Policy Research Working Paper 5099

Impacts of Policy Instruments to Reduce Congestion and Emissions from Urban Transportation

The Case of São Paulo, Brazil

Alex Anas
Govinda R. Timilsina

The World Bank
Development Research Group
Environment and Energy Team
September 2009

Abstract

This study examines impacts on net social benefits or economic welfare of alternative policy instruments for reducing traffic congestion and atmospheric emissions in São Paulo, Brazil. The study shows that expanding road networks, subsidizing public transit, and improving automobile fuel economy may not be as effective as suggested by economic theories because these policies could cause significant rebound effects. Although pricing instruments such as congestion tolls and fuel taxes would certainly reduce congestion and emissions, the optimal level of these instruments would steeply increase the monetary cost of travel per trip and are therefore politically difficult to implement. However, a noticeable finding is that even smaller tolls, which are more likely to be politically acceptable, have substantial benefits in terms of reducing congestion and emissions. Among the various policy instruments examined in the study, the most socially preferable policy option for São Paulo would be to introduce a mix of congestion toll and fuel taxes on automobiles and use the revenues to improve public transit systems.

This paper—a product of the Environment and Energy Team, Development Research Group—is part of a larger effort in the department to study climate change and clean energy issues. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at gtimilsina@worldbank.org.

Impacts of Policy Instruments to Reduce Congestion and Emissions from Urban Transportation: The Case of São Paulo, Brazil[1]

Alex Anas
Department of Economics
State University of New York at Buffalo
Amherst, New York 14260, USA

Govinda R. Timilsina[2]
Senior Research Economist, Environment & Energy Unit
Development Research Group
The World Bank, 1818 H Street, NW
Washington, DC 20433, USA

Key worlds: Transport externalities, emission reductions, urban planning

[1] We sincerely thank Mark Lundell, Mike Toman and Ashish Shrestha for their comments. The Knowledge for Change Program (KCP) Trust Fund is acknowledged for the financial support of this study. The views expressed are those of the authors and do not necessarily represent the World Bank and its affiliated organizations.

[2] Corresponding author; Tel: 202 473 2767; E-mail: gtimilsina@worldbank.org

Impacts of Policy Instruments to Reduce Congestion and Emissions from Urban Transportation: The Case of Sao Paulo, Brazil

Alex Anas, Govinda R. Timilsina

1. Introduction

São Paulo is the largest city in Brazil and in South America, and the São Paulo metropolitan area, with an estimated population of 21,616,060 in 2008 spread over an area of 7,944 square kilometers, ranks among the most populous metropolitan areas in the world (Instituto Brasileiro de Geografia e Estatística, 2002; 2008). Sao Paulo is a highly congested and polluted megacity due to the rapid increase in car ownership resulting from income growth without adequate expansion of road capacity. The city emitted more than 15 million tons of CO_2 in 2003, with 76% of that total from energy use, of which gasoline and diesel combustion accounted for 68%. The city government has identified climate change mitigation as a high priority and has presented an ambitious proposal to reduce CO_2 emissions by 30% by 2012, using both economic and fiscal incentives (Goldman, 2008). Moreover, it has been reported that particulate matter (PM) and the ozone precursors (nitrogen oxides and volatile organic compounds) from vehicular emissions pose a grave threat to local air quality in the São Paulo metropolitan area (Sánchez-Ccoyllo et al., 2009).

Vehicular transportation is the primary source of congestion and environmental externalities in Sao Paulo as one of the most desirable normal goods is the ownership of a private vehicle (Ingram, 1998). Income growth stimulates the demand for car ownership, and with more car ownership, congestion and pollution, including CO_2 emissions, increase. The increases in emissions that could come from some of the developing countries in the future can be staggering, due to the enormous urbanization and real income growth that potentially lies ahead. In particular, traffic congestion and pollution tends to be worst in the megacities of the rapidly developing nations such as China and Brazil (see, for example, Gurjar et. al. 2008).

With increased urbanization and higher incomes, suburban and exurban sprawl accelerates with urban areas spreading out in low density patterns that are believed to favor mobility over longer distances by private motorized vehicles rather than by public transit or by bicycle. Meanwhile, high densities that can be achieved in urbanized areas support potential investments in rail transit systems that could greatly reduce reliance on the automobile, than if the same population were spread over more but smaller cities, each unable to support the economies of scale inherent in rail mass transit. Although it is a widely held perception that sprawl in land use causes more aggregate car miles to be driven, recent results from modern theoretical models of the urban economy in which both jobs and residences decentralize with sprawl (e.g. Anas and Rhee, 2006) demonstrate that the total miles driven can actually decrease with sprawl as jobs can move closer to workers during the decentralization process. Anas and Pines (2008) have shown that pricing congestion can cause population to spread from larger to smaller cities reducing total congestion, while increasing the developed land area which corresponds to more sprawl.

Emissions and fuel use are curbed significantly not only by reductions in the distances traveled and in the number of trips made, but also by improving the speed of travel, which in turn is determined by the amount of road capacity available to accommodate the demand. Thus any policies which can improve the speed of travel in large and highly congested cities could be very beneficial in reducing fuel use and emissions, while raising tax revenues that can be used in a variety of complimentary ways such as adding mass transit capacity or subsidizing high density developments near mass transit lines. Beevers and Carslaw (2005) have studied by means of simulation tests whether the congestion charging scheme implemented in central London in 2003 has resulted in significant speed improvements of about 2.1 km/hour.[3]

We report on results from a highly aggregated model of mode choice in commuting representing Sao Paulo in 2002.[4] In that same year, the Brazilian flex-fuel vehicles that run on a mixture of alchohol and gasoline were introduced but these did not have a significant market share until later. In 2003, a year after the date of our data, flex-fuel vehicles made up only 3.7% of the light vehicle market. Therefore, our study is for conditions just before the introduction of these vehicles. In the model, the total number of trips per day is fixed as is also the representative (average) trip distance. The limited data does not allow us to develop a more sophisticated model such as the one for Beijing (see Anas, Timilsina, Zheng, 2009), where the substitutions between commuting and non-commuting (discretionary) travel and between travel and housing consumption were also included and the number of non-work trips was endogenous.

---

[3] Ultimately, reductions are also driven by fuel technology, and driving simulations of electric cars and hybrids show reductions in CO_2 emissions as documented by Saitoh et al (2005).

[4] Source: http://www.metro.sp.gov.br/empresa/pesquisas/afericao_da_pesquisa/afericao_da_pesquisa.shtml and related websites.

Our model incorporates a statistically verified relationship between car speed (km/hr) and fuel use per km, and the U.S.E.P.A.'s formula for converting fuel to grams of CO_2 per km. We use the model to compare the effects of: (1) a socially optimal toll per kilometer of car travel that internalizes the excess delays and excess fuel use caused by traffic congestion; (2) a toll on the excess delay externality only; and (3) an excise tax on the gasoline used by cars that raises the same revenue as the toll on excess delay only, but reduces fuel and emissions by about 1.6 times more. Having no data on average incomes by income group in Sao Paulo, we are limited to linear-in-income, trip-based indirect utility function. The value of time in commuting is made to increase by income group and is assumed proportional to the wage rate, while the sensitivity with respect to monetary cost decreases with income. Commuters' choices are described by a trinomial logit model of choice between: (i) private car; (ii) bus and other public transit; (iii) and all non-motorized modes such as walking and bicycling. Under these assumptions, we calibrate the model so that the elasticities of demand for travel by car with respect to travel cost and travel time are reasonable, while the mode-specific constants of the logit model are set to replicate the modal shares observed in 2002.

Because congestion affects speed directly and fuel use indirectly, two congestion externalities arise. First, the marginal trip imposes a time delay on other travelers; and second, by slowing them down, it causes them to use more fuel per trip. We use the model to calculate the socially optimal congestion toll, internalizing both externalities. The optimal toll dramatically raises the monetary cost of a car trip by 5.82 times and cuts CO_2 by 71%. A partial toll that internalizes only the time delay externality is also computed. We show that this partial toll is 59.2% of the optimal, achieves 75% of the fuel and CO_2 reduction, 90% of the revenue and 86.4% of the welfare gains of the optimal toll, the remaining 13.6% of the welfare gains being due to the fuel externality. These results are subjected to extensive sensitivity testing by varying the congestion technology, the within income group idiosyncratic taste heterogeneity and a coefficient that is used to scale the values of time up or down. Ignoring only the most extreme values in these coefficients, the optimal congestion toll reduces CO_2 by 60%-87%, and the partial toll internalizing only the delay externality reduces CO_2 by 40%-74%. Aggregate welfare, the sum of consumer surpluses and toll revenue, improves significantly. The consumer surplus of higher incomes improves because they value highly the time savings, while that of lower incomes decreases.

The paper is organized as follows. Section 2 presents the simulation model we developed for this study followed by presentation of data and key model parameters in section 3. This is followed by discussion of key simulation results in section 4. These results show that investing in highways, subsidizing transit or improving the fuel economy of cars are not as effective at reducing emissions as is taxing car travel. Then, section 5 discusses results of the policy simulations demonstrating the effectiveness of taxing car travel. In section 6, we present and discuss sensitivity tests on the key parameters, showing that the results on the emission reduction potential of taxing travel are quite robust. In section 7, we examine the presence of a "lock-in effect" due to highway investment. A lock-in effect exists if increasing road capacity reduces the effectiveness of improving transit on aggregate emissions. We do show that there is a lock-in effect and that it is indeed quantitatively significant reducing the effectiveness of a 20% transit travel time improvement on emissions by about one half, if road capacity is also expanded by 20%. Finally, Section 8 concludes the paper with some policy recommendations.

2. The model

We will treat the trip choice behavior of six income groups denoted by the subscript $f = 1,...,6$ and ordered in increasing income. The disutility function for choosing mode *m* is assumed to be of the form,

$$U_{fm} = -g_m - b_f \ln\left(G_m\right) + E_{fm} + u_{fm}, \tag{1}$$

where g_m is the monetary cost of travel by mode *m* (*m=1,2,3*), where *m=3* is auto, *m=2* is public bus or transit and *m=1* is non-motorized transport (bicycles, walking etc.). G_m is travel time by mode *m* and E_{fm} is the mode-specific systematic utility constant. The coefficient $b_f \equiv \frac{w_f}{2}$ (one half the wage rate) and reflects that higher income commuters are more sensitive to travel time changes than are lower income consumers. Thus, since the marginal utility of monetary travel cost is unity across all income groups, the marginal rate of substitution between travel time and monetary cost (or income) is $\frac{b_f}{G_m}$, which increases with the wage rate for the same travel time. Finally, u_{fm} is the idiosyncratic utility by mode *m* for a particular commuter of income group *f*. Utilizing the usual distributional assumptions about idiosyncratic utilities (namely, i.i.d. Gumbel), the mode choices of trips are then described by the following trinomial logit model,

$$P_{fm} = \frac{\exp\left[\lambda_f\left(-g_m - b_f\ln(G_m) + E_{fm}\right)\right]}{\sum_{n=1}^{3}\exp\left[\lambda_f\left(-g_n - b_f\ln(G_n) + E_{fn}\right)\right]}, \quad \sum_{n=1}^{3} P_{fn} = 1. \tag{2}$$

In (2), λ_f is the heterogeneity or (taste dispersion) coefficient of group *f* and is inversely proportional to the identical variance of the idiosyncratic utilities u_{f1}, u_{f2}, u_{f3}.

From (2) we can also calculate that the elasticity of choosing mode *m* with respect to monetary travel cost g_m and travel time G_m are respectively:

$$\eta_{P_m:g_m} = -\lambda_f g_m (1 - P_{fm}), \tag{3a}$$

$$\eta_{P_m:G_m} = -\lambda_f b_f (1 - P_{fm}). \tag{3b}$$

Taking the auto mode (*m=3*) as an example, and keeping λ_f the same by *f*, because the probability of choosing auto P_{f3} increases with *f*, the monetary cost elasticity decreases with income (i.e. lower income groups should be more sensitive to monetary cost since they have lower budgets) while because the wage rate rises with *f* (higher income groups should be more sensitive to travel time, since they have higher values of time).

The person trips by mode are calculated as,

$$T_m = \sum_{f=1}^{6} N_f P_{fm}, \quad m = 1, 2, 3. \tag{4}$$

In (4), N_f are the total number of person trips per day (commuting and non-commuting) by income group *f*. These daily person trips remain fixed in the model. Then, the sum of motorized vehicle traffic in car-equivalent units, is obtained as:

$$T = h\phi_2 a T_2 + \phi_3 T_3, \tag{5}$$

where ϕ_2, ϕ_3 are the inverse ratios of vehicle occupancies (car-equivalent buses per bus person-trip, and fractional cars per person-trip by car respectively) which are used to convert person trips to car equivalent vehicular trips after multiplying by *h* which is the car-equivalent capacity load of a whole bus. *a*=0.5 is the fixed share of bus in the public transit mode. Given an aggregate road capacity, *Z*, the congested round trip travel time per trip is then calculated from the Bureau of Public Roads type of congestion function:

$$G_3 = c_0\left(1 + c_1\left(\frac{T}{Z}\right)^{c_2}\right)d_3, \tag{6}$$

where d_3 is the given round-trip distance by car in kilometers that remains fixed in the model. Car speed then is calculated as $\hat{s}$:

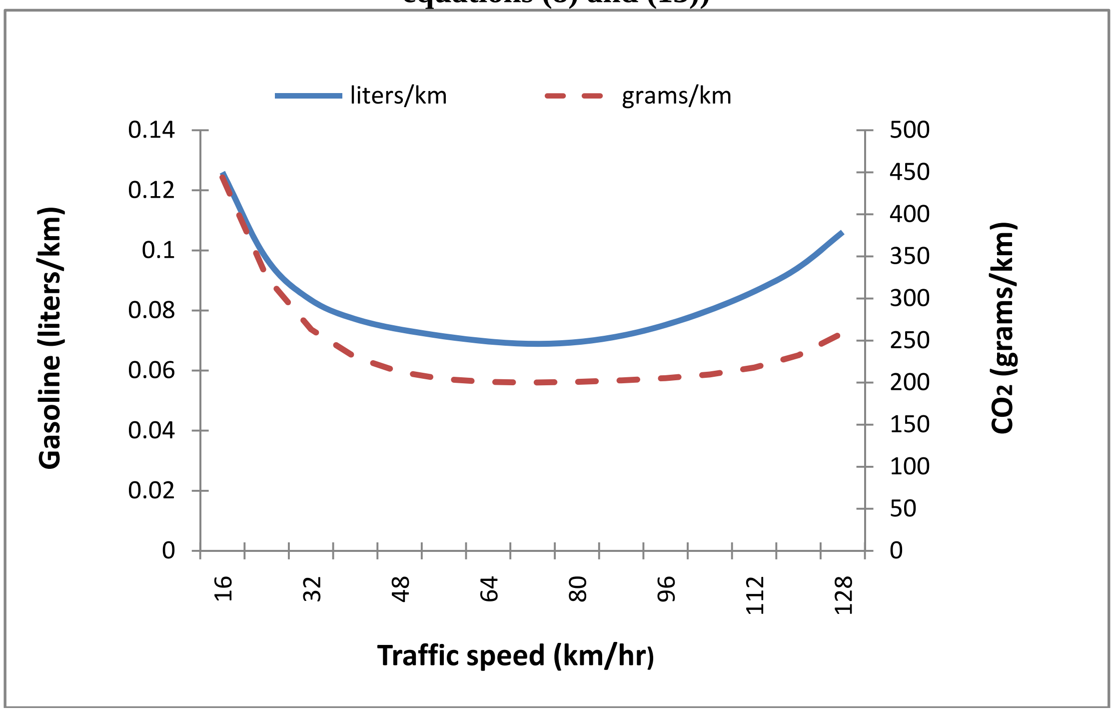
$$\hat{s} = \frac{d_3}{G_3} = \left\{c_0\left(1 + c_1\left(\frac{T}{Z}\right)^{c_2}\right)\right\}^{-1}. \tag{7}$$

The fuel expenditure in liters of gasoline per kilometer (assuming a vehicle efficiency of unity) is calculated from a polynomial fit to the following statistically verified equation reported by Davis and Diegel (2005) for a Geo Prizm (see Figure 1)[5],

$$f(\hat{s}) = (3.78541178/1.6093)\times\ [0.122619 - 0.0117211\times(\hat{s}) + 0.0006413\times(\hat{s})^2$$

$$-0.000018732\times(\hat{s})^3 + 0.0000003\times(\hat{s})^4 - 0.0000000024718\times(\hat{s})^5 + 0.000000000008233\times(\hat{s})^6\ ].\ (8)$$

FIGURE 1: Fuel consumed per car-km and grams of emissions per car-km. (Plots of equations (8) and (13))



[5] Davis and Diegel (2004) calculates fuel use in gallons/mile from speed in miles/hour for nine car brands. We converted the polynomial equation, that we fitted to their Geo Prizm curve, to the liters/km version by making the three required adjustments shown in (8). First, the speed in kilometers/ hour is divided by 1.6903 km/mile in order to get the speed in miles/hour. This is then used in the original equation to predict gas consumption in gallons/mile. Then, the result is multiplied by 3.785 liters/gallon to get fuel use in liters/mile and, lastly, that result is divided by 1.6903 to get the fuel use in liters/km.

The monetary fuel cost per passenger (including fuel taxes) is next calculated from,

$$FUELC = \phi_3(1+\tau_F)\times p_F \times e \times f(\hat{s})\times d_3, \tag{9}$$

where e is the vehicle efficiency, τ_F is the fuel excise tax rate and p_F is the price of gasoline.[6] Let *MTOLL* be a toll charged on each vehicle occupant (i.e. per person-trip) , per kilometer of travel, then the monetary cost of a car trip will be:

$$g_3 = FUELC + \delta \times MTOLL, \tag{10}$$

where $\delta = 1$ if a toll is levied, and $\delta = 0$ if no toll is levied. Tolls that improve economic efficiency can be calculated in a variety of ways. One way, for example, is to calculate a Pigouvian toll that internalizes the time delays car commuters impose on one another but ignores the fact that by slowing each other down, car commuters also decrease their travel speed, thus inducing higher fuel consumption. Such a toll per person-trip by car, levied only on the excess delay from congestion, would be calculated as

$$MTOLL = \phi_3(VOT)\times \underbrace{c_0 c_1 c_2\left(\frac{T}{Z}\right)^{c_2} d_3}_{=\left(T\frac{\partial G_3}{\partial T}+G_3\right)-G_3}, \tag{11}$$

where *VOT* is the (trip-weighted) average value of time of the commuters on the road and the remaining terms are the gap between the aggregate delay caused by one commuter $\left(T\frac{\partial G_3}{\partial T}+G_3\right)$, and the average travel time, G_3, of that commuter. For the utility function we are using, the *VOT* would be calculated as:

$$VOT = \frac{\sum_{f=1}^{6} N_f P_{f3}\left(\frac{b_f}{G_3}\right)}{\sum_{f=1}^{6} N_f P_{f3}} \tag{12}$$

[6] The monetary cost of travel depends also on the car's fuel inefficiency level. However, we have formulated the model as if everyone uses a standard efficiency vehicle since we could find no data on how car fuel inefficiency varied by income in Sao Paulo (or Brazil, more generally) . From Davis and Diegel (2004), a fuel efficiency of unity (e=1 in equation (9)) corresponds very closely to their curve for a Geo Prizm.

The second curve in Figure 1 is the CO_2 emissions in grams/km of car travel by taking the exponential of a polynomial equation that predicts log-CO_2 as a function of the speed in miles per hours (Barth and Boriboonsomsin, 2007). The divergence between the fuel and CO_2 curves occurs because the latter was calculated under cruising conditions in Southern California, whereas the curve we fit To Davis and Diegel's data for the GeoPrizm reflects actual conditions. Since cruising speeds are not realistic for Sao Paulo, we use the U.S.E.P.A.'s formula for converting a gallon of gasoline to kilograms of CO_2.[7]

CO2 grams/gallon = 2,421 grams x 0.99 x (44/12) = 8,788 grams/gallon. (13)

The above completes description of our model's demand structure and the equations that are used to calculate fuel use and emissions. We now turn to how a congested traffic equilibrium is calculated using the above equations.

A congested traffic equilibrium is calculated by the following iterative procedure which, if iterated properly and starting from a reasonable starting point, converges robustly and uniquely. The steps of the procedure are:

1. Set the value of the aggregate car equivalent traffic, T, and call it $\hat{T}$.
2. Calculate the auto travel time, G_3, from the congestion function (6).
3. Calculate or input any fuel tax rates and tolls, and then calculate the monetary cost of auto, g_3, using (9), (10) (and (11), (12) substituted into (9) in the case of a toll on excess delay only).[8]
4. Calculate the choice probabilities P_2, P_3, from (2).
5. Calculate T_2, T_3, from (4) and T from (5).
6. (a) If T and $\hat{T}$ (from steps 1 and 5) are close enough, stop and declare convergence. The criterion used for convergence is that $\frac{\left|\hat{T}-T\right|}{\left(\hat{T}+T\right)/2} < 1\times10^{-8}$.

[7] http://www.epa.gov/oms/climate/420f05001.htm#calculating. To calculate the CO_2 emissions from a gallon of fuel, the carbon emissions are multiplied by the ratio of the molecular weight of CO_2 (m.w. 44) to the molecular weight of carbon (m.w.12): 44/12. In our case, a further multiplication by 0.75 is required since we assume that 25% of the fuel was alcohol.

[8] In the case of the socially optimal toll, (11) and (12) are not used. Instead the value of *MTOLL* is varied and a congested equilibrium is calculated for each *MTOL*L until we find the value which maximizes social welfare (equation (15)).

(b) If $\hat{T}$ from step 1 is not sufficiently close to the T calculated in step 5, continue iterating by returning to step 1, with T from step 5 replacing $\hat{T}$ in step 1.

7. After convergence, calculate all outputs such as person trips by mode, vehicle kilometers traveled by car, total fuel consumption, total carbon emissions, total revenue from tolls or fuel taxes etc. and the aggregate value of consumer surplus and social welfare (see below).

Calculation of most model outputs is straightforward, but important outputs of the model are the per-capita consumer disutility measure for each income group and the aggregate social welfare measure. Note from equation (1) that the marginal utility of income of each income group is constant and equal to unity. Then, as it is well known from the properties of the logit model, the expected maximum utility (or expected least disutility) of a consumer in income group f is REAL valued and given by:

$$E\left[\max\left(U_{f1},U_{f2},U_{f3}\right)\right]=\frac{1}{\lambda_f}\ln\left(\sum_{m=1}^{3}e^{\lambda_f\left(-g_m-b_f\ln\left(G_{fm}\right)+E_{fm}\right)}\right). \tag{14}$$

Aggregate social welfare is calculated by summing all these BRL-valued consumer surpluses and adding the aggregate tax revenues:

$$\begin{pmatrix} Social \\ Welfare \end{pmatrix}=\left\{\sum_{f=1}^{6}\frac{N_f}{\lambda_f}\ln\left(\sum_{m=1}^{3}e^{\lambda_f\left(-g_m-b_f\ln\left(G_{fm}\right)+E_{fm}\right)}\right)\right\}+\left\{\begin{matrix} Aggregate \\ Tax\ Revenue \end{matrix}\right\}. \tag{15}$$

3. Data and calibration

Our data approximate 2002 conditions.[9] The geographic scope of São Paulo in our study is the entire metropolitan area, the highest level of aggregation possible. In 2002, the area had a population of 18,345,000 and 7,983,000 jobs (43.5% of the population), generating 7,615,000 commutes per day and 11,716,000 non-work trips per day for a daily total of 19,330,000 trips. In 2002, the typical car fuel used in Brazil was E25 gasohol (a mixture of 25% alcohol and 75%

[9] The source of the data on trips, wages and travel times by mode and income group is the official website of the city of São Paulo metropolitan planning agency, Compania do Metropolitano de São Paulo (Metro), in particular http://www.metro.sp.gov.br/empresa/pesquisas/afericao_da_pesquisa/afericao_da_pesquisa.shtml and related websites.

gasoline). The average prices of alcohol and gasoline were 0.89 and 1.70 BRL/liter respectively, resulting in a price for E25 gasohol of 1.50 BRL/liter. The congestion function coefficients were set as $c_0 = 1/80$, $c_1 = 0.15$, $c_2 = 2.0$. Aggregate highway capacity, *Z*, was calibrated in such a way that, given the 2002 car-equivalent traffic from (4) and (5), the car speed from (7) is the 24.10 km/hr reported for 2002.[10] The trips were distributed among the six income groups and the three modes, according to the data shares in Table 2. Tables 1 and 2 show all the basic data that was used in calibrating the model.

In the trinomial logit (2), the dispersion parameter, λ, was set to 0.25, the coefficient *b* in the value of time was set at 0.5, and the mode-specific constants for walking set to zero, while set to replicate the public transit and car mode shares for *f* in Table 2. This calibration resulted in car choice elasticity with respect to travel time and monetary cost shown in Table 3, the former elasticity increasing with income, the latter decreasing.

TABLE 1: Basic data and coefficients for the modes of travel

	Walk+ m=1	**Bus**	**Rail**	**Public transit** m=2	**Car** m=3
Average trip length (2 way km), d_n [a]	2.56	35.43	n/a	n/a	20.73
Trip times (2 way hrs.), G_n [a]	0.54	2.10	1.50	1.80 [d]	0.86
Speed (km/hr), $\hat{s}_3, \hat{s}_{bus} = 0.7\hat{s}_3$ [b]	4.83	16.87	n/a	n/a	24.10
Fuel expenditure, $f(\hat{s}_3)e$, miles/gallon	n/a	n/a	n/a	n/a	24.18
Fuel price, (BRL/liter), p_F (E25 gasohol: 25% alcohol at 0.89, 75% gasoline at 1.70 BRL/liter) [c]	n/a	n/a	n/a	n/a	1.50
Persons per vehicle, ϕ_n^{-1} [b]	n/a	37.36	n/a	n/a	1.75
Traffic load per vehicle, *h*	n/a	4 [e]	n/a	n/a	1.00
Average monetary cost of trip per worker (BRL/2-way trip), g_n	0	n/a	n/a	0.77 [b]	1.73

[10] Source: International Association of Public Transport (2007).

Notes: [a] Survey of the Companía do Metropolitano de Sao Paulo (Metro); [b] Millenium Cities Data Base. IAPT (2007); [c] The World Bank; [d] Public transit trip time, G_2 =0.5(2.10)+0.5(1.5); [e] Crude guess.

TABLE 2: Shares of trips and modes choices by income group

Income group		1	2	3	4	5	6	*Total or Average*
Shares of person trips per day, N_f [a]		0.17	0.25	0.27	0.19	0.09	0.03	1.00
Wage (BRL/hr), w_f [a]		1.8	3.6	7.2	13.8	27.0	48.0	9.75
Mode shares, P_{fm} [a]	*Walk*	*0.600*	*0.500*	*0.400*	*0.250*	*0.170*	*0.150*	0.402
	Public	*0.300*	*0.345*	*0.330*	*0.300*	*0.207*	*0.102*	0.305
	Auto	*0.100*	*0.155*	*0.270*	*0.450*	*0.623*	*0.748*	0.293
	Total	1.000	1.000	1.000	1.000	1.000	1.000	1.000

(Notes: [a] Survey of the Companía do Metropolitano de Sao Paulo (Metro))

To achieve a reasonable calibration of the mode choice probabilities, the dispersion parameters, λ_f , are all set to 0.25 and the mode-specific constants for the non-motorized mode in the utility function were set to zero ($E_{f1} = 0$), while E_{f2}, E_{f3} are set to replicate the mode shares for *f* shown in Table 2. This calibration resulted in car choice demand elasticities with respect to monetary travel cost and travel time that are shown in Table 3. As discussed earlier and now seen from Table 3, the model has been set up in such a way that the elasticity of the probability of choosing the car mode with respect to travel time (equation (3b)) increases with income, while the elasticity with respect to monetary travel cost (equation (3a)) decreases with income.

TABLE 3: Calibrated own and cross elasticity of mode choice with respect to monetary travel cost and travel time of a trip by car

Income group		1	2	3	4	5	6
Monetary cost (equations 3a)	*Own*	-0.39	-0.36	-0.32	-0.24	-0.16	-0.11
	Cross-bus/walk	0.04	0.07	0.12	0.19	0.27	0.32

Travel time (equations 3b)	*Own*	-0.20	-0.38	-0.66	-0.95	-1.27	-1.51
	Cross-bus/walk	0.03	0.07	0.24	0.78	2.10	4.49

4. Considering policies that reduce aggregate fuel use and emissions

The model ignored some important aspects because data were lacking. Although mode switching is treated, the total trips, and the trip distance by mode are fixed. Car fuel economy is uniform in the population and although fuel costs are a function of speed, the cost of car acquisition or non-fuel variable costs are not treated. Still, the model has enough detail to help us learn about how pricing policies affect congestion and emissions.

In Figure 1 where e=1, fuel and emissions per car km are high when speed is low, but each falls rapidly with speed, making relatively flat bottoms and rising again at higher speeds. Congestion mitigation policies would yield a high return in megacities where the speed is very low. A policy may improve emissions per car km (the intensive margin), but if car trips are increased by the policy (the extensive margin), then emissions also increase. We now examine this in the case of four policies treatable in the model.

(i) *Building more* roads: Increasing road capacity Z, reduces congestion. But as the speed increases, walkers switch to car and bus. The higher speed reduces fuel and emissions per car km. Whether aggregate fuel or emissions are reduced depends on the elasticity of car trips with respect to travel time. If this elasticity is sufficiently high (low), then trips by car and aggregate fuel and emissions increase (decrease). The elasticity shown in Table 3 is low enough that a 20% increase in base capacity cuts emissions by 0.95%.

(ii) *Subsidizing public transit*: A lower monetary cost of bus induces switching to bus and a higher car speed, but if a lot more trips switched to bus from walking than from cars, then speed can decrease. What happens to the aggregate fuel and emissions depends on the elasticity of the trips by bus with respect to the monetary cost of bus. In the present model, by the I.I.A.

property of logit equal percentages of car users and walkers switch to bus, but the number switching also depends on how many walkers versus how many car users there are to begin with. Reducing our base transit money cost by 20% reduces emissions by only 0.40%.

(iii) *Relaxing vehicle efficiency standards*: Making cars more fuel efficient by lowering e, keeping car ownership costs unchanged, reduces the per km cost, inducing switches to car which raises congestion, reducing speed. If the elasticity with respect to monetary cost is high (low) enough, aggregate fuel and emissions increase (decrease) as e is reduced. In our base case, improving e by 20% increases emissions by 5.4%, whereas worsening fuel economy cuts emissions by 4.98%.

(iv) *Reducing congestion by taxing car travel*: Taxing each car km, or gasoline per liter keeping e constant, induces switches to non-car modes, lowering congestion and increasing speed. Aggregate fuel and emissions are reduced unambiguously. An exception, unlikely in highly congested megacities, occurs only at very high speeds when the fuel and emissions curves rise (Figure 1). The effectiveness of taxing car travel is demonstrated next.

5. Comparing congestion pricing policies

Table 4 shows the results of the simulations with the calibrated (base) parameters. Column 1 is the base equilibrium in 2002, in which fuel taxes or congestion tolls are not used. Column 2 shows what would occur if an optimal toll were imposed that internalized both the excess congestion delay and the excess fuel consumption externality. This toll is specified as a charge per kilometer of car travel per person trip, and its value in (8) is varied until the value which maximizes social welfare (12) is found. Column 3 shows the results of the toll specified in (13a), (13b) that aims to internalize only the time delay externality of congestion. Column 4 shows the percentage of the improvement of the socially optimal toll that is captured by this partial toll internalizing only the congestion delay.

Fuel use per trip, the monetary cost of fuel and emissions are all reduced, as the tolls add a lot more than the fuel cost reduction. The socially optimal toll is 10.07 BRL/trip, but the fuel cost per trip falls to 1.29, thus the monetary cost at the optimum is 11.36 or 6.58 times its value without tolling. Under the toll on congestion delay only, the toll is 5.96 and the fuel cost 1.37 per trip. Thus, the monetary cost of 7.33 per car trip is 4.24 times the base. These large after-tax monetary costs cause significant substitution away from auto and in favor of the other modes. The optimal toll and the toll on congestion delay decrease emissions by a dramatic 71% and 53%, respectively. Figure 2 shows how speed increases and CO_2 decreases as the toll per km is increased. In the case of the optimal toll it takes a 567% increase in the monetary cost per km, to reduce CO_2 by 75%, a reduction of 0.13%, on average, for every 1% increase in the monetary cost per km. The toll on the delay externality only, achieves a 0.18% reduction on average, for every 1% increase in monetary cost. The toll internalizing the congestion delay is 59.2% of the optimal toll but captures 86.4% of the optimal welfare gain, and 77.3% of the optimal emission reduction, while raising 90% of the optimal toll's revenue. We also calculated that a fuel tax which raises the same revenue as the toll on excess delay (but is on the right side of the optimal toll, i.e. above optimal) reduces CO_2 by 1.6 times more than does the toll on excess delay only. This is not shown in the table.

For each policy, Table 4 shows the consumer surplus per trip for each group, as well as its trip weighted average over all groups. Tolling reduces the consumer surplus of the lower incomes, since they do not value time savings sufficiently after paying tolls. The two groups with the highest incomes, sufficiently value time savings for their consumer surplus to increase after paying the toll. The table also shows the social welfare, (12), divided by total trips and the toll revenue per day. Figure 3 plots social welfare per trip, average consumer surplus, and aggregate toll revenue against the toll per km. These three curves do not all peak at the same toll. The revenue from tolling is maximized at a toll of 9.52 BRL/trip, while the aggregate (or average) consumer surplus is maximized at a toll of 12.09 BRL/trip. Social welfare, the sum of total revenue and aggregate consumer surplus, peaks at a toll of 10.07 BRL/trip. The curves rise relatively steeply, but after peaking, they decline mildly. These properties insure that for any revenue level in Figure 3 that is below the peak revenue, the higher-than-optimal toll gives higher consumer surplus and social welfare than does the lower-than-optimal toll that raises the same revenue. Since CO_2 continually decreases with the toll (Figure 2), the higher-than-optimal toll also corresponds to lower CO_2. Therefore, if one had to choose between the two revenue neutral tolls, the higher-than-optimal toll is to be socially preferred. In the context of Figure 3, if we wanted to approach the social welfare maximizing optimal toll gradually by increasing the targeted revenue in steps, then the present value of social welfare would be maximized by starting at the higher-than-optimal toll and then gradually decreasing it towards its optimal value by following in reverse the falling part of the social welfare curve. Starting from a low toll and then gradually increasing it towards the optimum by following the rising part of the social welfare curve would yield a lower present value for social welfare.[11] From Table 4 we also note that the aggregate emissions predicted by the base equilibrium (without pricing policies) is 5.83 million tons per year (obtained by multiplying the daily prediction by 300 assumed average days of commuting) which is 78% of the 7.5 million tons per year reported. The difference we presume is from the diesel used in non-car vehicles (e.g. trucks etc) but we are unable to test this since we had no data on such vehicles.

[11] This fact that revenue peaks at a lower toll than does social welfare is a result of the presence of the bus mode that we assume is not tolled and the presence of the non-motorized modes. The presence of these modes cause car choices to be reduced more elastically, thus aggregate toll revenue reaches a maximum quickly, whereas the switching to the other modes continues to be welfare improving.

TABLE 4: The optimal toll and efficiencies captured by the toll on delay only

	2002 Base	**Optimal toll on excess delay and excess fuel** ($0.486\ BRL/pass.km.$ $or\ 6.679\ BRL/liter$)	**Toll on excess delay only** ($0.288\ BRL/pass.km.$ $or\ 3.723\ BRL/liter$)	**Percent of optimal change from base captured by toll on delay only**
Speed (km/hr)	24.10	51.27	39.89	58.1
Auto person trips ($\times 10^6$)	5.6745	2.2096	3.3528	67.0
Bus person trips ($\times 10^6$)	5.8861	9.0790	7.9720	65.0
Auto vehicle kil. ($\times 10^8$)	1.1761	0.4580	0.6949	67.0
Auto hours ($\times 10^6$)	4.8801	0.8933	1.7423	78.7
Auto fuel (liters) ($\times 10^7$)	1.1431	0.3331	0.5371	74.6
Auto CO_2 mill. tons/year	5.8330	1.6997	2.7407	74.6
Car equiv. traffic ($\times 10^6$)	3.5577	1.7487	2.3427	67.2
Fuel cost/trip (BRL)	1.7272	1.2928	1.3735	81.4
Toll/trip (BRL)	0	10.07	5.9636	59.2
Consumer surplus/trip	6.9595	7.2217	7.1459	71.1
f=1	2.5979	2.3469	2.3819	86.0
f=2	3.8817	3.6057	3.6345	89.5
f=3	5.8834	5.5607	5.5977	88.5
f=4	9.7969	9.8001	9.8151	105.7
f=5	15.4063	17.5192	17.1637	83.2
f=6	22.3769	30.4549	28.3891	74.4
Social welfare./trip	6.9595	8.3728	8.1803	86.4
Revenue /day ($\times 10^7$)	0	2.2251	1.9995	89.9

FIGURE 2: Car speed and emissions under tolling

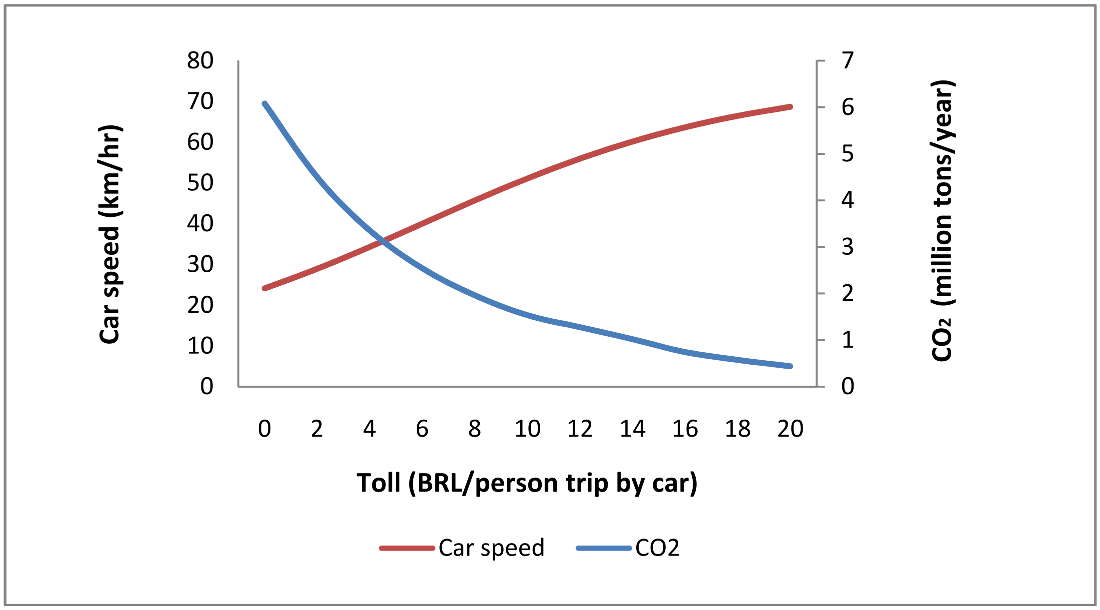
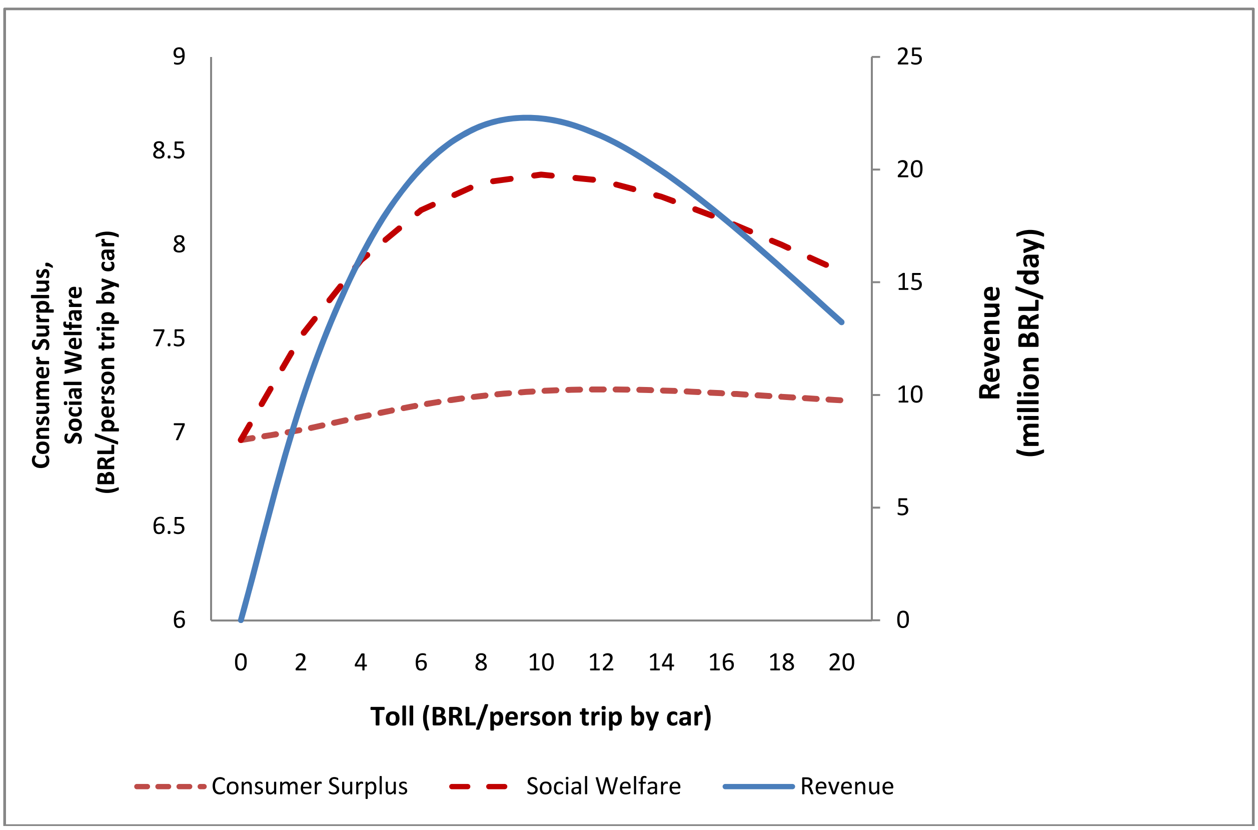


FIGURE 3: Consumer surplus, social welfare and revenue as functions of the congestion toll



6. Sensitivity analysis

We now want to see how robust the response of aggregate emissions to congestion pricing is under widely different parameter values. For this purpose, the most important parameters of the model are selected. These are: (i) c_2, exponent of the congestion technology, (6), which controls the sensitivity of traffic speed to the ratio of traffic to road capacity. The severity of the congestion externalities increases with c_2; (ii) the idiosyncratic taste heterogeneity parameter λ. From (3a), (3b), this parameter controls the sensitivity of the demands for mode to the monetary cost and to travel time. By decreasing λ toward zero, the sensitivity is reduced and demands become inelastic; (iii) the parameters $b_f = \frac{w_f}{2}$ controlling the increase in the values of time, $\frac{b_f}{G_m}$, with wage, are changed together for all f by changing the divisors from 2 to some other number. By reducing the divisor, travel time becomes more important in mode choice; (iv) increasing λ while changing $b_f = \frac{w_f}{2}$ to keep λb_f constant has the effect of increasing sensitivity to monetary cost relative to the sensitivity to travel time.

6.1 Congestion technology (c_2)

When c_2, congestion increases, speed decreases and, consequently, trips by car decrease. Since the lower speed works to raise emissions per kilometer, while the fewer trips by car work to decrease emissions, aggregate emissions can change non-monotonically with c_2, but have a decreasing trend in Table 5. This is interesting because it shows that a city with $c_2 = 0$ has no congestion but more emissions than an otherwise identical city with $c_2 = 4$. In the latter case, the higher road congestion induces more walking and transit trips causing lower aggregate emissions, even though emissions per car-km increase. Conversely, the emission reductions per car-km from the higher speed in the uncongested city are more than offset by the added emissions from more car trips. With higher c_2, the gap between the marginal social and average private cost of travel gets wider and a higher toll is required to internalize the externality due to the time delay or the time delay and the fuel use externalities together (optimum toll). Tolls increase and CO_2 decreases monotonically with c_2. CO_2 reductions from tolling increase with c_2 (reaching 88% when $c_2 = 4$). These results show that tolling is much more effective in cities with poor congestion technology. As c_2 increases, tolling only for the delay externality captures an increasing part of the optimal improvement in emissions (reaching 92.6% when c_2 =4), because with higher c_2, the time delay externality becomes large relative to the fuel cost externality.

TABLE 5: Effect of congestion technology changes on CO_2 reductions

Congestion severity δ	**No Congestion Tolling**			**Optimum Tolling**		**Tolling Delay Externality**	
	CO_2 ($mil.tons$)	**Speed** (km/hr)	**Auto Trips** ($\times 10^6$)	**% CO_2 Reduction**	**Toll** (BRL/trip)	**% CO_2 Reduction**	**Toll** (BRL/trip)
0.0	5.7492	69.56	5.674	0	0.00	0	0.00
1.0	5.4562	44.70	4.072	33.38	3.56	22.24	2.18
2.0	5.8830	24.10	3.2353	75.72	10.07	57.99	5.96
3.0	5.5230	17.00	4.0716	84.78	14.16	74.53	9.68
4.0	5.0663	13.97	3.2353	87.83	16.55	81.34	12.52

6.2 Taste heterogeneity (λ)

λ is the inverse of the variance of the idiosyncratic tastes within each income group. From (3a)-(3b) when $\lambda = 0$, all own and cross elasticity is zero. The idiosyncratic tastes are so large that choices are random and each mode has one third share. The demand for car is perfectly inelastic, so tolls cannot change behavior and cannot improve aggregate welfare because the decrease in consumer surplus caused by the toll is exactly offset by the toll revenue. The case of $\lambda = 0.001$ in Table 6 comes very close to $\lambda = 0$. The optimal toll reduces emissions by only 1.6%. As λ is increased, idiosyncratic heterogeneity decreases and choices become increasingly sensitive to monetary and time cost differences between the modes, making tolls more effective in reducing congestion and emissions. For example, when $\lambda = 1.25$ or 1.50, the optimal toll reduces emissions by 81%. At even higher λ, idiosyncratic variation diminishes and choices are concentrated on a single mode. When $\lambda \geq 10$ the lowest three income groups all walk, and the highest three all choose car. No one chooses public transit. To induce efficient changes in this extreme situation, the toll begins to rise with λ.

TABLE 6: Effect of taste heterogeneity changes on CO_2 reductions

Taste heterogeneity λ	No Congestion Tolling			Optimum Tolling		Tolling Delay Externality	
	CO_2 ($mil.tons$)	**Speed** (km/hr)	**Auto Trips** ($\times 10^6$)	**% CO_2 Reduction**	**Toll** (BRL/trip)	**% CO_2 Reduction**	**Toll** (BRL/trip)
0.001	7.6549	20.18	6.4348	1.62	12.21	0.56	4.17
0.01	7.4727	20.51	6.3600	14.35	11.89	5.27	4.18
0.15	6.1206	23.41	5.7731	68.66	10.21	45.86	5.16
0.25	5.8830	24.10	5.6744	75.72	10.07	57.99	5.96
0.50	5.8369	24.49	5.6994	79.48	9.83	67.93	6.67
0.75	5.9411	24.45	5.7950	80.83	9.79	71.14	6.85
1.00	6.0441	24.38	5.8821	81.29	9.82	72.79	6.93
1.25	6.1191	24.34	5.9473	81.44	9.86	73.72	6.97
1.50	6.1631	24.35	5.9921	81.47	9.90	74.21	6.99
5.00	5.9513	25.40	5.9756	81.40	10.30	73.56	7.07
10.00	5.9204	25.54	5.9691	81.61	10.41	73.40	7.0988

6.3 The value of time ($b_f G_m^{-1}$)

Table 7 shows that as b_f increases, car trips increase initially, decreasing after b_f = $0.75w_f$. Correspondingly, speed decreases as car trips increase and then begins to increase. Because the speed reduction is dominated by the initial increase in car trips, emissions roughly follow the car trips, rising initially and then falling. Why do car trips first rise and then fall, as the value of time increases? Since public transit is the slowest mode, increasing the value of time from zero induces marginal transit users to switch to car, and more so the higher their wage. At the same time, walking is faster than car on average (not because of speed but because those who walk choose shorter distances to their destinations). Hence, increasing the value of time induces marginal car users to switch to walking arrangements while marginal transit riders switch to walking or to car. Initially as b_f increases, more switch from transit to car than from car to walking and car trips increase, causing more congestion and emissions. As b_f increases more there is a reversal and more switch from car to walking than from transit to car, and car trips and emissions decrease. Looking at policy, when b_f=0, no one cares about time and the delay externality vanishes. The toll is 0.48 BRL/trip and is due to the fuel externality. As b_f increases, the delay externality grows and the optimal toll or the toll monetizing the delay increase dramatically. In all cases except b_f=0, optimal tolling reduces base emissions 63% to 76%. The reduction captured by monetizing the delay rises with b_f, reaching 90% when b_f =6.

TABLE 7: Effect of increasing the value of time on CO_2 reductions

Value of time $b_f G_m^{-1}$	**No Congestion Tolling**			**Optimum Tolling**		**Tolling Delay Externality**	
	CO_2 ($mil.tons$)	**Speed** (km/hr)	**Auto Trips** ($\times 10^6$)	**% CO_2 Reduction**	**Toll** (BRL/trip)	**% CO_2 Reduction**	**Toll** (BRL/trip)
$w_f \times 0.0$	3.4166	32.0	4.0570	12.43	0.48	0	0.00
$w_f \times 0.25$	5.0176	26.0	5.1352	66.18	5.40	40.08	2.46
$w_f \times 0.50$	5.8831	24.1	5.6743	75.72	10.07	57.99	5.96
$w_f \times 0.75$	5.7707	24.8	5.6909	75.28	14.18	62.50	9.36
$w_f \times 1.00$	5.4284	26.1	5.5638	73.94	17.78	63.22	12.37
$w_f \times 1.50$	4.8761	28.4	5.3164	71.57	24.28	62.68	17.66
$w_f \times 2.00$	4.4933	30.0	5.1398	69.69	30.61	61.50	22.54
$w_f \times 4.00$	3.9231	33.4	4.7892	65.78	57.17	59.33	40.90
$w_f \times 6.00$	3.6992	34.8	4.6400	63.65	84.65	58.17	46.45

6.4 Sensitivity to monetary cost (λ relative to λb_f)

When λ is increased while b_f is simultaneously lowered so that the monetary cost elasticity of car demand (equation (3a)) becomes larger relative to the time elasticity (equation (3b)), fewer trips choose car since time savings from switching to car are valued less, but the monetary savings from switching from car to walking and transit are valued more. Since the sensitivity to money cost is higher, a lower toll is required to achieve a certain reduction in CO_2 or to reach the optimum. With a 20% increase in the base λ (and a 20% decrease in b_f), the social welfare maximizing toll drops from 10.07 to 8.45 per person trip by car, but emissions are still cut by 75%, a 0.13% cut for every 1% increase in the monetary cost of driving. Doubling the base λ (while halving b_f), the optimal toll drops to 4.76, but emissions are still cut by 74%, a 0.2% cut for every 1% rise in monetary cost. In these two changes, the toll on the delay externality is 58.6% and 55% of the optimal toll respectively.

7. Observing the lock-in effect from expanding highway capacity

Table 8 shows the results of additional simulations to reveal some insights related to infrastructure lock-in effects. These simulations are designed to demonstrate that expanding highway capacity reduces the effectiveness of transit improvements. That is, expanding road capacity locks-in a situation in which subsequent expansions of transit level of service would not be as effective in reducing gasoline consumption or CO_2 emissions. Each simulation changes a key variable and compares the results to a base situation.

TABLE 8: Effects of the improvements on the aggregates and the lock-in effect

	Base	Expansion of road capacity (Z increased 20%)		Improvement of transit travel time (G_2 decreased 20%)		Road expansion and transit improvement	
Car equivalent traffic $,10^6$	2.4326	2.6729	*+9.98%*	2.359	*-3.0%*	2.5943	*+6.65%*
Car person trips, $,10^6$	3.603	4.0632	*+12.77%*	3.3725	*-6.40%*	3.8339	*+6.41%*
Bus person trips $,10^6$	3.7373	3.5112	*-6.05%*	4.3185	*+15.55%*	4.0348	*+7.96%*
Car travel time (hours)$,10^6$	3.0906	3.092	*+0.05%*	2.7724	*-10.30%*	2.8058	*-9.22%*
Car vehicle kilometers$,10^7$	7.4484	8.3998	*+12.77%*	6.9719	*-6.40%*	7.9258	*+6.41%*
Fuel used by cars (liters)$,10^6$	7.2393	7.6184	*+5.24%*	6.6064	*-8.74%*	7.0401	*-2.75%*
CO_2 emitted by cars (mil.tons/yr)	4.6059	4.7434	*+2.98%*	4.1732	*-9.40%*	4.3487	*-5.59%*
Speed of car travel (km/hr)	24.1	27.2	*+12.77%*	25.15	*+4.35%*	28.25	*+17.21%*

In order to properly capture the lock-in effect, four simulations needed to be done and the results are shown in the corresponding columns of Table 8. The first is a base simulation representing the calibrated state corresponding to 2002. The second column shows what happens if highway capacity is expanded by 20%. Speed, aggregate vehicle-kilometers and trips by car all increase by 12.7%, bus trips decrease by 6%, and car equivalent traffic increases by 9.8%. The aggregate travel time of those commuting by car increases only slightly by 0.04%, reflecting the effect of the road capacity expansion. Fuel consumption increases by 5.2% and CO_2 emissions by almost 3%. As expected, improving transit (column 3 of Table 8) has oppositely signed effects, except on car speed which increases as people switch from cars to the less congesting buses in larger numbers than walkers and cyclists switch to buses. In fact, the 20% improvement in bus travel time increases car speed by 1/3 as much as increasing road capacity by 20% did. Note that fuel consumption and emissions decrease by around nine percent. Now in the last step the road capacity improvement and the transit travel time improvement are introduced simultaneously. In this case note that emissions improve relative to base but by about 5.6% instead of the 9.4% improvement that occurred when transit had no competition from road improvements. In the case of fuel the reduction in the presence of highway improvements is

2.75% instead of 8.74%. It is clear, therefore, that the lock-in effect is important quantitatively for fuel use as well as for emissions.

8. Conclusions and further remarks

In this study we examined the welfare impacts of some policy instruments (e.g., tolls on excess delays, fuel tax, improving fuel economy standards, etc.) to reduce congestion and atmospheric emissions in Sao Paulo, Brazil. We employed a trinomial logit model that is capable of capturing the modal switching between walking, public transit and auto. The simulations reveal a number of findings that could help governments and city planners in designing policies to reduce transport sector externalities in megacities like Sao Paulo. These findings are summarized below:

- Expanding roads, subsidizing transit and improving car fuel economy may not be as effective as suggested by economic theories because these policies could cause rebound effects. For example, building more roads improves speed which reduces emissions from existing car trips, but it also attracts new car trips from the other modes which can cause the aggregate emissions to rise. Improving transit fares reduces emissions by attracting trips from cars, but it also attracts the trips of those who walked and if the latter effect is stronger, this may cause more buses on the roads increasing congestion. Improving the fuel economy of cars at least initially makes the fuel and therefore the monetary cost of driving a km cheaper. Although this cuts emissions from existing car trips, if more car trips are induced by the cheaper per km cost, then congestion and total emissions can increase. Our simulations reveal that a 20% expansion of road capacity would lower emissions by 0.95%; a 20% subsidy in public transit would result in 0.40% emission reduction and a 20% improvement in fuel economy standards would reduce emissions by approximately 5%.

- Pricing instruments such as tolls on excess delays, fuel tax etc. would certainly reduce congestion and emissions. But, optimal taxes on road travel steeply increase the monetary cost of travel per trip and are therefore politically difficult to implement. However, a noticeable finding is that even smaller tolls that are more likely to be politically acceptable, have substantial benefits in terms of reducing congestion and emissions.

- The toll, which is socially optimum for Sao Paulo, is estimated to be 10.07 BRL/trip. Although this raises the monetary cost of a trip by 6.57 times, it would reduce atmospheric emissions dramatically, 75%. The large increases in the monetary cost per trip due to these taxes induce significant substitutions away from auto and in favor of public motorized and other (non-motorized) modes, causing large decreases in fuel use and CO_2 emissions; a reduction of 0.13%, on average, for every 1% increase in the monetary cost per km.

- It would be socially preferable to use the daily revenue from these taxes which is about 14.6 million BRL per day, in improving transit rather than building roads. A 20% increase in highway capacity about halves the emission reductions that a 20% improvement in transit travel times could achieve in the absence of the highway improvment. Building roads would make future transit expansions less productive in reducing emissions. We refer to this finding as the lock-in effect or road investment.

Although the model used in the study is highly aggregated, it does show certain results that enrich our understanding of how pricing aimed at congestion and fuel externalities improve welfare, raise revenue and reduce fuel use and emissions. In a more disaggregated setting, such as one involving several geographic areas connected by a highway network, important differences between congestion tolling and fuel taxing are certain to emerge. In such a disaggregated setting, the optimal tolls will vary by highway, whereas the fuel tax would be a flat tax per liter regardless of the congestion and excess fuel externalities created by the driver. Therefore, in a more disaggregated analysis, the fuel tax would emerge as an inferior measure for controlling traffic externalities and could achieve significantly lower social welfare improvements. However, because the fuel tax is aimed directly at pricing fuel use, it will likely remain a more effective tool for reducing energy utilization and emissions, but at the cost of some welfare loss.

References

Anas, A. and Pines, D. "Anti-sprawl policies in a system of congested cities," *Regional Science and Urban Economics,* 38, 408-423, 2008.

Anas, A. and Rhee, H-J. "Curbing excess sprawl with congestion tolls and urban boundaries," *Regional Science and Urban Economics,* 36, 510-541, 2006.

Anas, A., Timilsina, G. and Zheng, S. "Effects of a Toll on Congestion versus a Tax on Gasoline On Car Travel, Fuel Consumption and CO2 Emissions in Beijing," World Bank Working Paper, February 2, 2009.

Barth, M. and Boriboonsomsin, K. "Real world CO_2 impacts of traffic congestion", Paper prepared for the 87th (2008) Annual Meetings of the Transportation Research Board, Washington, D.C. , November 15, 2007.

Beevers, S.D. and Carslaw, D.C. "The impact of congestion charging on vehicle speed and its implications for assessing vehicle emissions," *Atmospheric Environment*, 39, 6875-6884, 2005.

Davis, S.C. and Diegel, S.W. *Transportation Energy Data Book*, *Edition 24*, Oak Ridge National Laboratory, 2004.

Davis, S.C. and Diegel, S.W. *Transportation Energy Data Book*, *Edition 26*, Oak Ridge National Laboratory, 2007.

Goldman, F. 2008. The Experience of São Paulo. Presentation at Local Government Climate Sessions, Poznan, Poland. December, 2008.

Gurjar, B.R., Butler, T.M., Lawrence, M.G., Lelieveld, J. "Evaluation of emissions and air quality in megacities," *Atmospheric Environment*, 42, 1593-1606, 2008.

IAPT, 2007. *Millenium Cities Data Base for Sustainable Transport.* International Association of Public Transport, Brussels.

Ingram,G.K. "Patterns of metropolitan development: what have we learned?", *Urban Studies*, 35(7), 1019-1035, 1998

Instituto Brasileiro de Geografia e Estatística. 2002. Resolução No. 05, de 10 de Outubro de 2002, Área Territorial: UF – São Paulo – SP - 35

Instituto Brasileiro de Geografia e Estatística, 2008. Estimativas das Populações Residentes, em 1o. De Julho de 2008.

Saitoh, T.S., Yamada, N., Ando, D., Kurata, K. "A grand design of future electric vehicle to reduce urban warming and CO_2 emissions in urban area," *Renewable Energy*, 30, 1845-1860, 2005.

Sánchez-Ccoyllo, O.R., R.Y. Ynoue, LD. Martins, R. Astolfo, R.M. Miranda, E.D. Freitas, A.S. Borges, A. Fornaro, H. Freitas, A. Moreira, M.F. Andrade. 2009. Vehicular particulate matter emissions in road tunnels in Sao Paulo, Brazil. Environ Monit Assess (2009) 149:241–249.