4) over all her trips is greater then before the policy. That is one explanation for the low use of congestion tax as a policy for city congestion.[8]

As pointed out by Armelius and Hultkrantz Armelius and Hultkrantz [2006], a congestion tax tends to face opposition by the majority without further measures. The last two columns of the table show the percentage of citizens that gain, lose or are indifferent to the policy, conditional on initially being a car user or not. Non surprisingly, it faces more resistance from car users than from non users.

Table 8: Against, Neutral or in Favor of the Tax - total and percentage

	N	total (%)	car users (%)	non users (%)
Against	4,444,173	33	70	26
Neutral	7,122,215	54	1	64
Pro	1,736,782	13	29	10

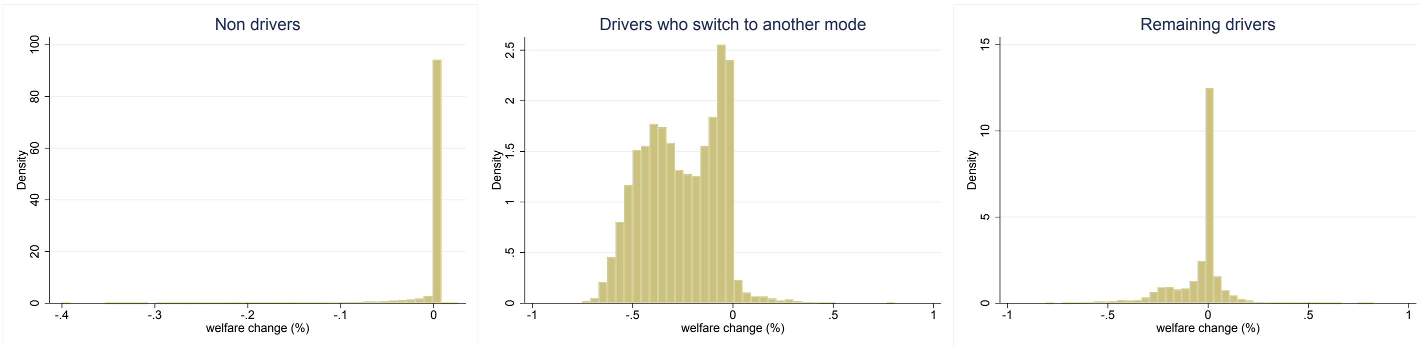
As discussed in the previous section, De Borger and Proost [De Borger and Proost, 2012] analyze the welfare and political effect of a congestion tax. Table 9 and figure 5 shows the welfare impact of a congestion tax on the three groups discussed: remaining drivers, those who switch to public transportation, and non drivers. The empirical results corroborate some of the theoretical findings, but also point to major differences in results. First, the tax has mixed impacts on remaining drivers: less than 10% of drivers suffer a welfare loss greater than 10%, while the others have a negligible impact, and the top 10% experience substantial welfare increase. The graph on the right hand side on figure 5 shows the histogram of welfare change for remaining drivers.

Second, as expected most non drivers are not affected by the policy, with 1% having a 5% reduction in welfare, as show on the graph on the left.[9]

[8]License plate restriction and similar policies are more common in developing countries. See Lucinda et al (2017).

[9]This is a feature of the logit model and its consumer surplus formula, the logsum. It is the log of the sum of all available alternatives. Hence, there is a utility reduction for non drivers which have driving as an available option.

Figure 5: Welfare change with congestion tax for different groups



And last, but perhaps the most interesting part, it is the fact that more than 95% of the citizens that switch from car to public transit are worse off. De Borger and Proost's model predicts that the remaining drivers is the group that suffers the greatest welfare loss. Our results show that the group that switch to the public transport is the one that carries the greatest policy burden. More than 50% of the individuals in this group experience a welfare loss of more than 20%. And the lowest decile has a welfare decrease of more than 50%. The histogram on the middle of figure 5 shows the welfare change for this group.

The reason is the following: value of time and value of money varies across citizens. Value of money decreases and value of time increases with income. Hence, the tax has little impact on the higher income consumers which at the end benefit from lower travel times. The group that switches is the one that really feels the impact of the tax on welfare, and has a lower value of time. Consumer heterogeneity plays a key role in the results, and shows that the cost of the policy falls heavily on this 'middle' group, which abandon the car and switch to public transportation. The tax is not high enough to impact the upper class, and it also does not affect the lower income groups since they do not drive anyway.

Table 9: Welfare Change for Different Groups

	p99	p95	p90	p75	p50	p25	p10	p5	p1
remaining-drivers	0.253	0.117	0.063	0.006	-0.000	-0.015	-0.088	-0.163	-0.335
switch-car-to-public	0.131	-0.004	-0.020	-0.074	-0.238	-0.405	-0.506	-0.553	-0.637
non-drivers	0.000	0.000	0.000	0.000	0.000	0.000	-0.000	-0.000	-0.055

Table 10 shows the regression of an indicator variable that takes value one for those who switch away from the car on explanatory variables. We can see that distance, income and their interaction have a negative effect on leaving the car, while crossing the downtown area has a strong positive effect, as expected. It shows that higher income individuals switch less away from the car than lower income ones. And this effects increases as the distance travelled increases. The results are in line with the stylized fact that wealthier people who live outside downtown are heavy car users.

Table 15 on the appendix shows the transition matrix across transportation modes when a congestion tax is

Table 10: Who stops driving?

	(1)
	Ex drivers
Distance	-0.00105***
	(-43.52)
Income	-0.0153***
	(-384.02)
Dist x Income	-0.0000138**
	(-2.77)
Downtown	0.735***
	(2566.41)
Constant	0.0640***
	(322.38)
Observations	3275940

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

introduced on car users. Before the tax, a total of more than 9 million trips were done by car. The introduction of the tax shifts nearly a million trips to other transportation modes. From these, 751 thousand switch to the "outside option" Other, which includes walking, cycling etc, while 162 thousand trips change to bus.

The changes in trip times for drivers caused by the congestion tax is shown in figure 8. The time reduction is rather small, with few trips having a reduction of more than one minute.

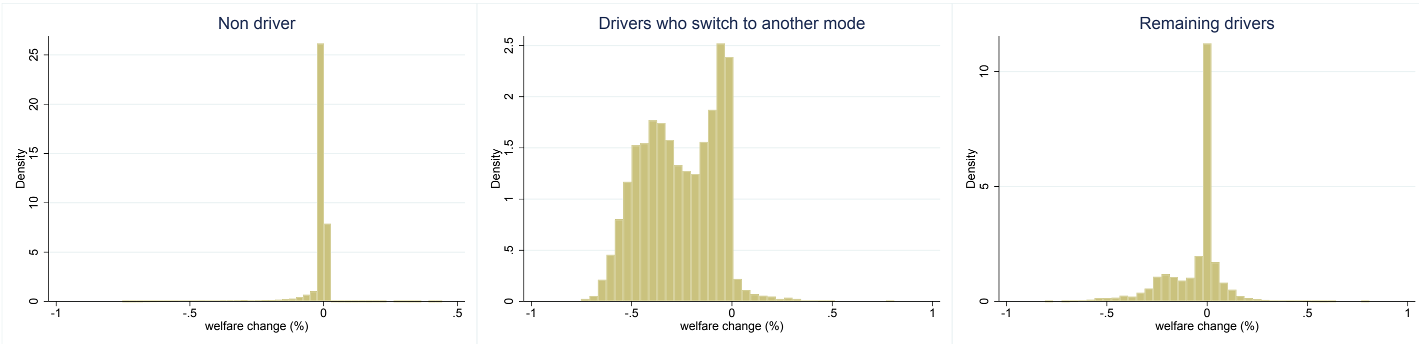
Scenario II: congestion tax with lump sum transfers In this case tax revenues are redistributed back to the whole population in a lump sum fashion. For this simulation, we added the rebate to the average (over trips) logsum of each citizen, since changes in the logsum are measures of equivalent variation changes from a given policy. A tax of 6.25 BRL was charged and the revenue returned back to the population as a rebate of 0.46 BRL to each citizen per day. As table 11 shows, the population majority now favors the policy, with all consumers who were neutral now supporting the policy and the share of individuals against it declining greatly. This is due to the rebate the population receives, especially the group that was at first indifferent between having the policy or not.

Table 11: Against or in Favor - tax with lump sum rebate

	N	total (%)	car users (%)	non users (%)
Against	1,800,404	14	39	8
Pro	11,502,766	86	61	92

Table 12 and figure 6 show the welfare changes from introducing a tax with a lump sum rebate of the total

Figure 6: Welfare changes with congestion tax and lump sum rebate for different groups



revenue. If one compares these results with the ones in table 9 all citizens which were indifferent before are now better off. However, it is important to note that the welfare change due to the lump sum transfer is small. For example, the 99th quantile of non drivers have a less than 0.5% change in welfare due to the transfer. Other non drivers have an even smaller change. Figure 6 is almost identical to figure 5.

Table 12: Welfare Change for Different Groups - lump sum redistribution

	p99	p95	p90	p75	p50	p25	p10	p5	p1
remaining-drivers	0.265	0.127	0.068	0.010	0.000	-0.023	-0.119	-0.200	-0.387
switch-car-to-public	0.132	-0.004	-0.020	-0.074	-0.238	-0.405	-0.506	-0.553	-0.637
non-drivers	0.004	0.002	0.002	0.001	0.001	0.000	0.000	0.000	-0.153

Since the lump sum rebate does not change preferences, the transition matrix among modes is the same as table 15.

Scenario III: congestion tax with transit subsidies Now the policy consists of taxing car drivers and subsidizing bus transportation. We choose the bus since this mode has more flexibility to adjust for a higher demand than rail, since bus lines can be deployed in response to increased demand much quicker than rail lines.

We compute the bus subsidy per bus trip as the equilibrium subsidy that would induce as much car and bus usage as to generate this amount of subsidy revenues. Given the optimal congestion tax, the product of the number of car users in downtown Sao Paulo and the tax gives the total revenue generated. However, when this amount is used to subsidize the bus, some drivers would switch to public transit, reducing the revenue from the tax and, consequently, the bus subsidy. The algorithm has to search for the equilibrium bus subsidy in which the tax and subsidy would imply as much car use as to generate a revenue that would give this exact subsidy for bus users.

The equilibrium subsidy is 1.95 BRL, which is a little less than the average bus fare (2.30 BRL in 2007). It would imply a 35 cents bus tariff. A very low value for any standard. Since buses are a relevant alternative for a large share of the population, subsidizing it has an important positive welfare effect. Table 13 below shows the percentage of winners and losers. Now, 87% of population supports the policy. As before, the support is stronger among non drivers.

Table 13: Against or in Favor - tax with bus subsidy

	N	total (%)	car users (%)	non users (%)
Against	1,669,248	12	33	9
Neutral	121,709	1	0	1
Favor	11,512,213	87	67	90

Table 14 shows the welfare change for the three groups - remaining drivers, people who switch from car to public and non drivers. Bus subsidies improved the situation of the three groups compared to the benchmark situation.

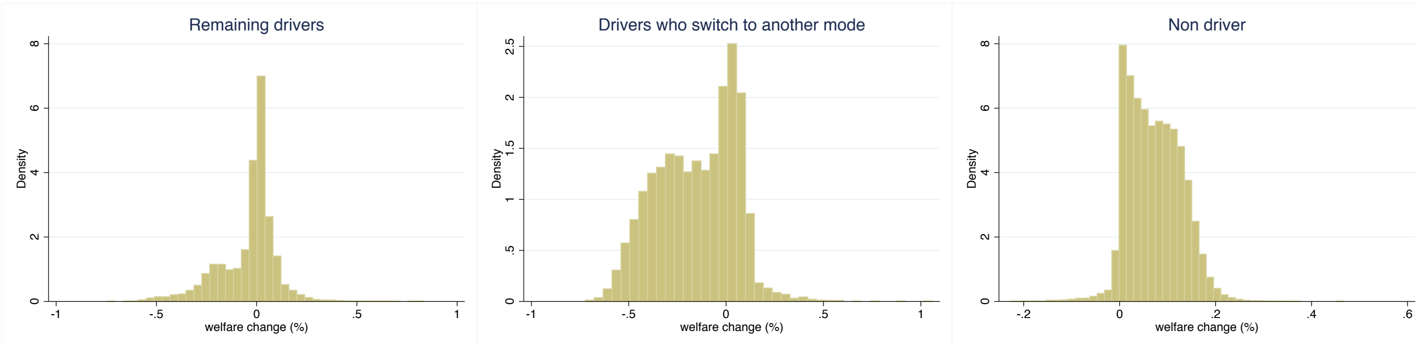
More interesting is to compare the impact of bus subsidies with those from the lump sum transfer. As expected, the largest impact is on the group that switched from car to public transportation. In the case of the lump sum transfer, more than 95% of the citizens in this group are negatively affected by the tax. When we move to bus subsidies, little more than 50% is negatively affected. However, the major difference is on the magnitude of the change. The top 10% in terms of welfare change experience a 10% or higher change in welfare with the subsidy. Under the lump sum transfer, only the top 1% do not lose with the implementation of the tax.

Non drivers are better off under both policies - lump sum transfer or bus subsidy. But the welfare increase is substantially larger under the bus subsidy. The top 25% of this population in terms of welfare change increase its welfare by 5% or more under the bus subsidy, and in less than 1% under the lump sum transfer. Although giving qualitatively similar results, the policies have very different quantitative impacts on welfare. Also, the results indicate that targeting the revenue usage to the group that suffers the most with the tax may help increase the political acceptance of the policy, as opposed to a uniform lump sum transfer for example.

Table 14: Welfare Change for Different Groups - public transit subsidies

	p99	p95	p90	p75	p50	p25	p10	p5	p1
remaining-drivers	0.286	0.159	0.102	0.047	0.011	-0.024	-0.124	-0.209	-0.363
switch-car-to-public	0.379	0.143	0.106	0.060	-0.039	-0.218	-0.352	-0.430	-0.535
non-drivers	0.198	0.163	0.146	0.116	0.070	0.028	0.008	0.002	-0.004

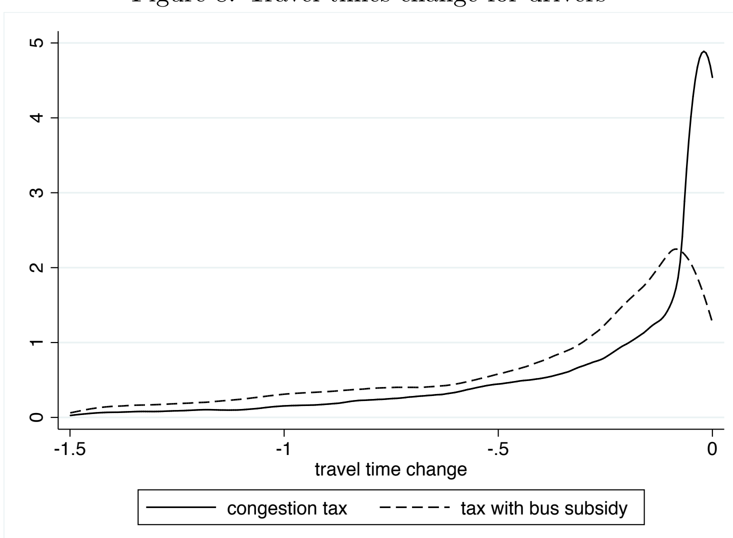
Figure 7: Welfare change with congestion tax and bus subsidy for different groups



If we compare the empirical results to the theoretical ones from section A.1, we can see that there is more disagreement within groups of citizens (remaining drivers, switch to public transportation and non drivers) than what theory predicts. In both policies with revenue recycling there is a massive support for the policy implementation.

Changes in trip times is displayed on picture 8 below. The time reductions now are substantially larger than in the case of a tax with no bus subsidy.

Figure 8: Travel times change for drivers



6 Conclusions

In this paper we develop a econometric structural model to estimate and compute the optimal congestion tax, and implement the estimation using origin and destination micro data for the city of Sao Paulo. Then, it uses the model estimates to conduct a welfare analysis to try to understand the political economy behind the low use of a congestion tax as a valid policy to alleviate traffic congestion. We compute an optimal congestion tax of 6.25 BRL (2 US Dollars) per trip in downtown Sao Paulo.

The political economy analysis shows that commuters that switch away from the car are bearing most of the tax burden. The reason is the following: there are strong heterogeneity across the population, with value of money decreasing and value of time increasing with income. Hence, the tax has little impact on the higher income consumers which at the end benefit from lower travel times. The group that switches - the middle group - is the one that really feels the impact of the tax on welfare, and has a lower value of time. The tax is not high enough to have a strong impact on the remaining drivers, and it also does not affect the lower income groups since they do not drive anyway.

Revenue recycling is a major issue when dealing with the acceptance of a congestion tax. In terms of general acceptance by the population, it does not make much difference in terms of welfare whether the recycling is about reducing bus fares or transferring lump-sum to individuals. However, bus subsidies have the greatest impact on the welfare of the new bus users, that stop using their cars due to the tax. This is the group that can potentially put up the strong opposition to the tax implementation.

There is a substantial difference in terms of traffic reduction, also. The bus subsidy induces a much stronger migration to the bus compared to the lump sum transfer or no transfer at all. There is a 30% reduction in car use using the revenues for transit subsidies, compared to a 10% reduction without it. However, such increases in bus usage could not be achieved considering existing capacity constraints in developing countries.

The scenario without revenue recycling mimics the situation of cities where the cost of the fiscal budget is too high or situations where there is low trust that revenue recycling is going to take place. Something familiar to large cities in the developing world. The results show that it is politically difficult to implement the congestion tax without revenue recycling, what may explain the seldom use of this policy. Since the general public may be much more favorable with revenue recycling, an important policy when trying to implement a congestion tax is to increase the credibility of the return of the tax revenue.

References

W L Adamowicz. Habit Formation and Variety Seeking in a Discrete Choice Model of Recreation Demand. *Journal of Agricultural and Rseource Economics*, 19(1):19–31, 1994.

Michael L Anderson. Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *The American Economic Review*, 104(9):2763–2796, 2014.

Hanna Armelius and Lars Hultkrantz. The politico-economic link between public transport and road pricing: An ex-ante study of the stockholm road-pricing trial. *Transport Policy*, 13(2):162–172, 2006.

Leonardo J Basso and Hugo E Silva. Efficiency and substitutability of transit subsidies and other urban transport policies. *American Economic Journal: Economic Policy*, 6(4):1–33, 2014.

Marco Batarce and Marc Ivaldi. Urban travel demand model with endogenous congestion. *Transportation Research Part A: Policy and Practice*, 59:331–345, 2014.

BUREAU OF PUBLIC ROADS. Traffic assignment manual. *US Department of Commerce*, 1964.

Bruno De Borger and Stef Proost. A political economy model of road pricing. *Journal of Urban Economics*, 71(1): 79–92, 2012.

Juan de Dios Ortuzar, Luis G Willumsen, et al. *Modelling transport*. Wiley New Jersey, 1994.

Edsger W Dijkstra. A note on two problems in connexion with graphs. *Numerische mathematik*, 1(1):269–271, 1959.

Anthony Downs. The law of peak-hour expressway congestion. *Traffic Quarterly*, 16(3), 1962.

Jonas Eliasson and Lars-Göran Mattsson. Equity effects of congestion pricing: quantitative methodology and a case study for stockholm. *Transportation Research Part A: Policy and Practice*, 40(7):602–620, 2006.

T Erdem. A dynamic analysis of market structure based on panel data. *Marketing Science*, 15(4):359–378, 1996. ISSN 0732-2399. doi: 10.1287/mksc.15.4.359.

Ronald Aylmer Fisher and Leonard Henry Caleb Tippett. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge Univ Press, 1928.

B. Gnedenko. Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire. *The Annals of Mathematics*, 44(3):423+, July 1943. ISSN 0003486X. doi: 10.2307/1968974. URL `http://dx.doi.org/10.2307/1968974`.

Gene M. Grossman and Elhanan Helpman. Protection for sale. *The American Economic Review*, 84(4):833–850, 1994. ISSN 00028282. URL `http://www.jstor.org/stable/2118033`.

Richard Layard. The Distributional Effects of Congestion Taxes. *Economica*, 44(175):297, aug 1977. ISSN 00130427. doi: 10.2307/2553654. URL http://www.jstor.org/stable/pdf/2553654.pdf?acceptTC=true http://www.jstor.org/stable/10.2307/2553654?origin=crossref.

Robin Lindsey. Do economists reach a conclusion on road pricing? *Econ Journal Watch*, 3(2):292–371, 2006. ISSN 1933-527X. URL <Go to ISI>://WOS:000203013600008.

Claudio Lucinda, Rodrigo Moita, Bruno Ledo, and Leandro Meyer. The economics of sub-optimal policies for traffic congestion. *Journal of Transport Economics and Policy*, 2017.

Jozsef Molnar and L Mesheim. A disaggregate analysis of demand for local bus service in great britain (excluding london) using the national travel survey. *Draft, University College London and Institute for Fiscal Studies*, 2010.

T. H. Oum and W. G. Walters II. *Transport Demand Elasticities*, volume 1, chapter 12. Elsevier Science Ltd., 2000.

Ian W.H. Parry. Comparing the efficiency of alternative policies for reducing traffic congestion. *Journal of Public Economics*, 85(3):333–362, 2002. ISSN 00472727. doi: 10.1016/S0047-2727(00)00163-8. URL http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-DP-00-28.pdf.

Ian W.H. Parry. Pricing Urban Congestion. *Annual Review of Resource Economics*, 1(1): 461–484, oct 2009. ISSN 1941-1340. doi: 10.1146/annurev.resource.050708.144226. URL http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.233.607&rep=rep1&type=pdf http://www.annualreviews.org/doi/abs/10.1146/annurev.resource.050708.144226.

Ian W.H. Parry and Antonio Bento. Revenue Recycling and the Welfare Effects of Road Pricing. *Scandinavian Journal of Economics*, 103(4):645–671, 2001. ISSN 14679442. doi: 10.1111/1467-9442.00264. URL http://www.feem-web.it/ess/ess06/files/parryandbentorevenuerecyclinginjose.pdf.

Elena Safirova, Kenneth Gillingham, Ian W.H. Parry, Peter Nelson, Winston Harrington, and David Mason. 8. Welfare and Distributional Effects of Road Pricing Schemes for Metropolitan Washington Dc. *Research in Transportation Economics*, 9(04):179–206, 2004. ISSN 07398859. doi: 10.1016/S0739-8859(04)09008-0. URL http://www.rff.org/files/sharepoint/WorkImages/Download/RFF-DP-03-57.pdf.

Kenneth a. Small. The incidence of congestion tolls on urban highways. *Journal of Urban Economics*, 13(1):90–111, 1983. ISSN 00941190. doi: 10.1016/0094-1190(83)90047-5.

Kenneth A. Small and Harvey S. Rosen. Applied welfare economics with discrete choice models. *Econometrica*, 49(1):105–30, January 1981.

Kenneth A Small, Erik T Verhoef, and Robin Lindsey. *The economics of urban transportation*. Routledge, 2007.

Kenneth Train. *Discrete choice methods with simulation*. Cambridge University Press, Cambridge New York, 2009. ISBN 0521747384.

Kenneth E. Train. Mixed logit. In *Discrete Choice Methods with Simulation*, chapter 6, pages 138–154. Cambridge University Press, Cambridge, MA, 1 edition, 2003.

William Vickrey. Some objections to marginal-cost pricing. *The Journal of Political Economy*, pages 218–238, 1948.

Alan A Walters. The theory and measurement of private and social cost of highway congestion. *Econometrica: Journal of the Econometric Society*, pages 676–699, 1961.

A De Borger and Proost (2011) Model

A.1 A Theoretical Political Economy Model of a Congestion Tax

We lay out a simplified version of De Borger and Proost (2012) model. It helps us understand how a congestion tax and different types of tax rebate affect the welfare of drivers and non-drivers. It will guide us in the empirical analysis of the next section.

Suppose there are two transportation modes, car and public transit - buses, to make it simple - and a continuum of commuters uniformly distributed on the $[0, 1]$ interval, such that for any driver indexed by n we have $n \in [0, 1]$. So that the total demand for transportation is inelastic but there is substitution between car and public transport. The average cost of car transport is $AC = cn$, increasing on the number of drivers. The total cost is $TC = cn^2$.

The generalized cost of public transport for individual i is $c_i = 1 - i$. The difference comes from access cost, for example. It implies that $c_0 = 1$ and $c_1 = 0$ are the highest and lowest access cost among all individuals. This generalized cost of public transport can also be understood as the willingness to pay for car use.

The equilibrium number of car users is given by the solution to $1 - n = cn$. Which is

$$n^o = \frac{1}{c+1} \tag{12}$$

And the equilibrium number of public transport users is $1 - n^o$.

Since $TC = cn^2$, the marginal social cost is $MSC = 2cn$. The social optimum number of drivers is given by the solution to $2cn = 1 - n$, which is

$$n^* = \frac{1}{2c+1} \tag{13}$$

Note that $n^o > n^*$.

There are two policy instruments: a congestion tax t and a subsidy to the public transport s. The user equilibrium with tax and subsidy is given by $1 - n - s = cn + t$. It implies that the optimal tax and subsidy is given by the following expression:

$$t + s = n^* c \tag{14}$$

We now are able to analyze three different policies. (i) a congestion tax alone, (ii) a congestion tax with a lump sum rebate to the whole population of commuters and (iii) a congestion tax with its revenue used to subsidize the bus fare.

A.1.1 Congestion tax

In this case, $t = cn^*$, $s = 0$ and there are no lump sum transfers.

Remaining drivers pay the tax and have a decrease in travel time, $-cn^* + c(n^o - n^*)$. Substituting equations 12 and 13, we have

$$\frac{-c}{(2c+1)(c+1)} < 0 \tag{15}$$

Therefore, remaining drivers are worse off in this case.

Ex-drivers that switch to public transport are also worse off. They gain some time, do not have the cost of the car trip but incur the cost of public transport: $c(n^o - n^*) + cn^* - (1-n)$. There is a commuter n' that is indifferent between driving car or taking the bus, such that

$$c(n^o - n^*) + cn^* - (1 - n') = 0 \tag{16}$$

It is easy to show that $n' = \frac{1}{c+1}$, and that is exactly n^o, given in equation 12. Since any individual n, such that $n < n'$, is worse off with the tax, it means that all commuters that switched from the cars to the buses lose with this policy.

Initial public transport users are better off since they pay nothing and save time on their trips.

Summarizing, remaining drivers and new public transport users are worse off, while the initial non drivers are better off. Majority voting implies that a tax is implemented if $1 - n^o > n^o$, which is true if $c > 1$.

A.1.2 Congestion tax with lump sum transfers

Now, $t = cn^*$ and $m = c{n^*}^2$, where m is the lump sum transfer to all commuters.[10]

Again, remaining drivers pay the tax, enjoy a decrease in travel time and now receive the lump sum transfer, $-cn^* + c(n^o - n^*) + c{n^*}^2$. It is straightforward to show that this expression is negative if $\frac{1-n^*}{n^*} > c$.

Ex-drivers that switch to public transport gain some time, do not have the cost of the car trip but incur the cost of public transport and receive the transfer $c(n^o - n^*) + cn^* - (1-n) + c{n^*}^2$. Now this group is better off, with the indifferent citizen being $n' = n^o - c{n^*}^2$. It means that not all new public transport users are worse off. Of course, initial bus users are even better since they save time on bus rides and receive the lump sum transfer.

[10]We assume everyone is a commuter and voter.

If we consider the more sensible and interesting case where remaining drivers lose with this policy, majority voting implies that the tax is implemented if $1 - n' > n'$.

A.1.3 Congestion tax with transit subsidies

In this case, $t + s = cn^*$, and the government budget restriction is $tn^* - s(1 - n^*) = 0$.

In this case remaining drivers are worse off than in the previous case with lump sum transfers. Commuters that switched to buses have a change in utility of $c(n^o - n^*) + cn^* - (1 - n) + \frac{cn^{*2}}{1-n^*}$, which is more than on the previous case. Original bus users receive a larger transfer now.

Among the three policies, this is the case where the policy is more likely to be implemented.

B Tables

Table 15 shows the transition matrix across transportation modes when a congestion tax is introduced on car users. Before the tax, a total of more than 9 million trips were done by car, as shown in the Total column at the far right of the table. This column shows the volume of trips before the tax. The bottom row shows the total number of trips after the congestion charge. The introduction of the tax shifts nearly a million trips to other transportation modes. From these, 751 thousands switch to the "outside option" Other, which includes walking, cycling etc. While 162 thousands trips switch to Bus and 98 thousands switch to Rail.

Table 15: Transition matrix - with a congestion tax

	Bus	Rail	Driving	Motorcycle	Taxi	Other	Total
Bus	11,133,748	0	24,598	0	0	35,815	11,194,161
Rail	291	4,710,151	4,566	0	0	11,169	4,726,177
Driving	162,124	98,857	7,996,957	84	0	751,554	9,009,576
Motorcycle	469	0	250	763,417	0	188	764,324
Taxi	0	0	0	0	169	0	169
Other	8,361	753	6,805	111	0	12,262,355	12,278,385
Total	11,304,993	4,809,761	8,033,176	763,612	169	13,061,081	37,972,792

Table 16 shows the transition matrix among modes when the policy consists of tax and bus subsidy. This policy creates a large reduction in car usage of approximately 2 million trips being done by bus instead of car. It is an attractive policy from this perspective. In practice, some caution must be taken: the bus system may not be able to handle a sharp and large demand increase, which would require substantial investments in capacity.

Table 16: Transition matrix - tax and bus subsidy

	Bus	Rail	Driving	Moto	Taxi	Other	Total
Bus	11,187,458	0	5,305	0	0	1,398	11,194,161
Rail	296,883	4,339,155	23,982	9,188	0	56,969	4,726,177
Driving	1,978,372	123,350	6,244,291	1,201	0	662,362	9,009,576
Moto	311,605	1,548	1,771	438,818	0	10,582	764,324
Taxi	0	0	0	0	169	0	169
Other	6,433,594	8,106	17,672	25,072	94	5,793,847	12,278,385
Total	20,207,912	4,472,159	6,293,021	474,279	263	6,525,158	37,972,792