TECHNICAL NOTE Nº IDB-TN-2798

The Efficiency of Urban Transport Policies in Latin-American Cities

Leonardo J. Basso Sotz
Hugo E. Silva Montalva

Inter-American Development Bank
Transport Division

October 2023

The Efficiency of Urban Transport Policies in Latin-American Cities

Leonardo J. Basso Sotz
Hugo E. Silva Montalva

Inter-American Development Bank
Transport Division

October 2023

Cataloging-in-Publication data provided by the Inter-
American Development Bank
Felipe Herrera Library
Basso, Leonardo.
The efficiency of urban transport policies in Latin-American cities / Leonardo J. Basso Sotz, Hugo E. Silva Montalva.
p. cm. — (IDB Technical Note ; 2798)
Includes bibliographical references.
1. Urban transportation-Latin America. 2. Transportation and state-Latin America. I. Silva M., Hugo. II. Inter-American Development Bank. Transport Division. III. Title. IV. Series.
IDB-TN-2798

Keywords: Public transportation; Urban transportation; Latin America and the Caribbean.
JEL Codes: L91; N76; R48.

http://www.iadb.org

Contents

1. Introduction

Urban sustainability requires reducing the environmental footprint of urban mobility. To achieve this goal public transport becomes key due to its lower impact on congestion, pollution, accidents and greenhouse gas emissions as compared to private transport. Thus, encouraging public transport use is a frequent goal among city authorities. For this, they often devote an important fraction of their budget to improve the quality of its service. Many cities subsidize not just the infrastructure for public transport, but also its operation based on (some of) the following reasons.

1. As a second-best policy since charging additional fees to automobile drivers to correct negative externalities may be impractical (if the externality is trip specific) or politically unfeasible.
2. Due to the *Mohring effect* in which the optimal frequency and network density of services grow with the demand. Thus, the waiting and access time drop for every user in the system when the demand grows, implying that the marginal social cost is lower than the average social cost (Mohring, 1972).[1]
3. Following distributional concerns, since often lower income people use public transport more intensively.

Subsidization figures vary widely around the globe. They are rather large in the developed world: 65% of operational cost on average for the largest 20 cities in the US, 45% for the main 26 European cities, 60% for top 5 Australian cities, 40% for Toronto. Moreover, a few cities have gone further and provide free public transport.[2] Subsidies are less common in the developing world, particularly in Latin-America. They reach, approximately 50% in Buenos Aires, 40% in Sao Paulo, 40% in Santiago and between 40 to 50% in Bogotá .

Parry and Small (2009) analyze optimal –in social welfare sense– fare schemes for public transport accounting for their impact in congestion, pollution, accident externalities, scale economies, and agency adjustment of transit service offerings. They found that fare subsidies beyond 50% of operating costs is welfare improving for Washingont DC. And Los Angeles, in the US, and for London. Gómez-Gelvez y Mojica (2022) applied Parry and Small's model to Bogotá finding that the subsidy vary widely, between 20% and 100% depending on modeling assumptions such as the assumed elasticities or the assumed change from private cars to buses as the fare diminishes. Importantly, these models do not allow to consider price changes in more than one mode at the time –not

[1] Note that the Morhing effect appears when the Social Cost Function includes users' resources. When this is the case, the Mohring effect contributes to the existence of scale economies, as defined by decreasing average costs along an output ray.

[2] Tallin, the capital of Estonia, provides the largest and most studied free fare case (Cats et al, 2017). Most recently in 2020 the country of Luxembourg set all its public transport services nationwide fare free, while Kansas City started rolling out a fare-free transit incremental policy.

allowing for the analysis of subsidies and congestion pricing together-- nor it considers long term changes in the bus system. And rather than having a modal choice model, they assume a diversion ratio from cars to public transport. In the case of Parry and Small, the simulations of increased subsidies stopped at 90% of operational costs. Basso and Silva (2014) report that transit subsidies are indeed welfare improving for Santiago, Chile and London, UK, at levels above what was in place. Other papers that look at optimal or near optimal transit subsidies are Glaister and Lewis (1978), Small (1983), Viton (1983), De Borger et al. (1996), Huang (2000), Proost and Van Dender (2008), Kutzbach (2009), Parry and Small (2009), Basso et al. (2011) and Basso and Jara-Díaz (2012).

There are two cases with well-known high quality transit systems that are not subsidized: Singapur and Hong-Kong. Why can they have good public transport systems without the need for subsidies? A common feature is high population density. For the case of Hong-Kong there is also a land-rent capture policy that allows the system to be funded through revenues that flow from the value the land gains when a subways station is built and operated. In the case of Singapore, the explanation may lie in the fact that it has one of the oldest congestion pricing systems in place. And, indeed, congestion pricing is another possible way to deal with congestion, reducing the environmental footprint of urban mobility by reducing the number of car trips and transferring travelers to public transport or to teleworking, while creating revenue that can be used to further improve the transit system. Congestion pricing has been advocated for decades by experts, but it has been applied only in a few cities, most possibly because it is perceived as regressive, in that only richer people will have the chance to drive, while the middle class, that had already a hard time buying a car, will be forcefully moved to the public transport. It is in fact because of this perceived unfairness that many countries have turned to rationing schemes or driving restrictions, in which cars are forbidden to circulate based on their license plate. One of the most stringent driving restrictions today is found in Bogotá, where restrictions were first introduced in 1998.[3] Since 2012 Bogota's driving restriction, better known as *Pico y Placa*, affects the vast majority of residential and commercial vehicles every other day of the week (excluding weekends) from 6:00 a.m. to 8:30 a.m. and then from 3:00 p.m. to 7:30 p.m.

These type of restrictions—which treat all cars the same—have been widely criticized for the perverse incentives they create on drivers to buy additional (often older and more polluting) vehicles, not only increasing the fleet size but also moving its composition toward higher emitting. vehicles, resulting in more congestion and pollution. The best documented evidence supporting this claim comes from Mexico City's Hoy No Circula program, as implemented in 1989 (e.g., Eskeland and Feyioglu 1997; Davis 2008; Gallego et al. 2013). Another, much less discussed, inefficiency of driving restrictions is that they act as a 'proportional rationing rule' rather than an efficient one,

[3] Other programs include, for example, Athens (where restrictions were first introduced in 1982), Santiago (1986), Mexico City (1989), Teheran (1991), São Paulo (1996), Manila (1996), Cali (2002), La Paz (2002), Medellín (2005), Beijing (2008), Tianjin (2008), Quito (2010), Hangzhou (2011), Chengdu (2012), Paris (2016) and Madrid (2019).

preventing some high value trips to take place, while allowing others that, should they face their social marginal cost, would not have been taken,

In response not only to this “second car” concern but also to help finance the public transport system, Bogota’s transport authority introduced a major reform to its Pico y Placa program in September of 2020: since then drivers have the option to pay a congestion fee to be exempted from the restriction, with the entire fee collection going to public transport. See Daganzo (2000) and Basso, Montero and Sepúlveda (2021) for a discussion about the benefits of restriction policies with exemption fees. We do want to add, though, that when the authority allow drivers to pay a (congestion) fee that exempts them from the restriction, despite the increase in traffic, the exemption fee restores some of those socially valuable car trips that were inefficiently rationed in the first place, increasing welfare, something that will be going on in our simulations.

There is an additional alternative, however, which is far less demanding financially than subsidies and less contentious that congestion pricing: using part of the existing road capacity exclusively for buses. When buses are physically separated, through some investment, from cars, this has become to be known as a Bus Rapid Transit (BRT) system. A BRT system is defined, for example by the Institute for Transportation and Development Policy as “... a high-quality bus-based transit system that delivers fast, comfortable, and cost-effective services at metro-level capacities. It does this through the provision of dedicated lanes, with busways and iconic stations typically aligned to the center of the road, off-board fare collection, and fast and frequent operations”. The same organization states that “Because BRT contains features similar to a light rail or metro system, it is much more reliable, convenient and faster than regular bus services. With the right features, BRT can avoid the causes of delay that typically slow regular bus services, like being stuck in traffic and queuing to pay on board”. From the first system in Curitiba, Brazil, in 1977, the penetration of BRT systems has been increasing fast, mostly because of the promise of better, faster and cheaper public transport at a fraction of the cost of what a subway or heavy rail would cost. According to Global BRT Data Report from October 2018, in the year 2000 there were 40 cities with BRT systems, for a total constructed length of 1,100 km. By 2018, the numbers exploded to 170 cities around the world, for a total of 376 corridors and 5,046 km, while 121 additional cities are either building or have plans to build BRT systems. A regional panorama shows that Latin America leads with 171 corridors in 55 cities, followed by Asia, which has 94 BRT corridors in 43 cities, and Europe that has 58 corridors in 44 cities. North America has 37 corridors in 19 cities. Mohring (1979), Small (1983) , Huang et al. (2007), Kutzbach (2009), Basso et al. (2011), Gonzales and Daganzo (2012), Basso and Silva (2014) and Basso, Feres and Silva (2019) have studied the effects of moving from mixed traffic conditions to dedicated lanes for transit, using a wide variety of models.

Thus, for the goal of reducing the environmental footprint of urban mobility one may think of many different policies, such as improving vehicle efficiency (e.g. promoting electric vehicles and buses) or to improve land use and transportation integration (e.g. transport oriented development). If in addition managing congestion is also a goal, there seems to be three additional tools: subsidies, congestion pricing or driving restrictions, and dedicated bus lanes / BRT. And, for the particular case of Latin-America, there seems to be that subsidies have not been the preferred tool, while BRTs lead the way. The obvious question that arises then is, what is the efficiency of these tools for Latin-American cities? How do they 'mix and match'? when is one preferable over the other or should they be mixed? How much subsidy is required and how does the resulting transit system look like? These are the questions that this study attempt to answer, by providing ad-hoc models for Santiago, Chile, Bogotá, Colombia, and Sao Paulo, Brasil, and that are indeed aligned with the Terms of Reference that gave origin to this study.

Our work is firmly based on Basso and Silva (2014), hereafter also denoted as BS2014, and Basso, Montero and Sepúlveda (2021), hereafter also denoted as BMS2021. They both propose transport equilibrium models that can be calibrated with real data from a city, and then used in simulation mode to study different transport policies alone or in combination.

BS2014 is the base model that enables us to analyze the efficiency and substitutability of three urban congestion management policies: congestion pricing, bus lanes / BRT policies, and public transport subsidization. It also provides us with a suitable initial framework to modify, in order to analyze the case of Santiago, where Metro (subway) appears sometimes as an alternative to buses, sometimes as a complement. BMS2021, on the other hand, allows us to accommodate license plate restriction policies instead of congestion pricing such as '*pico y placa*' in Bogotá and Medellín. Moreover, it can accommodate Bogota's newest version of the policy, allowing users restricted on any given day to pay a pass and drive despite the restriction. However, Basso, Montero and Sepúlveda do not analyze the design and subsidization of the public transport system, which requires bringing in some of the machinery from BS2014.

Now, both BS2014 and BMS2021 consider 'representative networks' which are somewhat an average case for Santiago and Bogotá, respectively. Hence, a major contribution of our work here is to expand both models to a spatial setting that better represents the geographical differences in income and public transport alternatives, and which will allow us to analyze a wide array of possible urban transport policies and their combinations, looking at both efficiency and distributional aspects, with a particular spatial emphasis. This novel spatial analysis is only feasible in Santiago and Bogotá but not Sao Paulo, however, given data availability. For the latter city, then, a more direct application of BS2014 is followed.

We summarize our main results now. Using the most actual data to calibrate the Santiago model, the analysis shows that, when the city center is modeled alone, as has been usually the case in the literature and was the case in BS2104 and BMS 2021, we observe that with optimal subsidies, congestion pricing or bus lanes, each implemented separately, frequencies increase and the car market share decrease. When bus and metro fares cannot be too different, they are both very low and the optimum subsidy reaches 90%. If the price of the subway can be very different, then the subsidy is decreased optimally to 80%, the subway has fares like the current ones, and buses are free. Congestion pricing and bus lanes, without subsidies, have fares larger that the reference scenario, that has 40% subsidy. Results show that the best stand-alone policy welfare wise in the city center network –where Metro is present--is to build bus lanes. However, they create heavy congestion for cars, which might make them implausible to implement. Congestion pricing, on the other hand achieves the largest speed increase for cars in the peak and the second largest speed increase for buses. Adding subsidies to congestion pricing or bus lanes decrease both welfare and consumer surplus. In that sense, if the subsidy level is not to be touched for distributional reasons, bus lanes and congestion pricing will improve the city center performance.

When we add space, and consider corridors that feed the city center, what is optimal to do depends on the way the corridor is structured: with the existence or not of subway/BRT. Yet there are general insights to be gained. First, in all cases, implementing congestion pricing in both the corridors and the city center will improve welfare. In fact, the welfare optimum is with congestion pricing and no subsidies, but that comes at the expense of higher fares and decreased consumer surplus. Increasing or maintaining subsidies will decrease welfare but increase consumer surplus. In terms of mixed traffic vs bus lanes, building bus lanes in corridors are always welfare increasing, adding additional benefits over any level of subsidy. When there is metro available in the corridor, bus lanes generate benefits, but are somewhat milder, possibly given the existence of an already high-speed mode.

We then conducted the simulation analysis to determine the most effective policies for reducing car usage and improving public transportation in Bogotá, Colombia. The simulation included several scenarios, including a "reference" scenario representing the situation in 2019, that is, with a 33% subsidy for public transportation, a two-day restriction without the possibility of paying to avoid it. Note that the current situation has indeed the possibility of paying a pass to avoid the restriction. Our results showed that the optimal fare for public transportation required a subsidy of 100% or making the system free for riders. The combination of driving restrictions and passes was also found to be effective, with the best results coming from restricting driving five days per week and allowing the purchase of an exemption pass; in short, implementing congestion pricing. All the policies resulted in a decrease in the market share of cars, with the pass being the most effective. Additionally, both the optimal subsidization and driving restriction policies resulted in higher speeds for both cars and buses.

Setting the appropriate price for the exemption pass is key to maximizing the benefits of the policy, yet the spatial analysis showed that the average optimal daily pass may hide wide variations. This indeed stresses the system, as an average pass which is independent of space will indeed hurt the poor for the benefit of the rich. Finally, unlike in Santiago, public transport subsidization and car congestion pricing complement each other to reduce car ridership, congestion and to maximize social welfare. Yet, the analysis of the marginal benefit of increased subsidization showed that, even if efficiency increases, it does so at smaller rate above 80% when congestion pricing is not there, and quite slowly in general after congestion pricing has been implemented. This implies that the social return to an investment in increased subsidies may not be attractive to the decision maker at all subsidy levels, despite it increases welfare.

Finally, we performed a non-spatial analysis of Sao Paulo. This analysis is less deep than the previous two due to data restrictions, but it does provide us with results that, by now, seem to appear often. First, it shows that bus lanes/BRT policies perform very well. They here reach high levels of service for buses, without a huge cost on car congestion, while inducing important cost savings through fleet size, which allows fares to be keto at bay, without the need for subsidies. All this implies the highest possible welfare gain. Congestion pricing is a good policy as well, which keeps the best level of service for the car. If only subsidies are used, we again reach high values, 90%, which imply very low fares for the subway and the bus if fares are tied together or, as in Santiago, fare free buses and positive Metro fares if they are allowed to be different. The logic is that full fares are more alike: one has high speeds and a positive fare, the other is free but slower as it rolls in mixed traffic conditions.

The rest of the report is structured as follows: In section 2, we review the most recent and relevant papers in the literature, particularly Basso and Silva (2014) and Basso, Montero and Sepúlveda (2021). Appendix A in Section 14 contains summaries of other relevant references. Section 3 describes our spatial network approach which expands both papers to allow for different network configurations to capture socio-geographical differences and transport alternatives, and that is used for Santiago and Bogotá. Section 4 contains detailed explanations of the model we propose for Santiago, while section 5 provides information on the calibration process and the spatial network definition for Santiago. The simulation results are then presented in Section 6. Section 7 details the proposed model for Bogotá, with section 8 providing information on calibration and the spatial network chose, with simulation results coming in section 9. Section 10 provides information on the calibration process for Sao Paulo and its simulation results are presented in Section 11. Section 12 concludes, while Section 13, 14 and 15 contain references and appendices.

2. Literature review

Basso and Silva (2014)

They analyze the efficiency and substitutability of three urban congestion management policies: transit subsidization, car congestion pricing, and dedicated bus lanes. They model a representative kilometer of a city's road network where bus service is offered and look at one day of operation. Travelers choose whether to travel in one of the two possible periods, peak and off-peak, or not to travel at all; furthermore, if they travel, they choose between the two modes available in both periods: car and bus.

As it is clear from the setup, they are only considering two always-available modes. They also assume that households are uniformly distributed across the corridor and that the trip length is constant. The following Figure summarizes the model of the representative network.

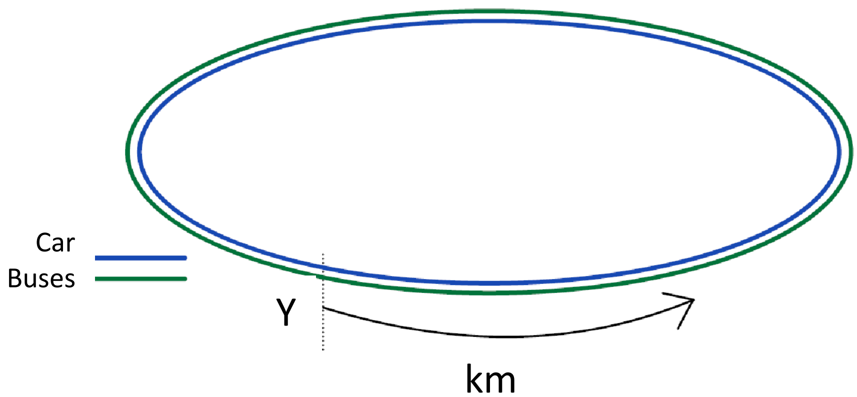


Figure 1: Basso and Silva's (2014) representative network.

The Figure represents a dense city center with uniform conditions and two available modes: car and bus. There are Y people per kilometer that may travel l kilometers by any of the two modes. As the distribution of individuals is uniform and the trip length constant, the traffic conditions and flow will also be constant along the corridor.

They calibrate the model using data from London and Santiago and simulate different policy scenarios. In each scenario, social welfare is maximized subject to different policy conditions. So, for example, one scenario will search for budget covering bus fares and congestion charges that, considering mixed traffic conditions and equilibrium con, maximize the sum of consumer surplus, revenues from the cars and buses minus expenses, all subject to modal spit equilibrium conditions. Another scenario will add searching for the best percentage of the capacity to be dedicated to bus lanes.

The authors find that, in terms of total social welfare, congestion pricing and dedicated bus lanes far more efficient in London than subsidizing bus fares. The additional contribution of subsidized bus fares would therefore be small. In Santiago, however, bus lanes yield a much higher benefit than congestion pricing and optimal subsidization.

They also concluded that there is large efficiency substitutability among the three policies; once one is implemented, adding another does not increase welfare as much. Bus lanes are an attractive way to increase frequencies and decrease fares without injecting public funds. Finally, they point out that congestion pricing and subsidies are not equivalent, which contradicts previous studies.

Kilani et al. (2014)

They explore reforms in the pricing of private and public transport in Paris. The Paris transport network is represented as a stylized concentric city with the choice between car, rapid rail, metro, and buses, two-income classes, and different transport motives. The following Figure summarizes their framework.

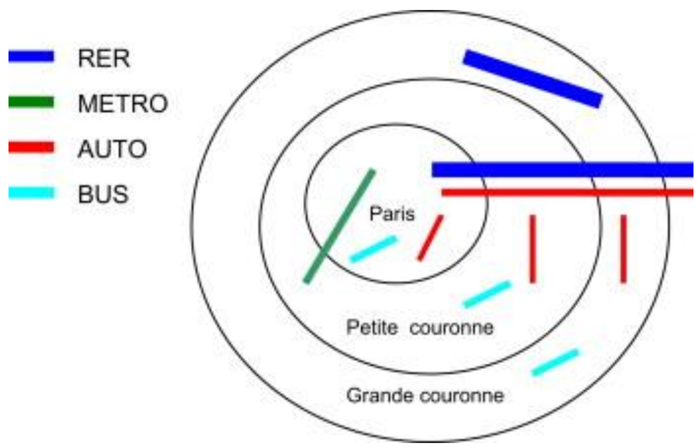


Figure 2: Kilani et al. (2014) representative network.

They model a situation where not all modes are available for all trips: the car can be used for all trips, the bus can only be used for trips within a specific zone, and the metro is available for trips within Paris or between Paris and PC, while the RER can be used for trips between any two zones but not within a zone, except for some trips within the Grande Couronne.

They find that a zonal pricing scheme for the center of Paris combined with higher public transport fares in the peak performs best (are complements in addressing inefficiencies in the transport sector). The low-income earners are not necessarily worse off for several reasons, including their more intensive bus use. The benefits of an overall capacity extension of public transport supply are much lower than the benefits of pricing reforms and could very well not pass the cost-benefit test.

Börjesson et al. (2017)

They derive optimal bus pricing, bus frequency, bus size, and bus lanes for a corridor in Stockholm, taking advantage of available revealed modal choice data.

They model one corridor that links two suburban areas to the city center of Stockholm. The model is generally similar to Basso and Silva (2014). Passengers can use either the car or the bus during peak or off-peak periods. All transport is from either the suburb to the CBD or back. In this corridor, only buses are available as public transport, and there is currently a dedicated bus lane.
Stockholm introduced congestion pricing in 2006, and the authors use that price change as a quasi-natural experiment to estimate price and cross-price elasticities. They find that Stockholm's subsidies for peak bus trips are too high. Lowering the off-peak frequency increases welfare more than reducing subsidies. Using larger buses also increases welfare.

Börjesson et al. (2018)

They model an urban area around a given bus corridor. Like in Basso and Silva (2014), the population is homogenously spread along the corridor. The difference is that users can demand two different separable goods: long and short trips. Moreover, travelers can choose between three short and long journey modes: car, bus, and bicycle. They find that the number of bus stops is already close to optimal. Welfare would increase if the peak frequency were increased if the bus fares were lowered and differentiated between long and short trips, and the toll for longer car trips was increased. The optimal toll for cyclists, and its welfare benefit, is small and does not compensate for the transaction costs. The distributional effects of bus fare changes and higher car tolls are small because, on the one hand, high-income groups place more value on travel time gains, but on the other hand, low-income groups travel less frequently by car. They find that in the welfare optimum, the bus service only requires a small subsidy due to congestion in the bus lane, crowding in the buses, and extra boarding and alighting time per passenger; that is, all known negative externalities in the bus system. The results suggest that the higher tolls are optimal for long car trips at the peak.

Börjesson et al. (2019)

Using a similar approach to the previous two studies, they compare the optimal public transport subsidies for a representative bus corridor in a small city and a big city in Sweden. The subsidy is computed by assuming optimal pricing, frequency, bus stop spacing, and bus lane policies. The high crowding cost dominates in the big city, approaching full cost recovery in the first-best optimum. In the small town, the waiting time dominates, implying more significant optimal subsidies. The subsidy is also more effective as a redistribution policy in a small city (the main reason for the different degrees of subsidization is the importance of crowding costs relative to waiting and schedule delay costs).

Basso, Montero and Sepúlevda (2021)

When authorities have decided to deal with the congestion externality, in a few cases, they have turned to price schemes (notable exceptions include London, Stockholm, Singapore, Milan, and Gothenburg). Instead, they increasingly rely on rationing schemes, which are known as driving restrictions or license-plate bans. These policies seem to be preferred by authorities because, as opposed to rationing schemes based on prices are seen as regressive and unfair to those with limited capacity to pay. Driving restrictions would allow authorities to strike a better balance between efficiency and equity. Importantly, driving restrictions have been more prevalent in Latin America and the developing world: Atenas (1982), Santiago (1986), Ciudad de México (1989), Teheran (1991), Sao Paulo (1996), Manila (1996), Bogotá (1998), Cali (2002), La Paz (2002), Medellín (2005), Beijing (2008), Tianjin (2008), Quito (2010), Hangzhou (2011), Chengdu (2012), Nueva Delhi (2016), Paris (2016), y Madrid (2019).

On the other hand, these types of restrictions—which treat all cars the same—have been widely criticized for the perverse incentives they create for drivers to buy additional (often older and more polluting) vehicles. This increases the fleet size and moves its composition toward higher emitting cars, resulting in more congestion and pollution. The best-documented evidence supporting this claim comes from Mexico City's Hoy No Circula program, as implemented in 1989 (e.g., Eskeland and Feyioglu 1997; Davis 2008; Gallego et al. 2013).

To cope with the resistance to congestion pricing, Carlos F. Daganzo advanced an ingenious scheme that combines pricing and driving restrictions (Daganzo, 2000; Daganzo and Garcia, 2000). Daganzo's basic idea is for people to take turns having unpaid access to the road. Thus, an individual who travels daily would have to pay a toll only on those days of the week in which her car is restricted from circulation, say, those days in which the car's license plate ends in a particular digit. Daganzo's premise is that this hybrid scheme leaves everybody better off while providing the necessary public support for the policy.

The story would go like this: higher-income individuals would benefit from the plan as they continue commuting by car daily (and paying the toll the day or days of restriction) but faster. On the other hand, lower-income individuals would incur a loss during the day or days of restriction as they could not afford to pay the toll and have no choice but to either switch to public transport or cancel the trip altogether. This loss, however, would be more than compensated by the gain from faster car travel during the rest of the week, i.e., days of no restriction. In addition, Daganzo's scheme would possess two other advantages that should ease its implementation. One is that it builds around a policy that authorities increasingly rely on to curb congestion and local air pollution, as discussed above. Second, it would deal with the perverse incentive of buying additional vehicles, as it would be much cheaper to pay the toll.

Basso Montero and Sepúlveda (2021) test the Pareto-improving property of the Daganzo's hybrid scheme with a simple model that uses Santiago, Chile, as a case study. They allow commuters to choose between two modes of transportation: private vehicles and public transit, as in Basso and Silva (2014). Travelers have no choice but to commute every day of the week, so only those who own a car can switch to a different transportation mode (i.e., public transport). Importantly, and differently from Basso and Silva (2014), travelers decide how many days of the week they will drive,

as a function of the price of the daily pass and the number of days they are restricted, thus including the possibility of driving restrictions and daily passes. Commuters are heterogenous about income, preferences for transportation modes, and their vehicles (if they own one). In particular, and as in Basso and Silva (2014), they divide individuals into five income groups, following SECTRA's (2013) value-of-time criteria, and characterize the assortment of cars in each group by classes (e.g., SUVs, compact cars), fuel types (gasoline, diesel), and vintage, according to information from different databases. They also extend Daganzo's restriction scheme to incorporate local pollution considerations.
A transport authority has control over four variables: (i) the number of days per week in which a car is restricted from circulation, (ii) the value of the daily toll, (iii) the vintage threshold above which car owners can have their cars exempted from the restriction by paying the toll, and (iv) the destiny of toll revenues which, in the paper, can be recycled to the same income group or destined to cheaper bus fares and better frequencies (but without changing the design of the fleet).

One of the main results of Basso, Montero and Sepúlveda is that Daganzo's Pareto-improving premise that all income groups would benefit from a taking-turns scheme with one- or two-day-a-week restrictions does not hold. Individuals in lower-income groups (remarkably those few who own a car) are strictly worse off, and more so as the number of days of restriction increases. This negative result calls for two seemingly contradicting measures. The first is that all toll revenues should be recycled into the public transit system through some combination of lower fares and better service. The second measure is that authorities should aim for the most ambitious restriction format. A more ambitious goal not only contributes to welfare with lower travel times and pollution levels, but it also contributes to more toll revenues to be spent in public transit, leaving lower-income groups increasingly better off as well. Note that recycling is not subsidizing public transport. Subsidization was not studied.

The closest situation in reality to driving restrictions with toll exemptions is the current case of Bogota. In September 2020, the Bogotá authority resumed the Pico and Placa policy after a halt induced by the Covid-19 crisis. Still, this time they allowed those who pay a new congestion fee to circulate (as had been discussed in previous years). This measure was called Pico y Placa Solidario. At first, the toll was the same for everyone. Still, later, in August 2021, the authority implemented that the congestion fees began to vary according to a vehicle's characteristic, more precisely, its value and pollution rate.

3. Spatial Network Approach

We base our work on a mixture of the models presented in Basso and Silva (2014) –also noted BS2014-- and Basso, Montero and Sepúlveda (2021) –also noted BMS2021. We use a modified version of the former to model demand in of Santiago, because it is well suited to incorporate a third transport mode, Metro, which sometimes is a substitute for buses while in occasions it is a complement. We use a modified version of BMS2021 to model demand for private and public transport trips in Bogotá, as it can accommodate typical driving restrictions, toll exemptions, and congestion pricing (five days restriction with toll exemption is a full fledge congestion pricing) and telework. We use BS2014 for both cities to model supply, and to consider subsidization, public transport optimal design, bus lanes and Bus Rapid Transit, and the overall analysis of efficiency, distributional impacts and substitutability.

We describe the main elements of demand and supply below, yet, the main methodological challenge of this study is to extend the network modeling approach of BS2014 and BMS 2021 to a spatial setting that better represents the geographical differences in income and public transport alternatives. In particular, we need to allow for the following key characteristics:

a) Include Metro as a travel alternative
b) Provide the flexibility to have parts of the city where metro and buses are substitutes and others in which they may be complements,
c) To model that in some parts of the city, buses may run in mixed traffic conditions while in others, they run in BRT mode.
d) To capture that different parts of the city have different mixtures of socio-economic groups

The idea is to capture the main trip patterns of a city straightforwardly, and the framework must be flexible enough to accommodate different types of cities. The spatial representation in BS2014 and BMS2021 represents a dense area –such as a city center– where conditions are homogeneous. In this study, we aim to generalize it to include trips that originate in the suburbs and have the center as the destination.

The trips from the suburbs provide the flexibility to incorporate multimodal trips into the model. For example, they can be considered corridors where buses feed a metro-based trunk system. We adopt different ad-hoc configurations for each city based on the available mobility reports and the existing public transport network.

We propose the following network configuration to construct representative networks for different cities.

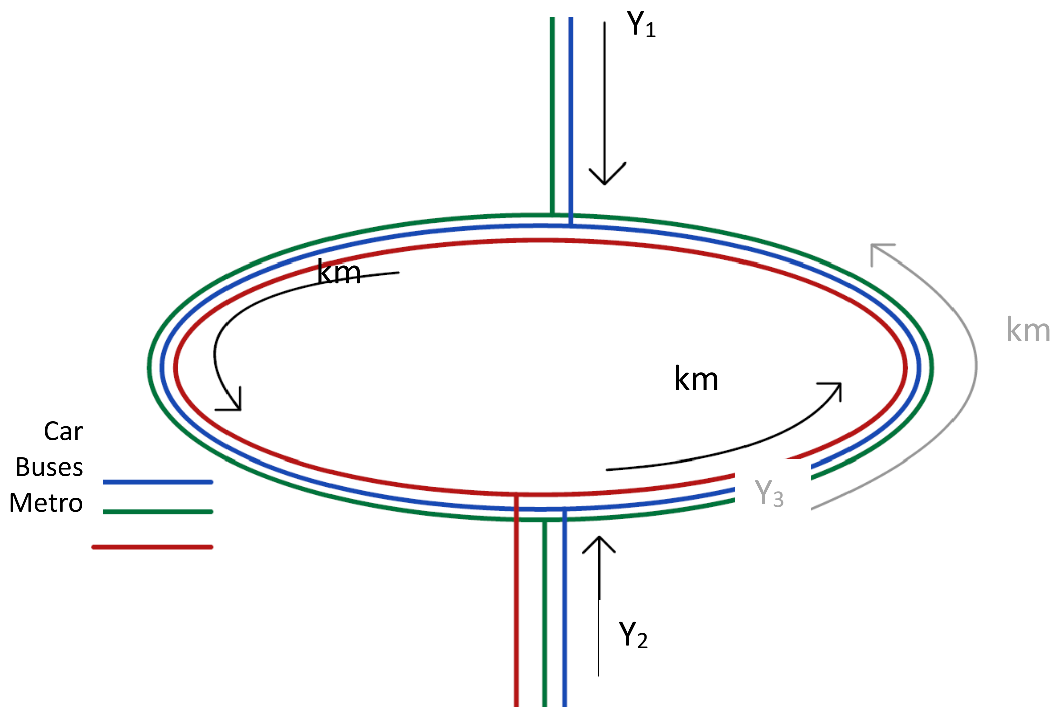


Figure 3: Possible network configuration

Figure 3 encompasses the following types of trips:

1) A group of individuals, Y_1, who travel from the periphery (North in the Figure) and only have car and bus as options. Note that this could be further divided into two: buses may run in mixed traffic conditions or BRT conditions. These individuals may need to travel further within the center (ring) to reach their destination. For this purpose, they can switch to the metro or continue using the bus if it does enter the city center, or they may need/decide to change to different bus lines. Of course, if they travel by car, they will finish the journey by car.
2) A second group, Y_2, who travel from the periphery (South in the Figure) and can choose between the three modes for the entire trip. Again, buses may or may not run in BRT conditions, and people may need/decide to change mode once reaching the city center loop to finish their journey.
3) A third group, Y_3, that travels only within the city center and has three options. Buses may or may not run in BRT conditions.

Thus, we propose to have one dense area that resembles the city center, where all modes may be available. Some trips may begin and end in this area, following the original model in Basso and Silva (2014) and Basso, Montero and Sepúlveda (2014), with origins distributed uniformly over the corridor/ring. Note that, any of the groups of people, Y_1, Y_2 and Y_3 have a common origin and destination but encompasses different income groups. That is, Y_1 may be the result of grouping wealthy and less-wealthy people that travel from the periphery with only car and bus options.

We propose to arrange the trips spatially so that the traffic conditions are homogenous within the center. The relative importance of the suburbs, corridors, and the center would thus be given by the actual relative size of the populations and trip lengths. Ultimately, the decision about the importance of each type of trip depends on the features of the city in question, implying that the specific network model needs to be discussed for each application. That is, calibration will require specifically designing a network for each city, which depends on the current situation. Later, though, all public policies can be analyzed from obvious ones, such as changing fares, to others, like switching from mixed traffic conditions to BRT conditions on a given corridor.

The rest of the methodology follows Basso and Silva (2014) for Santiago and Basso, Montero and Sepúlveda (2021) for Bogotá. We will then discuss each city separately: first the model, including demand, supply and the planner's problem. Then we will discuss calibration and the proposed spatial network. We will finally simulate to obtain relevant insights.

4. The Santiago model

Demand

In each corridor, there are up to three available modes to travel: car, bus, and subway. There are several groups of individuals who differ in income, and therefore on values of time savings and demand elasticities. Of course, all these will also differ across corridors. There are two periods in which people travel, peak and off-peak hours, and people may choose not to travel.

Thus, everyone chooses whether to travel in one of the two possible periods, peak and off-peak, or not to travel at all; furthermore, if they do travel, they choose between two or three modes available in both periods: car, bus and, if available in their part of the city, subway. As in Basso and Silva (2014), we use the Nested Logit model introduced by Ben-Akiva (1973), well-rooted in the random utility theory framework. The following Figure illustrates the decision-making process when all the modes are available in one corridor of the city.

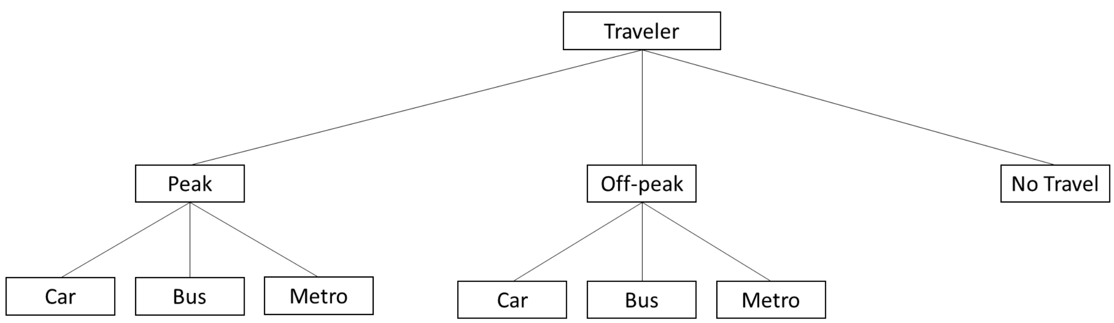


Figure 4: Traveler's decision process

Consider an individual in group i that travels by mode m in period q. Her utility depends on monetary costs, time costs, and an alternative-specific constant in the following way:

$$U_{qm}^{i} = \theta_{qm}^{i} + \gamma^{i} \cdot cost_{qm} + \beta_{qm}^{i} \cdot gt_{qm} \tag{1}$$

Where θ_{qm}^{i} is the alternative-specific constant, γ^{i} is the cost parameter (the absolute value of the marginal utility of income), β_{qm}^{q} is the marginal (dis)utility of generalized travel time gt_{qm}. Thus, the quality of the two modes is measured by the generalized travel time and price, while different constant utility that reflects differences in tastes.

The monetary cost of any of the public transport modes is the fare, while for car travel is the sum of a per-kilometer constant expenditure, which is mostly fuel, and the congestion charge, if any. The generalized travel time is a weighted sum of in-vehicle time, waiting time, and walking time, where weights capture the fact that people perceive these times differently (for example, waiting is always more unpleasant). We discuss them in detail next.

The number of people that chooses mode m in nest n is given by:

$$Y_{nm} = \sum_{i=1}^{I} \quad Y^i \cdot P_n^i \cdot P_{m\vee n}^i, n \in N, m \in M_n \tag{2}$$

where Y^i is the number of people per kilometer of income group i, P_n^i is the probability of choosing nest n, and $P_{m\vee n}^i$ the probability of choosing alternative m conditional of choosing nest n. The following equations define the probabilities within a nest:

$$P_{m\vee n}^i = \frac{exp(\lambda \cdot U_{nm}^i)}{\sum_{r\in M_n} \quad exp(\lambda \cdot U_{nr}^i)}, n \in N \tag{3}$$

$$M_{peak} = M_{offpeak} = \{Car, Subway, Bus\} \qquad M_{notravel} = \{Notravel\}$$

$$q \in \{Peak, Offpeak\} \quad U_{notravelnotrave}^i = \theta_{notravel}^i$$

where λ is a scale parameter related to the degree of substitutability of alternatives within a nest. The no-travel nest does not have multiple alternatives, so its utility is given by a constant.

The formulas are similar for the choice probabilities for nests but using each nest's expected utility. Formally,

$$P_n^i = \frac{exp(\mu \cdot A_n^i)}{\sum_{r\in N} \quad exp(\mu \cdot A_r^i)}, n \in NA_n^i = ln\left(\sum_{r\in M_n} \quad exp\left(\lambda \cdot U_{nr}^i\right)\right) \tag{4}$$

Where $\mu < \lambda$ is a scale parameter related to the substitutability of nests.

Transport times

The generalized travel time for a bus or subway user in each period is the weighted sum of in-vehicle time, waiting time, and constant access times, which are normalized to zero. The weights are obtained empirically and detailed in the calibration section. The waiting time is simply half the headway, which is $\frac{1}{f_{qm}}$ for mode m and period q.

For the in-vehicle time, things depend on whether there is dedicated-bus lanes/BRT or mixed traffic conditions. For the latter, the complexity is that there are several sources of congestion externalities: cars and buses congest each other while in motion, buses delay each other while operating at bus stops, and the bus stop operations cause delays on cars.

Under mixed traffic conditions, the travel time by bus in the period q is given by:

$$t_{bus\ MT}^{q} = T_b + \alpha \cdot \left(\frac{b \cdot f^{q} + V_{car}^{q}}{C}\right)^{\beta} + t_{sb} \cdot Y_{qb} \cdot p + T_S\left(f^{q}, p, Y_{qb}, t_{sb}\right) \quad (5)$$

where T_b is the uncongested travel time, the second term on the right-hand side is the variable time while buses are in motion and represents congestion between vehicles on the road. It is a power function of the flow-capacity ratio known as the BPR function, commonly used in transportation analyses to model congestion (see Small and Verhoef, 2007). The frequency (f^{q}) is the flow of buses, and it is multiplied by a factor b that transforms buses into equivalent vehicles, V_{car}^{q} is the flow of cars, and the capacity for buses is C, and α and β are parameters. The third term is the dwell time due to boarding, which is the product of the time each passenger takes to board t_{sb}, the number of passengers boarding, Y_{qb}, and the number of stops, p. In the fourth term, $T_S(\cdot)$, is a non-linear function of frequency, the number of stops that a bus makes, the time each passenger takes to board and alight, and the number of passengers boarding a bus at each stop, Y_{qb} and, importantly, the per-passenger boarding time. It also considers bus congestion at the bus stop. The specific functional forms are from Tirachini et al. (2014).

To model the travel time for cars under mixed traffic conditions, we follow Basso and Silva (2014) and add to the car travel time a fraction of the time a bus needs for bus stop operations. The fraction ϵ depends on the frequency of buses. Denoting T_c the uncongested travel time by car, the total travel time by car is:

$$t_{car\ MT}^{q} = T_c + \alpha \cdot \left(\frac{b \cdot f^{q} + V_{car}^{q}}{C}\right)^{\beta} + \epsilon(f^{q}) \cdot T_S\left(f^{q}, Y_{qb}\right) \quad (6)$$

When buses have dedicated bus lanes, they no longer share the capacity with cars. Instead, they can use a fraction η of the capacity. Thus travel time is

$$t_{bus\ DL}^{q} = T_b + \alpha \cdot \left(\frac{b \cdot f^{q}}{\eta\ C}\right)^{\beta} + t_{sb} \cdot Y_{qb} \cdot p + T_S\left(f^{q}, p, Y_{qb}, t_{sb}\right) \quad (7)$$

And with bus lanes, cars use the remaining capacity, while no longer interacting with bus stops:

$$t_{car\ DL}^{q} = T_c + \alpha \cdot \left(\frac{V_{car}^{q}}{(1-\eta)C}\right)^{\beta} \quad (8)$$

Finally, travel time by subway is considered to be uncongested and equal to T_m, which is equal to the inverse of the metro mean speed.

<u>Operating costs</u>

We model the operating costs of the bus system (G, in $/day) as the sum of expenses that are proportional to the bus fleet (B), which are mainly labor and vehicle-capital expenses, and costs that are proportional to the total number of vehicle-kilometers of each period VKT_q to capture operational expenses:

$$G = G_b \cdot B + \sum_q \quad G_v \cdot VKT_q \tag{9}$$

where G_b and G_v are the cost per bus per day and cost per vehicle-km, respectively. The fleet required is $B = \max_q f^q \cdot t_{bus}^q \cdot L$, where L is the total distance that a bus covers before starting a new cycle; t_{bus}^q is the is travel time per kilometer in period q. This reflects that there is one period (usually the peak) that defines the number of buses required for operation. At the same time, there will be idle capacity (spare buses) in the other period. The daily vehicle-kilometers are the sum of the vehicles-kilometers of both periods, which are directly the product of frequency, length of the cycle, and hours of operation.

The operations costs of the subway system follow the same logic. Some expenses are proportional to the number of subway cars (labor and vehicle-capital) and the car-kilometers driven in each period. We do not include subway track expenses in the analysis, as we focus on an existing network. The final component of the costs, which is relevant for considering fare-free buses, is the on-board fare collection system's cost. For a system based on contactless smartcards, these expenses include a fixed cost and card readers on buses. The variable cost is thus proportional to the fleet size B.

Planner's Optimization problem

The objective function we consider is unweighted social welfare for a day of operation, summing over all corridors. Each corridor's welfare has the same structure: it includes consumer surplus (CS), the financial result of the public transport system in that corridor, including buses and the subway if available, and the implementation costs of any policy in place.

Consumer surplus in the nested-logit model is the maximum expected utility and is obtained through what is known as the log-sum formula:

$$CS = \sum_{i \in I} \quad Y^i \cdot \frac{1}{\gamma^i} \cdot \frac{1}{\mu} \cdot ln\left(\sum_{n \in N} \quad \mu \cdot A_n^i\right) \tag{10}$$

Since consumer surplus is an unweighted sum of each individual commuter surplus, the Marshallian measure will value more the time savings of those with higher willingness to pay, which is related to higher income levels through smaller marginal utilities of income. Therefore, the measure could be considered regressive. An alternative is to assign different weights on individual consumer surplus according to income, but this departs from pure efficiency analysis.

Therefore, the welfare function for one corridor is:

$$SW = CS + mcpf \cdot \left(\sum_q \quad Y_{qbus} \cdot P_{qb} - G\right) + mcpf \cdot \left(\sum_q \quad Y_{qsubway} \cdot P_{qs} \cdot l - C_m\right) \tag{11}$$

Where Y_{qm} is the number of passengers of mode m that board per kilometer and P_{qm} is the fare. C_m is the subway system cost. We multiply the aggregate financial result of both public transport modes by the marginal cost of public funds, $mcpf$, to reflect that public funds are costly. We do not consider the possibility that subsidies induce cost inefficiencies on the transit system because the extent of these inefficiencies, if they exist, depend on the contract between the transit operator and the regulator (see, e.g., Gagnépain and Ivaldi (2002)). However, considering the cost of public funds works as a proxy.

To compare benefits and service levels of the different transport policies, we build scenarios defined as maximizing social welfare subject to various constraints. For instance, we may start by looking for the optimal subsidy level and fares in each corridor, while later imposing that the bus fare has to be the same everywhere. We may then, alternatively, study the effects of building BRTs in all corridors that do not have one, or study the effects of congestion pricing, where the tax per kilometer may or may not be the same in each corridor. Of course, we will mix policies, looking for the best combination of them. As a result of these simulation exercises, we will be able to obtain social welfare and consumer surplus per income group differentiated spatially.

5. Santiago calibration and network

Calibration

As explained before, the model parameters are calibrated to reflect Santiago's traffic and pollution reality, as captured by the most recent available data. The model can then be used in simulation mode to produce predictions about the possible outcomes of different transport policies. Let us then talk about calibration of the model parameters. All of them are summarized in Appendix B.

For the BPR-type congestion function (equations (5)–(8)), we assume that the speed at capacity is reduced to one third, that the free flow speed is 60 km/hr, and that $\beta = 4$.[4] The equivalence factor between buses and cars, which enters travel time equations (5) and (6) is a function of bus size estimated with a linear regression using values that are common in project appraisal in Chile. We obtain figures such as 1.6 and 2.5 cars for buses ranging from 40 to 120 passengers respectively.

The parameters for the functions describing time at bus stops are obtained from a microsimulation study (Fernández, Valenzuela, and Gálvez (2002)) and from empirical surveys (e.g., Transportation Research Board (1985)). The time a passenger takes to board is set to 2.5 seconds, which is consistent with a system of contactless cards as the one used in Santiago. Car operating costs are about one third of a dollar per kilometer.

The functions Gb(k) and Gv(k) in equation (9) are assumed to be linear. In the case of Santiago, the functions are estimated with a linear regression over some cost studies made for four different firms with different bus size varying from 40 passengers to 160 passengers. Because of this, we use the cost parameters estimated for Santiago multiplied by a factor that makes the bus fare in the reference scenario similar to the observed fare.

Regarding the marginal cost of public funds (MCPF), Parry and Small (2009) state that a typical estimate is 1.15, which is what we use. This value is consistent with actual estimates: Harrison, Rutherford, and Tarr (2002) find a MCPF for Chile that is between 1.08 and 1.18 depending on the tax considered. Ballard, Shoven, and Whalley (1985) estimate a range of 1.17 to 1.33 for the United States, while Auriol and Warlters (2012) find an average MCPF for 38 African countries of 1.2. The share of congestion pricing revenues that is spent operating the system is set to the average of the reported values by Transport for London for the period 2004–2008. We believe this is a sensible assumption because, despite the changes during the mentioned period, the share has been fairly constant (between 0.42 and 0.49). The cost of operating dedicated bus lanes is estimated by Tirachini, Hensher, and Jara-Díaz (2010) for Australia, and it includes the operation and maintenance of track, right-of-way, signaling, communications, and so on. We use a value equivalent to 30

[4] Verhoef and Small (2004) and Kutzbach (2009) use $\beta = 4$; Parry and Small (2009) use 3.7.

percent of that cost for Santiago, to account for differences in cost in developing countries (e.g., labor).

In the case of traffic data, that is speeds, we rely on the Origin- Destination survey of 2012 (ODS-2012) and the congestion-pricing simulations conducted by Chile's Transport Planning Office (SECTRA), which are reported in SECTRA (2013).

Following SECTRA (2013), we divide commuters in five income groups. Tables 1 and 2 indicates some relevant characteristics of these groups

Group	Strata	Share of total	Average monthly income per household	Car ownership	Value of travel time savings ($US/hr.)
1	E	12%	<$368	16%	1.36
2	D	27%	368–734	34%	3.11
3	C3	34%	735–1,468	54%	5.89
4	C2	19%	1, 469–2,935	77%	10.47
5	ABC1	8%	>$2,935	95%	28.20

Table 1. Socioeconomic characteristics - Santiago.

Shares peak	**ABC1**	**C2**	**C3**	**D&E**
Car	70%	49%	40%	31%
Bus	14%	31%	40%	50%
Metro	15%	20%	20%	18%
Shares Off-peak	**ABC1**	**C2**	**C3**	**D&E**
Car	76%	57%	43%	31%
Bus	11%	26%	39%	51%
Metro	13%	17%	18%	18%

Table 2. Modal choice by socioeconomic group - Santiago.

The parameters required to specify the logit models are marginal utilities of income (the cost parameter) and time, and the modal constants. Marginal utilities can be obtained directly from estimated modal choice logit models or derived from observed elasticities and values of time; modal constants, however, must be calibrated for each particular case (corridor). In the case of Santiago, data enables us to use a demand model that includes heterogeneity. EOD 2012 provides a good source of information for trips in Santiago based on revealed preferences. One main feature is that different socioeconomic groups differ significantly in values of time: the ratio between the highest and the lowest is 2.5 for peak travel and 1.5 for off-peak. The car is heavily used mainly by the two groups with higher income while the two groups with lower income have a large use of transit and negligible car trips (see Table 2). We use the main demand parameters from EOD 2012, including marginal utilities of income for each income group.

Fares, speeds, and frequencies, needed for calibration, are reported in the tables below:

Bus fare peak [US $]	1.20
Bus fare off-peak [US $]	1.20
Car toll peak [US $]	0.00
Car toll off-peak [US $]	0.00
Metro fare peak [US $]	1.32
Metro fare off-peak [US $]	1.32

Table 3. Calibration prices for Santiago.

Bus frequency peak [bus/hr]	15
Bus frequency off-peak [bus/hr]	8
Bus size [pax]	160
Metro frequency peak [bus/hr]	15
Metro frequency off-peak [bus/hr]	7

Table 4. Public transport frequencies and bus size for Santiago calibration.

Car speed peak – mixed traffic [km/hr]	16
Bus speed peak – mixed traffic [km/hr]	13
Car speed off-peak – mixed traffic [km/hr]	35
Bus speed off-peak – mixed traffic [km/hr]	25

Table 5. Mixed traffic modal speeds for calibration.

Car speed peak – dedicated bus lanes [km/hr]	14
Bus speed peak – dedicated bus lanes [km/hr]	20
Car speed off-peak – dedicated bus lanes [km/hr]	30
Bus speed off-peak – dedicated bus lanes [km/hr]	28

Table 6. Dedicated lanes speeds for calibration.

This calibration situation captures that there is no congestion pricing and that the operational subsidy to the public transport system is 40%. Some corridors do have bus lanes.

Network

According to the different travel options observed in reality, we propose to model Santiago through four type of corridors and a central destination were trips end. We now describe each corridor, in terms of the available modes and transport alternatives, and the socioeconomic composition in each corridor. Modal split within each corridor is not known, so we will use the modal split per income group reported in table 2.

- *Santiago transport corridor 1 (MTNM)*

This corridor will feature mixed traffic conditions, that is, cars and buses share the capacity, and there is no Metro. Two examples of this type of corridors are Pedro de Valdivia and Vitacura. Given the conditions, the alternatives available to people are

i. Bus in mixed traffic all the way to the city center (with no transfer)
ii. Bus in mixed traffic first + subway at the center
iii. Car in mixed traffic (no transfer)

The socioeconomic composition of this corridor is:

ABC1	C2	C3	D&E
31%	22%	31%	16%

Table 7. Socioeconomic composition in corridor 1-MTNM.

- *Santiago transport corridor 2 (DLNM)*

This corridor features dedicated bus lanes, that is, buses have an exclusive share of the capacity, and there is no Metro. An example of such a corridor is Grecia. Given the conditions, the alternatives available to people are

i. Bus in BRT infrastructure (corridor) & mixed traffic (center). No transfer.
ii. Bus in BRT infrastructure (corridor) + transfer to subway (center)
iii. Car in segregated traffic (no transfer)

The socioeconomic composition of this corridor is:

ABC1	C2	C3	D&E
9%	14%	19%	58%

Table 8. Socioeconomic composition in corridor 2-DLNM.

- *Santiago transport corridor 3 (MTYM)*

This corridor features dedicated mixed traffic, that is, buses and cars share the capacity, and there is Metro available. Examples of such corridors are Recoleta and Vespucio Sur. Given the conditions, the alternatives available to people are

i. Bus in mixed traffic (no transfer)
ii. Bus in mixed traffic (corridor) + transfer to subway (center)
iii. Subway all the way to the center
iv. Car in mixed traffic (no transfer)

The socioeconomic composition of this corridor is:

ABC1	C2	C3	D&E
1%	2%	29%	68%

Table 9. Socioeconomic composition in corridor 3-MTYM.

- *Santiago transport corridor 4 (DLYM)*

This corridor features dedicated bus lanes, that is, buses have an exclusive share of the capacity, and there is also Metro. Examples of such corridors are Independencia and Vicuña Mackenna. Given the conditions, the alternatives available to people are

i. Bus in BRT infrastructure (corridor) & mixed traffic (center). No transfer
ii. Bus in BRT infrastructure (corridor) + transfer to subway (center)
iii. Subway all the way to the center
iv. Car in segregated traffic (no transfer)

The socioeconomic composition of this corridor is:

ABC1	C2	C3	D&E
6%	12%	25%	57%

Table 10. Socioeconomic composition in corridor 4-DLYM.

- *Central city*

Central city hosting destinations possibly with intra center trips. There is mixed traffic for buses and cars and the subway is available. The socio-economic composition depends on the number of transfers from the previous corridors and is, therefore, a result.

6. Santiago results

The strategy we pursue to produce results is as follows. First, we present the results for a representative network of the city which can be understood as the city center. This is an approach similar to what both BS2014 and BMS2021 did, and it is in fact closer to what the literature does in general. We will then show results for each of the four corridors which adds a suburb that connects with the city center As explained in Section 5, the corridor brings people to the edge of the city center, where people connects with the city center network, possibly changing mode.

In each case, we show what can be attained by using subsidies, congestion pricing, or by building bus lanes/BRT when not there, and by combinations of these policies. In all cases, corridors are considered to have three lanes and when a bus lane/BRT is considered, we devote one lane to it. The city center network is a useful benchmark because it will allow us to understand insights gained by using our spatial network framework, while looking at corridors individually will allow us to understand where that particular corridor pushes the results.

We of course focus on efficiency of every set of transport policies, but also on the distribution of surplus across income groups and corridors.

6.1 City center corridor

In table 11 below, we show simulation results for Santiago's city center. In all cases but Sub80*, the Metro fare is restricted to be not larger than 120% that of the bus; we consider this a political constraint. The 'Reference' column represents a situation where there is a 40% subsidy, no congestion pricing, and no bus lanes. The 'Sub90' column corresponds to a situation where subsides are bumped up to 90% (the optimum). The 'Sub80*' corresponds to 80% subsidy but bus and metro fares are no longer tied together. The last two columns correspond to an implementation of congestion pricing and bus lanes/BRT respectively, when nothing else is in place, that is, when subsidy is downturned to 0% (yet the Metro fare is constrained).

We observe that with 90% subsidies, congestion pricing or bus lanes, each implemented separately, optimal bus frequencies increase between 15 to 20% in the peak. They do not change much in the off-peak. The Metro frequencies vary, but not strongly. The optimal bus size remains rather unchanged as different policies are implemented.

Regarding prices, when bus and metro fares cannot be too different, they are both very low and the optimum subsidy reaches 90%. If the price of the subway can be very different, then the subsidy is decreased optimally to 80%, the subway has fares like the ones in the reference scenario, and buses are now free. Congestion pricing and bus lanes, without subsidies, have fares larger that the reference scenario, that has 40% subsidy.

With respect to the reference scenario, all policies achieve a decrease in car market share, with congestion pricing being the most effective at this. Congestion pricing also achieves the largest speed increase for cars in the peak and the second largest speed increase for buses. Interestingly, bus lanes achieve a very large increase in speed for buses, at the cost of a strong increase in car congestion.

	Reference	SUB90	SUB80*	CON+	DL
Social benefit [$/day-km]	0	1208	3511	1977	6090
CS change [$/day-km]	0	9515	9648	-5582	-1211
Bus fare peak [$/km]	0,11	0,01	0,00	0,14	0,12
Bus fare off-peak [$/km]	0,11	0,01	0,00	0,14	0,12
Car toll peak [$/km]	0,00	0,00	0,00	0,29	0,00
Car toll off-peak [$/km]	0,00	0,00	0,00	0,00	0,00
Metro fare peak [$/km]	0,13	0,02	0,11	0,17	0,14
Metro fare off-peak [$/km]	0,13	0,02	0,11	0,17	0,14
Bus frequency peak [bus/hr]	17,4	20,5	22,7	21,8	22,4
Bus frequency off-peak [bus/hr]	13,3	15,0	16,7	12,3	14,2
Bus size [pax]	163,0	163,0	163,0	163,0	148,1
Metro frequency peak [bus/hr]	14,7	17,0	12,3	15,5	14,2
Metro frequency off-peak [bus/hr]	9,0	8,2	6,7	9,0	8,8
Car speed peak [km/hr]	16,2	26,5	23,9	36,9	7,4
Bus speed peak [km/hr]	13,5	20,0	18,5	25,3	35,8
Car speed off-peak [km/hr]	34,9	50,9	48,5	38,9	21,8
Bus speed off-peak [km/hr]	27,6	36,7	35,1	29,8	42,1
Peak share [percent]	48,4%	49,6%	49,5%	47,2%	48,6%
Off-peak share [percent]	48,8%	47,9%	48,0%	49,9%	48,6%
No-travel share [percent]	2,7%	2,5%	2,5%	2,9%	2,8%
Car modal share peak [percent]	41,6%	33,3%	34,8%	29,0%	36,0%
Bus modal share peak [percent]	37,7%	43,4%	48,3%	48,5%	44,0%
Metro modal share peak [percent]	20,6%	23,4%	16,9%	22,5%	20,0%
Car modal share off-peak [percent]	50,5%	41,9%	44,8%	54,0%	50,8%
Bus modal share off-peak [percent]	34,8%	39,9%	46,0%	32,4%	34,7%
Metro modal share off-peak [percent]	14,8%	18,1%	9,2%	13,7%	14,4%

Table 11. Simulation results for an average public transport corridor in Santiago.

The results discussed above might be difficult to digest without a simple way to measure the goodness of a policy. We therefore now look at the welfare implications of the different scenarios. These are presented in Figure 5 below, which contain all the information concerning social benefit.

A line represents the benefit of a policy together with an X percent of subsidization; thus, the welfare of congestion pricing and dedicated bus lanes as stand-alone measures are given by the intercept of their respective curves. All figures are here represented as change with respect the reference scenario.

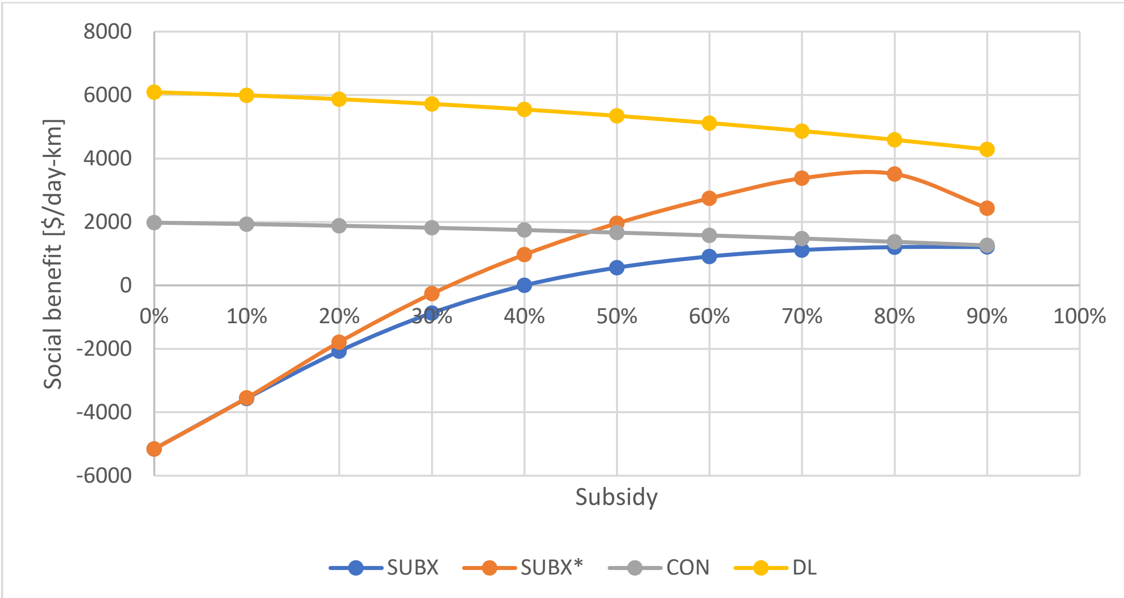
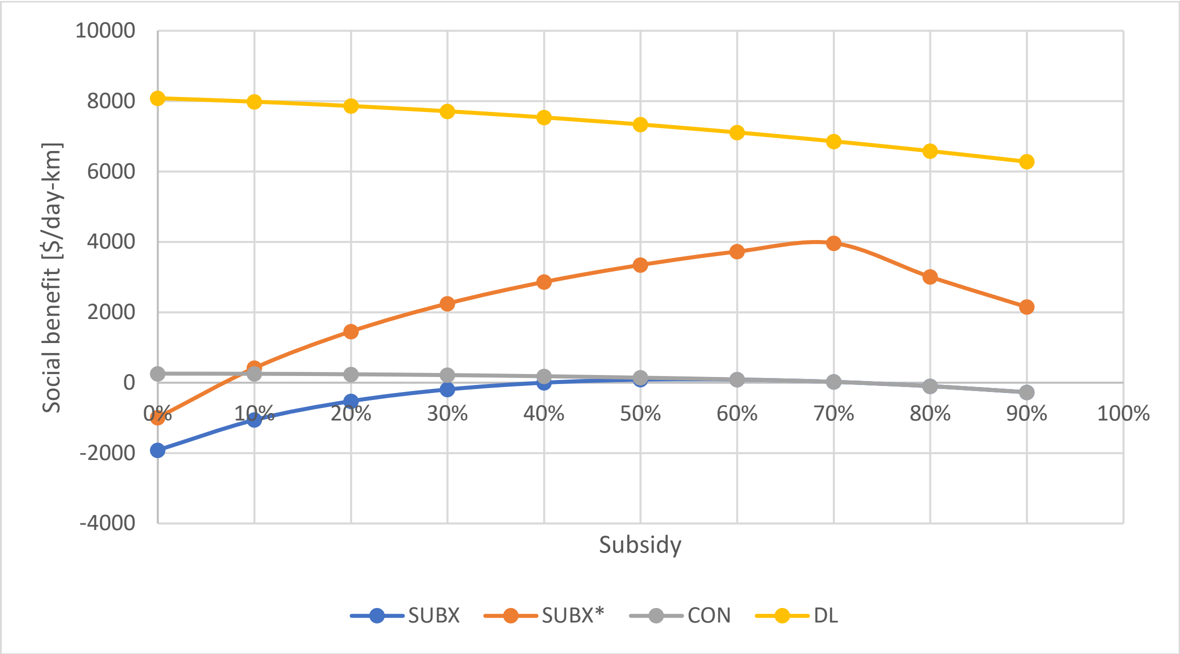


Figure 5. Welfare for different transport policies -- average public transport corridor in Santiago

These results show that the best stand-alone policy in this city center network is building a BRT line, like what was obtained by BS2014. In this exercise, however, bus lanes create heavy congestion for cars. Adding subsidies to bus lanes diminish welfare, although it increases consumer surplus. notably, the second-best stand-alone policy is an 80% subsidy, targeted at making buses free, while having a subway at almost current fares. Note that this alternative, as opposed to bus lanes without subsidies, have a positive change in consumer surplus. Congestion pricing and 90% subsidy with bus and subway fare 'not too far', come next. Finally, something that is worth noticing is that, fixing the 40% subsidy of the reference scenario, both adding congestion pricing and bus lanes increase welfare.

In the previous exercise we did not consider the car capital expenses (capex). We replicate the city center corridor's analysis when adding the capex as a fixed cost of using the car, which is independent of the kilometers travelled. To provide a reference, including the capex results in a situation similar to increasing the operating cost by 20%, which represents slightly less than 40% of the optimal car congestion toll. The results are summarized in the following Figure.



As in this new base scenario car use becomes more expensive, the optimal congestion toll is very small and yields little effect on modal split and welfare. On the other hand, putting a BRT line in place is still the best stand-alone policy.

6.2 Specific corridors

We now show results for optimization of the four types of suburban corridors that feed the city center. Here, making a full optimization makes less sense than in the average city or the overall-integrated city. What we do, then, is parametrically increase and decrease the bus fare and subway fare and optimize the rest. So, for example, if the bus fare is increased from its current value of 1,2 dollars by 10%, for that new fare, we calculate the optimal fleet size, frequencies and so on. When we consider congestion tolls, they apply to both the corridor and the city center. And, for congestion pricing, we also change the bus and metro fare parametrically. We will not consider building bus lanes in the city center, given the rather poor car speed that bus lanes induce, but will study the effects of building bus lanes in the corridor. We show figures to summarize results, leaving tables for Appendix C.

- *Santiago transport corridor 1 (MTNM)*

Recall that in this corridor, there were mixed traffic conditions, and no metro. The socioeconomic configuration of the calibrated case was closer to Vitacura, that is, skewed to high-income population.

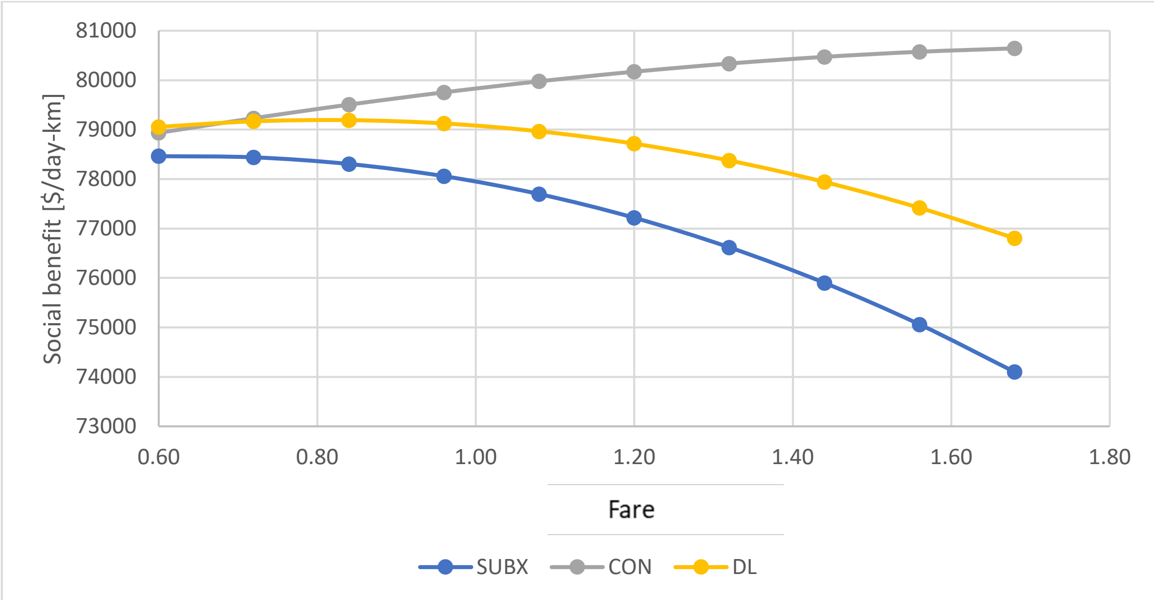


Figure 6. Welfare for different transport policies -- corridor 1 (MTNM) in Santiago

What the figure shows is that there are welfare gains if –everything else absent-- the bus fare is decreased and, actually, halved from its actual value of $1.2, as shown by the blue line. We can also see that the largest achievable welfare, in this case, is through congestion pricing and increase in fares, as the grey line shows. Bus lanes (in the corridor, not in the city center) do always better than subsidization but here, as opposed to the city center, they are dominated by congestion pricing. In summary, if there is mixed traffic conditions both in the corridor and in the city center, congestion pricing reaches higher welfare, but not by much, than approximately doubling the subsidy. Bus lanes in the corridor are a sound policy as well, which here do not induce heavy car congestion.

- *Santiago transport corridor 2 (DLNM)*

Recall that in this corridor, there were bus lanes and no metro. The socioeconomic configuration of the calibrated case was closer to Grecia avenue.

What the figure shows is that there are welfare gains if –everything else absent-- the bus fare is decreased from its actual value of $1.2 to $0.84. That is, the presence of a bus lane in the corridor helps to decrease the optimal subsidy. We can also see that the largest achievable welfare, in this case, is through congestion pricing and increase in fares, as the grey line shows.

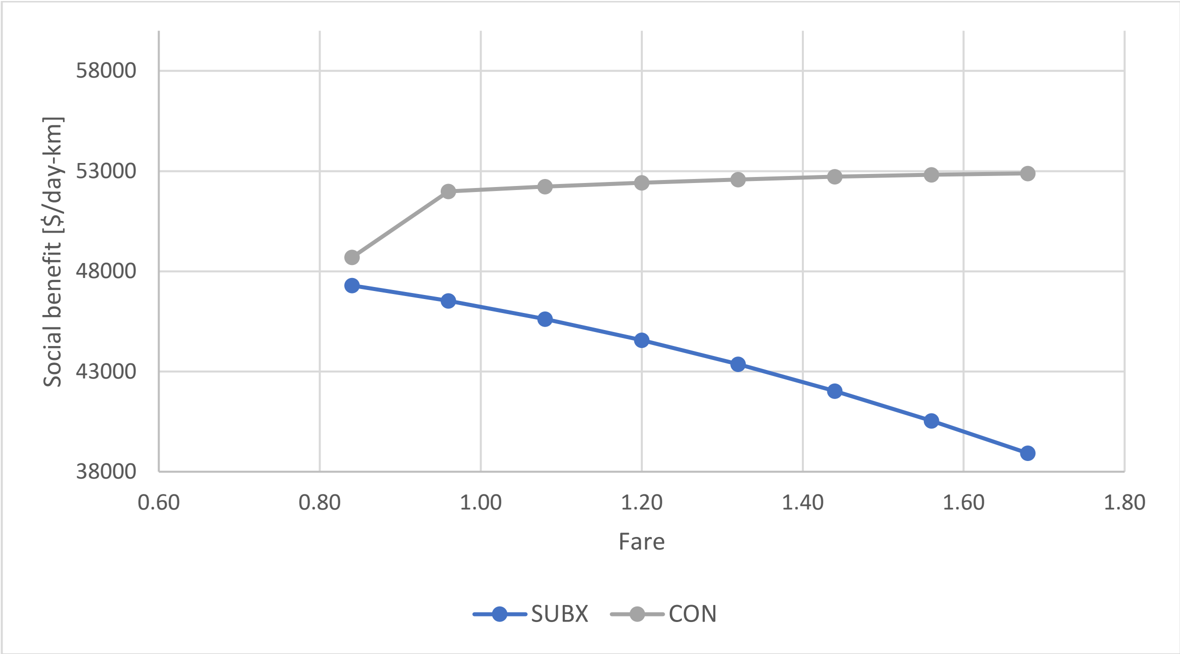


Figure 7. Welfare for different transport policies -- corridor 2 (DLNM) in Santiago

- *Santiago transport corridor 3 (MTYM)*

Recall that in this corridor, there is mixed traffic for buses and cars, but it is the first corridor with a subway. The socioeconomic configuration of the calibrated case was closer to Recoleta avenue.

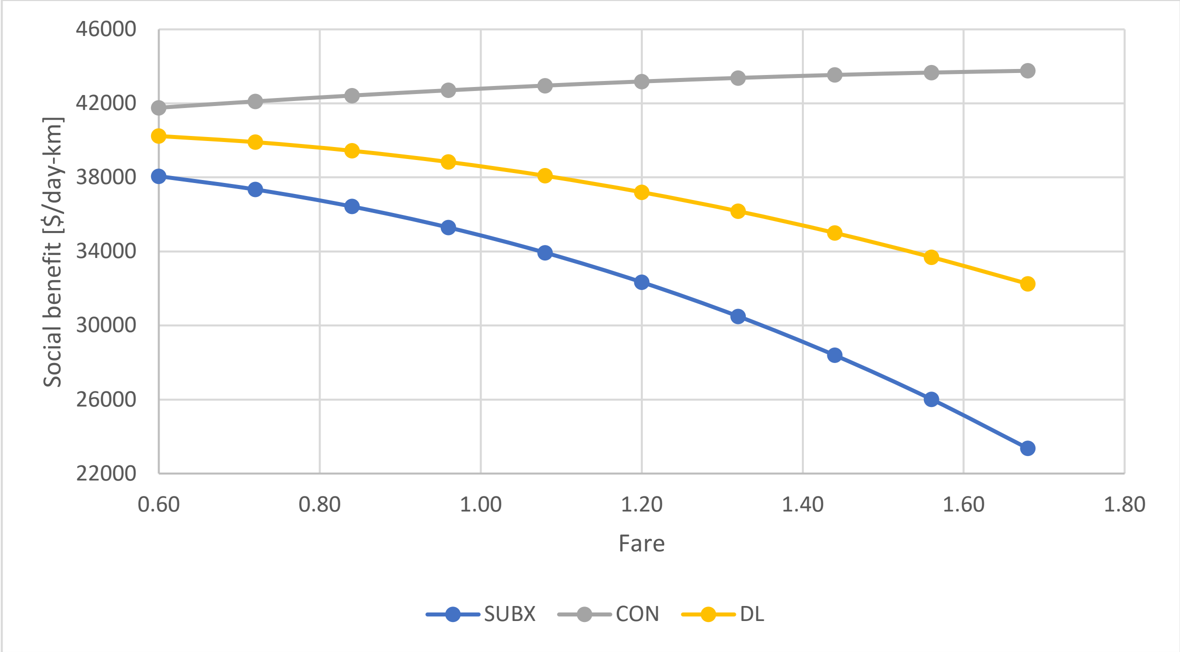


Figure 8. Welfare for different transport policies -- corridor 3 (MTYM) in Santiago

What the figure shows is that there are welfare gains if –everything else absent-- the bus fare is decreased and halved from its actual value of $1.2, as shown by the blue line. Note that this is similar corridor 1, where there was mixed traffic as well. Building a bus lane in the corridor does little more than subsidization, most likely because there is already a subway, which performs with high speeds. We can also see that the largest achievable welfare, again, is through congestion pricing and increase in fares, as the grey line shows.

- *Santiago transport corridor 4 (DLYM)*

Recall that in this corridor, there were both bus lanes and metro. The socioeconomic configuration of the calibrated case was closer to Vicuña Mackenna Avenue.

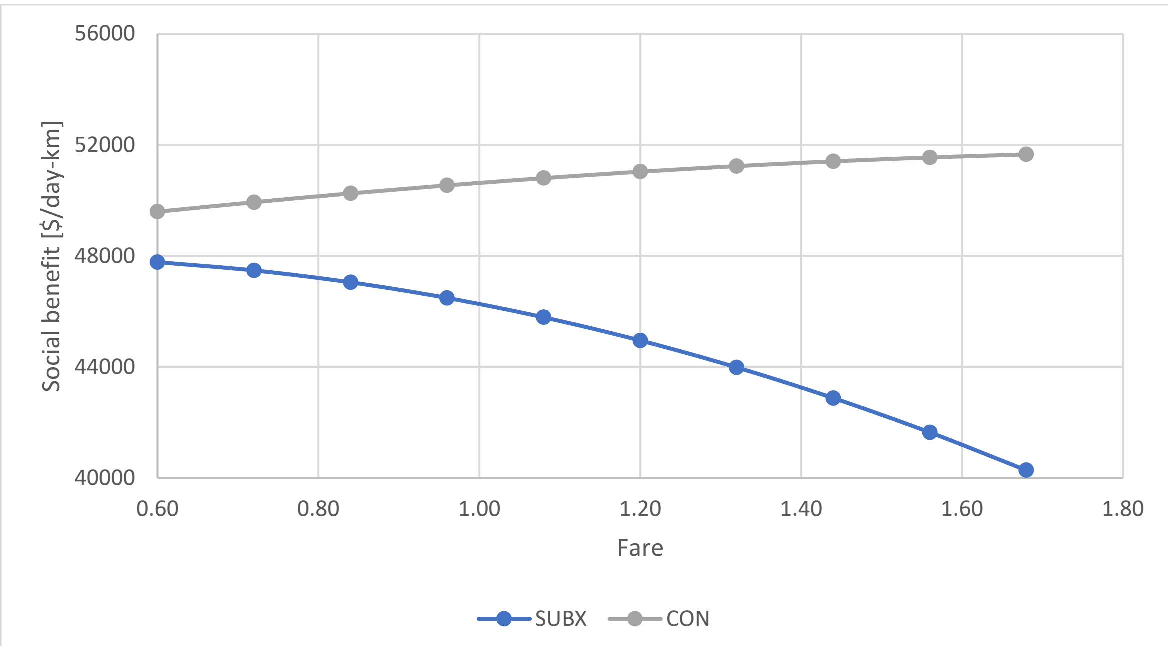


Figure 9. Welfare for different transport policies -- corridor 4 (DLNM) in Santiago

What the figure shows is that there are welfare gains if –everything else absent-- the bus fare is decreased from its actual value of $1.2 by 50%. We can also see that the largest achievable welfare, in this case, is through congestion pricing and increase in fares, as the grey line shows. Bus lanes are already in place.

6.3 Summary of results

When the city center is modeled alone, as has been usually the case in the literature and was the case in BS2104 and BMS 2021, we observe that with optimal subsidies, congestion pricing or bus lanes, each implemented separately, frequencies increase and the car market share decrease. When bus and metro fares cannot be too different, they are both very low and the optimum subsidy reaches 90%. If the price of the subway can be very different, then the subsidy is decreased optimally to 80%, with the subway having fares like the ones in the reference scenario while buses are now free. This difference is explained because for the case of buses, additional demand induced by lower fares can be met with increased capacity through larger frequencies, of course up to the point at which congestion on the road kicks in while, on the other hand the subway is at capacity at rush times: larger trains cannot be accommodated in stations and frequencies have achieved their technical maximum, implying that additional demand induced by potential lower fares cannot be met.

Results show that the best stand-alone policy in the city center network is building a BRT line, like what was obtained by BS2014. However, bus lanes in the city center create heavy congestion for cars, which might make them implausible to implement. Congestion pricing, on the other hand achieves the largest speed increase for cars in the peak and the second largest speed increase for buses. Adding subsidies to congestion pricing or bus lanes decrease both welfare and consumer surplus. In that sense, if the subsidy level is not to be touched for distributional reasons, bus lanes and congestion pricing will still improve the performance.

When we add space, and consider corridors that feed the city center, what is optimal to do depends on the way the corridor is structured. Yet there are general insights to be gained. First, in all cases, implementing congestion pricing in both the corridors and the city center will improve welfare. In fact, the welfare optimum is with congestion pricing and no subsidies, but that comes at the expense of higher fares and decreased consumer surplus. Increasing or maintaining subsidies will decrease welfare but increase consumer surplus. In terms of mixed traffic vs bus lanes, building bus lanes in corridors are always welfare increasing, adding over any level of subsidy. When there is metro in available in the corridor, bus lanes generate benefits, but are somewhat milder, possibly given the existence of an already high-speed mode.

7. The Bogotá model

Demand

We consider a single period –the peak period– and a single corridor; the model is then extended to deal with several corridors and the downtown ring according to their specificities. We build upon BS2014 standard origin-destination transport model (2014) and BMS2021, because the latter allows us to accommodate the *Pico y Placa* driving restrictions system.

On a daily basis, many individuals, say *n*, must decide whether to commute to the city center to work/study either by car or public transport (bus) or to work/study from home. There is no subway in Bogotá. Let d_i be the number of days of the week (excluding weekends) that $i = 1, \dots, n$ commutes by car, h_i the number of days that she works from home, and $5 - d_i - h_i$, the number of days that uses public transport. The (transport) surplus that individual $i = 1, \dots, n$ obtains after a week of travel is given by

$$S_i = \Omega_i(d_i, h_i) - C_i(d_i, h_i, r_i) - T_i(d_i, h_i) \tag{12}$$

where r_i = 0, …, 5 measures the extent of the restriction, i.e., the number of days her car, provided she owns one, is restricted from entering the city center during the week, $\Omega_i(d_i, h_i)$ captures the gross benefit of travel, $C_i(d_i, h_i, r_i)$ is the financial cost of travel, and $T_i(d_i, h_i)$ is the time cost of travel. Note that $\Omega_i(d_i, h_i)$, $C_i(d_i, h_i, r_i)$ and $T_i(d_i, h_i)$ are all measured in dollars and vary across individuals according to their income levels, which will be divided in income groups according to available data (it was five in the case of the base papers). We use g = 1, …, 5 to denote an income group.

The gross benefit of travel for i ∈ g depends on her intrinsic (relative) preferences for each transport mode as follows

$$\Omega_{i\in g}(d_i, h_i) = \lambda_i^{-1}\phi_0\left[d_i + (5 - d_i - h_i)\theta_{i\in g} + H_{i\in g}(h_i)\right] \tag{13}$$

where λi corresponds to i's marginal utility of income, ϕ_0 is a constant, $\theta_i \in g$ captures i's taste for public relative to private transport and $H_i \in g(h_i)$ corresponds to the benefit of remote work relative to private transport, which we capture with the linear demand $H_i'(h_i) = \vartheta_i - \xi_i h_i$.

On the other hand, i's weekly financial travel cost is given by

$$C_i(d_i, r_i) = d_i c + p\max\{0, d_i + r_i - 5\} + (5 - d_i - h_i)f \tag{14}$$

where c is the daily cost of using a car (set to infinity for those individuals that do not own one), including expenses on fuel, parking, lubricants, tires, and so on, *p* is the exemption fee, and *f* is the daily expense on public transit (i.e., the product of single-ride fare and the average number of daily rides). Two observations regarding how the restriction enters into the model are in order. The first is that we allow the restriction's extent to vary for individuals with different access to cars. In particular, and following the evidence documented by Gallego et al. (2013), we let individuals in households with two or more cars face a milder restriction, more precisely, one less day of restriction a week than the nominal level. The second is that individuals have ample flexibility to I to all possible driving restrictions. For example, an individual that faces two days of restriction (r_i = 2) would need to spend nothing on exemption fees if she is planning to use the car for three days (d_i = 3); the days of restriction would be those in which she either worked from home or took public transit.

Transport Times and costs

i's time travel cost per week is expressed as follows

$$T_i(d_i, h_i) = \lambda_i^{-1}\left[d_i\gamma_i^c t^c l + (5 - d_i - h_i)\left(\gamma_i^b t^b l + \gamma_i^w w^b\right)\right] \quad (15)$$

where γ_i^m is i's marginal utility of time when using transport mode $m \in \{c, b\}$; $t_m \equiv t_m(n_c, n_b)$ is time (in hours per kilometer) spent on transport mode m on any given day, l is the average distance traveled in a round trip from home to work including any shorter trips during the day, γ_i^w is the marginal utility of time when waiting at the bus station, and w^b is the average waiting time at the station which, as usual, is simply half the headway (inverse of frequency).

Because crowding in Bogotá's public transport system is a known phenomenon (see Basso, Feres & Silva, 2019 for some discussion), we allow γ_i^b to capture any inconvenience that may result from increasing public-transport use without the corresponding adjustment in service frequency. Following Tirachini et al. (2017) we let

$$\gamma_i^b(n_b) = \gamma_i^c\left(1 + \zeta\,\frac{n_b l}{f_b s q L}\right) \quad (16)$$

where ζ is a crowding penalty, f^b is the bus frequency, s is the average bus size, q is the duration of the peak period and L is length of the road network.

To model travel times t^c and t^b we follow the same logic as in the Santiago model: things depend on whether there is dedicated-bus lanes/BRT or mixed traffic conditions. Under mixed traffic conditions, the travel time by mode is given by:

$$t^m = T^m + \alpha^m \cdot \left(\frac{b \cdot f^b + n_c\ l/aqL}{K}\right)^{\beta^m} \quad (17)$$

where T is the uncongested travel time. The second term on the right-hand side is the variable time while vehicles are in motion and represents congestion between vehicles on the road; it is a power function of the flow-capacity ratio known as the BPR function, commonly used in transportation analyses to model congestion (see Small and Verhoef, 2007). The frequency (f^b) is the flow of buses, and it is multiplied by a factor b that transforms buses into equivalent vehicles, which is added to the flow of cars; K is the capacity of infrastructure, and α and β are parameters.

The presence of dedicated lanes/BRT that use a fraction η of the capacity eliminates the interaction between vehicles. Travel times now are:

$$t^b = T^b + \alpha^b \cdot \left(\frac{b \cdot f^b}{\eta\, K}\right)^{\beta^b} \tag{18}$$

$$t^c = T^c + \alpha^c \cdot \left(\frac{n_c\ l/aqL}{(1-\eta)\, K}\right)^{\beta^c} \tag{19}$$

<u>Operating costs</u>

We model the operating costs of the bus system (G, in $/day) as the sum of expenses that are proportional to the bus fleet (B), which are mainly labor and vehicle-capital expenses, and costs that are proportional to the total number of vehicle-kilometers VKT to capture operational expenses:

$$G = G_b \cdot B + G_v \cdot VKT \tag{9}$$

where G_b and G_v are the cost per bus per day and cost per vehicle-km, respectively. The fleet required is $B = f^b \cdot t_{bus} \cdot L$, where L is the total distance that a bus covers before starting a new cycle; t_{bus} is the is travel time per kilometer.

<u>The individual and planner problem</u>

The decision problem of individual i is to choose d_i and h_i to maximize (12), while taken as given the equilibrium choice of the remaining individuals, that is, taken as given n_c, n_p and n_h. The latter matters because of all the congestion externalities between vehicles and the positive externality that increased transit ridership has on frequencies (the so-called Mohring effect).

Note that the transport surplus function is linear with respect to the number of days that the individual commutes by car and public transportation. Therefore, for any given parameters, the solution to the individual's problem is a corner solution. However, this does not imply that all commuters within an income group will choose the same because there is heterogeneity in the value of time across and within each income group. Consequently, the model allows for variation in the mode choice for the different income groups.

The objective function of the planner is, as in the case of Santiago, unweighted social welfare for a day of operation, summing over all corridors. Each corridor's welfare has the same structure: it includes consumer surplus (CS), the financial result of the public transport system in that corridor, and the implementation costs of any policy in place. Since consumer surplus is an unweighted sum of each individual commuter surplus, the Marshallian measure will value more the time savings of those with higher willingness to pay, which is related to higher income levels through smaller marginal utilities of income. Therefore, the measure could be considered regressive. An alternative is to assign different weights on individual consumer surplus according to income, but this departs from pure efficiency analysis. We multiply the aggregate financial result of both public transport modes by the marginal cost of public funds, $mcpf$, to reflect that public funds are costly. We do not consider the possibility that subsidies induce cost inefficiencies on the transit system because the extent of these inefficiencies, if they exist, depend on the contract between the transit operator and the regulator (see, e.g., Gagnépain and Ivaldi (2002)). However, considering the cost of public funds works as a proxy.

The planner has a large number of decisions she can make, in Stackelberg leader fashion. On each corridor, she can decide on the design of the public transport system, particularly frequency, fare and level of subsidization, and whether buses will run in mixed traffic conditions or a BRT will be build if not present. The planner can also decide on the number of days each license plate will be forbidden to circulate and the level of the exemption toll. Recall that if driving restrictions are positive but smaller than five days, and there is no exemption toll (or it is set to infinity), we obtain the original “pico y placa” scheme, as it was the case in Bogotá. If there is an exemption toll, we recover Bogota´s “pico y placa solidario” case. If, on the other hand, driving restrictions are set to five, and the exemption toll is finite, we recover the congestion pricing case. With no driving restrictions, we recover the case where there is no fare over private cars, as in Santiago.

Also note that, while decisions by the transport authority may differ on different corridors, financial constraints are possibly city-wide. In a sense, cross-subsidization among corridors, both in terms of revenues and costs, is possible.

8. Bogotá calibration and network

Calibration

The model is parametrized to capture Bogotá's traffic and air pollution reality by 2019 with the most recent available data. All calibration parameters are summarized in Appendix B.For traffic information, Bogotá's 2019 Mobility Survey (MS-2019) and ProBogota GSD+ (2021) is used. The following table summarizes the socioeconomic strata, the number of people that travel in a typical day and the share of the total population in the strata that they represent, where the lowest income strata is 1 and the highest income is 6.

Strata	Number of people who travel	Share of strata population
1	716,137	86.6%
2	2,420,863	85.3%
3	2,192,789	84.7%
4	649,861	84.8%
5	179,880	84.6%
6	127,577	82.5%
Total	6,287,107	

Source: Unión Temporal Steer - CNC - Encuesta de Movilidad, 2019.

Table 12: Number and percentage of travelers within each starat

Commuters are then divided in five income groups following the characterization of MS-2019, where groups 5 and 6 are collapsed into a single high-income group.

Strata	Share of total	Average monthly income per household	Car ownership	Value of travel time savings ($US/hr)
Low (1)	12%	$184	11%	1.86
Middle-low (2)	40%	$288	21%	1.86
Middle (3)	34%	$502	39%	2.0
Middle-high (4)	10%	$1,027	66%	2.5
High (5-6)	5%	$1,564	82%	3.1

Source: Built from information form Bogotá's 2019 Mobility Survey (MS-2019) and ProBogota GSD+ (2021).

Table 13 Socioeconomic characteristics - Bogotá.

The next step is calculating the adjusted modal share, excluding the modes that will not be modeled. First, we do not consider inter-municipal buses. They are not part of the city's integrated public transport system as they are designed to connect peripheral municipalities with Bogotá. Second, TransMilenio is the trunk system that operates under dedicated BRT infrastructure. A set of bus routes feeds it called SITP, explicitly designed for this purpose which has fare integration with the TransMilenio buses. Depending on the origin, people may travel only by SIPT or transfer.

To simplify the analysis, we model public transportation in a reduced-form fashion by considering two modes: TransMilenio (or BRT) and SITP (or simply buses). In what we denote, the TransMilenio mode, we aggregate all the public transportation trips that use TransMilenio as the primary mode. It includes the trips that use SITP first and then transfer to TransMilenio. We let the mode SITP (or bus) represent the trips made by bus and do not have any stage by TransMilenio.

The adjusted modal share, which is obtained excluding the modes that will not be modeled, is described in the following table:

Mode	**Stratum 1**	**Stratum 2**	**Stratum 3**	**Stratum 4**	**Stratum 5**
Car	9%	16%	31%	53%	71%
Transmilenio	37%	39%	33%	26%	19%
SITP	54%	43%	36%	21%	10%
Total	100%	100%	100%	100%	100%

Source: own calculations

Table 14. Modal choice by socioeconomic group - Bogotá.

Then, based on these modal shares, we divide the population of each corridor into two groups that differ in their access to public transportation: Those who live in an area where it makes sense to use TransMilenio and those who use SITP only. Each group faces the choice between two modes: car and public transportation.

Income-related parameters include the marginal utility of time, the marginal utility of income and the intrinsic preference for different transportation modes and remote working. On one hand, Bogota's Mobility District Secretary has estimated a general value for the marginal utility of time; to generate heterogeneity within each group, this marginal utility is assumed to be normally distributed (and truncated at zero) with a mean value equal to Bogota's Mobility District Secretary's numbers and a standard deviation of 20%. On the other hand, marginal utilities of income for calibration come from MS-2019.

Regarding transport parameters, some of them were provided by the Bogota's Mobility District Secretary, such as the average distance (l), the daily cost of using a car to *c* \$/day, the duration of the peak period, the car occupancy, the network length, the free-flow speed of cars and the free-flow speed of buses. On the other hand, the public-transit (daily) fare is set at its 2019 value, the conversion factor of buses to cars is calculated as in the case of Santiago, and values for α and β come from BMS2021.

Fares, speeds, and frequencies, needed for calibration, are reported in the tables below:

Bus fare [US $]	0.7
Car toll [US $]	0.0

Table 15. Calibration prices for Bogotá (in 2019 values).

Transmileno frequency [bus/hr]	10
SITP frequency [bus/hr]	10

Table 16. Public transport frequencies for Bogotá calibration.

Car speed [km/hr]	18.9
Transmilenio speed [km/hr]	18.5
SITP speed [km/hr]	14.4

Table 17. Bogotá modal speeds for calibration.

Network definition

To build a representative network for Bogotá, we study the mobility patterns in the city and choose a structure that covers most, but not all, of the trips. The number of trips by origin and destination are condensed in the following Figure.

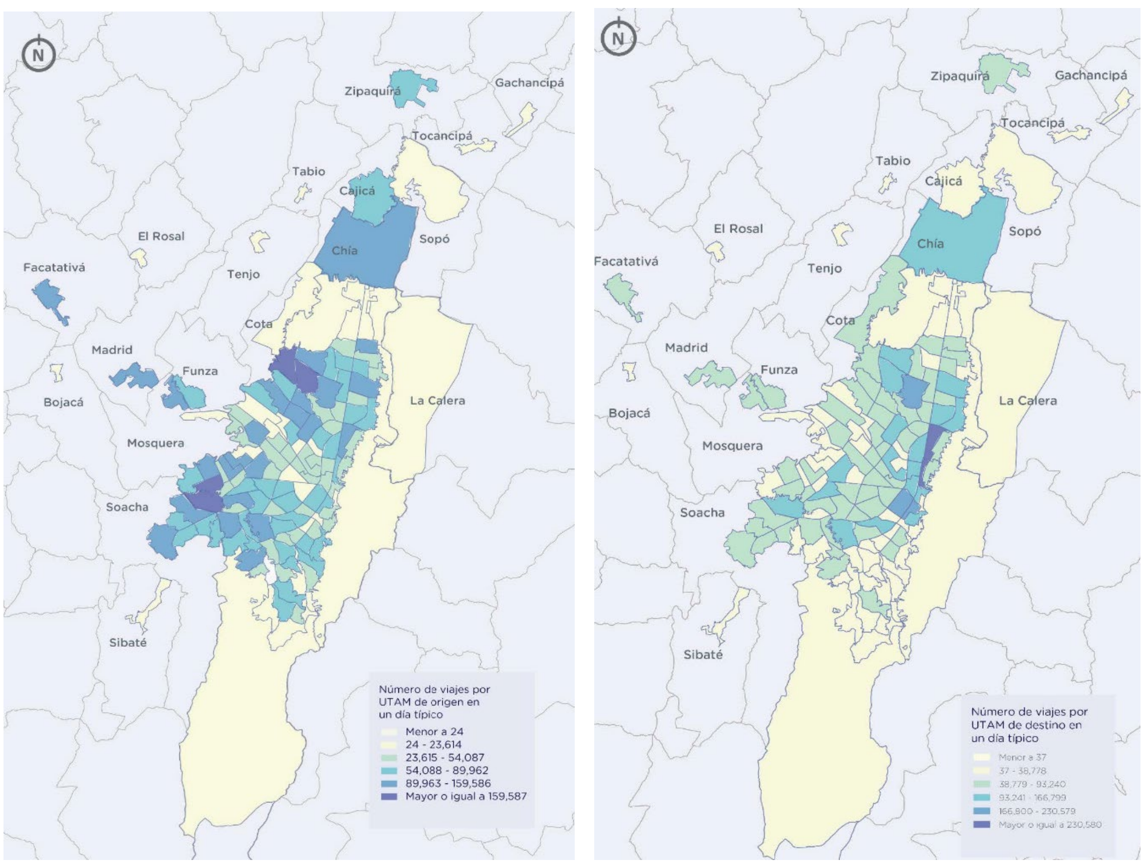


Figure 10: Origin (left) and destinations (right) in a typical day - Bogotá.

The Figure reveals that most of the trips' destinations are concentrated in a small central area. Based on these patterns, we model the city as a system of 7 zones where trips originate and one city center that concentrates all destinations. We will therefore model Bogotá as a system of seven corridors, consisting of seven homogenous origin areas and one destination. We abstract from considering trips within the city center because they represent a small share of the total trips but also because the transport network does not change significantly in the center. The following Figure summarizes the location of the origin zones.

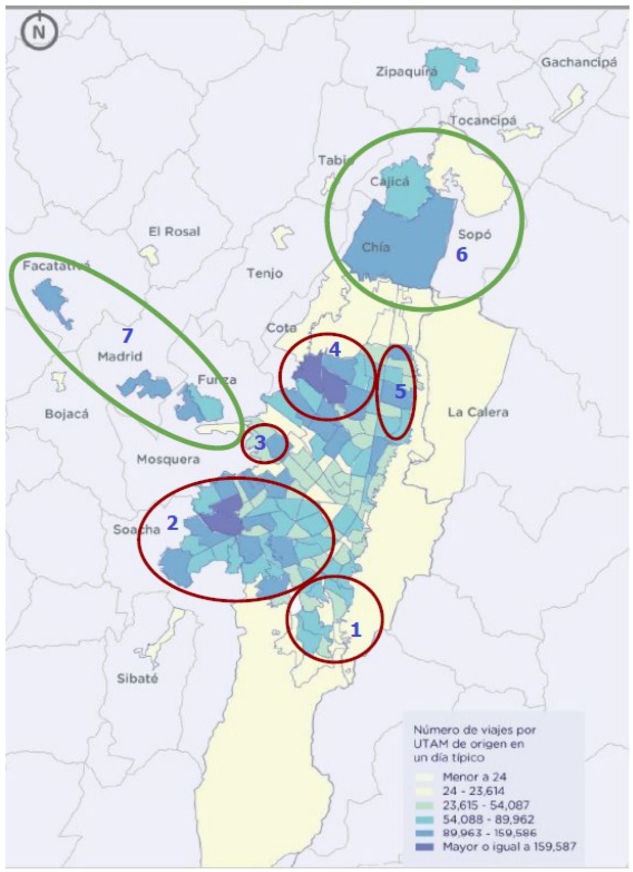


Figure 11: Zone definition - Bogotá.

The five zones circled in red are entirely urban areas with access to most of the transportation modes, while the zones circled in green are suburbs that do not have access to the integrated public transport system of Bogotá. To model trips with similar characteristics, we excluded the zones and municipalities that are far from the area that concentrates the origin of the trips. For this reason, we do not include Zipaquirá as an origin for zone 6.

- *Zone 1*

Zone 1 is in the south part of the city and its residents belong to the two strata with the lowest income. The units of transport analysis (UTAM) included are 52, 55, 56, 57, 58, 59, 60, 61, 63, UPR2, and UPR3. The following tables summarize the information on the zone regarding households, the share of trips, modal shares and essential characteristics of the daily trips.

Stratum	Number of households	Share
1	63,426	45.48%
2	76,018	54.52%
Total	139,444	100%

Table 18: Households and share of trips by period by socioeconomic stratum in Bogotá's Zone 1.

Variable	Value
Trip length	12 km
Car speed	8.0 km/hr. (peak) 14.4 km/hr. (off-peak)
Bus speed	10.3 km/hr. (peak) 12 km/hr. (off-peak)
Modal shares group 1	Car: 9% Transmilenio: 37% SITP: 54%
Modal shares group 2	Car: 16% Transmilenio: 39% SITP: 43%

Table 19: Trip characteristics and modal shares in Bogotá's Zone 1.

- *Zone 2*

Zone 2 has residents belonging to the four strata with the lowest income and is one of the largest. The units of transport analysis (UTAM) included 38, 39, 40, 41, 42, 43, 44, 44, 46, 46, 47, 49, 53, 62, 65, 66, 67, 69, 80, 82, 83, 84, 85, 86, 87, 570, 571, 572, 573, 574, 575. The following tables summarize the information on the zone regarding households, the share of trips, modal shares, and essential characteristics of the daily trips.

Stratum	Number of households	Share
1	128,882	14%
2	411,972	46%
3	355,776	40%
4	2,094	0%
5	0	0%
6	0	0%
Total	898,724	100%

Table 20: Households and share of trips by period by socioeconomic stratum in Bogotá's Zone 2

Variable	Value
Trip length	15km
Car speed	12.9 km/hr. (peak) 30 km/hr. (off-peak)
Bus speed	10 km/hr. (peak) 18 km/hr. (off-peak)
Modal shares group 1	Car: 9% Transmilenio: 37% SITP: 54%
Modal shares group 2	Car: 16% Transmilenio: 39% SITP: 43%
Modal shares group 3	Car: 31% Transmilenio: 33% SITP:36%

Table 21: Trip characteristics and modal shares in Bogotá's Zone 2.

- *Zone 3*

Zone 3 is relatively small and includes the following units of transport analysis (UTAM): 75, 76, 77. The following tables summarize the information on the zone regarding households, the share of trips, modal shares, and essential characteristics of the daily trips.

Stratum	Number of households	Share
1	0	0%
2	27,100	44.80%
3	33,449	55.20%
4	0	0%
5	0	0%
6	0	0%
Total	60,549	100%

Table 22: Households and share of trips by period by socioeconomic stratum in Bogotá's Zone 3.

Variable	Value
Trip length	13.5km
Car speed	18 km/hr. (peak) 32.4 km/hr. (off-peak)
Bus speed	11.6 km/hr. (peak) 13.5 km/hr. (off-peak)
Modal shares group 2	Car: 16% Transmilenio: 39% SITP: 43%
Modal shares group 3	Car: 31% Transmilenio: 33% SITP:36%

Table 23: Trip characteristics and modal shares in Bogotá's Zone 3.

- *Zone 4*

Zone 4 comprises the following units of transport analysis (UTAM): 27, 28, 71, 72, and 610. The following tables summarize the information on the zone regarding households, the share of trips, modal shares, and essential characteristics of the daily trips.

Stratum	Number of households	Share
1	0	0%
2	121,358	46.44%
3	136,164	52.10%
4	3,809	1.46%
5	0	0%
6	0	0%
Total	261,331	100%

Table 24: Households and share of trips by period by socioeconomic stratum in Bogotá's Zone 4.

Variable	Value
Trip length	15km
Car speed	18 km/hr. (peak) 36 km/hr. (off-peak)
Bus speed	15 km/hr. (peak) 20 km/hr. (off-peak)
Modal shares group 2	Car: 16% Transmilenio: 39% SITP: 43%
Modal shares group 3	Car: 31% Transmilenio: 33% SITP:36%
Modal shares group 4	Car: 53% Transmilenio: 26% SITP: 21%
Modal shares group 5	Car: 71% Transmilenio: 19% SITP: 10%

Table 25: Trip characteristics and modal shares in Bogotá's Zone 4.

- *Zone 5*

Zone 5 is one of the few with people from all strata and includes the following units of transport analysis (UTAM): 9, 10, 11, 12, 13, 14, 15, 16. The following tables summarize the information on the zone regarding households, the share of trips, modal shares, and essential characteristics of the daily trips.

Stratum	Number of households	Share
1	4,546	2.3%
2	14,089	7.2%
3	43,498	22.2%
4	72,014	36.7%
5	22,479	11.5%
6	39,414	20.1%
Total	196,040	100%

Table 26: Households and share of trips by period by socioeconomic stratum in Bogotá's Zone 5.

Variable	**Value**
Trip length	15 km
Car speed	15 km/hr. (peak) 36 km/hr. (off-peak)
Bus speed	15 km/hr. (peak) 18 km/hr. (off-peak)
Modal shares group 4	Car: 53% Transmilenio: 26% SITP: 21%
Modal shares group 5	Car: 71% Transmilenio: 19% SITP: 10%
Modal shares group 6	Car: 80% Transmilenio: 8% SITP: 12%

Table 27: Trip characteristics and modal shares in Bogotá's Zone 5.

- *Zones 6 and 7*

These two zones are located outside the Bogotá DC municipality and are connected by interurban buses that are not integrated with the city's system. The mobility survey provides much less detailed information than the internal zones, which we summarize below. Zone 6 includes the following UTAMs: 590, 600, 620, while zone 7 comprises UTAMs 500, 501, 520, 540, 563.

Stratum	Number of households	Share
1	0	0%
2	33,484	62.10%
3	13,120	24.30%
4	3,740	6.90%
5	3,566	6.60%
6	0	0%
Total	53,910	100%

Table 28: Households and share of trips (peak) by socioeconomic stratum in Bogotá's Zone 6.

Variable	**Value**
Trip length	30 km
Car speed	26 km/hr. (peak) 40 km/hr. (off-peak)
Bus speed	26 km/hr. (off-peak)

Table 29: Trip characteristics in Bogotá's Zone 6.

Stratum	Number of households	Share
1	13,386	11.8%
2	82,320	72.6%
3	12,323	10.9%
4	5,398	4.8%
5	0	0%
6	0	0%
Total	113,427	100%

Table 30: Households and share of trips (peak) by socioeconomic stratum in Bogotá's Zone 7.

Variable	**Value**
Trip length	30 km
Car speed	33 km/hr. (peak) 40 km/hr. (off-peak)
Bus speed	23 km/hr. (off-peak)

Table 31: Trip characteristics in Bogotá's Zone 7.

9. Bogotá results

The strategy we pursue is similar to the method for Santiago. First, we present the results for an ‘average’ corridor of the city and then show results for each corridor as if they were independent. Finally, we will integrate them into an overall city planning problem.

We show what can be attained by using subsidies and by implementing a daily pass to obtain an exemption from the driving restriction in place. The ‘average’ corridor is a valuable benchmark because it will allow us to understand insights gained using our spatial network framework. Looking at corridors individually will allow us to understand where that corridor pushes the results.

We, of course, focus on the efficiency of every set of transport policies and the distribution of surplus across income groups and corridors.

9.1 Average corridor

The table below presents the simulation results for an average city in Bogotá. The "Reference" column represents the current situation, which includes a 52% subsidy for public transportation, no congestion pricing, and the implementation of the BRT system Transmilenio. Regular buses, which serve as alternative modes of transportation to the BRT, operate on a separate road alongside cars. As a result, there are two types of public transport systems: the BRT and regular buses.

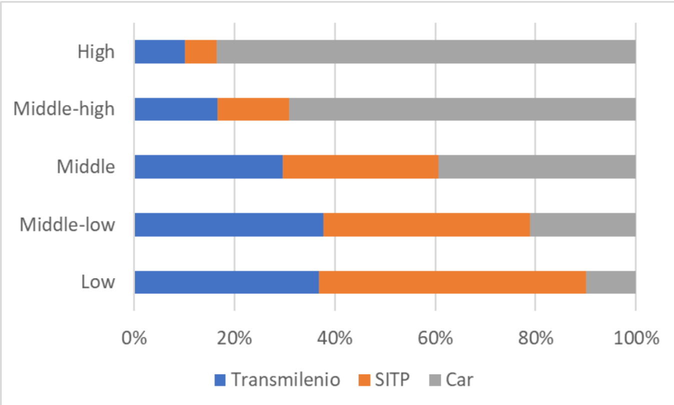
The "SUB" column represents the combination of the optimal fare and frequency for both types of public transport. The "DR + Pass" scenario combines the optimal number of days per week when vehicles are not allowed to circulate with the price of a daily pass that allows for driving on one of those restricted days, while keeping public transport fares at reference levels. The last column combines the optimal driving restriction policy with subsidization.

We find that the optimal fare requires a subsidy of 100%, which is equivalent to making the public transportation system free for riders. The table also shows that the best combination of driving restrictions and passes includes five days of restriction per car. This effectively serves as a form of congestion pricing, as purchasing a daily pass is the only way to use a car. All policies in the table decrease the market share of cars, with the pass being the most effective. In contrast to Santiago, the optimal subsidization policy results in the most significant speed increase for cars and the second largest speed increase for buses. Interestingly, both approaches involve an optimal reduction in the amount of subsidy, as higher speeds reduce the operating costs of public transportation.

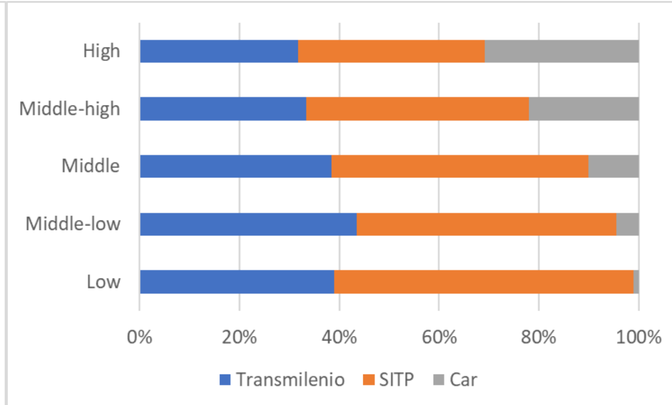
Overall, we find that congestion pricing in the form of driving restrictions with exemption passes complements subsidization. The highest welfare is achieved by making all buses free for riders and restricting driving all weekdays, with the option to pay for an exemption pass. Note that frequencies are always adjusted optimally in our simulations.

	Reference	SUB100	DR + pass	DR + pass + SUB100
Social benefit [$/week]	0	17,085	25,531	31,934
CS change [$/week]	0	30,720	13,888	31,378
Driving restriction [days]	2	2	5	5
Bus fare [$]	0.7	0.0	0.70	0.00
Driving restriction exemption fee [$/day]	-	-	0.80	0.20
Bus (SITP) frequency peak [bus/hr]	10	16	16	16
TransMilenio frequency peak [bus/hr]	13	16	12	12
Bus size [pax]	160	160	160	160
Car speed [km/hr]	13.9	29.3	29.2	28.3
Bus speed [km/hr]	10.2	24.3	24.2	23.2
BRT speed [km/hr]	19.4	16.0	17.5	17.8
Car modal share	27.2%	9.1%	9.3%	12.0%
Bus modal share	41.4%	51.3%	51.3%	51.2%
BRT modal share	31.4%	39.6%	39.3%	36.8%
Subsidy	48%	100%	50%	100%

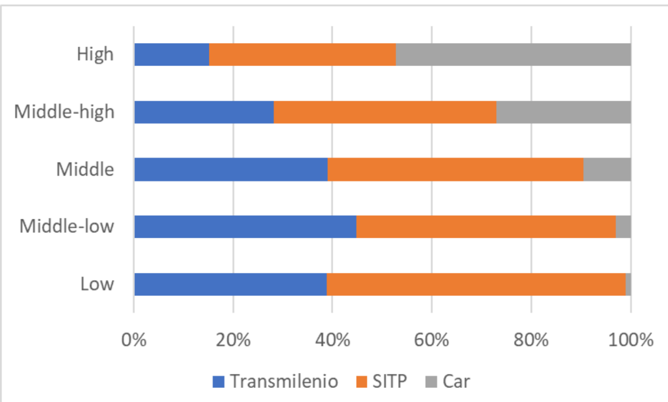
Table 32: Simulation results for an average public transport corridor in Bogotá.



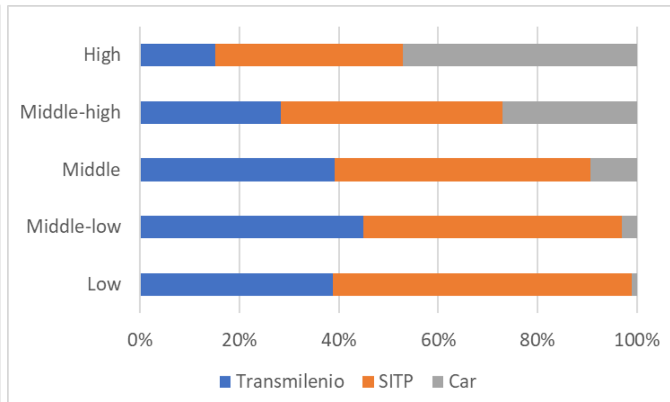
(a) Reference



(b) SUB100



(a) DR + pass



(b)DR + pass + SUB100

Figure 12: Simulation results for an average public transport corridor in Bogotá - Modal shares by income group.

The figure below illustrates the impact of driving restrictions on welfare, while keeping public transport fares at reference levels. It demonstrates the effects of restricting driving from two to five days per week and the impact of various daily exemption pass prices. The key takeaway from this analysis is that setting the appropriate price for the exemption pass is crucial in maximizing the positive welfare effects of driving restrictions. There are significant social benefits of increasing the cost of the pass, up to a certain point - around 0.9 US dollars - beyond which the changes become minimal.

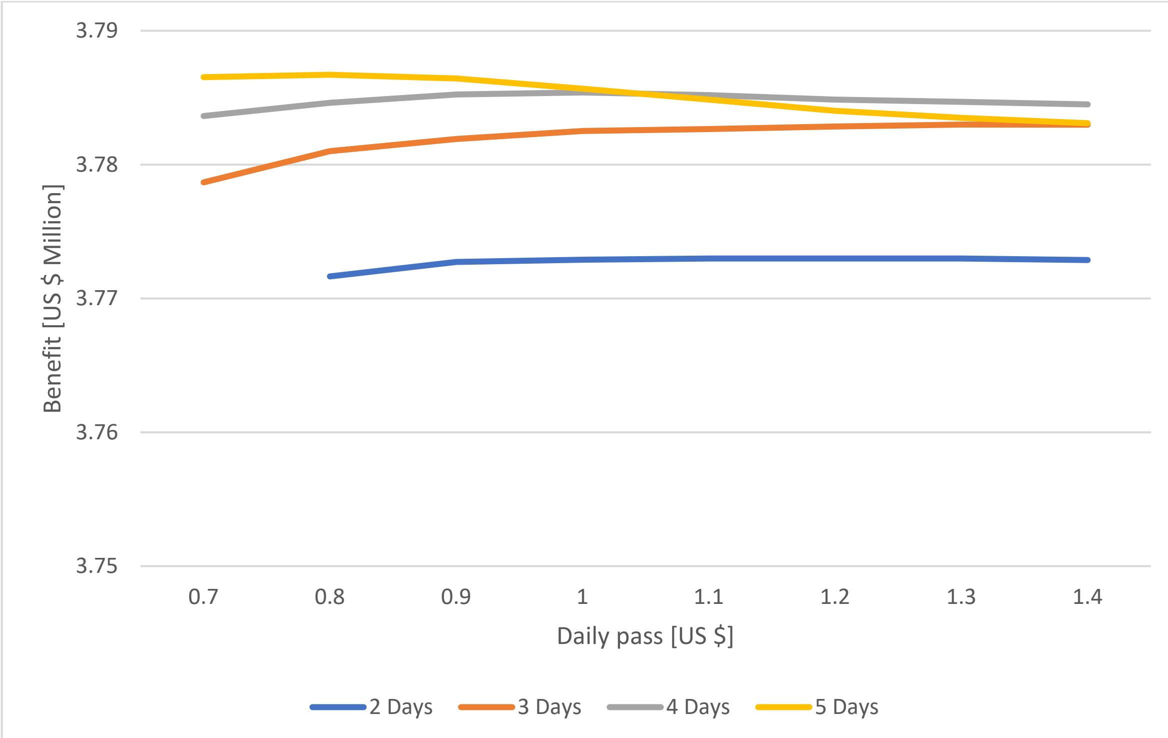
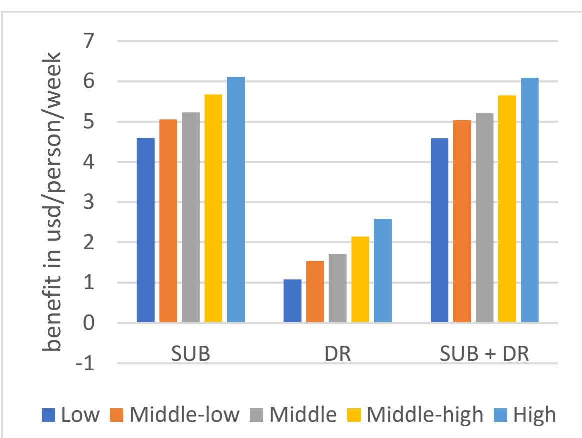


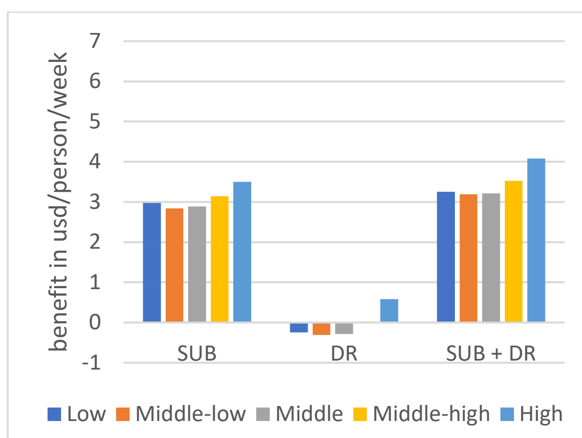
Figure 13: Driving restriction policies - Bogotá.

To conclude the analysis of the average corridor, we look at the distributional impacts of the three policies. This is, we study how each policy impacts the welfare of individuals belonging to different income groups. Furthermore, we differentiate the impact according to the access to the BRT or lack thereof.

Figure 14 summarizes the results of the distributional analysis. Panel (a) shows the impacts on individuals who do not have the BRT as an option and panel (b) for those who do.



(a) Individuals without access to BRT



(b) Individuals with access to BRT

Figure 14: Distributional effects of policies - Bogotá.

The only scenario in which there are losers is when the optimal driving restriction is implemented without public transport subsidization. Figure 13 shows that implementing the five-day restriction with an exemption fee of $ US 0.8, benefits people without access to the BRT regardless of their income. However, except for the wealthiest group, individuals with access to the BRT are worse off with the implementation of the policy.

The intuition behind this result is as follows. The driving restriction effectively reduces car usage through substitution to public transportation. Those who pay to continue driving face decreased road congestion. For them, the losses from paying may be more than compensated by the reduced travel times if their value of time is sufficiently high, which is more likely in the wealthiest groups. The two types of public transportation face different situations. Regular buses in mixed traffic move faster because of reduced congestion, but the BRT system becomes more crowded if frequency is not increased and does not face improved travel times. The result that most of the individuals with access to the BRT are worse off with the driving restriction is explained by the increased crowding from former car users. Those using regular buses are better off as the speed gains are substantial.

The distributional assessment is useful to understand the support that each policy could get but is also key as a basis for extending the analysis to the spatial setting and different corridors.

9.2 Specific corridors

Just as in Santiago, we move to replicate the analysis above but for the different types of suburban corridors that feed the city center. Recall that, as we explained in section 8, these corridors are widely different between them in terms of socioeconomic composition and transport alternatives. The results are summarized in the following tables, that are a short version of the analogous Table 32, and the following figures, that mimic Figure 12. Overall, the results are fully consistent with the previous exercise. As a stand-alone policy, subsidization improves welfare significantly and leads to fare-free buses in all corridors. The optimal driving restriction policy is with five days and an exemption fee between US$ 0.7 and US$ 0.9. It leads to welfare gains that are like those brought by the optimal subsidy.

The analysis of corridors reveals two main new insights. First, it may be that subsidizing optimally the public transport system benefits society more than implementing the driving restriction. This is the case in the wealthiest areas of the city, as illustrated by the corridor 5 that connects the northern part of Bogotá with the center. Naturally, as the average income and value of time is higher, more people are willing to pay the fee that allows them to continue using the car and the congestion relief benefits of the policy are reduced.

As in the average corridor, the largest social benefit is achieved with a combination of driving restrictions and subsidization in corridors 1 to 5. However, in those cases, the optimal exemption fee varies significantly across corridors. As the income composition varies substantially within the city, the price needed to decrease car usage varies substantially as well. Therefore, a driving restriction policy that is uniform across the city entails losses. This is especially true for low-income corridors, where the driving restriction is almost absent (very low exemption fee), with an extreme case in corridor 7. The other side of the coin is corridor 6, where without driving restrictions, subsidization does not work. This happens because this is the richest area of the metropolitan region, where the car use is high and therefore subsidies are not enough to induce modal change.

- *Corridor 1*

	SUB100	DR + pass	DR + pass + SUB100
Social benefit [$/week]	26,360	37,619	43,311
CS change [$/week]	74,268	57,607	75,146
Driving restriction [days]	2	5	5
Bus fare [$]	0.00	0.70	0.00
Driving restriction exemption fee [$/day]	-	0.70	0.03
Subsidy [percent]	100%	47%	100%

Table 33: Simulation results for Bogotá – Transport corridor 1.

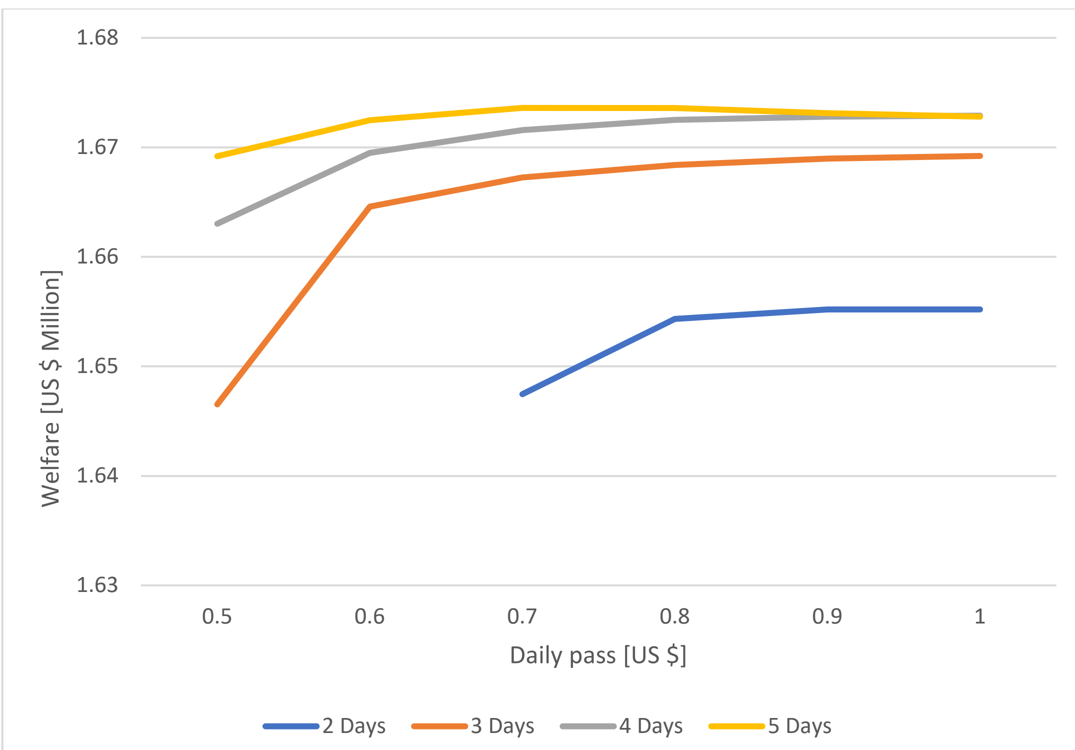


Figure 15: Driving restriction policies – Corridor 1 Bogotá.

- *Corridor 2*

	SUB100	DR + pass	DR + pass + SUB91
Social benefit [$/week]	26,354	37,752	49,389
CS change [$/week]	31,914	14,827	29,103
Driving restriction [days]	2	5	5
Bus fare [$]	0.00	0.70	0.10
Driving restriction exemption fee [$/day]	-	0.90	0.25
Subsidy [percent]	100%	46%	91%

Table 34: Simulation results for Bogotá – Transport corridor 2.

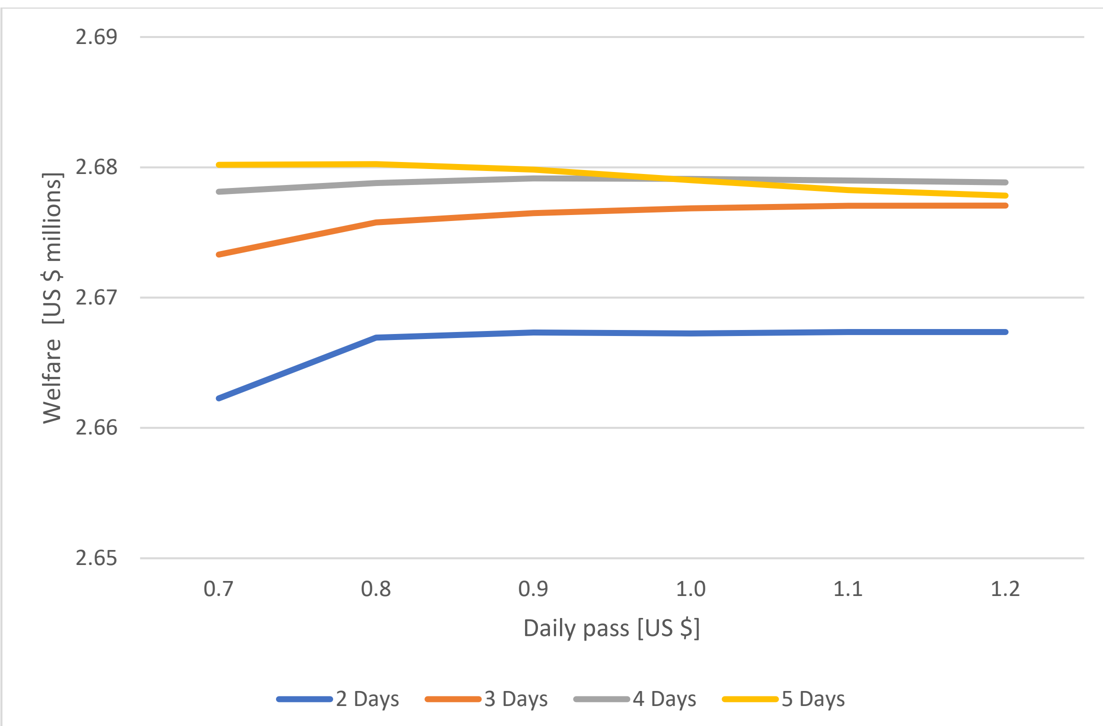


Figure 16: Driving restriction policies – Corridor 2 Bogotá.

- *Corridors 3 and 4*

	SUB100	DR + pass	DR + pass + SUB100
Social benefit [$/week]	29,336	40,577	46,452
CS change [$/week]	34,547	17,253	35,448
Driving restriction [days]	2	5	5
Bus fare [$]	0.00	0.70	0.00
Driving restriction exemption fee [$/day]	-	0.90	0.20
Subsidy [percent]	100%	48%	100%

Table 35: Simulation results for Bogotá – Transport corridors 3 and 4.

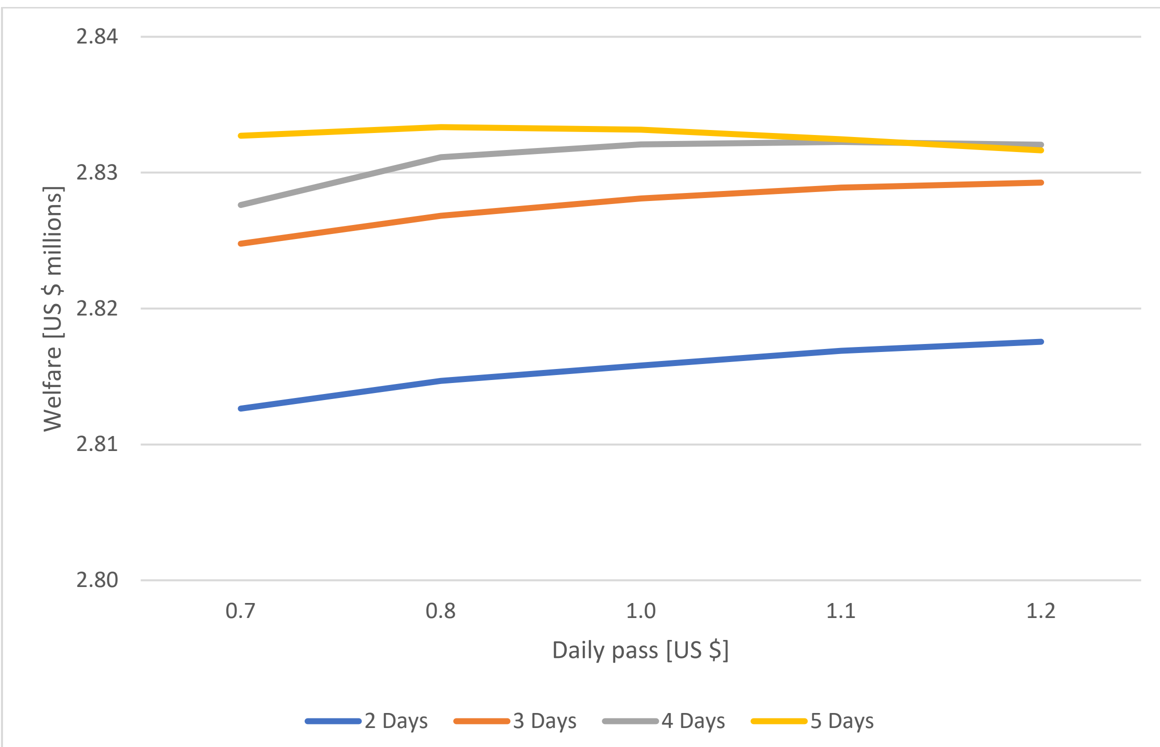


Figure 17: Driving restriction policies – Corridors 3 and 4 Bogotá.

- *Corridor 5*

	SUB100	DR + pass	DR + pass + SUB100
Social benefit [$/week]	34,761	35,525	37,416
CS change [$/week]	30,639	11,837	27,390
Driving restriction [days]	2	5	5
Bus fare [$]	0.00	0.70	0.00
Driving restriction exemption fee [$/day]	-	0.90	0.65
Subsidy [percent]	100%	51%	100%

Table 36: Simulation results for Bogotá – Transport corridor 5.

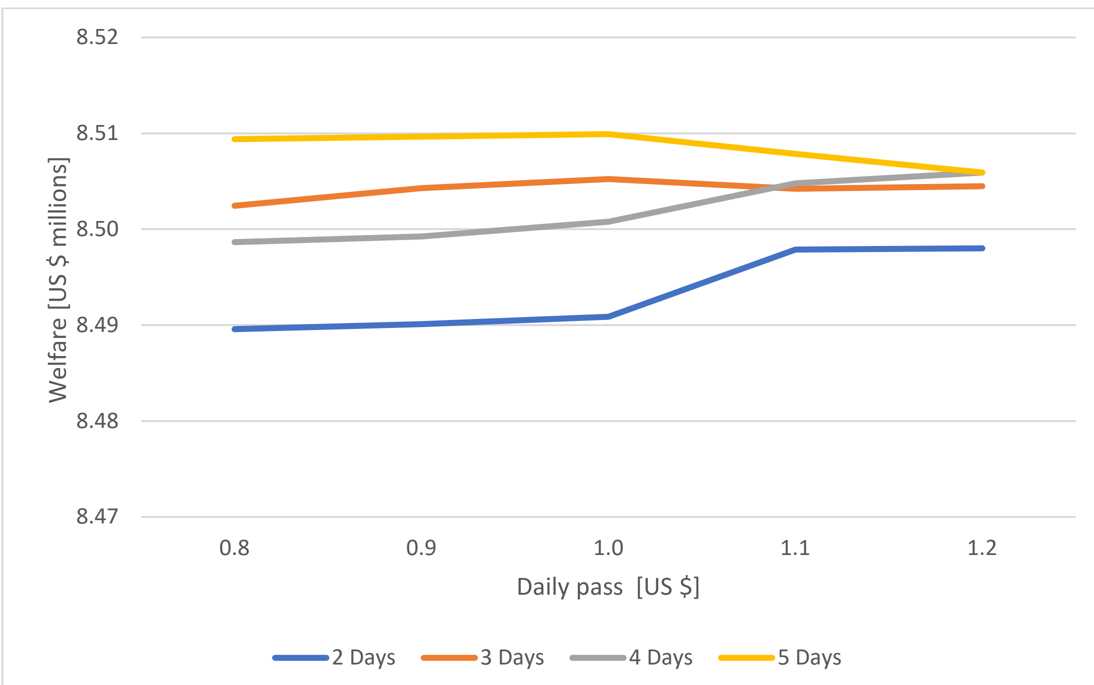


Figure 18: Driving restriction policies – Corridor 5 Bogotá.

- *Corridor 6*

	SUB	Pass	SUB + DR
Social benefit	(5,245)	1,260	7,534
CS change	15,535	3,061	20,435
Driving restriction [days]	2	5	5
Bus fare [US $]	0.2	0.7	0.0
Driving restriction exemption fee [$/day]	0.0	0.7	0.1
Subsidization	88%	47%	100%

Table 37: Simulation results for Bogotá – Transport corridor 6.

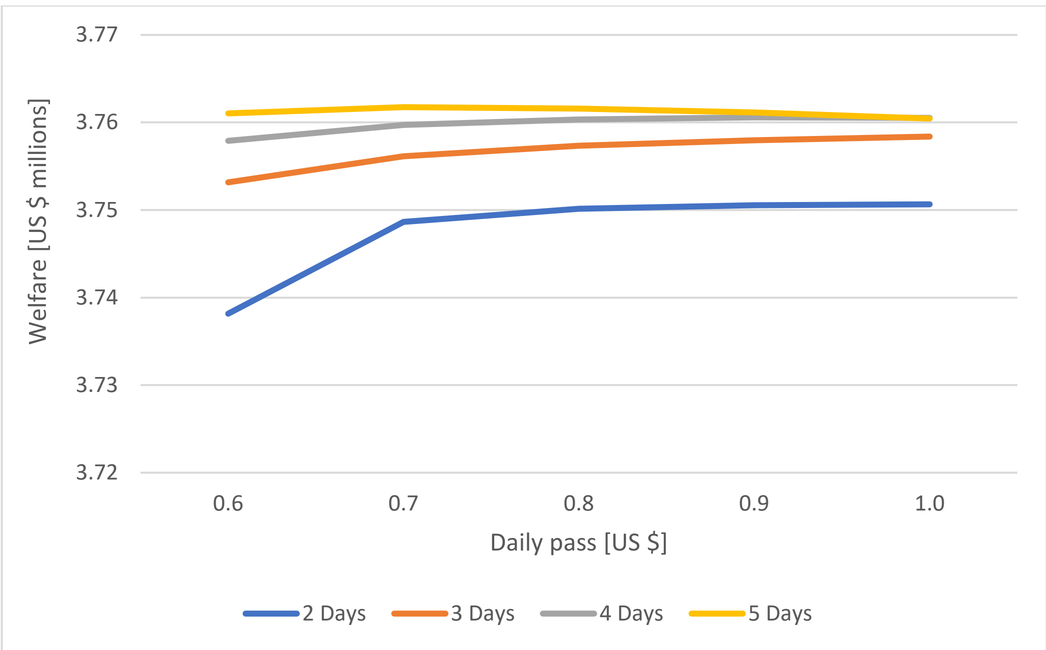


Figure 19: Driving restriction policies – Corridor 6 Bogotá.

- *Corridor 7*

	SUB	Pass	SUB + DR
Social benefit	10,864	3,205	9,249
CS change	23,727	6,228	23,757
Driving restriction [days]	2	5	5
Bus fare [US $]	0.0	0.7	0.0
Driving restriction exemption fee [$/day]	0.0	0.7	0.0
Subsidization [percent]	100%	46%	100%

Table 38: Simulation results for Bogotá – Transport corridor 7.

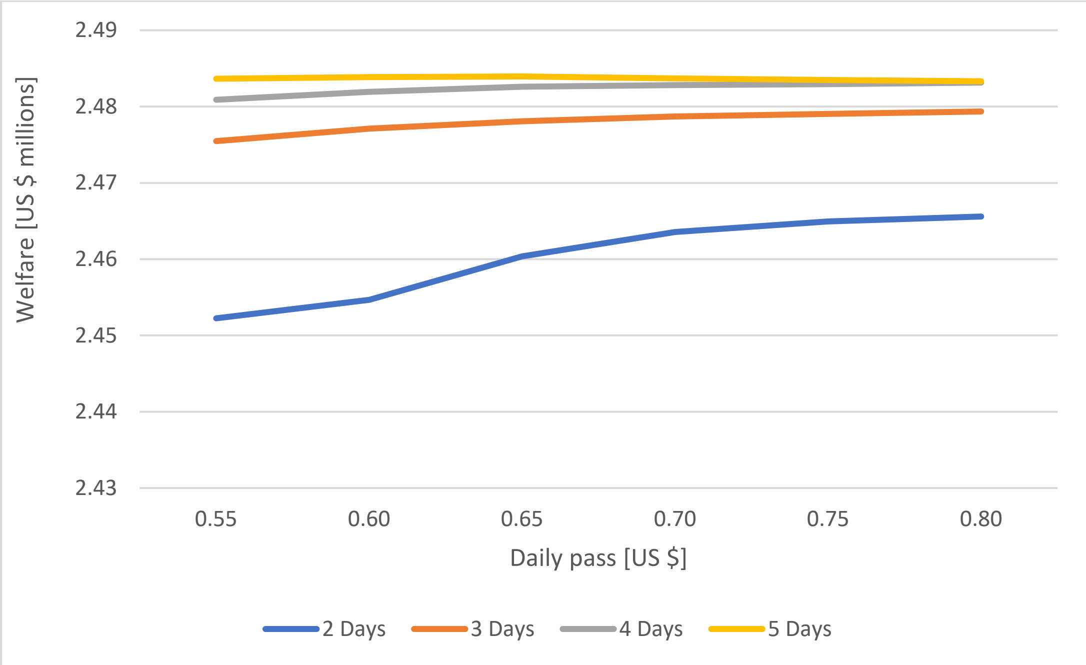


Figure 20: Driving restriction policies – Corridor 7 Bogotá.

The following figure summarizes the social benefit of implementing each of the policies across corridors. The differences in size between corridors is due to differences in population and other characteristics.

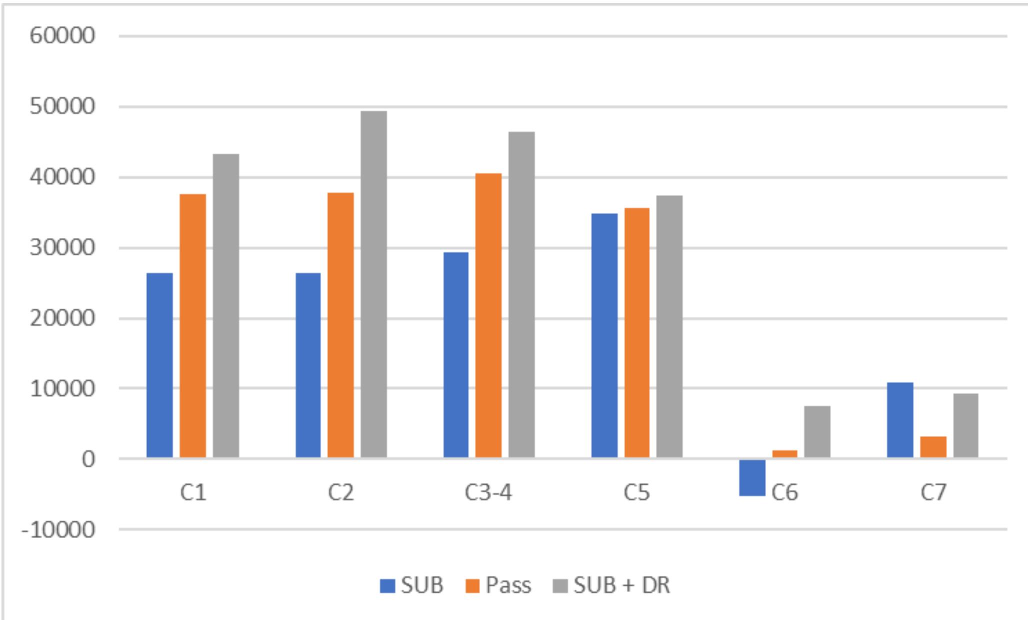


Figure 21: Summary of social benetis across policies and corridors.

9.3 About full subsidization

A perhaps striking result for both the average network and the corridors, is that in all cases and for all combinations, full subsidization –fare free buses—is efficient. Note that, as opposed to Santiago, in Bogotá free buses are optimal even when congestion pricing, that is, five days restriction and with daily passes, is present. This is a good place then to recall that, when it comes to subsidization levels, two issues are important to be brought to light: first, that when a certain amount of subsidization is efficient it does not imply that reaching it is a good project in terms of the alternative use of those public funds; and second, that the actual value is indeed dependent on parameter values.

Regarding the first point, what matters at some point is how much social welfare is added by an extra dollar in subsidization. What authorities want is dollars that produce large increments in welfare while, at some point, increasing subsidies may not induce as much welfare and, from a social point of view, may be better spent elsewhere. What we do to address this issue is, much like in the case of Santiago, increase the subsidy marginally for two cases: the reference case where there is two days of restriction and no daily pass, and the case with full congestion pricing. We do this for the average network of Section 9.1 as an example. Note that we do now that full subsidization for both cases is optimal. The simulation results are shown in Figure 20.

The X-axis shows percentage of subsidization. The Y-axis shows change in welfare. Let us start by considering the 'SUBX' series, that corresponds to the Reference scenario (two days restriction, no pass). Since the reference scenario already considers a 30% subsidy, the blue curve has a 0 value in that point. First, it is obvious that in the reference scenario it is best to fully subsidy, as already discussed in section 9.1. But the slope of the curve shows us the marginal contribution of an additional percentage point of subsidization. The curve then shows that, if the subsidy is increased to 80%, there is a huge gain in welfare, but that the additional 20 percentage points, that lead to full subsidization, have a quite smaller performance, as the curve becomes flatter. Note also that increasing the subsidy up to 60% does not bring in much either, it is only above 65% that the additional benefits start to come in.

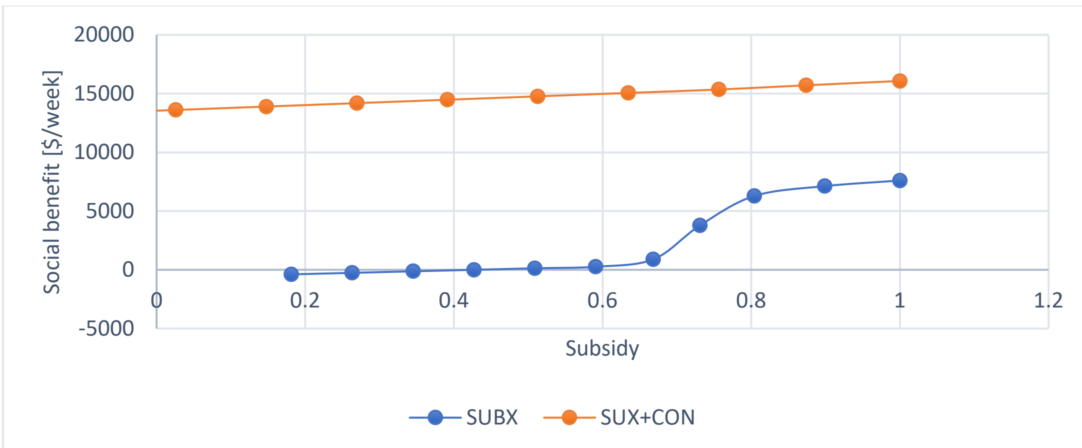


Figure 22: Marginal benefit of increasing subsidies - Bogotá.

We can now look at the 'SUX+CON' series, that corresponds to the congestion pricing case, that is, when a five days restriction is in place, but daily passes are available. Note that at each point of the curve, the value of the daily pass is estimated. What the curve shows is that, as explained before, full subsidization is efficient. However, note that the curve is quite flat, implying the marginal benefit of increasing subsidies is quite low. If one imagines that starting from the reference scenario, full congestion pricing is implemented, the change in welfare is large: start at the blue curve, at 30% subsidy, and then move vertically to the orange curve. The change in welfare is larger than what any increase in subsidy may achieve in the reference scenario, but then, increasing the subsidy will only achieve a small amount of welfare increase. To put it numerically, starting from the reference scenario, implementing congestion pricing achieves an increase in welfare that is seven times larger than what is additionally achieved if full subsidization is added. The social return to that investment may not be attractive to the decision maker.

The second issue to consider is that the actual optimal subsidization is dependent on parameter values. In all cases we used the most sensible parameter values possible but, most likely, the one that is harder to pinpoint is the marginal cost of public funds (MCPF). For the simulation analysis we presented, we used a value of 1.15 but, what if it was actually higher? We repeated the analysis for a MCPF of 1.5. In this case, for the reference scenario, the optimal subsidy is no longer 100% but it is now 85%. One way to picture this is to go back to Figure 20; there, after an 80% subsidy the curve became flatter but was still increasing; what happens as the MCPF grows is that the curve will reach a maximum at 85% and then will actually start decreasing. And for the case when we allow five days of restriction and a daily pass, with MCPF 1.5 the optimal subsidy is actually now nil: once congestion pricing is put in place, any subsidy becomes (slightly) welfare decreasing.

10. Sao Paulo model and calibration

For the analysis of Sao Paulo, we use the same overall model as for Santiago but with a different calibration to reflect Sao Paulo's reality. As explained earlier, due to data limitations we will not be able to perform a spatial network approach. Instead, we can only perform the analysis for what we refer to as a "city center corridor" abstracting from looking at specific corridors and the actual distribution of population in the network.

Calibration

The model parameters are calibrated to reflect Sao Paulo's traffic reality, as captured by the most recent available data. The model can then be used in simulation mode to produce predictions about the possible outcomes of different transport policies. As the model is the same as in Santiago, the calibration also is, and we briefly summarize the process highlighting the main differences if any.

The demand model will be exactly the same, as described in Section 4 and equations (1) to (4). For the BPR-type travel time function (equations (5)–(8)) and for the bus stop delay, we use the same parameters as for Santiago. The car operating costs and the bus operating costs are the cost parameters estimated for Santiago multiplied by the factor that makes the bus fare in Santiago's reference scenario like the observed fare in Sao Paulo in 2019. Finally, observed travel times and speeds come mainly from the Origin-Destination survey of 2017 (STP-2019).

We also bring in the choice of parameters for the value of the marginal cost of public funds (MCPF), the share of congestion pricing revenues that is spent operating the system, and the cost of operating dedicated bus lanes.

Following the Origin-Destination Survey of 2017, commuters are divided in five income groups. We aggregate groups four and five for simplicity. Tables 39 and 40 indicates some relevant characteristics of these groups, including values of ime.

As in BS0214 and the Santiago model, the parameters required to specify the logit models are marginal utilities of income (the cost parameter) and time, and the modal constants. We replicate Santiago's calibration procedure for the logit model but using the values of time reported in table 39. Fares, speeds, and frequencies, needed for calibration, are reported in the tables 41 and 42.

Group (faixa)	Share of total	Average monthly income per household	Value of travel time savings ($US/hr.)
1	21%	<$382	0.8
2	46%	382−763	2.4
3	25%	763−1,526	4.7
4	5%	1,526−2,290	7.8
5	3%	>$2,290	11.5

Source: STP-2019

Table 39. Socioeconomic characteristics – Sao Paulo.

Shares	Group 1	Group 2	Group 3	Groups 4 and 5
Car	29%	40%	60%	74%
Bus	57%	45%	26%	13%
Metro	14%	15%	15%	13%

Table 40. Modal choice by socioeconomic group - Sao Paulo.

Bus fare peak [US $]	1.1
Bus fare off-peak [US $]	1.1
Car toll peak [US $]	0.0
Car toll off-peak [US $]	0.0
Metro fare peak [US $]	1.1
Metro fare off-peak [US $]	1.1

Table 41. Calibration prices for Sao Paulo.

Bus frequency peak [bus/hr]	15
Bus frequency off-peak [bus/hr]	15
Bus size [pax]	160
Metro frequency peak [bus/hr]	7
Metro frequency off-peak [bus/hr]	7

Table 42. Public transport frequencies and bus size for Sao Paulo calibration.

Car speed peak – mixed traffic [km/hr]	16
Bus speed peak – mixed traffic [km/hr]	13
Car speed off-peak – mixed traffic [km/hr]	35
Bus speed off-peak – mixed traffic [km/hr]	25

Table 43. Mixed traffic modal speeds for calibration.

Note that the calibration situation we use is one with no congestion pricing, an operational subsidy to the public transport system of 40%, and no bus priority infrastructure. This last fact may seem contradictory with the fact that Sao Paulo has more than 120 kilometers of BRT. We use mixed traffic speeds for calibration, though, because most BRTs are in a few corridors to the city center or peripheral, and it therefore makes more sense to analyze the policy of 'adding' bus lanes rather than 'eliminitaing' them.

11. Sao Paulo Results

The strategy we pursue is like the strategy applied for Santiago's city center corridor. We begin by discussing stand-alone policies and then turn to the analysis of complementarity or lack thereof between them. The objective function in each case is the same as in Santiago, and given by equation (11).

Table 44 below, summarizes the simulation results for Sao Paulo. In all cases but Sub90*, the Metro fare is restricted to be the same as of the bus; we consider this a political constraint based on the observed situation.

The 'Reference' column represents a situation where there is a 40% subsidy, no congestion pricing, and no bus lanes. The 'Sub90' column corresponds to a situation where subsides are bumped up to 90% (the optimum). The 'Sub90*' corresponds to 90% subsidy but bus and metro fares are no longer tied together. Again, it is welfare-maximizing to increase the subsidy to its maximum feasible value. The last two columns correspond to an implementation of congestion pricing and bus lanes/BRT respectively, when nothing else is in place, that is, when subsidy is downturned to 0% (but the Metro fare remains tied to the bus fare).

We observe that with 90% subsidies optimal bus frequencies increase around 35%, but with congestion pricing or bus lanes, each implemented separately, the increase is substantially larger. The Metro frequencies vary, but not strongly, and sometimes decrease. This is because the coverage of the system is low. The optimal bus size remains rather unchanged as different policies are implemented.

Regarding prices, when bus and metro fares cannot differ, they are both very low and the optimum subsidy reaches 90%. If the price of the subway can be different, then the subsidy is also optimally set to 90%, but the subway fare doubles while buses are now fare-free. Congestion pricing and bus lanes, without subsidies, have fares that are very similar as in the reference scenario, that has 40% subsidy. In other words, the speed increase that these two policies achieve, save enough fleet costs that allows fares to be as if there was nothing else than the current subsidy.

With respect to the reference scenario, all policies achieve a decrease in car market share, with congestion pricing being the most effective at this. Congestion pricing also achieves the largest speed increase for cars in the peak and the second largest speed increase for buses. Interestingly, bus lanes achieve a very large increase in speed for buses –comparable to that of congestion pricing--, at a minumu cost in decreased car speed. In other words, bus lanes as a stand alone policy achieves the highest welfare level and bus speeds, without the need to externally affect prices.

Are the policies complements or substitutes? Figure 23 sheds light. First it indeed shows that, without subsidies, bus lanes are a better policy by itself than congestion pricing (vaoue at x=0), but we can further see that, if you start adding subsidies to either policy, the marginal benefit is close to negligible. In that sense, once either is implement, the welfare return to a dollar of subsidy is close to zero. Also, the curve of congestion pricing plus bus lanes sits right on top of that of bus lanes, proving that once bus lanes have caused their effect, an additional price instrument does not seem that necessary. The figure also shows the decreasing returns of extra dollars of subsidy if only subsidization is in place.

	Reference	SUB90	SUB90*	CON	DL
Social benefit [$/day-km]	0	10,465	12,438	16,639	19,144
CS change [$/day-km]	0	18,707	19,735	8,363	11,053
Bus fare peak [$/km]	0.09	0.01	0.00	0.10	0.09
Bus fare off-peak [$/km]	0.09	0.01	0.00	0.10	0.09
Car toll peak [$/km]	0.00	0.00	0.00	0.05	0.00
Car toll off-peak [$/km]	0.00	0.00	0.00	0.00	0.00
Metro fare peak [$/km]	0.09	0.01	0.10	0.10	0.09
Metro fare off-peak [$/km]	0.09	0.01	0.21	0.10	0.09
Bus frequency peak [bus/hr]	22.4	29.7	30.7	42.2	37.7
Bus frequency off-peak [bus/hr]	22.4	29.7	30.7	31.0	37.5
Bus size [pax]	163	160	163	151	145.0
Metro frequency peak [bus/hr]	7.5	6.5	3.8	2.5	1.8
Metro frequency off-peak [bus/hr]	7.5	6.5	3.8	2.5	1.8
Car speed peak [km/hr]	21.6	34.2	33.3	58.4	19.8
Bus speed peak [km/hr]	18.9	17.6	17.2	27.1	27.8
Car speed off-peak [km/hr]	32.2	55.9	55.2	51.0	41.2
Bus speed off-peak [km/hr]	28.5	29.6	29.0	28.0	32.8
Peak share [percent]	45.0%	47.9%	47.8%	49.2%	49.4%
Off-peak share [percent]	52.6%	50.1%	50.2%	48.6%	48.4%
No-travel share [percent]	2.4%	2.0%	2.0%	2.2%	2.2%
Car modal share peak [percent]	36.2%	26.7%	27.1%	9.7%	23.6%
Bus modal share peak [percent]	52.4%	63.9%	67.5%	83.3%	71.3%
Metro modal share peak [percent]	11.4%	9.3%	5.4%	7.0%	5.1%
Car modal share off-peak [percent]	34.4%	27.1%	28.2%	35.5%	33.2%
Bus modal share off-peak [percent]	55.0%	61.6%	67.5%	55.7%	58.8%
Metro modal share off-peak [percent]	10.6%	11.2%	4.3%	8.7%	8.0%

Table 44. Simulation results for an average public transport corridor in Sao Paulo.

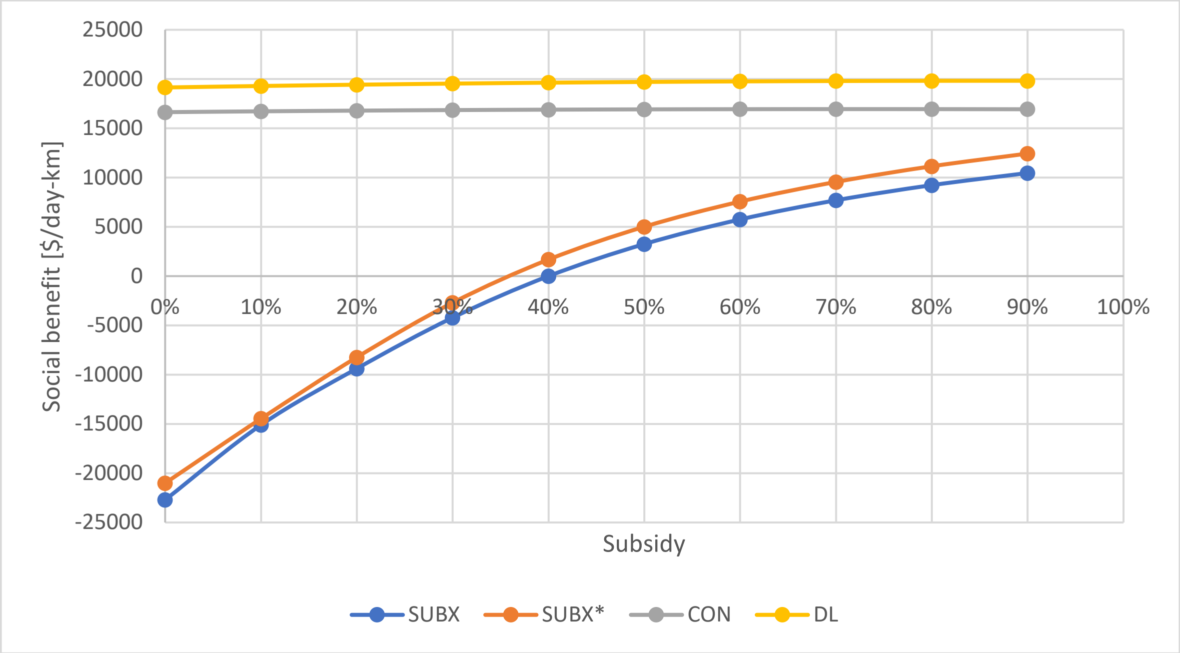


Figure 23. Welfare for different transport policies -- average public transport corridor in Sao Paulo.

12. Discussion and conclusions

Urban sustainability requires reducing the environmental footprint of urban mobility. To achieve this goal, public transport becomes key due to its lower impact on congestion, pollution, accidents, and greenhouse gas emissions as compared to private transport. Thus, encouraging public transport use is a frequent goal among city authorities. For this, they often devote an important fraction of their budget to improve the quality of its service. Many cities subsidize not just the infrastructure for public transport, but also its operation. Subsidization figures vary widely around the globe: they are rather large in the developed world (with an overall average well above 55%) while subsidies are less common in the developing world, particularly in Latin-America,

Congestion pricing is another possible way to deal with congestion, reducing the environmental footprint of urban mobility by reducing the number of car trips and transferring travelers to public transport or to teleworking, while creating revenue that can be used to further improve the transit system. Congestion pricing has been advocated for decades by experts, but it has been applied only in a few cities, most possibly because it is perceived as regressive, in that only richer people will have the chance to drive, while the middle class, that had already a hard time buying a car, will be forcefully moved to the public transport. It is in fact because of this perceived unfairness that many countries have turned to rationing schemes or driving restrictions, in which cars are forbidden to circulate based on their license plate. This is the case of Bogota's *Pico y Placa*, or Mexico City's *Hoy No Circula* program.

And then, an additional alternative, which is far less demanding financially than subsidies and less contentious that congestion pricing is using part of the existing road capacity exclusively for buses. These bus lanes which, through some investment, physically separate buses from cars, are sometimes known as Bus Rapid Transit (BRT) systems. A BRT system is "a high-quality bus-based transit system that delivers fast, comfortable, and cost-effective services at metro-level capacities. It does this through the provision of dedicated lanes, with busways and iconic stations typically aligned to the center of the road, off-board fare collection, and fast and frequent operations". From the first system in Curitiba, Brazil, in 1977, the penetration of BRT systems has been increasing fast: by 2018, 170 cities around the world had BRT systems, for a total of 376 corridors and 5,046 km, while 121 additional cities are either building or have plans to build BRT systems. A regional panorama shows that Latin America leads with 171 corridors in 55 cities.

Thus, for the goal of reducing the environmental footprint of urban mobility one may think of many different policies, such as improving vehicle efficiency (e.g. promoting electric vehicles and buses) or to improve land use and transportation integration (e.g. transport oriented development). If in addition managing congestion is also a goal, there seems to be three additional tools: subsidies, congestion pricing or driving restrictions, and dedicated bus lanes / BRT. And, for the case of Latin-America, there seems to be that subsidies have not been the preferred tool, while BRTs lead the way.

This study attempts to provide and assessment on the efficiency of these tools for Latin-American cities, and how they ‘mix and match’. For this, we create urban/transport equilibrium models for Santiago, Chile, and Bogotá, Colombia. Our work is firmly based on Basso and Silva (2014) and Basso, Montero and Sepúlveda (2021), who proposes transport equilibrium models that can be calibrated with real data from a city, and then used in simulation mode to study different transport policies alone or in combination. But, as opposed to the work in these two papers, where ‘representative networks’ (city center like) are considered, we expand both models to a spatial setting that better represents the geographical differences in income and public transport alternatives, and which allows us to analyze a wide array of possible urban transport policies and their combinations, looking at both efficiency and distributional aspects, with a particular spatial emphasis.

Using the most actual data to calibrate the Santiago model, the analysis shows that, when the city center is modeled alone, as has been usually the case in the literature and was the case in BS2104 and BMS 2021, we observe that with optimal subsidies, congestion pricing or bus lanes, each implemented separately, frequencies increase and the car market share decrease. When bus and metro fares cannot be too different, they are both very low and the optimum subsidy reaches 90%. If the price of the subway can be very different, then the subsidy is decreased optimally to 80%, the subway has fares like the ones in the reference scenario, and buses are now free. Congestion pricing and bus lanes, without subsidies, have fares larger that the reference scenario, that has 40% subsidy.

Results show that the best stand-alone policy in the city center network is building a BRT line, like what was obtained by BS2014. However, bus lanes in the city center create heavy congestion for cars, which might make them implausible to implement. Congestion pricing, on the other hand achieves the largest speed increase for cars in the peak and the second largest speed increase for buses. Adding subsidies to congestion pricing or bus lanes decrease both welfare and consumer surplus. In that sense, if the subsidy level is not to be touched for distributional reasons, bus lanes and congestion pricing will still improve the performance.

When we add space, and consider corridors that feed the city center, what is optimal to do depends on the way the corridor is structured. Yet there are general insights to be gained. First, in all cases, implementing congestion pricing in both the corridors and the city center will improve welfare. In fact, the welfare optimum is with congestion pricing and no subsidies, but that comes at the expense of higher fares and decreased consumer surplus. Increasing or maintaining subsidies will decrease welfare but increase consumer surplus. In terms of mixed traffic vs bus lanes, building bus lanes in corridors are always welfare increasing, adding over any level of subsidy. When there is metro available in the corridor, bus lanes generate benefits, but are somewhat milder, possibly given the existence of an already high-speed mode.

We then conducted the simulation analysis to determine the most effective policies for reducing car usage and improving public transportation in Bogotá, Colombia. The simulation included several scenarios, including a "reference" scenario representing the situation in 2019, that is, with a 33% subsidy for public transportation, a two day restriction without the possibility of paying to avoid it, and with the existence of both the BRT system and a mixed-traffic bus system; a "SUB" scenario that combines the optimal fare and frequency for both types of public transportation; and another scenario combines driving restrictions with the price of a daily pass that allows for driving on one of the restricted days. The final scenario combines the optimal driving restriction policy with subsidization.

The results of the simulation showed that the optimal fare for public transportation required a subsidy of 100% or making the system free for riders. The combination of driving restrictions and passes was also found to be effective, with the best results coming from restricting driving five days per week and allowing the purchase of an exemption pass. All the policies resulted in a large decrease in the market share of cars, much larger than in Santiago, with the pass being the most effective. We believe that the difference is because the Santiago model has more margins for substitution. Unlike in Bogotá, in Santiago we model intertemporal substitution (peak vs. off-peak) and the total demand is imperfectly elastic. For the same reason, both the optimal subsidization and driving restriction policies resulted in higher speeds for both cars and buses.

Setting the appropriate price for the exemption pass is key to maximizing the benefits of the policy, yet the spatial analysis showed that the average optimal daily pass may hide wide variations. This indeed stresses the system, as an average pass which is independent of space will indeed hurt the poor for the benefit of the rich.

Unlike in Santiago, public transport subsidization and car congestion pricing complement each other to reduce car ridership, congestion and to maximize social welfare. Congestion pricing and transit subsidies are perfect substitutes when the marginal cost of public funds is one, there are only two modes, the total demand is inelastic and there is only one income group (as in Basso and Jara-Díaz, 2012). When any of those modeling assumptions is lifted, the two policies may become imperfect substitutes or even complements. What we believe happens in Bogotá is that, on one hand, the 1.15USD spent to provide a 1USD in subsidy has large returns because it not only decreases the fare, but also induces large congestion relief and higher frequencies given the larger effect on modal shift, already discussed. But, on the other hand there are several income groups that differ in their marginal valuations of time and income, and therefore using more than one price is useful to increase welfare.

Additionally. note that the analysis of the marginal benefit of increased subsidization showed that, even if efficiency increases, it does so at a smaller rate above 80% when congestion pricing is not there, and quite slowly in general after congestion pricing has been implemented. This implies that the social return to an investment in increased subsidies may not be attractive to the decision maker at all subsidy levels, despite it increases welfare. Finally, a word of caution on the sensitivity of results to parameter values is always warranted.

Finally, we performed a non-spatial analysis of Sao Paulo. Evidently, this analysis is less deep than the previous two due to data restrictions, but it does provide us with results that, by now, seem to appear often. First, it shows that bus lanes/BRT policies perform very well. They here reach high levels of service for buses, without a huge costs on car congestion, while inducing important cost savings through fleet size, which allows fares to be kept at bay, without the need for subsidies. All this implies the highest possible welfare gain. Congestion pricing is a good policy as well, which keeps the best level of service for the car. If only subsidies are used, we again reach high values, 90%, which imply very low fares for the subway and the bus if fares are to tied together or, as in Santiago, fare free buses and positive Metro fares if they are allowed to be different. The logic is that full fares are more alike: one has high speeds and a positive fare, the other is free but slower as it rolls in mixed traffic conditions.

As a general conclusion and paths for future research, we believe that the spatial netowrk approach has unveiled that many of these transport policies, which are usually thought and implement at the level of the city, hide huge distributional impacts that need to be studied. Furthermore, spatial economic transport models like the ones used here may be used to assess the impacts of spatially differentiated policies, be them through prices or capacity decisions. Indeed, though, any spatially differentiated policy may have impacts on lthe medium to long run interms of land prices and location. Urban economic models mixed with models such as the ones used here seem to be a very relevant path for future work

13. References

Barahona, N., F. Gallego and J.-P. Montero (2020). Vintage-specific driving restrictions. Review of Economic Studies, 87(4): 1646-1682.

Basso, LJ., Guevara, CA., Gschwender, A. & Fuster, M. (2011). Congestion pricing, transit subsidies and dedicated bus lanes: Efficient and practical solutions to congestion. Transport Policy 18 (5): 676–84.

Basso, Leonardo J., and Sergio R. Jara-Díaz, SR. (2012). Integrating congestion pricing, transit subsidies and mode choice. Transportation Research Part A: Policy and Practice 46 (6): 890–900

Basso, L. J., & Silva, H. E. (2014). Efficiency and substitutability of transit subsidies and other urban transport policies. American Economic Journal: Economic Policy, 6(4), 1-33.

Basso, L. J., Feres, F., & Silva, H. E. (2019). The efficiency of bus rapid transit (BRT) systems: A dynamic congestion approach. Transportation Research Part B: Methodological, 127, 47-71.

Basso, L.J., Montero, J.P. Sepulveda, F. (2021). A practical approach for curbing congestion and air pollution: Driving restrictions with toll and vintage exemptions. Transportation Research Part A, 148, 330–352.

Ben-Akiva, Moshe Emanual. 1973. Structure of Passenger Travel Demand Models. PhD dissertation. Massachusetts Institute of Technology.

Börjesson, M., Fung, C. M., & Proost, S. (2017). Optimal prices and frequencies for buses in Stockholm. Economics of transportation, 9, 20-36.

Börjesson, M., Fung, C. M., Proost, S., & Yan, Z. (2018). Do buses hinder cyclists or is it the other way around? Optimal bus fares, bus stops and cycling tolls. Transportation research part A: policy and practice, 111, 326-346.

Börjesson, M., Fung, C. M., Proost, S., & Yan, Z. (2019). Do small cities need more public transport subsidies than big cities?. Journal of Transport Economics and Policy (JTEP), 53(4), 275-298.

Daganzo, C. (2000). Taking turns: Rx for Congestion, Access 17, 14-19.

Davis, L. (2008), The effect of driving restrictions on air quality in Mexico City, Journal of Political Economy 116, 38-81.

De Borger, Bruno, Inge Mayeres, Stef Proost, and Sandra Wouters. 1996. “Optimal Pricing of Urban Passenger Transport: A Simulation Exercise for Belgium.” Journal of Transport Economics and Policy 30 (1): 31–54.

Eskeland, G. S. and T. Feyzioglu (1997), Rationing can backfire: The day without a car in Mexico City, World Bank Economic Review 11, 383-408.

Gagnepain, Philippe, and Marc Ivaldi. 2002. "Incentive regulatory policies: The case of public transit systems in France." RAND Journal of Economics 33 (4): 605–29.

Gallego, F., J.-P. Montero and C. Salas (2013), The effect of transport policies on car use: Evidence from Latin American cities, Journal of Public Economics 107, 47-62.

Glaister, Stephen, and Davis Lewis. 1978. "An integrated fares policy for transport in London." Journal of Public Economics 9 (3): 341–55.

Gonzales, E.J., Daganzo, C.F. (2012). Morning commute with competing modes and distributed demand: user equilibrium, system optimum, and pricing. Transportation Research Part B 46 (10), 1519–1534.

Huang, HJ. (2000). Fares and tolls in a competitive system with transit and highway: the case with two groups of commuters. Transportation Research Part E: Logistics and Transportation Review 36 (4): 267–84.

Huang, H.J., Tian, Q., Yang, H., Gao, Z.Y. (2007). Modal split and commuting pattern on a bottleneck-constrained highway. Transp. Res. Part E 43 (5), 578–590.

Gómez-Gelvez, J.A. y Mojica, C.H. (2022) " Subsidios al Transporte Público en América Latina desde una Perspectiva de Eficiencia: Aplicación a Bogotá, Colombia", Documento de trabajo del BID, 1352.

Kilani, M., Proost, S., & Van der Loo, S. (2014). Road pricing and public transport pricing reform in Paris: complements or substitutes?. Economics of Transportation, 3(2), 175-187.

Kraus, M., Yoshida, Y. (2002). The commuter's time-of-use decision and optimal pricing and service in urban mass transit. Journal of Urban Economics 51 (1), 170–195.

Kutzbach, MJ. (2009). Motorization in developing countries: Causes, consequences, and effectiveness of policy options. Journal of Urban Economics 65 (2): 154–66.

Mohring, H. (1979). The Benefits of Reserved Bus Lanes, Mass Transit Subsidies, and Marginal Cost Pricing in Alleviating Traffic Congestion. In: Mieszkowski, P., Straszheim, M. (Eds.), Current Issues in Urban Economics. The Johns Hopkins University Press, Baltimore, pp. 165–195 .

Parry, I. W. H., and K. A. Small. 2009. "Should Urban Transit Subsidies Be Reduced?" *American Economic Review,* 99 (3): 700–724.

Pavón, N., & Rizzi, L. I. (2019). Road infrastructure and public bus transport service provision under different funding schemes: A simulation analysis. Transportation Research Part A: Policy and Practice, 125, 89-105.

Proost, Stef, and Kurt Van Dender. 2008. "Optimal urban transport pricing in the presence of congestion, economies of density and costly public funds." Transportation Research Part A: Policy and Practice 42 (9): 1220–30.

Small, Kenneth A. 1983. "Bus priority and congestion pricing on urban expressways." Research in Transportation Economics 1: 27–74.

Small, K.A., and Verhoef, E.T., 2007. The economics of urban transportation. Routledge.

Tirachini, A., Hensher, D. A., & Rose, J. M. (2014). Multimodal pricing and optimal design of urban public transport: The interplay between traffic congestion and bus crowding. Transportation Research Part B: Methodological, 61, 33-54.

Tirachini, A., Hurtubia, R., Dekker, T. & Daziano, RA. (2017) Estimation of crowding discomfort in public transport: Results from Santiago de Chile. Transportation Research Part A: Policy and Practice 103, 311-326.

Viton, Philip A. 1983. "Pareto-optimal Urban Transportation Equilibria." Research in Transportation Economics 1: 75–101.

14. Appendix A – Short summaries of relevant papers

Multimodal Pricing and Optimal Design of Urban Public Transport: The Interplay between Traffic Congestion and Bus Crowding (Alejandro Tirachini, David A.Hensher, and John M.Rose, 2014)

- Journal: Transportation Research Part B: Methodological
- Setting: The interplay between congestion and crowding externalities in the design of urban bus systems is identified and analyzed.
- Data: Single transport corridor in Sydney, Australia.
- Strategy: A multimodal social welfare maximization model with spatially disaggregated demand is developed, in which users choose between travelling by bus, car or walking in a transport corridor. Optimization variables are bus fare, congestion toll, bus frequency, bus size, fare collection system, bus boarding policy and the number of seats inside buses. The authors introduce two trade-offs within the microeconomic modelling framework of optimal supply levels and pricing of an urban bus route, that are crucial to the level of service offered to public transport users: (i) the interaction between road congestion and passenger crowding externalities when setting supply levels of public transport, and (ii) the decision on the number of seats that public transport vehicles should have.

Results

They find that optimal bus frequency results from a trade-off between the level of congestion inside buses, i.e., passengers' crowding, and the level of congestion outside buses, i.e., the effect of frequency on slowing down both buses and cars in mixed-traffic roads.

Optimal bus frequency is quite sensitive to the assumptions regarding crowding costs, impact of buses on traffic congestion and the overall congestion level. In particular, if the planner takes into account that crowding matters to users, our numerical application shows that bus frequency should increase (for a given bus size) with demand even under heavy congestion, however that might not be the case if the crowding externality is not accounted for, in which case an increase of total demand might be met by a decrease of both frequency and number of seats per bus, at the expense of crowding passengers inside buses and making more passengers stand while travelling.

The consideration of crowding externalities (on both seating and standing) imposes a sizeable increase in the optimal bus fare, and consequently, a reduction of the optimal bus subsidy.

Contributions

Their model considers that reducing the number of seats increases bus capacity by allowing more standees; thus, a seat implies a trade-off between comfort and capacity.

In contrast to other social welfare maximization models, their approach provides a more comprehensive modelling of the bus mode, including bus frequency, bus size, fare collection system, bus boarding policy, number of bus seats and fare level as decision variables.

They also show that the inclusion of a non-motorized mode (walking) as an alternative to choosing bus and car for short trips may have a significant role when the transport system is optimized in highly congested scenarios.

Economic and distributional effects of different fare schemes: Evidence from the Metropolitan Region of Barcelona (Anna Matas, Josep-Lluis Raymond, and Adriana Ruiz, 2020

- Journal: Transportation Research Part A
- Setting: The authors provide evidence for how switching from flat to distance fares or from integrated (to promote public transport use) to non-integrated tickets affects both the ridership and the financial situation of public transport companies.
- Data: Metropolitan Region of Barcelona (Transport Authority in Barcelona introduced a multimodal integrated fare system in 2001, with a zone fare structure).
- Strategy: They estimate a probabilistic modal choice equation between public and car transport and using the estimated equation to simulate the consequences of alternative fare systems. For each individual in the sample, they calculated the fare under the current and the simulated schemes, as well as the corresponding subsidy in each scenario. The subsidy was measured as the difference between the fare paid and the operating costs per passenger of the transport modes used.

Results

They show that different pricing structures (flat fares, distance-based and integrated tickets) have only a moderate effect on ridership, while the potential for revenue changes is higher.
The distributional profiles of alternative pricing strategies are quite homogeneous. However, there appears to be a mild regressive effect when an integrated fare system is removed.

Contributions

This paper confirms that travelers are more sensitive to changes in quality –mainly waiting and access time – than to changes in prices.

Transit reforms in intermediate cities of Colombia: An ex-post evaluation (Andrés Gómez-Lobo, 2020)

- Journal: Transportation Research Part A: Policy and Practice
- Setting: They use monthly data on transit supply and ridership to evaluate the impact of BRT type reforms in intermediate cities in Colombia.
- Data: Monthly data on transit supply and ridership (Colombia).
- Strategy: Since not all cities implemented a transit reform nor, in the cities that did, at the same time, we are able to use a staggered difference in difference panel data model to estimate the impact of these reforms.

Results

They find that these reforms are associated with a decrease in aggregate transit ridership. They show that reform reduced fleet size and commercial kilometers supplied and they conjecture that this, together with additional transfers required in the new systems, raised the generalized cost of transport for transit services.

Contributions

This paper is the first attempt to evaluate the transit reforms in intermediate cities of Colombia using rigorous econometric techniques.

Different result from Basso and Silva (2014). However, in Basso and Silva (2014), bus only lanes also increase welfare in their analysis of Santiago and London but in this case bus fares are endogenous and therefore adjust to fund the optimal fleet size.

Distributional Effects of Public Transport Subsidies (Maria Börjesson, Jonas Eliasson, and Isak Rubensson, 2020)

- Journal: Journal of Transport Geography.
- Setting: This paper analyze the distribution of transit subsidies across population groups in Stockholm.
- Data: Stockholm
- Strategy: The authors develop a novel methodology that considers that the subsidy per passenger varies across transit links, since production costs and load factors vary. With this, they calculate the subsidy per trip in the transit network and analyze the distribution of subsidies across population groups. Then, they calculate the concentration index to explore the distribution of subsidies across income groups.

Results

The transit subsidies are mildly progressive in Stockholm, to a large extent due to discounts for students and retired, but also because the citizens in the top income quintile make fewer transit trips per person. Still, the progressivity is weak because a wide range of income groups get roughly equal subsidies.Transit subsidies is hence not effective as a redistribution policy in Stockholm.

The largest systematic variation they find is across residential areas: the average subsidy per person is five times higher in the peripheral areas of the region compared to the regional core, and the subsidy per trip is ten times higher.

Contributions

Basso and Silva (2014) find that low-income groups gain from optimal transport subsidies compared to the baseline. However, the perspective of Börjesson et al. (2020) differs to the prior authors because their starting point is the subsidy to each individual service. Therefore, services with high occupancy levels need lower subsidies (or even generate profits), while services with low occupancy need higher subsidies. This might well mean that high-income people residing in single-family houses with lower densities receive higher subsidies than average.

The main contribution of this paper is to develop a methodology for empirically computing the actual distribution of subsidies across different groups and individuals, which is different from the distribution of transit supply, fare structure or trip frequencies.

The efficiency of bus rapid transit (BRT) systems: A dynamic congestion approach (Leonardo J.Basso, Fernando Feres, and Hugo E.Silva, 2019)

- Journal: Transportation Research Part B.
- Setting: Dynamic congestion approach, which is equipped to model queuing endogenously, both on the road and at BRT stations.
- Data: Numerical simulations.
- Strategy: Commuters travel from a single residential area to the city center and must choose to either drive or take public transportation, together with the departure time, which makes schedule delays important. If a commuter decides to travel by car, she may face road congestion. If she chooses to use public transport instead, she will need to go to a station where she may face boarding delays caused by queues. The bus will then go into the road, where it may join the queue of cars if traffic is mixed (without BRT), or it may not face road queuing if part of the capacity is devoted to a BRT (which decreases the capacity for cars). They focus on second-best policies, both for mixed traffic conditions and BRT, meaning that we consider that fares and tolls are time-invariant, and the public transport system operates at a constant headway

<u>Results</u>

They show analytically that, if capacity is perfectly divisible, implementing a BRT is always efficient (it decreases total social cost), while if capacity is not perfectly divisible, a BRT is in most cases efficient. Moreover, BRT can induce a Pareto Improvement where both time costs and public transport operating costs decrease.

In a second best-world where fares cannot vary perfectly with time, BRTs are efficient and have the potential to provide a Pareto improvement of the transport system: in equilibrium both bus users and car users can be better off, while the costs of providing public transport decrease.
With a BRT, fares will be lower because the peak hours of operation of the system are shorter. This better-for-all situation features more boarding delays, that is, queues at bus stops will be longer than under mixed-traffic conditions.

Contributions

First microeconomic analysis of BRTs in the context of dynamic congestion, where queuing and congestion delays are endogenous because of individual schedule of departures.

The main difference with previous literature that investigates bimodal (car and public transport) systems is that, rather than focusing on crowding and assuming from the outset that capacities of each mode are independent, we focus on modeling boarding delays in equilibrium, and comparing mixed traffic conditions with what would arise from dedicating part of the road capacity to a BRT.
This paper is close to Basso and Silva (2014) who study the efficient and substitutability of bus lane and pricing measures but do so in a static congestion framework.

Economic and Environmental Effects of Public Transport Subsidy Policies: a Spatial CGE Model of Beijing (Ping Xu, Weiyu Wang, and Chunxia Wei, 2018)

- Journal: Mathematical Problems in Engineering
- Setting: Spatial Computable General Equilibrium (SCGE) model to examine the economic and environmental effects of public transport subsidy policies
- Data: Statistical data from Beijing were used in calibration to obtain benchmark equilibrium.
- Strategy: SCGE model containing firms, consumers, and transport modules in one framework to investigate the effects of public transport subsidies.
- Using a benchmark equilibrium calibrated for Beijing, the model simulated social welfare, population distribution, and travel-related CO2 emission effects under different subsidization levels with four forms of subsidy policies: fare subsidy, cash grant, road expansion, and public transport speedup.

Results

Public transport subsidies can enhance overall social welfare, regardless of what form the policy takes. Moreover, public transport speedup has the strongest effect on social welfare, followed by fare subsidy, cash grant, and road expansion, respectively.

Different forms of public transport subsidies can exert varied influences on city job-housing population distribution.

Cash grant policy and road expansion construction encourage urban agglomeration, and residential populations aggregate more densely than employed populations.
Fare subsidy policy affects employed population distribution only slightly but stimulates residential population diffusion to suburban areas.
Public transport speedup suburbanizes both residential and employed populations, and residential populations show a stronger suburbanization, which can alter population convergence on the downtown area.

Most public transport subsidy policies give rise to modestly higher public transport split rates. Comparatively, fare subsidies have the most apparent effect, followed by public transport speedup, while cash grants have no influence on public transport share. Road expansion, however, acts in the opposite way, slightly reducing public transport share.

Except for road capacity expansion, public transport subsidy policies do not reduce travel-related CO_2 emissions. In fact, fare subsidy, cash grant, and public transport speedup policies all stimulate higher travel frequency among consumers and therefore aggravate total travel-related CO_2 emissions. CO_2 emissions can only be reduced by investing subsidies in road expansion construction. In conclusion, the social welfare, spatial, and environmental effects of the four subsidy policies are quite different.

<u>Contributions</u>

Basso and Silva (2014) compared the efficiency and substitutability of three different policies. However, they only considered effects on the labor market without considering the effect of subsidies on consumers' choices for job-housing locations. Xu et al. (2018) introduced the concept of space into the research framework. Thus, the function of subsidies can be studied based on their action mechanisms in the spatial distribution of populations.

They added job-housing spatial choice as endogenous variable.

Transport taxes and subsidies in developing countries: The effect of income inequality aversion (Alejandro Tirachini and Stef Proost, 2021)

- Journal: Economics of Transportation
- Setting: The authors propose a marginal tax reform model that includes both formal and informal sectors in the economy, traffic externalities (congestion, pollution, crashes and noise) and distributional concerns.
- Data: Santiago, Chile. Data from the transport sector, the labor market, and the budget shares of different income groups.
- Strategy: In this paper, a transport taxation reform model for developing countries is formulated. This transport taxation reform takes a general equilibrium approach and uses a public finance budget constraint that includes transport taxes as well as non-transport taxes. This is relevant as the ultimate efficiency and redistribution effects of any transport tax or subsidy depends on the other tax instruments that might be changed to balance the public budget.

<u>Results</u>

If marginal cost of public funds are computed only taking economic efficiency into account, a revenue neutral reform within the transport sector suggests increasing the car cost and reducing the bus fare in peak periods, and reducing the car cost and increasing the bus fare in off-peak periods.

Including income inequality aversion leads to suggesting lower bus fares and higher car costs in both peak and off-peak periods, significantly changing the economic assessment of current tax and subsidy instruments.

The inclusion of traffic externalities has a large effect on the marginal cost of public funds for fuel tax and a mild effect on marginal cost of public funds for bus subsidy.

An increase in the size of the informal sector (i.e., a lower rate of workers in the formal sector) reduces the MCF for all tax and subsidy instruments, because in such a case increasing transport taxes has a lower effect on income tax revenue losses. The relative order of instruments does not change.

Contributions

The model explicitly considers the existence of formal and informal sectors in the labor market and includes multiple externalities within the transport sector. Including efficiency considerations only (without inequality aversion), as Basso and Silva (2014), they estimate that the bus fare should be lower in the peak than in the off-peak in Santiago. Appendix B – calibration parameters

15. Appendix B.1 – Calibration parameters for Santiago

Demand parameters

Modal constant θ_{qm}^{i}

	ABC1	C2	C3	D&E
Car peak	0.000	0.000	0.000	0.000
Bus paid peak	-4.778	-3.124	-2.683	-2.605
Metro peak	-4.846	-4.388	-4.529	-5.184
Car off-peak	-0.051	0.121	0.094	-0.031
Bus paid off-peak	-6.473	-4.410	-3.391	-3.091
Metro off-peak	-6.207	-5.641	-5.547	-5.875
No travel	-19.632	-22.672	-23.906	-25.741

Marginal utilities γ^{i} and β_{qm}^{i}

	ABC1	C2	C3	D&E
Monetary cost (γ^{i})	-0.5479	-0.8581	-1.0435	-1.3419
Car peak	-0.5786	-0.5763	-0.5754	-0.5741
Bus peak	-0.9943	-0.9905	-0.9888	-0.9867
Metro peak	-0.9943	-0.9905	-0.9888	-0.9867
Car off-peak	-0.2042	-0.2654	-0.2975	-0.3446
Bus off-peak	-0.5099	-0.6628	-0.7431	-0.8609
Metro off-peak	-0.5099	-0.6628	-0.7431	-0.8609

Scale parameters

λ	0.4
μ	0.15

Transport times

T_b[hr/km]	0.01666
T_c [hr/km]	0.01666
α	2
β	4
b [veq/bus]	3
C [veq/hr]	3600
t_{sb} [sg]	2.7
p	2

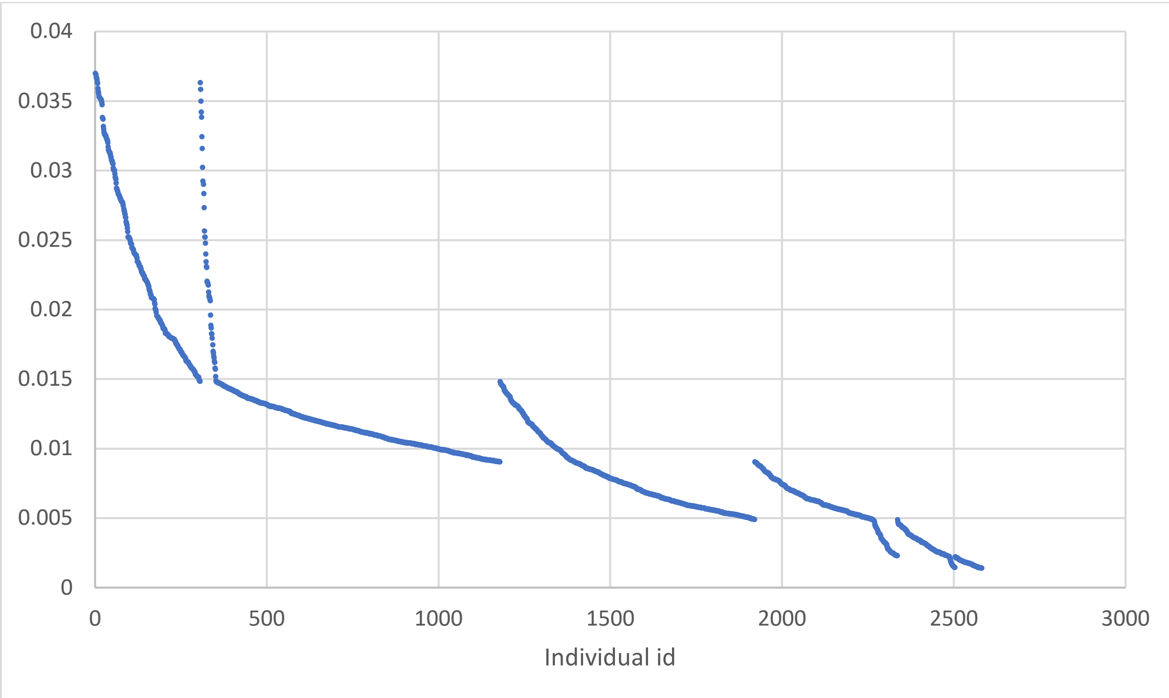
Operating costs

G_b[$/veh-day]	703.1
G_v [$/veh-km]	13.1
G Metro [$/veh-day]	7412.7
G Metro [$/veh-km]	17.7

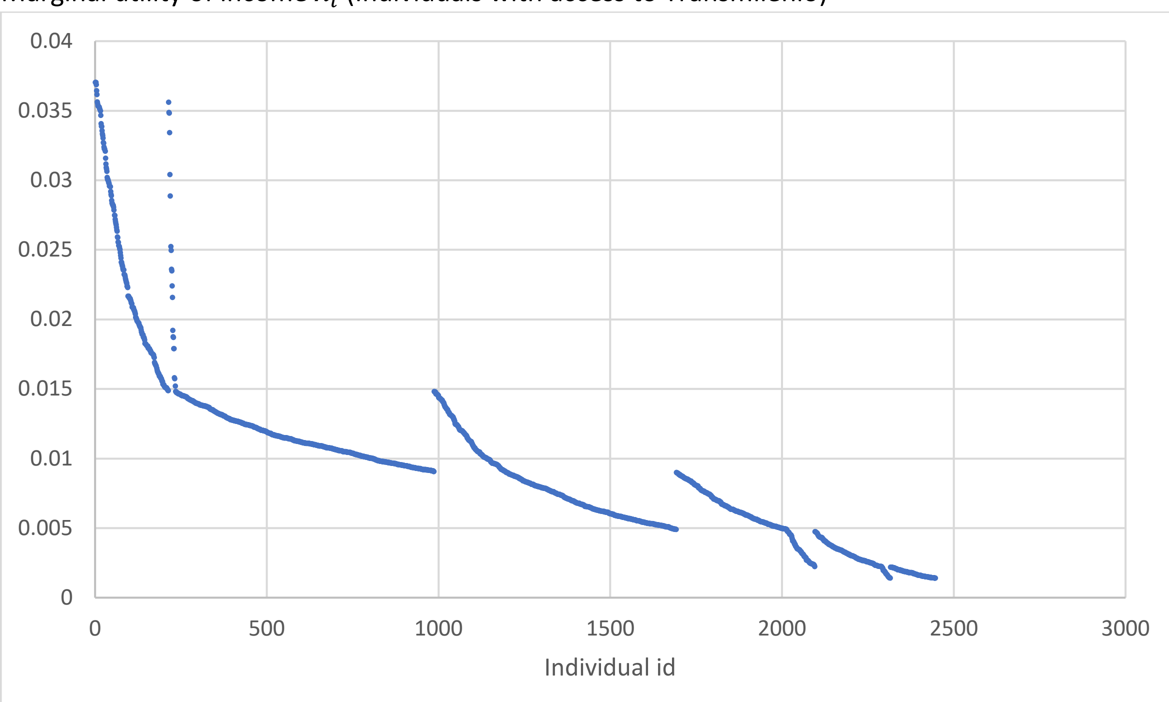
16. Appendix B.2 – Calibration parameters for Bogotá

Demand parameters

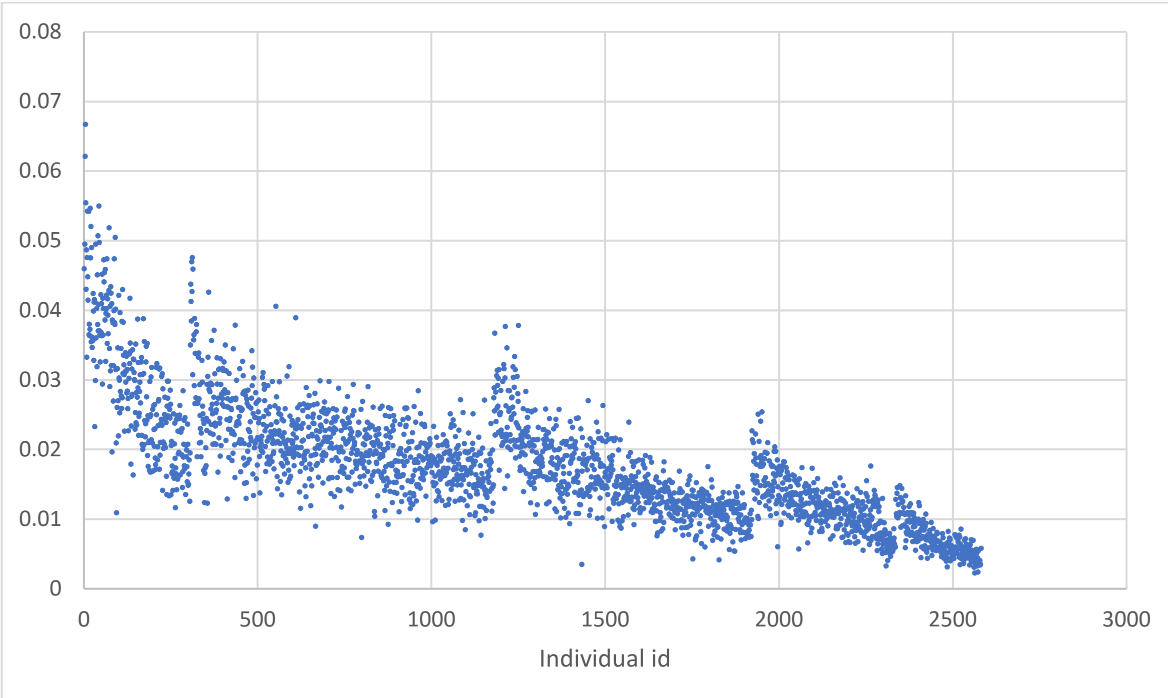
Marginal utility of income λ_i (Individuals with access to SITP)



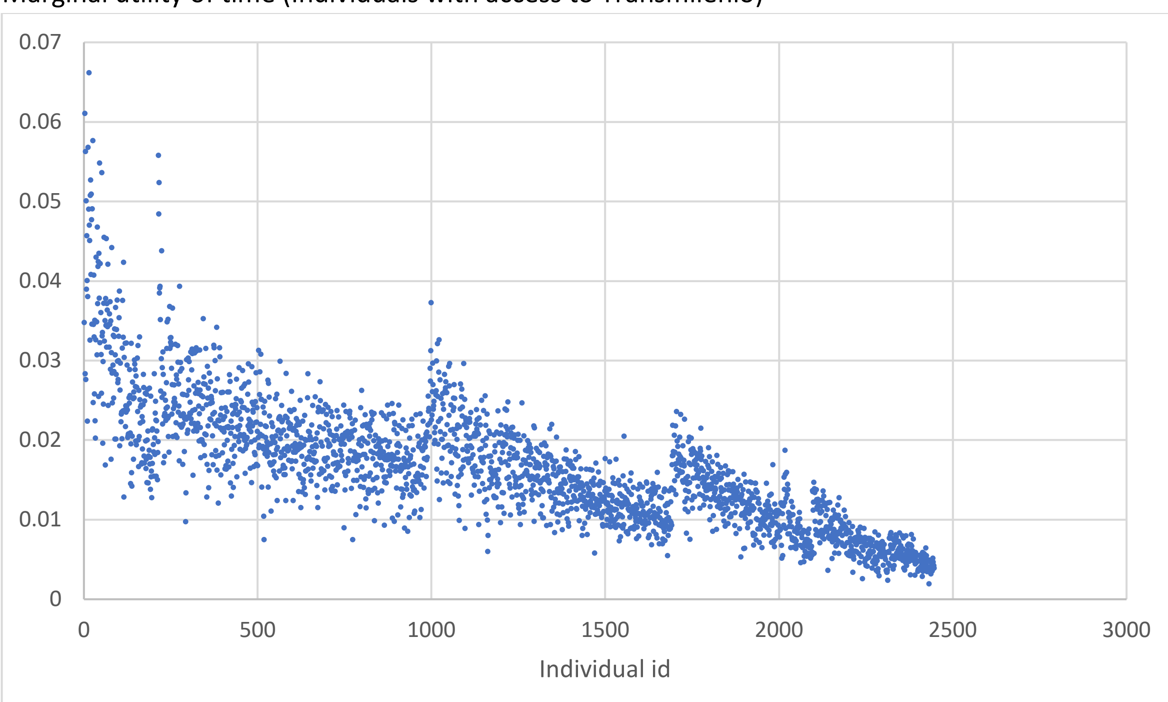
Marginal utility of income λ_i (Individuals with access to Transmilenio)



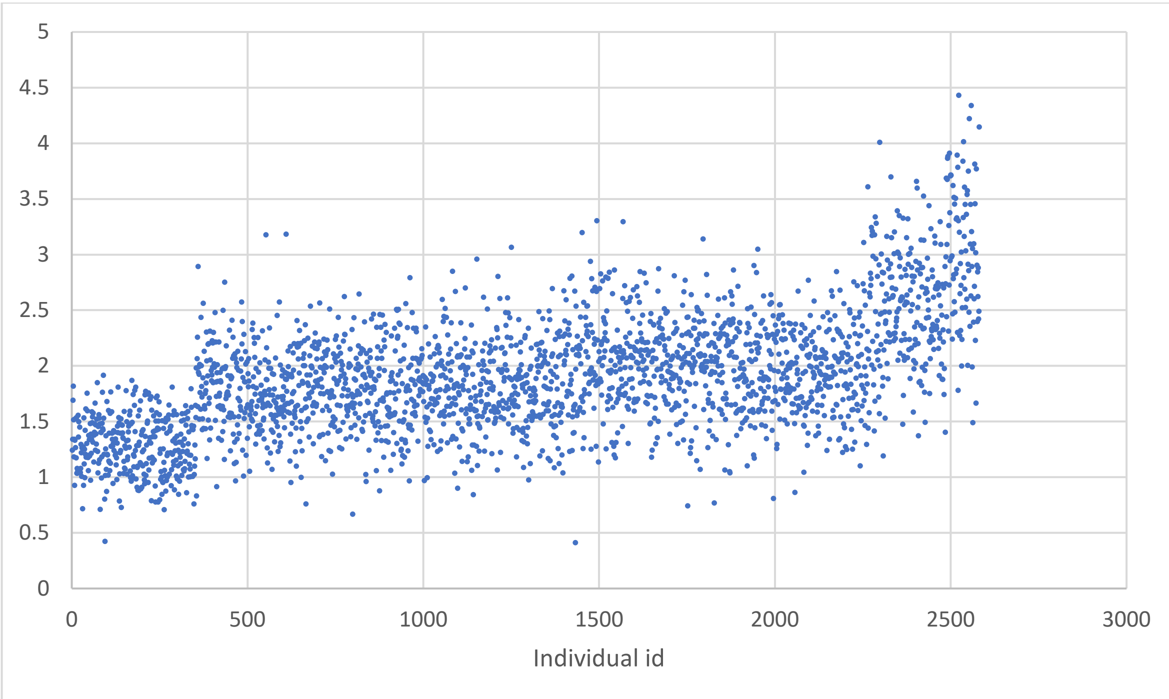
Marginal utility of time (Individuals with access to SITP)



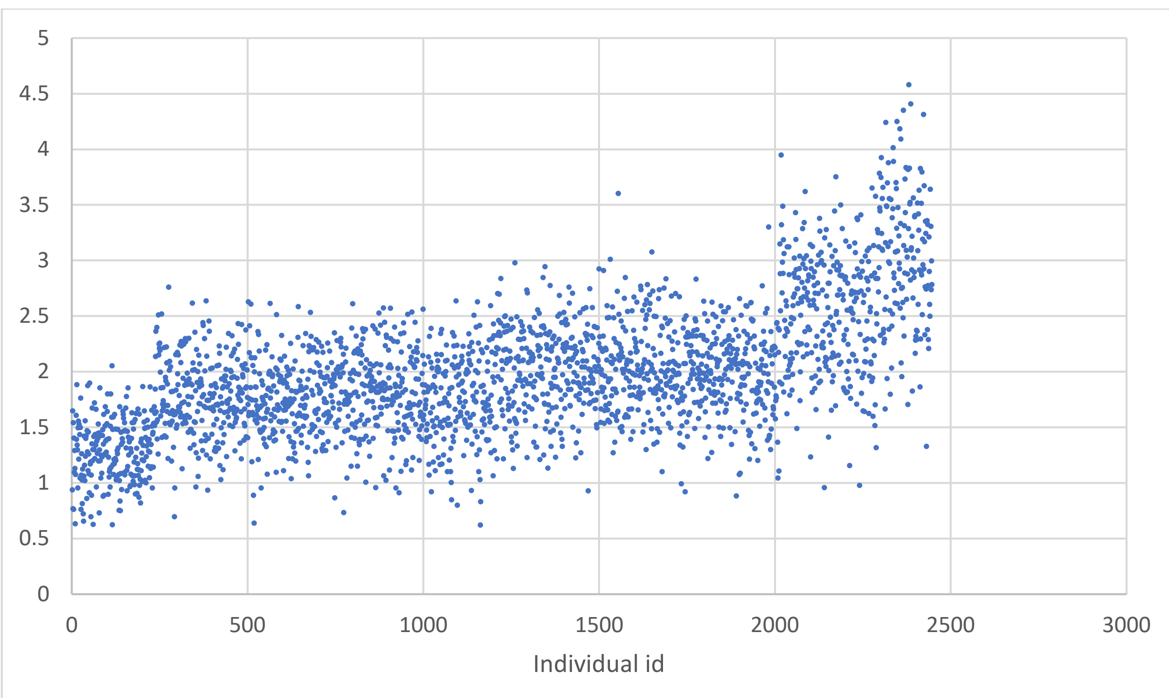
Marginal utility of time (Individuals with access to Transmilenio)



Value of travel time savings (Individuals with access to SITP)



Value of travel time savings (Individuals with access to Transmilenio)



Transport times

T_{BRT} [hr/km]	0.0333
T_c [hr/km]	0.0333
T_{SITP} [hr/km]	0.04
α^b	2.5
α^c	2.0
β	3.7
b [veq/bus]	2.0
C [veq/hr]	3600

Operating costs

G_b [$/veh-day]	119.3
G_v [$/veh-km]	6.6

17. Appendix C – simulation results for individual corridors in Santiago

- *Santiago transport corridor 1 (MTNM)*

	Reference	FARE	CON	DL
Social benefit [$/day-km]	0	1246	3426	1974
CS change [$/day-km]	0	1599	-5728	1005
Bus fare peak [$]	1,20	0,60	1,68	0,84
Bus fare off-peak [$]	1,20	0,60	1,68	0,84
Car toll peak [$/km]	0,00	0,00	0,29	0,00
Car toll off-peak [$/km]	0,00	0,00	0,00	0,00
Metro fare peak [$]	1,32	0,66	1,85	0,92
Metro fare off-peak [$]	1,32	0,66	1,85	0,92
Bus frequency peak [bus/hr]	15,0	15,0	15,0	15,0
Bus frequency off-peak [bus/hr]	8,0	8,0	8,0	8,0
Bus size [pax]	160,0	160,0	160,0	160,0
Metro frequency peak [bus/hr]	7,7	8,7	8,3	8,5
Metro frequency off-peak [bus/hr]	2,8	3,0	3,0	3,0
Car speed peak [km/hr]	18,8	18,6	25,0	18,4
Bus speed peak [km/hr]	17,2	48,7	39,5	47,3
Car speed off-peak [km/hr]	47,2	40,0	34,6	39,4
Bus speed off-peak [km/hr]	39,9	24,1	33,7	14,3
Peak share [percent]	48,5%	48,9%	45,8%	48,8%
Off-peak share [percent]	49,3%	49,0%	51,9%	49,1%
No-travel share [percent]	2,2%	2,1%	2,3%	2,1%
Car modal share peak [percent]	52,5%	47,1%	43,4%	47,6%
Bus modal share peak [percent]	32,5%	36,3%	39,7%	36,0%
Metro modal share peak [percent]	4,3%	4,9%	4,5%	4,7%
Bus + Metro modal share peak [percent]	10,7%	11,7%	12,4%	11,7%

- *Santiago transport corridor 2 (DLNM)*

	Reference	FARE	CON
Social benefit [$/day-km]	0	3831	8323
CS change [$/day-km]	0	14664	-22960
Bus fare peak [$]	1,20	0,60	1,68
Bus fare off-peak [$]	1,20	0,60	1,68
Car toll peak [$/km]	0,00	0,00	0,45
Car toll off-peak [$/km]	0,00	0,00	0,00
Metro fare peak [$]	1,32	0,66	1,85
Metro fare off-peak [$]	1,32	0,66	1,85
Bus frequency peak [bus/hr]	20,0	20,0	20,0
Bus frequency off-peak [bus/hr]	8,0	8,0	8,0
Bus size [pax]	160,0	160,0	160,0
Metro frequency peak [bus/hr]	7,9	8,8	8,9
Metro frequency off-peak [bus/hr]	2,9	3,3	2,6
Car speed peak [km/hr]	14,6	24,7	44,3
Bus speed peak [km/hr]	25,5	22,3	37,1
Car speed off-peak [km/hr]	42,0	51,3	40,6
Bus speed off-peak [km/hr]	43,8	41,8	35,3
Peak share [percent]	48,6%	49,0%	44,6%
Off-peak share [percent]	49,2%	48,9%	53,0%
No-travel share [percent]	2,2%	2,1%	2,3%
Car modal share peak [percent]	43,7%	38,4%	28,4%
Bus modal share peak [percent]	41,2%	45,0%	53,6%
Metro modal share peak [percent]	4,2%	4,8%	4,8%
Bus + Metro modal share peak [percent]	10,9%	11,8%	13,2%
Car modal share off-peak [percent]	12,7%	14,6%	15,0%
Bus modal share off-peak [percent]	34,7%	29,7%	17,0%
Metro modal share off-peak [percent]	49,1%	52,8%	63,6%

- *Santiago transport corridor 3 (MTYM)*

	Reference	FARE	CON	DL
Social benefit [$/day-km]	0	5727	11422	7896
CS change [$/day-km]	0	16343	-22077	16580
Bus fare peak [$]	1,20	0,60	1,68	0,60
Bus fare off-peak [$]	1,20	0,60	1,68	0,60
Car toll peak [$/km]	0,00	0,00	0,49	0,00
Car toll off-peak [$/km]	0,00	0,00	0,00	0,00
Metro fare peak [$]	1,32	0,66	1,85	0,66
Metro fare off-peak [$]	1,32	0,66	1,85	0,66
Bus frequency peak [bus/hr]	20,0	20,0	20,0	20,0
Bus frequency off-peak [bus/hr]	10,0	10,0	10,0	10,0
Bus size [pax]	160,0	160,0	160,0	160,0
Metro frequency peak [bus/hr]	8,9	9,9	10,4	9,9
Metro frequency off-peak [bus/hr]	3,3	3,8	3,0	3,8
Car speed peak [km/hr]	26,4	26,0	41,4	26,4
Bus speed peak [km/hr]	24,0	53,9	44,1	53,9
Car speed off-peak [km/hr]	52,8	45,0	38,9	45,1
Bus speed off-peak [km/hr]	45,2	35,6	54,6	29,2
Peak share [percent]	48,7%	49,0%	44,3%	49,0%
Off-peak share [percent]	49,2%	48,9%	53,3%	48,8%
No-travel share [percent]	2,2%	2,1%	2,3%	2,1%
Car modal share peak [percent]	41,3%	35,2%	23,5%	34,9%
Bus modal share peak [percent]	41,8%	46,0%	55,7%	46,4%
Metro modal share peak [percent]	4,2%	4,8%	4,8%	4,8%
Bus + Metro modal share peak [percent]	12,8%	14,0%	15,9%	13,9%
Car modal share off-peak [percent]	12,7%	14,6%	15,4%	14,6%
Bus modal share off-peak [percent]	31,1%	25,0%	10,7%	24,6%
Metro modal share off-peak [percent]	49,8%	54,2%	66,0%	54,7%

- *Santiago transport corridor 4 (DLYM)*

	Reference	FARE	CON
Social benefit [$/day-km]	0	2822	6703
CS change [$/day-km]	0	14976	-22492
Bus fare peak [$]	1,20	0,60	1,68
Bus fare off-peak [$]	1,20	0,60	1,68
Car toll peak [$/km]	0,00	0,00	0,44
Car toll off-peak [$/km]	0,00	0,00	0,00
Metro fare peak [$]	1,32	0,66	1,85
Metro fare off-peak [$]	1,32	0,66	1,85
Bus frequency peak [bus/hr]	20,0	20,0	20,0
Bus frequency off-peak [bus/hr]	8,0	8,0	8,0
Bus size [pax]	160,0	160,0	160,0
Metro frequency peak [bus/hr]	9,0	10,0	10,3
Metro frequency off-peak [bus/hr]	3,3	3,8	3,0
Car speed peak [km/hr]	16,6	26,7	45,6
Bus speed peak [km/hr]	26,8	23,9	38,1
Car speed off-peak [km/hr]	44,4	52,3	42,1
Bus speed off-peak [km/hr]	44,3	42,5	36,5
Peak share [percent]	48,6%	49,0%	44,7%
Off-peak share [percent]	49,2%	48,9%	53,0%
No-travel share [percent]	2,2%	2,1%	2,3%
Car modal share peak [percent]	42,3%	37,0%	27,4%
Bus modal share peak [percent]	40,4%	44,1%	52,0%
Metro modal share peak [percent]	4,2%	4,8%	4,8%
Bus + Metro modal share peak [percent]	13,0%	14,1%	15,8%
Car modal share off-peak [percent]	12,7%	14,6%	15,0%
Bus modal share off-peak [percent]	32,7%	27,6%	15,4%
Metro modal share off-peak [percent]	47,9%	51,4%	61,3%