Working Paper

A practical approach for curbing congestion and air pollution: Driving restrictions with toll and vintage exemptions

Leonardo Basso[1], Juan-Pablo Montero[2], and Felipe Sepulveda[3]

November 2020

[1] Civil Engineering Department, Universidad Chile, and Sistemas Complejos de Ingenieria (ISCI)
[2] Department of Economics, Pontificia Universidad Catolica de Chile and ISCI
[3] FK Economics

La Chaire Energie et Prospérité

La chaire Energie et Prospérité a été créée en 2015 pour éclairer les décisions des acteurs publics et privés dans le pilotage de la transition énergétique. Les travaux de recherche conduits s'attachent aux impacts de la transition énergétique sur les économies (croissance, emploi, dette), sur les secteurs d'activité (transport, construction, production d'énergie, finance), aux modes de financement associés ainsi qu'aux problématiques d'accès à l'énergie. Hébergée par la Fondation du Risque, la chaire bénéficie du soutien de l'ADEME, de l'Agence Française de Développement, de la Caisse des Dépôts, d'Engie et de la SNCF.

Les opinions exprimées dans ce papier sont celles de son (ses) auteur(s) et ne reflètent pas nécessairement celles de la Chaire Energie et Prospérité. Ce document est publié sous l'entière responsabilité de son (ses) auteur(s).

Les Working paper de la Chaire Energie et Prospérité sont téléchargeables ici :

http://www.chair-energy-prosperity.org/category/publications/

Chair Energy and Prosperity

The Energy and Prosperity Chair was created in 2015 to inform decisions of public and private actors in managing the energy transition. The Chair research deals with the impacts of energy transition on national economies (growth, employment, debt...), on specific sectors (transportation, construction, energy , finance), on acces to energy and with the associated financing issues. Hosted by the Risk Foundation, the chair has the support of ADEME, the French Development Agency, the Caisse des Dépôts, Engie and SNCF.

The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the position of Chair Energy and Prosperity. It is therefore published under the sole responsibility of its author(s).

Chair energy and Prosperity working paper can be downloaded here:

http://www.chair-energy-prosperity.org/en/category/publications-2

A practical approach for curbing congestion and air pollution: Driving restrictions with toll and vintage exemptions

Leonardo Basso, Juan-Pablo Montero and Felipe Sepúlveda *

November 25, 2020

Abstract

Congestion and local air pollution continue to be a serious problem in many cities around the world, partly because of an increasing and ageing car fleet. Unfortunately, the use of pricing schemes for handling these externalities, such as congestion and pollution charges, still face much resistance. To cope with it, Carlos F. Daganzo advanced an ingenious hybrid scheme that supposedly leaves everybody better off: driving restrictions with toll exemptions. We extend Daganzo's idea to include vintage exemptions in an effort to also control for the pollution externality. We then test for its Pareto-improving property using Santiago as a case study. We find the latter not to hold in that low-income drivers are strictly worse off: the gain from faster car travel in days of no restriction is not enough to compensate the loss from switching to public transport in days of restriction. To make all individuals better off, all toll revenues ought to be recycled back into the public transport system, lowering its fares and improving its quality. If so, the most ambitious hybrid restriction format —a 5-day-a-week restriction with vintage thresholds during fall and winter— reports per-year net benefits of around 1.2 billion dollars (or 0.5% of the country's GDP), 58% of which comes from lighter traffic and the remaining 42% from cleaner air.

Key words: congestion, local pollution, driving restrictions, road pricing

*Basso (lbasso@ing.uchile.cl): Civil Engineering Department, Universidad Chile, and Sistemas Complejos de Ingeniería (ISCI); Montero (jmontero@uc.cl): Department of Economics, Pontificia Universidad Católica de Chile and ISCI; Sepúlveda (fsepulveda@fkeconomics.com): FK Economics. We thank Jorge Bonilla, Francisco Cabezón, Andrés Gómez-Lobo, Luis Rizzi, Hugo Silva, Alan Thomas and participants at the 10th Economics for Energy Workshop (Galicia, Spain), Public Policy Workshop (PUC-Chile), WCERE 2018 (Gothenburg), and Ecole Polytechnique for many comments and discussions; and Pablo Valenzuela for research assistance. Basso and Montero thanks ISCI for financial support, and Montero also thanks similar support from the Chair in Energy and Prosperity at the Risk Foundation.

1 Introduction

Congestion and local air pollution continue to be a serious problem in many cities around the world, partly because of an increasing and ageing car fleet.[1] While transport scholars have long advocated for the use of economic instruments to correct for these externalities, only a few cities around the world have adopted them. Arguably, an important reason for the slow emergence of economic instruments such as congestion pricing and pollution taxes is the lack of widespread public support (e.g., Harsman and Quigley 2010; Baranzini et al. 2018; Fageda et al. 2020).[2] To cope with this resistance, in particular with regard 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 in having unpaid access to the road. Thus, an individual who travels every day would have to pay a toll only on those days of the week in which his or her car is restricted from circulation, say, those days in which the car's license plate ends in a certain digit. Daganzo's premise is that this "taking turns" scheme leaves everybody better off, providing the necessary public support for the scheme. While it is easy to see why identical individuals would benefit from it (Daganzo and Garcia 2000), the story for heterogeneous individuals would go like this. Higher-income individuals would benefit from the scheme as they continue commuting by car every day (and paying the toll the day or days of restriction) but faster. Lower-income individuals, on the other hand, would incur a loss the day or days of restriction as they could not afford paying 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 with the gain from faster car travel during the rest of the week, i.e., days of no restriction.

In addition to this supposedly Pareto-improving property, Daganzo's scheme possesses two other advantages that should ease its implementation. One is that it builds around a policy that authorities are increasingly relying upon to curb congestion and local air

[1]Cars are major contributors of carbon monoxide (CO), hydrocarbons (HC), nitrogen oxides (NO_x), and fine particles (PM2.5). HC and NO_x are precursors to ground-level ozone (also known as smog) and also contribute to the formation of PM2.5. At least in Santiago, vehicles are responsible for 30 and 36% of PM2.5 and O_3 concentrations, respectively (Rizzi and De la Maza 2017). These local pollutants, unlike global pollutants such as carbon dioxide (CO_2), are characterized as having a local impact, at the city level, that lasts for a short time, sometimes only a few hours. The adverse health effects of these local pollutants are well documented (e.g., Currie and Neidell 2005).

[2]In many cities, including Santiago, plans to introduce congestion pricing were seriously considered but finally abandoned. A congestion charge was rejected at the ballot box in Edinburgh and Manchester. Voters in Gothenburg rejected it in a non-binding referendum, but it was nevertheless enacted. Voters in Stockholm approved it in a close referendum, but only after a trial period. In London, it also faced significant opposition at the beginning (Leape 2006). And New York City's recently enacted congestion fee, which supposedly goes into effect in 2021, was no different in that regard.

pollution, which is to impose limits to car use, typically implemented on the basis of some combination of the last digit of a vehicle's license plate and colored stickers displayed on its windshield. Good examples of these rationing schemes or driving restrictions, as they are typically called, include, among many others, Athens (where restrictions were introduced in 1982), Santiago (1986), Mexico City (1989), Teheran (1991), São Paulo (1996), Manila (1996), Bogotá (1998), Cali (2002), La Paz (2002), Medellín (2005), Beijing (2008), Tianjin (2008), Quito (2010), Hangzhou (2011), Chengdu (2012), New Delhi (2016), Paris (2016), and Madrid (2019).

Another advantage of Daganzo's hybrid scheme is that it deals with a problem often associated with driving restrictions, namely, the perverse incentives they create for drivers to buy additional vehicles. This "second-car" effect would not only increase fleet size but also move its composition toward higher-emitting vehicles, resulting in more congestion and pollution in the long run. The best documented evidence supporting the second-car effect 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).[3] However, by giving drivers the option to pay a toll to get their cars exempt from the restriction, the taking-turns scheme eliminates these perverse incentives, unless of course the toll is set extremely high. As far as we know, some of Colombia's *Pico-y-Placa* programs (e.g., Cali since 2017 and Bogotá since 2019) are the only restriction programs in the world that include toll exemptions (Ramos et al. 2017).

The objective of this paper is to test the Pareto-improving property of the Daganzo's hybrid scheme with a simple model that uses Santiago as a case study. In doing so, we depart from Daganzo and Garcia (2000) in four ways. First, we allow commuters to have a choice between two modes of transportation: private vehicles and public transit. Second, we allow commuters to be heterogenous with regard to income, preferences for transportation modes, and their vehicles (if they own one). In particular, we divide individuals in five income groups, following SECTRA's (2013) value-of-time criteria,[4] 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.

Third, we extend Daganzo's restriction scheme to incorporate local pollution considerations. We do so by following recent driving-restriction programs which exempt cleaner cars from the restriction. Thus, the option to pay the toll to have the car exempted from the restriction is only available to owners of relatively clean cars, i.e., younger than some vintage threshold; threshold that may well vary during the year depending on air quality

[3]See also Cantillo and Ortuzar (2014), Wang et al. (2014), Viard and Fu (2015), and Nie (2017).

[4]SECTRA is the office for transport planning of Chile's Ministry of Transportation.

conditions. As documented by Barahona, Gallego, and Montero (2020) (hereafter BGM), an increasing number of restriction programs now differentiate cars by vintage, or more precisely, by pollution rates. This vintage differentiation is observed not only in the current restriction programs in Santiago and Mexico City but also in many cities in Europe, where authorities have been adopting low-emission zones (LEZs) since 2008.[5] Unlike the partial circulation bans in Santiago and Mexico City, LEZs completely ban certain higher-emitting vehicles from entering city centers (e.g., Wolff 2014).[6]

Our fourth departure from Daganzo and Garcia (2000) is that we replace their "bottleneck model" with a "static congestion" model with a time-invariant (i.e., daily) toll. Notwithstanding the scheme's Pareto-improving property supposedly applies to either type of model (Daganzo 2000), we believe a static-congestion model fits better with regulatory schemes that build upon driving restrictions. Time-invariant tolls are not only present in the very few driving-restriction programs that have toll exemptions —the ones in Colombia— but also in many congestion-charge programs, including London and the upcoming program in New York City.

Daganzo advances his restriction scheme in a static context, absent of potential changes in car ownership. While we retain this short-run focus in our analysis, we do discuss how the introduction of toll and vintage exemptions are sufficient to rule out any second-car concern associated to the taking-turns policy.[7] Our study begin by extending the static-congestion model of Basso and Silva (2014) to pollution considerations. We then calibrate the model to capture Santiago's current traffic and pollution reality. It is a short-run model in that individuals' only decision is whether to commute by car, provided they have one, or by public transport. The model also considers a transport authority with 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.[8]

[5] LEZ programs have also been introduced in China; for example, in Beijing in 2009 and Nanchang in 2013.

[6] This "complete-ban" structure is also in the restriction introduced in Paris in 2016 (where any car built before 1997 is banned permanently from circulation within the city limits weekdays from 8 am to 8 pm) and in recent announcements made by several European cities, including London, Paris and Rome, to completely ban diesel vehicles from entering city centers in the coming decade.

[7] BGM provide further evidence on how vintage exemptions eliminate second-car effects.

[8] Many important aspects are left outside our simple model: the availability of additional transportation modes (e.g., bicycle, Uber-like platforms, etc), the possibility to substitute trips within the week, the possibility to substitute peak with off-peak travel, downward sloping demand for daily trips, etc. Expanding individuals' action space to accomodate for all these considerations would greately complicate the equilibrium search, specially if the transportation game may now accept multiple fixed points (see David and Fourcat (2014) for a discussion of multiple equilibria in these type of games). Moreover, unless these considerations are shown to be correlated with income, something that is beyond the scope of this paper to evaluate, our

The main result of the paper 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 (particularly those few who own a car) are strictly worse off, and more so as we increase the number of days of restriction. This negative result calls for two, seemingly contradicting measures.

The first is that all toll revenues should be recycled back into the public transit system through some combination of lower fares and better service (e.g., higher frequency). Doing so ensures that all income-groups are better off.[9] And given this need of recycling for handling distributional impacts, no matter the restriction format, the second measure is that authorities should aim for the most ambitious restriction format (except perhaps during consultation and/or transition phases). A more ambitious goal not only contributes to welfare with lower travel times and pollution levels, it also contributes with more toll revenues to be spent in public transit, hence, leaving lower-income groups increasingly better off as well.

Because Santiago's pollution problem is predominantly present during fall and winter (see, e.g., Gallego, Montero, and Salas 2013), the most ambitious restriction format is to have all cars restricted every day of the week throughout the entire year while at the same time preventing owners of polluting cars from paying the daily toll during fall and winter. If toll revenues are devoted entirely to reduce transit fares, the optimal restriction design predicted by our model involves: (i) daily tolls of \$10.6 and \$8.3 for spring/summer and fall/winter, respectively,[10] (ii) vintage thresholds of 1998 and 2003 for gasoline and diesel cars, respectively, during fall and winter, and (iii) a reduction in public-transit fares of 70%. This ambitious design amounts to overall benefits somewhere between \$1.19 and \$1.27 billion annually (or 0.49-0.52% of the country's GDP in 2015),[11] 58% of which comes from lighter traffic and the remaining 42% from cleaner air.[12]

Our paper contributes to different strands of the literature on vehicle externalities. First,

distributional results should not change, at least qualitatively.

[9]Lump-sum transfers to lower income groups, while distributionally effective in theory, are rarely used in practice. One way or another, all existing congestion-charge schemes allocate part of the congestion revenues to investments in public transport.Our proposal is unique, perhaps, in that we recommend, for distributional reasons, that a good fraction of these congestion revenues, if not all, go to reduce current public-transit fares.

[10]The currency used throughout the paper is 2015 U.S. dollars. Note that without revenue recycling (and fare reduction), (optimal) daily passes would be higher, \$12.9 and \$11.4, respectively.

[11]Note that overall welfare can increase at most by 3% if part of the toll revenues (17%) are allocated to improve service quality.

[12]Note, however, these contributions vary widely throughout the year. In spring and summer congestion allevation contribute with approximately 86% to total welfare and pollution reduction with 14%, while in fall and winter these contributions reverse, with 41% and 59%, respectively.

it contributes to the literature on driving restrictions (e.g., Davis 2008; BGM) by simultaneously considering toll and vintage exemptions and by studying their distributional implications. Second, it adds to the literature on congestion management with a practical proposal that pays close attention to distributional concerns.[13] Perhaps surprisingly, we find that these concerns are better handled, at least in a city like Santiago, by allocating the majority of the toll revenues to transit-fare reductions and less to better transit services.[14]

And third, it contributes to an increasing literature that simultaneously looks at pollution and congestion externalities. According to Parry and Small (2005), for example, the external congestion cost of an additional kilometer traveled in the United States and the United Kingdom is approximately 1.8 and 3.5 times larger, respectively, than the local pollution external cost associated to that extra kilometer. Rizzi and De la Maza (2017) also find marginal external congestion costs in Santiago to be much larger than marginal pollution external, going from 3.1 times larger, during off-peak hours, to 15 times larger, during peak hours.[15]

We find these average numbers to be potentially misleading to policy makers. While our marginal external costs are largely consistent with those in Rizzi and De la Maza (2017), our pollution-alleviation benefits are far more important than these average numbers suggest. In fact, in fall and winter they are significantly larger than the congestion-alleviation benefits.[16] There is a simple reason for this apparent contradiction. While all cars congest the same, old cars pollute a lot more than newer cars. So targeting old cars first, as vintage restrictions do, yields greater benefits than targeting the average car, which is what looking at average external (pollution) costs implicitly does.

The rest of this paper is organized as follows. The model is developed in Section 2. Model parameters, some of which are calibrated to fit the model to existing data, are presented in Section 3. The scheme's Pareto-improving property is tested in Section 4. Distributional implications of alternative revenue-recycling options are covered in Section 5. Vintage exemptions are added in Section 6. Conclusions are offered in Section 7. Additional material is relegated to the online Appendix.

[13] Distributional concerns also studied in Bento et al. (2009), but in the context of fuel taxes.

[14] We only consider quality improvements on surface public transit (buses running more often) and not on the expansion of the underground network.

[15] See also Parry and Strand (2012) for earlier estimates for Santiago, when congestion was less of a problem.

[16] And they would be even larger if we could replace the vintage restrictions for a pricing instrument (see concluding section).

2 A simple traffic-pollution model

We consider a standard origin-destination transport model following Basso and Silva (2014).[17] On a daily basis, a large number of people, say n, commute to the city center to work and study either by car or public transport. Since the majority of public-transport rides in Santiago, above 75%, involve some combination of subway and buses, we treat public transport in a "reduced form" combining both modes. Travelers differ in several dimensions, most notably income and whether they own a car or not, and if they do, how much their cars pollute per kilometer traveled.

The transport authority can intervene the existing transport equilibrium in two ways. The first is by imposing a limit on the number of days a car can circulate within the boundaries of the city center,[18] which is done, for example, by placing restrictions according to the last digit of the car's license plate. A restricted car can still circulate, but only after paying a toll and provided its emission rate e (i.e., grams of pollution per kilometer traveled) is below some pollution threshold $\bar{e}$.

We vary the extent of the restriction from one day a week to full restriction, i.e., five days a week. Given the option to pay a toll, a 5-day restriction is close, if not equivalent, to a pure road-pricing scheme for those individuals who own cars with emission rates $e < \bar{e}$, in that they face a price each time they decide to use their cars to enter the city center.[19] On the other hand, setting the toll to infinity is equivalent to a pure driving-restriction policy.

The second way in which the transport authority can intervene the market is by making public transport more attractive. We allow the authority to use part or all the revenue collected from toll payments to either reduce the existing public-transport fare, improve its quality (e.g., service frequency), or both.

Travelers have no choice but to commute every day of the week, so only those who own a car have the option to switch to a different transportation mode (i.e., public transport) for some days of the week in response to the government intervention. The (transport) surplus that individual $i = 1, ..., n$ obtains after a week of travel (excluding weekends) is given by

$$S_i = \Omega_i(d_i) - C_i(d_i, r) - T_i(d_i) \tag{1}$$

where $d_i = 1, ..., 5$ is the number of days of the week that i commutes by car, provided she

[17] See also David and Foucart (2014) for a similar model.

[18] We define the city center vaguely, enough to capture most trips that take place in the city of Santiago. Adopting this ample view is particularly relevant from a pollution perspective.

[19] It may not be exactly equivalent to a road-pricing scheme if the latter involves time-varying prices, which neither London nor New York City consider.

has one, $r = 0, ..., 5$ is the number of days that cars are restricted from entering the city center, $\Omega_i(d_i)$ captures the gross benefit of travel, $C_i(d_i, r)$ is the financial cost of travel, and $T_i(d_i)$ is the time cost of travel. Note that $\Omega_i(d_i)$, $C_i(d_i, r)$, and $T_i(d_i)$ are all measured in dollars and vary across individuals according to their income levels, which, as in Basso and Silva (2014), we divide in five income groups: low, middle-low, middle, middle-high and high.

The gross benefit of travel depends on an individual's intrinsic preferences for each transport mode as follows

$$\Omega_i(d_i) = \lambda_i^{-1}\phi_0[d_i + (5 - d_i)\theta_i]$$

where λ_i corresponds to i's marginal utility of income, ϕ_0 is a constant, and θ_i is a dimensionless parameter drawn from a uniform distribution that captures i's public vs private transport relative taste.[20] Their values are obtained by calibrating the model to the data. Calibration results indicate that θ does not vary much on average across income groups but that lower income groups exhibit much wider dispersion. This may explain why some low-income individuals appear to have much stronger preferences for car use (e.g., negative values of θ) relative to others in their income group.

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

$$C_i(d_i, r) = d_i c + p_c \max\{0, d_i + r - 5\} \times \mathbf{1}\{e_i \leq \bar{e}\} + (5 - d_i)p_p \tag{2}$$

where c is the daily cost of using a car, including expenses on fuel, parking, lubricants, tires, and so on, p_c is the daily toll to be paid to have a car exempted from the restriction (note that subscript "c" denote cars and subscript "p" public transport), $\mathbf{1}\{e_i \leq \bar{e}\}$ is an indicator function that takes the value of 1 when i owns a car and is entitled to pay the toll, and p_p is the daily expense on public transit (i.e., the product of single-ride fare and the average number of daily rides). Note that $d_i > 5 - r$ whenever an individual who owns a car with emissions rate $e_i \leq \bar{e}$ and subject to r day(s) of restriction per week decides to use her car not only the $5 - r$ day(s) of no restriction but also the day(s) of restriction. Furthermore, since there are no income effects in our model, if that individual decides to pay the toll to get her car exempted from the restriction, she will do so for all days of restriction. In other words, for someone that owns a car, $d_i \in \{0, 5 - r, 5\}$. Thus, the case of $d_i = 5 - r$ is when i leaves her car at home the days of restriction either because she finds p_c too high or $e_i > \bar{e}$.

[20]Note that θ_i is fixed for each individual. There is neither aggregate uncertainty at the income-group level since each group is composed by a large number of individuals.

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

$$T_i(d_i) = \lambda_i^{-1}\left[d_i\gamma_i^c t^c l + (5-d_i)(\gamma_i^p t^p l + \gamma_i^w w^p)\right] \tag{3}$$

where γ_i^m is i's marginal utility of time when using transport mode $m \in \{c, p\}$, t^m is time (in hours per kilometer) spent on transport mode m on any given day, l is the average travel-distance (in kilometers) involved 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 station, and w^p is the average waiting time at the station. Following Basso and Silva (2014), we also assume that $\gamma_i^w = 2\gamma_i^c$.

We allow γ_i^c and γ_i^p to differ as a way to control for any inconvenience that may result from increasing public-transport use without the corresponding adjustment in service frequency. We assume this inconvenience to be similar in both public-transit modes —buses and subway—, so we model this inconvenience following Tirachini et al. (2017) for surface public-transport

$$\gamma_i^p = \gamma_i^c\left(1 + \zeta\frac{\gamma l D_p}{f_b s q K}\right) \tag{4}$$

where ζ is a "crowding" penalty, D_p is the total number of individuals using public transport in any given day (since n is very large, $D_p = \sum_{i=1}^{n}(5-d_i)/5$), $\gamma < 1$ accounts for the fact that only a fraction of the distance l is covered by the surface mode (while $1-\gamma$ is covered by the underground mode), f_b is buses' frequency, s is buses' average size (in m^2), q is duration (in hours) of the peak period,[21] and K is the network length (in kilometers of road lane).[22]

We also need to distinguish between the two forms of public transport —surface and underground— when it comes to estimate the travel times t^c and t^p that enter into (3). Only surface travel, where cars and buses share the road, is affected by government intervention. We model this surface travel time following a standard BPR function (Basso and Silva 2014) of the form

$$t^s = t_f^s\left(1 + \alpha_s\left(\frac{f_b\kappa + lD_c/aqK}{C}\right)^{\beta_s}\right) \tag{5}$$

where t_f^s is the free-flow travel time of surface mode $s \in \{c, b\}$ and where $s = b$ corresponds to buses (note that $t_f^s = 1/v_f^s$, where v_f^s is the free-flow speed), α_s is a parameter related to speed reduction caused by congestion, β_s captures the sensitive of travel time to changes in congestion, κ is an equivalence factor between buses and cars, C is the capacity of a

[21] Since l is the round-trip average distance, q includes duration of both morning and evening peaks.

[22] The difference between γ_i^b and γ_i^c is similar, on average, to the difference in Basso and Silva (2014), i.e., about two times larger.

road lane (maximum number of cars per hour a road lane can absorb without affecting travel time and taking into account traffic signals), a is the (constant) car occupancy, and D_c is the number of individuals commuting by car in any given day (again, since n is very large $D_c = \sum_{i=1}^{n} d_i/5 = n - D_p$). We let surface travel times differ between the two relevant modes —buses and cars— because bus stop operations and traffic signals affect modes differently. Indeed, buses need to stop at stations while cars do not, added to the fact that at traffic signals buses have smaller acceleration rates.[23] Finally, to control for the fact that t^p responds less than t^b to government interventions because underground speed is not affected by them, we simply make $t^p = t^b$ and increase γ_i^p accordingly, by adjusting ζ.

In addition to (transport) surplus S_i, individual i's welfare is also affected by air pollution. But unlike transport, here we only need an estimate of the overall pollution harm, which, after a week of travel (excluding weekends), is given by

$$H = \sum_{i=1}^{n} lhe_i d_i + H_0 \tag{6}$$

where he_i is the harm per kilometer generated by a car with an emissions rate of e_i (so h is a measure of harm in dollars per gram of pollution emitted) and H_0 is some background harm unaffected by the restriction policy.[24] Thus, the overall welfare effect of a policy intervention reduces to

$$\Delta W = \sum_{i=1}^{n} \Delta S_i - \Delta H = \sum_{i=1}^{n} (\Delta S_i - lhe_i \Delta d_i) \tag{7}$$

where ΔS_i is the change in transport surplus due to the policy and Δd_i is the change in car use.

The decision problem of individual i who owns a car with an emissions rate of e_i is to chose d_i so as to that maximize (1) subject to $e_i \leq \bar{e}$ in days of restriction, while taken as given the equilibrium choice of the remaining car owners, that is, taken as given D_c and D_p. We compare welfare and distributional implications of the equilibrium outcome for different government interventions, which include different values of $r \in \{0, 1, ..., 5\}$, $\bar{e}$ and p_c, and alternative uses of toll revenues, whether to reduce the cost of using public transport (e.g.,

[23] Depending on frequency and road capacity, bus stop operations may also affect traffic, of both cars and buses. Basso and Silva (2014) proposed a model where these effects are explicitly modeled. Here, given the calibration with real data approach that we will follow, it seems better to let the parameters, particular the "beta" values, to capture those effects.

[24] Background harm includes pollution from other sources (e.g., industrial sources) but it may also include pollution generated by the public-transport system. We make the implicit assumption that emissions from public transit are not affected by the restriction policy. In the absence of public-transport expansion this is clearly a conservative assumption given the increase in buses speed.

to reduce p_p) and/or to improve its quality (e.g., to increase f_b).

Note that according to David and Fourcat (2014), these type of games, with network externalities, may accept multiple equilibria. There are two reasons, however, this potential multiplicity is less of a problem here than in David and Fourcat (2014). One is the fact that public-transit quality is exogenous (i.e., determined outside the game), so Morhing's (1972) positive externality from public-transit use is absent in our setting. And the second reason is the "crowdiness " penalty, which makes public transit less attractive as more people switch to it. We only share with David and Fourcat (2014) the fact that buses run faster as more people switch to public transport, leaving behind less congested roads. Whether this network externality alone is enough to generate multiplicity is something we do not formally explore in the paper; although none of our simulations supports it.[25]

3 Parameter values

The model is parameterized to accommodate Santiago's traffic and pollution reality as captured by the most recent available data. In the case of traffic, we rely on the Origin-Destination survey of 2012 (ODS-2012) and the congestion-pricing simulations carried out by Chile's Transport Planning Office (SECTRA), which are reported in SECTRA (2013).[26] In the case of pollution, we combine three sources of information: car's ownership characteristics are obtained from the 2015 Vehicle Survey of the National Statistics Bureau (NSB) and the 2015 circulation-permit dataset,[27] car's emission rates are obtained from the 2015 smog-check dataset,[28] and estimates of pollution harm specific to Santiago are taken from BGM. Some parameter values are obtained directly from these sources, while others from calibrating our model to replicate relevant traffic patterns.

3.1 Preference parameters

Following SECTRA (2013), we divide commuters in five income groups. As shown in Table 1, groups are of different sizes (they are not quintiles) since SECTRA's classification criteria

[25]For example, no matter the initial modal shares we adopt in searching for an equilibrium, we always arrive at the same no-intervention equilibrium benchmark.

[26]The ODS-2012, demanded by SECTRA, was carried out by the Observatorio Social of the Universidad Alberto Hurtado and published in 2014. It was applied to 18,000 households during July 2012 and November 2013. Households were chosen randomly from 45 municipalites in Santiago's Metropolitan Region.

[27]In March every year, each car owner is required to obtain a circulation permit upon payment of an annual fee to her home municipality. Among other things, this data specifies the number of cars of each vintage by municipality.

[28]With the exception of new vehicles, which are exempt for two years, all vehicles are required each year to pass emission inspections before their circulation permit is renewed for the following year.

is based on individuals' values of time. Cars are heavily used only by higher income groups, while the majority of individuals (84%) in the lowest-income group rely on public transport.

Table 1: Income-group characteristics and preferences

Group No.	Income level	Fraction of total	Range of monthly income per household	Car ownership	Marginal utility of income ($/hr)	Relative transport preferences
1	Low	12%	<$368	16%	1.36	[-2.4 , 3.7]
2	Middle-low	27%	368−734	34%	3.11	[-1.2 , 2.1]
3	Middle	34%	735−1,468	54%	5.89	[-0.8 , 1.8]
4	Middle-high	19%	1,469−2,935	77%	10.47	[0.1 , 1.4]
5	High	8%	>$2,935	95%	28.20	[0.2 , 1.2]

Note: This table contains household characteristics for five income groups based on information from SECTRA (2013), ODS-2012, and our own model calibration.

For each of the income groups in Table 1, SECTRA (2013) provides a value for the marginal utility of time, γ_i^c. To introduce some heterogeneity within each group we let γ_i^c be normally distributed (and truncated at zero) with a mean value equal to SECTRA's numbers (see the fifth column of Table 1) and a standard deviation of 20%. To obtain γ_i^b we use (4) and the following parameters with exception of D_p, which is endogenously provided by the model's equilibrium: $\zeta = 0.2$, the crowding penalty, which builds on the crowding multipliers of Tirachini et al. (2017) for buses and controls for the fact that underground travel time is unaffected by government interventions; $\gamma = 1/2$, which is roughly consistent with the way individuals, at least on average, combine surface and underground public transport; $f_b = 15$ buses/hr, which is from Basso and Silva (2014); and $s = 26.4$ m^2, which is the average bus size currently in the system.

A commuter's marginal utility of income, λ_i, is obtained in two steps. First, we let the income distribution of our sample of n commuters replicate the actual income distribution observed in the ODS-2012, which is very similar to the one in CASEN 2011, a national socioeconomic survey.[29] Then, if Y_i is household i's income, we let $\lambda_i = \lambda_0/Y_i$, where λ_0 is a scaling factor to be obtained from the model calibration (which we will explain shortly).

Intrinsic preferences for public transport (or modal constants), θ_i, are also obtained from the model calibration along with the constants ϕ_0 and λ_0. In particular, for each income group $g = 1, ..., 5$ we assume $\theta_{i\in g}$ to be uniformly distributed between $\bar{\theta}^g - n_g^{nc}/n_g^c$ and $\bar{\theta}^g + n_g^{nc}/n_g^c$, where $\bar{\theta}^g$ is the distribution mean to be obtained from the model calibration, n_g^c is the number of individuals in group g who own a car, and n_g^c is the number of individuals in group g who do not own a car. Calibration results show that intrinsic preferences for public transport of low-income individuals who own a car are lower than those of higher-income

[29]CASEN is the Characterization Survey of Socioeconomic status in Chile and has a much larger number of observations than ODS-2012.

individuals who own a car. As observed in the ODS-2012, the reason is that low-income individuals use their cars, provided they own one, more often than their higher-income counterparts.

3.2 Transport parameters

The remaining parameters concern those that enter in (2) and (5). Following Basso and Silva (2014) and ODS-2012, the round-trip average distance is set to $l = 27.8$ km, the daily cost of using a car to $c = 16.40$ \$/day (the product of the cost per km, \$0.59, and l), the duration of the peak period to $q = 7$ hrs, the car occupancy to $a = 1.5$, the equivalence factor to $\kappa = 2.06$, the capacity of a lane to $C = 900$ car/hr (which, as we explain in the next section, is calibrated along other preference parameters; note also that C takes into account the presence of traffic signals), the network length to $K = 2154$ km, the free-flow speed of cars to $v_f^c = 45$ km/hr and the free-flow speed of buses to $v_f^b = 30$ km/hr. The public-transit (daily) fare is set at its current value of $p_p = \$3.17$ (the product of a single-ride fare, \$1.14, and the average number of daily rides, 2.78).

Values of α_s and β_s, on the other hand, are obtained by fitting function (5) to SECTRA's (2013) congestion-pricing simulations. SECTRA (2013) provides information on changes in transport modes (D_c and D_p) and travel times (t^b and t^c) from three pricing scenarios for the city of Santiago. The fitted values are in Table 2, along the other parameter values with their corresponding sources. While values of α_b and α_c are similar to those found in the literature, values of β_s, particularly β_b, may seem small. The explanation for this is that these parameters are calibrated from actual data that pools in-motion times with stopping times at traffic signals, bus stops and so on. All these operations tend to linearize total travel time, eliminating much of the convexity of the BPR function.

3.3 Calibrating the remaining preference parameters

We still need to find values for seven parameters of the model: λ_0, ϕ_0, and $\bar{\theta}_g$, where $g = 1, ..., 5$. We do this by calibrating the model's predictions to match the modal shares (i.e., fraction of individuals using public transport) observed in the ODS-2012. The loss function we minimize is

$$L = (M_0 - \hat{M}_0)^2 + \sum_{g=1}^{5}(M_g - \hat{M}_g)^2$$

Table 2: Summary of transport parameters[(a)]

Parameter (units)	Symbol	Value	Source
Trip length (km)	l	27.8	ODS-2012
Network length (km)	K	2,154	SECTRA (2013)
Average daily trips		2.78	ODS-2012
Passenger car equivalence factor for buses	κ	2.06	Basso and Silva (2014)
Public transport fare ($)	p_p	3.17	Transantiago[(b)]
Average waiting time at station (min)	w_p	2	Basso and Silva (2014)
Car operating cost ($/day)	c	16.4	following SECTRA (2003)
Car occupancy	a	1.5	ODS-2012
Lane capacity (car/hr)	C	900	Own calibration
Free-flow speed – cars (km/hr)	v_f^c	45	Own estimation
Free-flow speed – buses (km/hr)	v_f^b	30	Transantiago[(c)]
Bus frequency (bus/hr)	f_b	15	Ours, following Basso and Silva (2014)
Bus average size (m^2)	s	26.4	Transantiago[(c)]
Fraction of public transport on surface	γ	0.5	Transantiago[(c)]
Crowding penalty	ζ	0.2	Own estimation[(d)]
Parameters of BPR function – cars	α_c	0.15	Own calibration
	β_c	1.8	Own calibration
Parameters of BRP functions – buses	α_b	0.225	Own calibration
	β_b	1.05	Own calibration

[(a)]More detail on parameter estimation can be found in the online Appendix.
[(b)]EMBARK (2017): *Informe de la evaluación externa al sistema de transporte público remunerado de pasajeros de la provincia de Santiago y de las comunas de San Bernardo y Puente Alto.* Technical report, prepared by the Centro de Transporte Sustentable de México A.C. and the World Resources Institute, México City, February.
[(c)]Transantiago 2016: Boletín de velocidades de servicio para el Sistema, Troncales y Alimentadores.
[(d)]Estimation follows Tirachini et al. (2017).

where M_0 is the observed modal share for the entire population in the ODS-2012, $\hat{M}_0$ is the corresponding model's prediction, M_g is the observed modal share for income group $g = 1, ..., 5$, and $\hat{M}_g$ is the corresponding model's prediction and is given by

$$\hat{M}_g = \frac{1}{5n} \sum_{i=1}^{n} (5 - d_i) \times \mathbf{1}\{i \in g\}$$

where $\mathbf{1}\{i \in g\}$ is an indicator function that takes the value of 1 when individual i belongs to group g and 0 otherwise.

The calibration proceeds as follows. According to SECTRA (2013) and the ODS-2012, Santiago's roads must currently accommodate 2.87 million commuters every weekday during peak hours in a network of 2154 kms of road lane. We assume individuals make a uniform use of the network and rather than working with 2.87 million individuals directly (which would make the computation of any equilibrium extremely slow), we work with a representative sample of $n = 2870$ individuals that is then scaled up by a factor of 1000. Thus, we first take 2870 random draws from ODS-2012's income distribution, which are then allocated according to SECTRA's income grouping. Second, we get random draws for γ_i^c. Third, we assign a value to ϕ_0 and get random draws for λ_i and θ_i after assigning values to λ_0 and to the boundaries of the θ's uniform distributions of each income group. Fourth, we compute the equilibrium based on these parameters and compare the equilibrium modal shares to actual shares. Steps three and four are repeated until L is minimized. As shown in Table 3, calibrated parameters (which are shown in the last column of Table 1) provide a close match between actual modal shares and those predicted by the model for each group and overall. Table 3 also shows small differences in the fraction of individuals in each group predicted by our model and by the ODS-2012, which is the result of working with a small sample.

Table 3: Model fit

	Modal share		Elasticities
Income groups	Observed	Model prediction	
Low	85%	85%	-4,88
Middle-low	72%	69%	-3,83
Middle	61%	62%	-0,43
Middle-high	44%	42%	-0,27
High	25%	20%	0,23
Overall	55%	59%	-0,31

Note: This table shows how our model calibration matches observed data and existing literature. The first and second columns contrast the observed modal shares to the predictions of our model. And the third columns report the average elasticity at the income-group level of increasing the cost of driving in $7.6.

Before moving to policy simulations we run a robustness check to see how the "elasticities" predicted by our model compare to those in the literature. Following Heckman and Vytlacil's (2005) marginal-treatment approach, we compute the equilibrium effect of an increase in the cost of using the car on commuting-mode decisions as the weighted average of the effects of a collection of small cost increases from zero to $7.6, which is when all individuals in the lowest-income group give up their cars.[30] The last column of Table 3 reports that while individuals in the lowest-income group are remarkably sensible to such cost increase (with an "equilibrium elasticity" of -4.88) individuals in the highest-income group actually end up increasing their use of the car, presumably in response to a relatively faster car travel. Our

[30]The weight is the number of new individuals that give up the car at each price increase divided by the total number individual that give up their cars after the complete price increase.

average equilibrium elasticity is about -0.31 (as reported in the last row of the table), and that of the middle-income group is -0.43. Both are within the (long-run) elasticity range reported by Litman (2017, p. 22) for car travel, from -0.10 to -0.63.

3.4 Car ownership and pollution parameters

Cars emit all sorts of pollutants, some with global effects (e.g., CO_2), others with local effects (e.g., CO, HC, NO_x), i.e., effects at the city level that last for a short time, sometimes only a few hours. The focus of this paper is on local pollutants, although there is nothing that prevents it from extending the analysis to include global pollutants.

Since in our short-run analysis car owners only respond along the intensive margin, i.e., by adjusting how much they drive their existing cars but not by adjusting the type of cars they drive, our model only requires information on current car ownership. In particular, we need information at the income group-level on the type of cars individuals drive and their emission rates. The circulation-permit database contains car details (e.g., make, model, vintage, fuel type) for all cars registered at a given municipality in a given year. Unfortunately, their owners' income is not available and is not feasible to cross-check it with other databases.

We know, however, that some municipalities in Santiago are much richer than others and, not surprisingly, they tend to concentrate a larger fraction of newer and larger cars.[31] We exploit this income heterogeneity to build a "portfolio of cars" for each income group in our model that vary by class, vintage and fuel type. To do that, we first characterize the actual portfolio of cars in each municipality according to the circulation-permit database and the NSB vehicle survey (step 1). We then divide each income group in 10 income brackets of equal width but of different weights (step 2). Next, we take municipalities' average income, variance and size to compute the probability that an individual of a certain income is drawn from a particular municipality (step 3). Note that these probabilities may need to be scaled up/down to add to the unity. Finally, we create the income-group vehicle portfolio by weighing each municipality portfolio (step 1) by the weights from step 2 and the probabilities from step 3. The result of this exercise is summarized in Tables 4 and 5.

Table 4 shows, for example, that the most popular class in the 2015 fleet is subcompact (35.2%) followed by SUV (17.1%).[32] The table also shows a large class of other cars (24.8%)

[31] There are 40 municipalities in Santiago with income per-household varying from an average of \$940 per month (s.e. = \$72) in Cerro Navia to \$4422 (s.e. = \$171) in Las Condes.

[32] According to ANAC, the national association of car dealers, SUV has been the best the selling class in recent years, with 37% of the market for new units (https://www.anac.cl/wp-content/uploads/2019/10/09-ANAC-Mercado-Automotor-Septiembre-2019.pdf).

Table 4: Car ownership as a function of income and vehicle class and fuel

	Income group					
Vehicle class/fuel type	Low	Middle-low	Middle	Middle-high	High	Total
Subcompact	2.4%	3.6%	5.8%	9.4%	14.0%	35.2%
Gasoline	1.6%	2.8%	5.1%	8.80%	13.0%	31.5%
Diesel	0.8%	0.8%	0.7%	0.6%	0.8%	3.7%
Compact	0.3%	0.9%	2.1%	3.9%	6.6%	13.8%
Gasoline	0.1%	0.7%	1.8%	3.7%	6.2%	12.5%
Diesel	0.2%	0.2%	0.3%	0.3%	0.4%	1.4%
Midsize	0.1%	0.1%	0.5%	1.1%	2.1%	3.9%
Gasoline	0.0%	0.0%	0.4%	1.0%	2.0%	3.4%
Diesel	0.1%	0.1%	0.1%	0.1%	0.1%	0.5%
Pick-up truck	0.8%	0.9%	1.0%	1.2%	1.4%	5.3%
Gasoline	0.3%	0.3%	0.3%	0.4%	0.4%	1.7%
Diesel	0.6%	0.6%	0.7%	0.8%	1.0%	3.7%
SUV	0.6%	1.3%	2.5%	4.5%	8.2%	17.1%
Gasoline	0.2%	0.7%	1.7%	3.4%	6.4%	12.4%
Diesel	0.4%	0.5%	0.7%	1.1%	1.9%	4.6%
Other models	2.1%	2.5%	3.5%	5.4%	11.3%	24.8%
Gasoline	1.7%	2.1%	3.1%	5.2%	10.8%	22.9%
Diesel	0.4%	0.4%	0.3%	0.3%	0.5%	1.9%
Total	6.3%	9.3%	15.4%	25.5%	43.6%	100%

Sources: 2015 Vehicle Survey of the National Statistics Bureau and 2015 circulation-permit dataset.

that for the most part include minivans, luxury and fullsize models, but also a few other models that we were not able to classify. More interesting for our analysis, however, is that the distribution of classes varies greatly among the different income groups. In fact, for each SUV in the lowest-income group (i.e., group 1) there are 14 in the highest-income group (i.e., group 5); whereas, for each subcompact in group 1 there are only 6 in group 5. The more even distribution of subcompact models is largely explained by their lower prices.

From Table 4 we also infer that the majority of cars in Santiago run on gasoline (84.3%) as opposed to diesel (15.7%).[33] This is an important distinction to keep in mind when it comes to estimate emission rates, e_i. Another important distinction is the age of vehicles, which is summarized in Table 5. As already documented by Barahona et al. (2020), lower income groups tend to own a larger fraction of older cars. These two factors, vehicle age and fuel type, together with vehicle class explain much of the difference in emission rates

[33]According to ANAC's report of the previous footnote, electric and hybrid vehicles accounted for less than 0.1% of the total fleet in the country by 2018.

Table 5: Car ownership as a function of income and vehicle class and age

Income group No.	Class	≤1990	1991-1995	1996-2000	2001-2005	2006-2010	2011-2015	Fraction of total
1	Subcompact	0.15%	0.49%	0.70%	0.20%	0.23%	0.00%	1.77%
	Compact	0.02%	0.11%	0.35%	0.00%	0.00%	0.05%	0.53%
	Midsize	0.08%	0.06%	0.00%	0.00%	0.00%	0.00%	0.14%
	SUV	0.10%	0.26%	0.27%	0.00%	0.00%	0.05%	0.68%
	Pick-up truck	0.01%	0.05%	0.15%	0.19%	0.20%	0.01%	0.61%
	Other models	0.32%	0.16%	0.22%	0.00%	0.11%	0.72%	1.52%
				⋮				
5	Subcompact	0.16%	0.69%	1.74%	3.18%	4.00%	3.18%	12.96%
	Compact	0.03%	0.26%	1.11%	1.40%	1.71%	2.08%	6.59%
	Midsize	0.10%	0.25%	0.40%	0.35%	0.35%	0.28%	1.74%
	SUV	0.11%	0.39%	1.09%	1.42%	2.45%	3.96%	9.43%
	Pick-up truck	0.01%	0.04%	0.14%	0.28%	0.48%	0.66%	1.59%
	Other models	0.58%	0.45%	1.20%	1.45%	2.87%	12.83%	19.38%
Total		3.86%	7.48%	16.36%	16.06%	22.21%	34.03%	100%

Note: Information on remaining income groups can be found in the online Appendix.

across vehicles. To estimate these emission rates we need to convert the gas readings in the smog-check database to emissions in grams per kilometer traveled using relevant vehicle information such as vintage, weight and fuel type.[34] As an illustration of the emission rates we use in our simulations, Table 6 below reports HC and NO_x emission rates for selected classes and vintages.[35,36] For example, a year-2000 SUV that runs on diesel emits 54 times more HC per kilometer traveled than a year-2010 subcompact that runs on gasoline.[37]

There are three remaining parameters in the model that require our attention. One is the allocation of cars within each group, since not all individuals own a car (see fifth column of Table 1). We allocate cars to those that prefer them most, i.e., with lower θ_i, so that if $i \in g$ owns a car and $j \in g$ does not, then $\theta_i < \theta_j$. The second parameter is the possibility of a within-group relationship between car-usage decisions and car characteristics, owner income (Y_i) and her transport preference (θ_i). We assume there is none. This implies that if a toll p_c and threshold $\bar{e}$ result in that a fraction of car owners in group g with relatively clean cars (i.e., $e_i < \bar{e}$) decide in a day of restriction not to pay the toll and leave their cars at home that day, any of those relatively clean cars, regardless of age, fuel type, class, and

[34]For gasoline vehicles we use the conversion equations reported in Morrow and Runkle (2005) and for diesel cars we use the conversion (diesel/gasoline) rule developed by Cifuentes (2018).

[35]Emission rates for remaning classes and years are in the online Appendix.

[36]By being precursors to ground-level ozone and contributing to the formation of PM2.5, HC and NO_x are responsible for major health problems in Santiago (Rizzi and De la Maza 2017) as well as in other cities (Fullerton and West 2010).

[37]That older cars (locally) pollute significantly more than newer cars is also documented by Molina and Molina (2002) and Knittel and Sandler (2018).

Table 6: Emission rates for selected classes and vintages

	Selected classes					
	Subcompact		Compact		SUV	
Vintage	Diesel	Gasoline	Diesel	Gasoline	Diesel	Gasoline
			Panel A: HC			
1995	0.535	0.388	0.589	0.396	0.641	0.444
2000	0.282	0.181	0.297	0.148	0.325	0.229
2005	0.108	0.041	0.104	0.035	0.123	0.058
2010	0.020	0.006	0.021	0.007	0.024	0.008
2015	0.009	0.003	0.006	0.002	0.006	0.002
			Panel B: NO_x			
1995	0.496	0.487	0.476	0.469	0.737	0.579
2000	0.253	0.200	0.222	0.190	0.372	0.252
2005	0.085	0.035	0.033	0.029	0.132	0.044
2010	0.012	0.004	0.012	0.004	0.015	0.005
2015	0.001	0.000	0.002	0.001	0.003	0.001

Note: Emission rates (in gr/km) are obtained by converting smog-check gas readings from the 2015 smog-check dataset, which are in ppm, to grams per kilometer using formulas in Morrow and Runkle (2005) for gasoline cars, and the proportional (diesel/gasoline) rule in Cifuentes (2018) for diesel cars.

owner's Y_i and θ_i, are equally likely to be left at home. And the third parameter is the value of h in (6), i.e., the harm per gram of pollution. We borrow from BGM that a percentage point increase in vehicle emissions generates a pollution cost in Santiago of \$6.8 million in fall and winter and a fifth of that in spring and summer.[38]

4 Testing the taking turns' Pareto-improving property

We now use our model and the parameter values estimated in the previous section to test for the Pareto-improving property of the taking-turns scheme and run additional policy simulations. To facilitate the presentation, in this and the following section we concentrate exclusively on toll exemptions, that is, we set $\bar{e}$ sufficiently high so that no car owner is prevented from paying the toll in days of restriction. It turns out this to be optimal during spring and summer, when emissions are not nearly as harmful as they are in fall and winter. We extend the simulations to include vintage exemptions in Section 6, when pollution calls

[38]We arrived at these numbers in two steps. First, we update BGM's annual cost of \$3.4 million per percentage point from 2006 to 2015 dollars. And second, we decompose their cost figure in spring/summer and fall/winter periods using Gallego, Montero, and Salas (2013), who show that Santiago's atmospheric pollution concentrations —the ones responsible for health problems— in spring and summer are approximately a fifth of those in fall and winter despite the same level of vehicle emissions. Thus, $\$6.8 = \$3.4 \times 1.2 \times 2/1.2$.

for a binding $\bar{e}$.

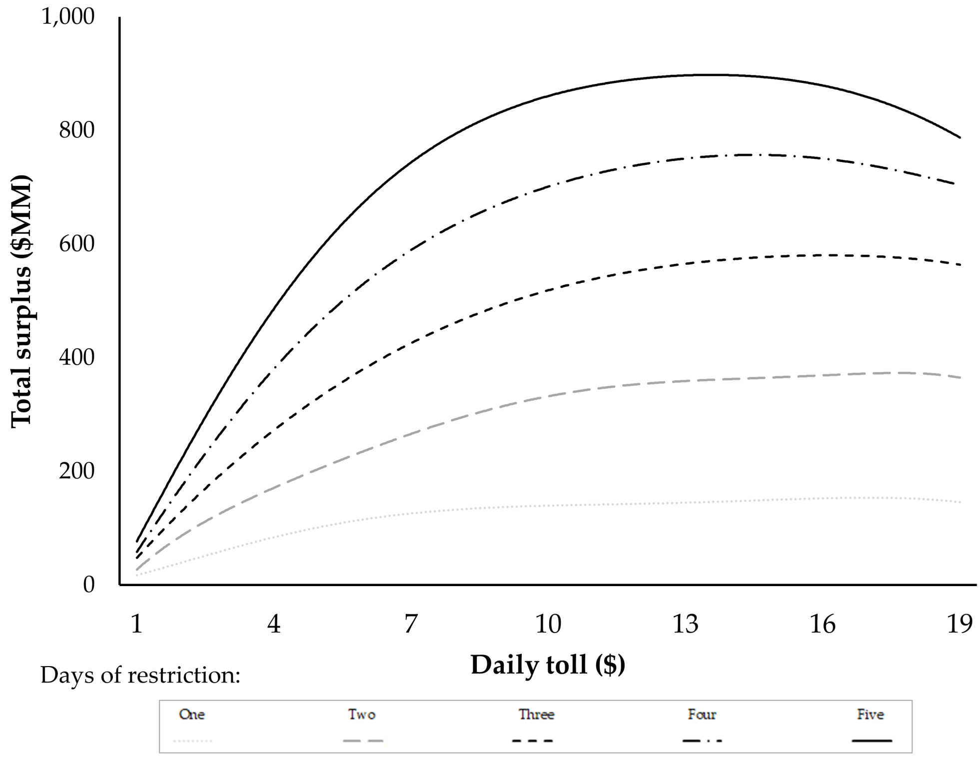
Our first set of simulations considers a distributionally neutral recycling policy as envisioned by Daganzo: all the revenue collected from toll payments is returned in a lump-sum fashion to individuals while preventing any transfers between income groups and from individuals that own a car to those that do not. In other words, the entire toll collection coming from individuals that own a car in group $g = 1, ..., 5$ is returned to that same subset of individuals. Using this neutral revenue-recycling criteria, we exclusively concentrate on efficiency and distributional implications of allocating road capacity among heterogenous users. It is as if a perfectly informed (surplus-maximizing) planner could directly inform each individual that owns a car whether she can use it or not.

Although Daganzo's scheme considers 1 or 2-day restrictions, in what follow we present results for all possible restriction formats, from a 1-day restriction per week ($r = 1$) to a 5-day restriction ($r = 5$). Aggregate results are summarized in Figure 1. Each curve depicts changes in aggregate consumer (transport) surplus (i.e., $\sum_{i=1}^{n} \Delta S_i$) as a function of the toll set by the planner. The upper curve indicates that the surplus-maxizing toll under a 5-day restriction is around \$12.9 (similar to the daily congestion charge in London, £11.5). Implementing this surplus-maximizing policy leads to important increases in travel speed: 43% in the case of cars (from the benchmark speed of 23.7 km/hr to 33.8 km/hr) and 20% in the case of buses (from the benchmark speed of 18.4 km/hr to 22.0 km/hr). These savings in travel time explain the substantial surplus gain from setting the toll at its optimal level: \$899 million per year, or 0.37% the country's GDP.[39]

It may be argued that a toll of \$12.9 may prompt some individuals to by-pass the restriction not paying the toll but buying a second car. Since our model cannot handle this possibility by construction (individuals own at most one car), we offer a separate analysis in the online Appendix showing this to be unlikely. Interestingly, lowering the toll in the 5-day restriction to cope with this "second-car" concern, say, by about 20% (from its optimal level to \$10.6), does not entail much of an efficiency loss.[40] The reason is because few commuters change their decisions at these price levels (see last column of Table 3). Indeed, at the surplus-maximizing toll all individuals in groups 1 and 2 commute by public transport

[39]Interestingly 0.4% of GDP (or \$87 billion) is the congestion cost (i.e., hours lost in traffic) estimated by INRIX (2019) for the entire United States in 2018. Similar estimations are in De Palma and Lindsey (2011). Also, note that by dividing our congestion benefit figure (\$899 million) by the reduction in kilometers traveled by cars (9.57 million kilometers per day for 260 days of the year) we arrive at an average congestion externality of \$36.1 per kilometer, which is very close to the \$44.9 figure that Rizzi and De la Maza (2017) obtained during peak hours for Santiago but following a different approach.

[40]Even if we reduce the toll by 50%, to \$6.5, the efficiency loss is not that great, of 24%. Interestingly, these numbers are comparable to the numbers in Braid's (2018) bottleneck model, who reports a reduction of 25% in benefits if the maximum toll is reduced by 50%.



Note: This figure depicts changes in total transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels. Toll revenues are given back to individuals in a lump-sum neutral fashion (i.e., all revenues generated in group $g = 1, ..., 5$ are returned to car owners in that same group).

Figure 1: Total transport surplus change under neutral recycling

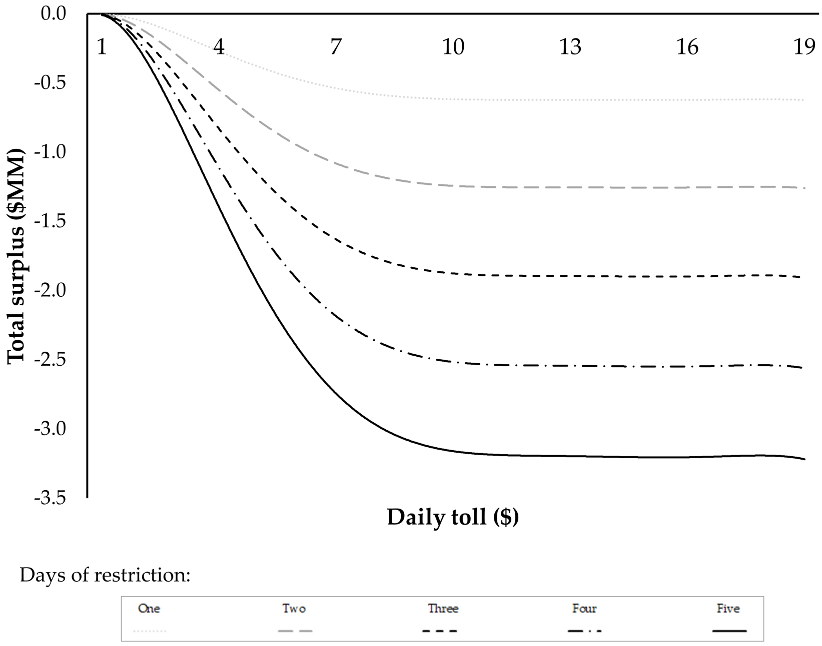
(compared to 85% and 69%, respectively, in the benchmark scenario), 78% in group 3 (compared to 61% in the benchmark scenario), 53% in group 4 (compared to 35% in the benchmark scenario), and 7% in group 1 (compared to 20% in the benchmark scenario).

The remaining curves in Figure 1 show how efficiency is greatly compromised as we reduce the number days in which commuters must pay a toll every time they use their cars. In fact, the maximum surplus to be achieved with a 4-day restriction (i.e., when the toll is set at its optimal level of $13.6) is the same as the one achieved under a 5-day restriction with a toll set at almost half of its optimal level, $7.1. Moreover, if the planner decides to implement a 1-day restriction, the maximum surplus gain, which can be obtained with a toll anywhere between $10 and $20, is equivalent to that under a 5-day restriction with a toll as low as 10% of its optimal value. Strictly speaking, as the planner opts for restriction formats with fewer days of restriction, the optimal toll goes up in an effort to compensate for fewer car owners facing a price. At the same time, however, there is a wider range of toll values that deliver similar outcomes (see Figure 1).

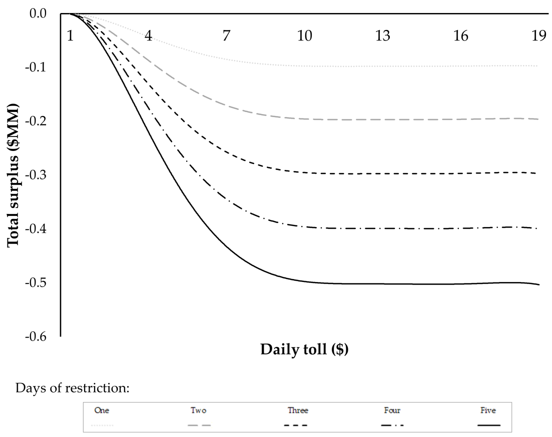
Despite 1- or 2-day restrictions may appear quite inefficient, one can imagine a situation in which a regulator is willing to sacrifice efficiency in favor of more equitable outcomes that leave individuals in all income groups better off, as Daganzo has argued. If these more equitable designs have a much better chance at being implemented in practice, this is a compromise worth considering. In case 1 or 2-day restrictions end up leaving all individuals better off, it would be for different reasons. High-income individuals who own a car would benefit from these restrictions as they will continue commuting by car every day (and paying the toll the day or days of restriction) but faster. Low-income individuals who own a car, on the other hand, would incur a loss the day or days of restriction since they would have no choice but to switch to public transport. This loss, however, would be more than compensated with the gain of faster (car) travel during the rest of the week, i.e., during days of no restriction.

While this Pareto-improving possibility cannot be discarded in theory, our simulations tell otherwise. As shown in Figure 2 (panel a), all individuals in the lowest-income group (i.e., group 1), regardless of whether they own a car or not (panels b and c), are worse off in any of the policy designs and for any toll level. There are two reasons for this. For individuals who own a car (panel b), the benefit of driving faster the days of no restriction is not enough to compensate for the loss of leaving the car at home the days of restriction and riding the public-transport system, which is now more crowded. The latter explains why individuals who do not own a car (panel c) are also worse off: the time savings in public transport (from slightly faster buses) are not enough to compensate for the inconvenience of riding a more crowded system.

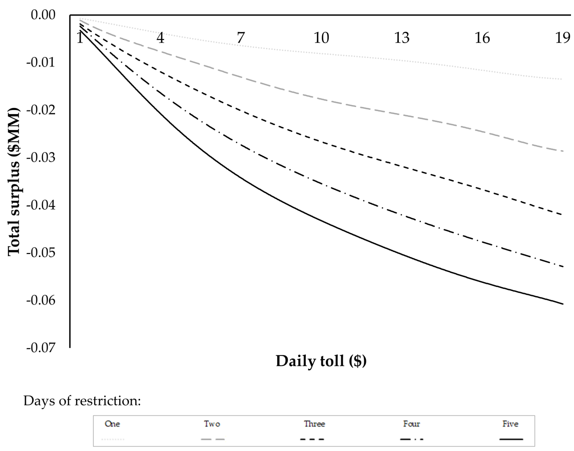
As shown in Figure 3, individuals in the highest-income group (i.e., group 5) that commute by public transport are also worse off for the same reasons that individuals in group 1 are (panel c). The difference with group 1 is that the majority of individuals in group 5 commute by car, so overall they benefit greatly from any of the restriction designs, particularly from the 5-day restriction. The impact upon the other income groups can be found in the online Appendix. An interesting difference with what we see in Figures 2 and 3 is that car owners in groups 3 and 4 benefit from the restrictions, in any of their formats, as long as the toll is set not too high, so that many of them continue using their cars on days of restriction.



(a) Total



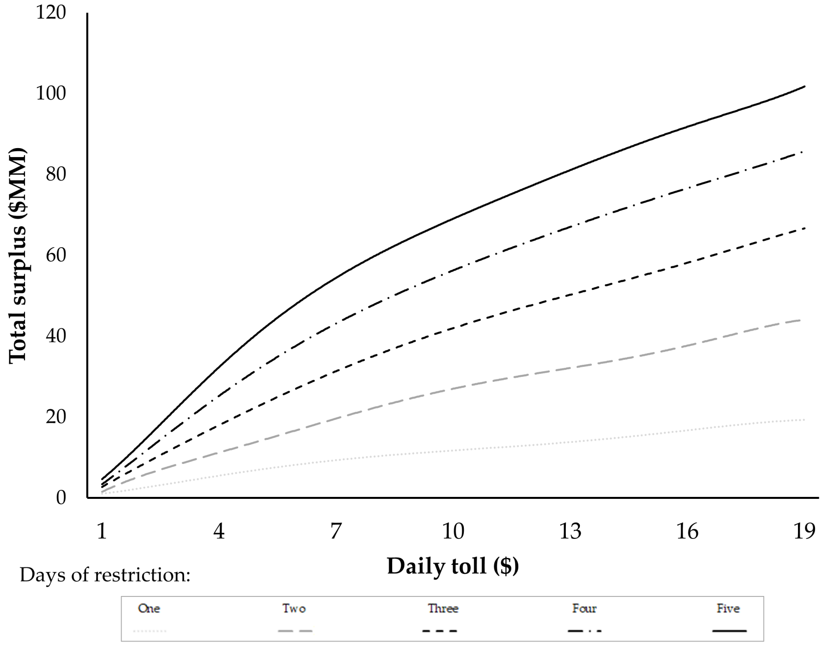
(b) Car owners



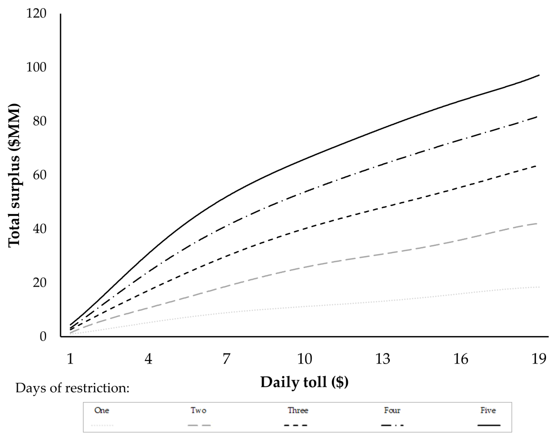
(c) Public transit users

Note: This figure depicts changes in group 1's transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels. All toll revenues collected from this group are given back to car owners in that group in a lump-sum fashion.

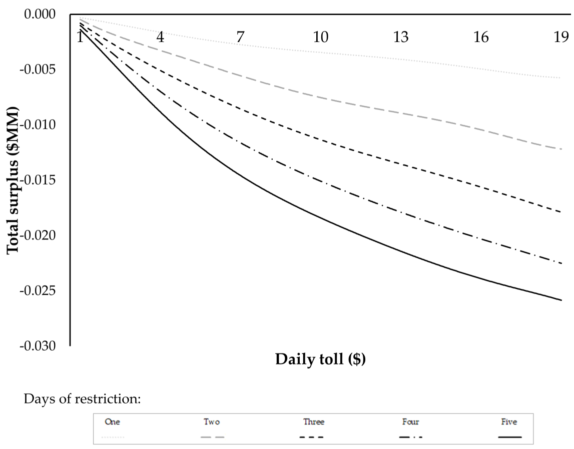
Figure 2: Group 1's surplus change under neutral recycling



(a) Total



(b) Car owners



(c) Public transit users

Note: This figure depicts changes in group 5's transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels. All toll revenues collected from this group are given back to car owners in that group in a lump-sum fashion.

Figure 3: Group 5's surplus change under neutral recycling

Given that anything less than a 5-day restriction inflicts large inefficiencies without alleviating any distributional concerns —groups 1 and 2 are strictly worse off in any of the restriction designs—, an important lesson from this discussion is that any authority currently constrained, for practical reasons, to implement anything less than a 5-day restriction should only see it as a necessary first step towards the implementation of a full road-pricing scheme in the future. Another important lesson is that distributional concerns should (and can, as we will see next) be handled separately without compromising efficiency.

5 Addressing distributional concerns

Given the failure of Daganzo's premise —that 1- or 2-day restriction could leave everyone better off— we need to find ways to more evenly distribute the aggregate benefits of any of the policy designs described above. In theory, this should help reach wider public support for their implementation.

With that aim, in this section we consider two alternative uses of toll collection.[41] In the first case, the totality of the collection is directed to reduce the public-transport fare p_p.[42] In the second case, the collection is divided between reductions in p_p and improvements in service quality, more precisely, in surface service frequency f_b.

5.1 Public-transport fare reduction

Consider the most ambitious restriction format —5-day restriction— and that the complete toll collection is used to reduce the fare p_p. The surplus-maximizing allocation is the same as before, only now the optimal toll p_c drops to \$10.6 and p_p to \$0.97, a 70% reduction from its current level, so that the driver who is indifferent to taking the car every day or riding public transit remains unchanged: in either case the difference $p_c - p_p$ is the same, about \$10.[43]

Though fare reduction has little impact on the overall surplus going to group 5 (see Figure 5, panel a), and on how this surplus is split between those who own a car and those who do not (panels b and c), it has a great impact on the overall surplus going to group 1 (see Figure 4, panel a). All individuals in group 1 who do not own a car benefit greatly from the fare reduction (panel c) whereas those who own a car do not benefit nearly as much, if at all (panel b). Since these latter individuals give up their cars rather quickly in response to a toll increase, they only benefit from the 5-day restriction when the toll is high

[41]The need for revenue recycling to alleviate adverse distributional impacts was already mentioned by Basso and Jara-Díaz (2012).

[42]Note, as in most cities around the world, p_p is already a subsidized fare, so the assumption here is that toll revenues come as additional resources and *not* to replace existing subsidies.

[43]For more detail see the online Appendix, where we show changes in the price of a single ride as a function of different restriction formats and toll levels.

enough, i.e., when is set at its optimal level of \$10.6, so the fare-reduction compensation is accordingly large.

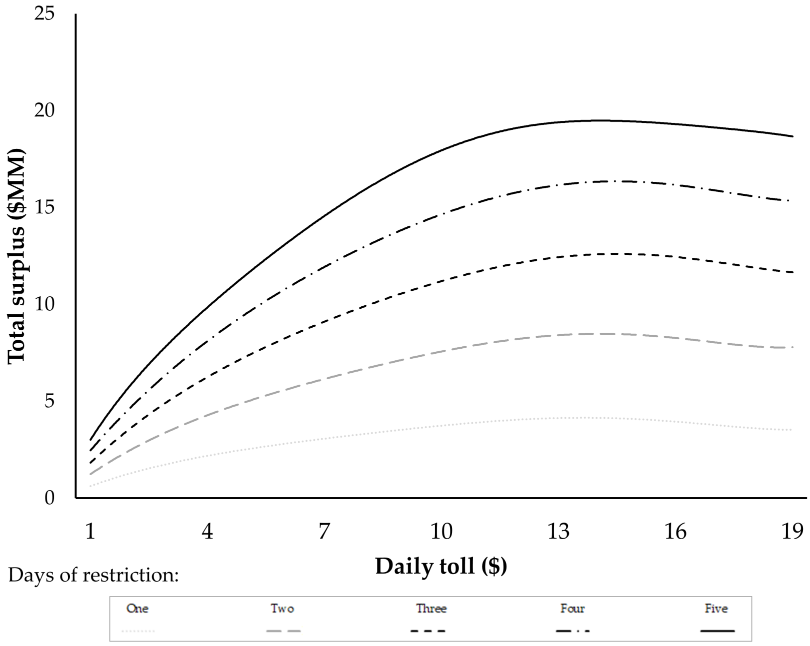
Extending this analysis to other income groups and to less ambitious restriction formats is also informative. Individuals who do not own a car are all better off with the fare reduction except those in group 5 (the online Appendix contains equivalent figures for groups 2, 3 and 4). For these high-income individuals, the fare reduction is not enough to compensate for a more crowded public-transit system.

One important difference with the less ambitious formats of the previous section (see Figure 1) —where toll payments are recycled within each group— is that the maximum attainable overall surplus could be much higher now (see Figure 6). This is particularly evident in the one-day restriction, where the maximum attainable surplus goes up by 52%, from \$152 million to \$231 million (in the two, three and four-day restrictions, it goes up by 37%, 19% and 8%, respectively). The reason for this increase is simple: in days of no restriction, some drivers may decide to leave their cars at home because public transit is cheaper now.

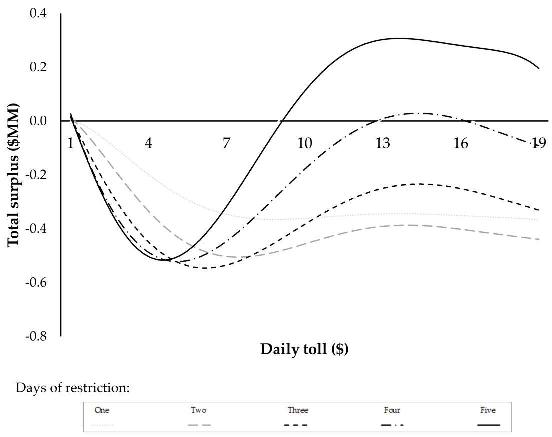
5.2 Mixed recycling: fare reduction and frequency increase

It is clear that if the toll collection can be used not only to reduce the public transit fare p_p but also to increase its surface service frequency f_b, overall surplus can only go up (unless, of course, f_b is already at a suboptimal high level, which is not the case). Assuming that the (annualized) cost of increasing f_b by one unit (from its current level of 15 bus/hr) is \$270 per kilometer of road lane (SECTRA 2013), the surplus-maximizing solution under a 5-day restriction consists of increasing f_b to 19 bus/hr while keeping p_c at \$10.6 and increasing p_p to \$1.43.[44]

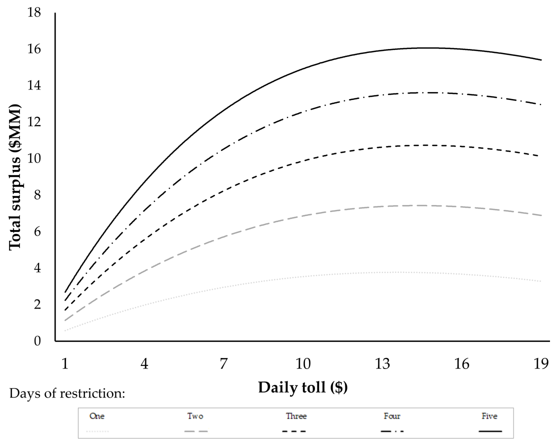
[44]Note that this not only reduces crowdiness in the public-transit system but also waiting times from the baseline level of $w^p = 2$ min to $w^p = 1.6$ min.



(a) Total



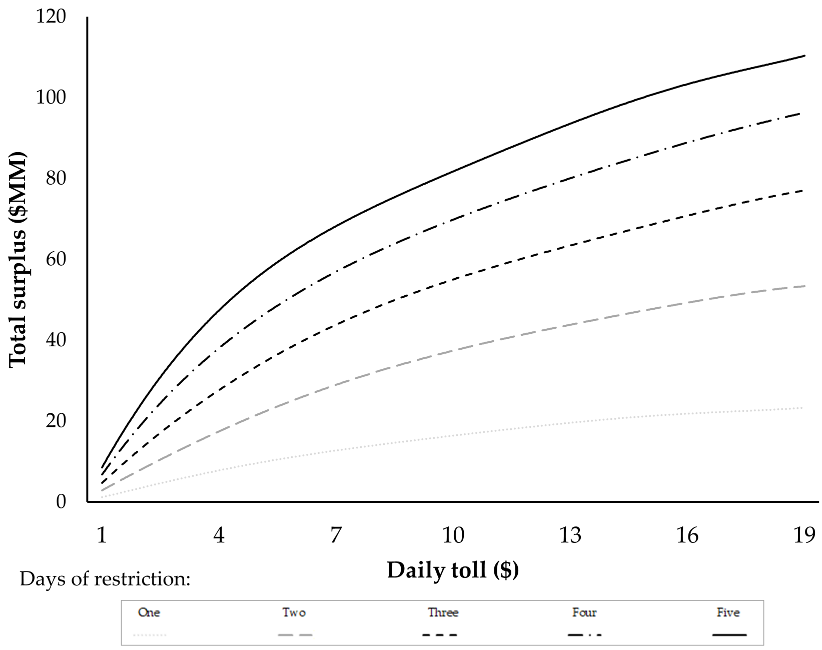
(b) Car owners



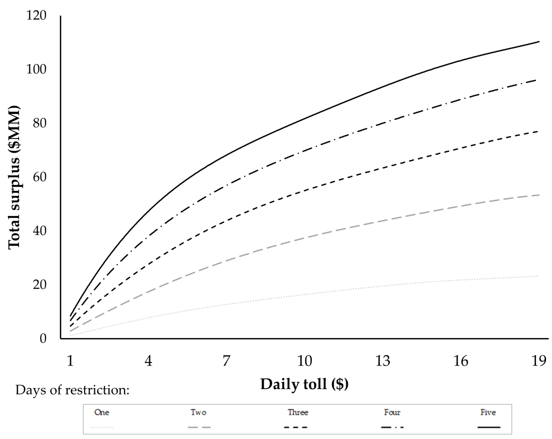
(c) Public transit users

Note: This figure depicts changes in group 1's transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels. All toll revenues collected from this group are used to reduce the transit fare.

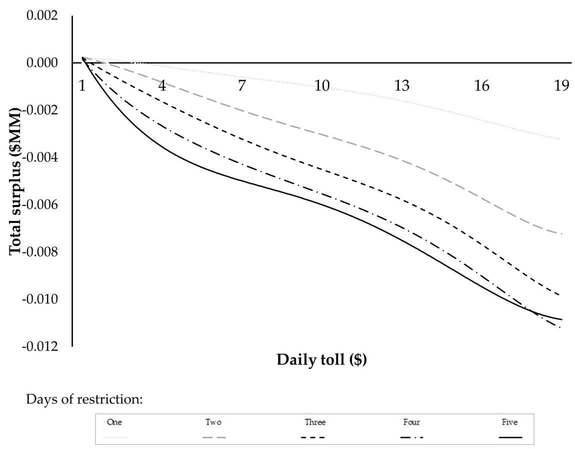
Figure 4: Group 1's surplus change under fare-reduction recycling



(a) Total



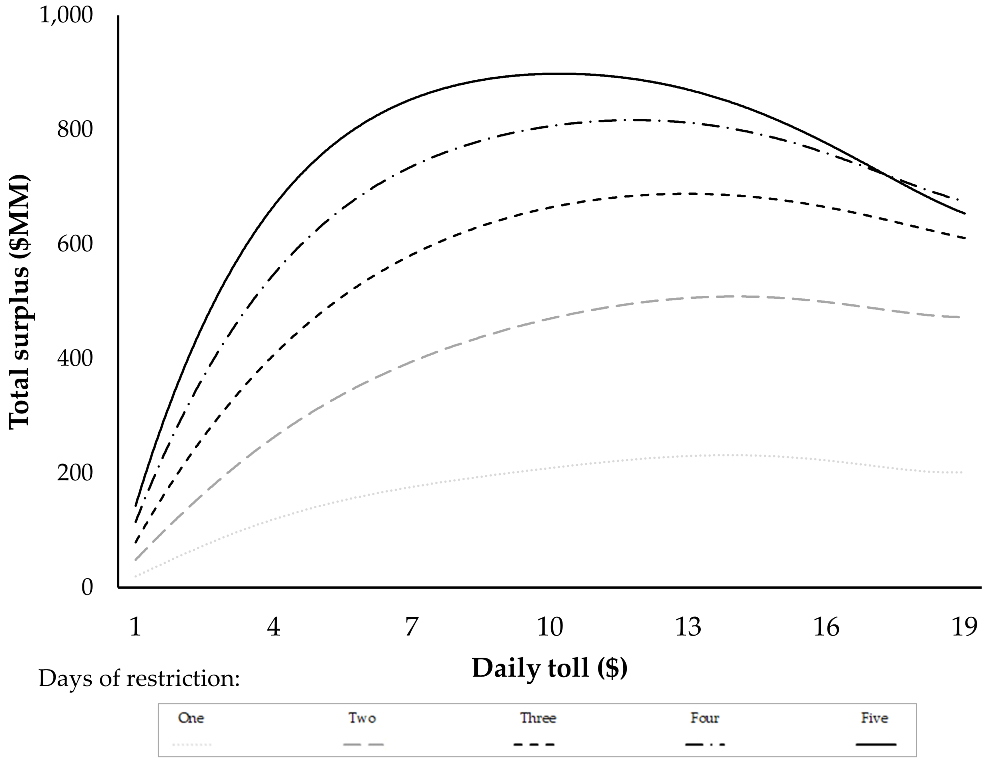
(b) Car owners



(c) Public transit users

Note: This figure depicts changes in group 5's transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels. All toll revenues collected from this group are used to reduce the transit fare.

Figure 5: Group 5's surplus change under fare-reduction recycling



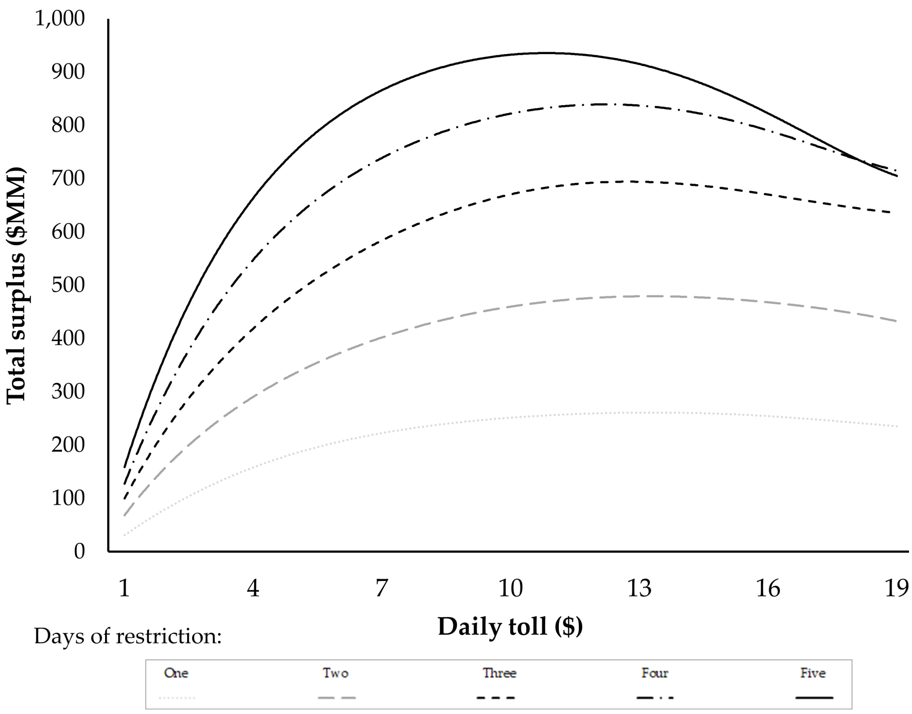
Note: This figure depicts changes in total transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels when all toll revenues are used to reduce the public-transit fare.

Figure 6: Total surplus change under fare-reduction recycling

The cost of (optimally) increasing f_b takes up 17% of the toll collection, leaving less for fare reduction, explaining why p_p cannot be reduced as much. Less obvious is why p_c remains unchanged, at the level when toll revenues are entirely allocated to transit-fare reduction. It is the work of two opposing forces that happen to cancel out. One the one hand, a better service (i.e., higher f_b) calls for a lower p_c in an effort to switch drivers to public transit. And, on the other, a higher p_p calls for a higher p_c, so as to preserve such switching effort.

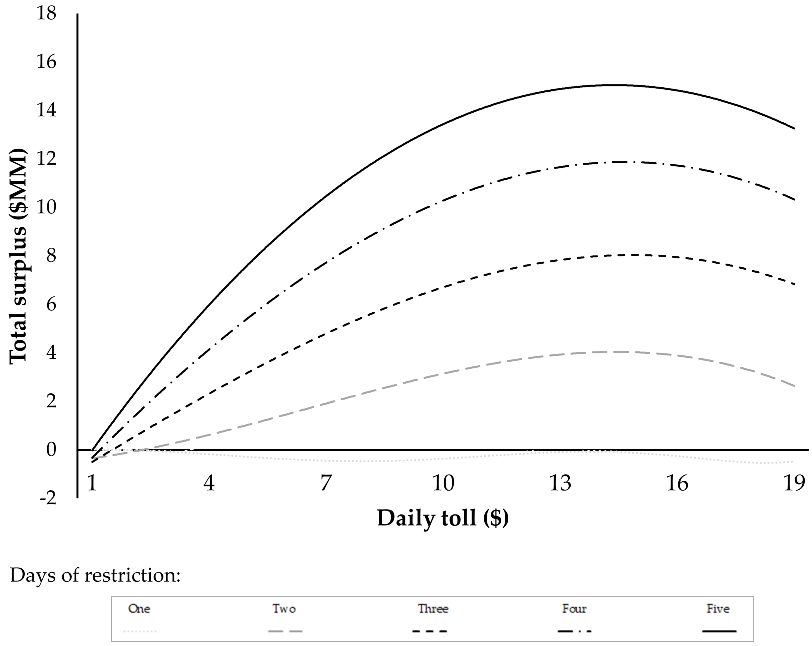
Although this mixed recycling option has a positive impact on overall surplus (see Figure 7), increasing it by 3.9%, it is not necessarily better for all individuals. It favors individuals for whom savings in travel time are relatively more important than savings in fare. This is particularly noticeable in individuals in group 5 who do not have a car, who are now strictly better off (see Figure 9, panel c); on the contrary, individuals in group 1 who own a car are now worse off in all scenarios, even under the 5-restriction format (see Figure 8, panel b). For the remaining groups the situation is not much different as in the previous section (see online Appendix). For instance, individuals in groups 2 and 3 who do not have a car are still better off than without the policy, but not as much as with a recycling option targeted exclusively to fare reduction.

Figure 7 also illustrates how groups are impacted as we move to less ambitious restriction formats. Note that as we reduce the number of days of restriction the (overall) surplus-maximizing frequency may go up a bit, provided the toll collection is sufficiently high to pay for it (in a one-day restriction it is not possible to reach the optimal frequency level despite the entire toll collection goes to pay for it). Perhaps the most interesting finding coming from these less ambitious formats is that distributional impacts are non monotonic. All individuals benefit the most from a 5-day restriction with the toll set at its optimal level, except those car owners in groups 2 and 3 (see the online Appendix). The former would prefer no restriction whatsoever and the latter a 5-day restriction but with a toll set at a lower level.

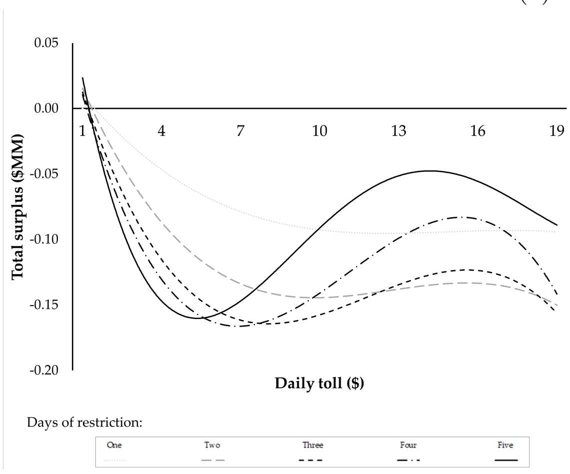


Note: This figure depicts changes in total transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels under mixed recycling, i.e., when toll revenues are used both to: (i) reduce the public-transit fare, and (ii) increase surface public-transit frequency.

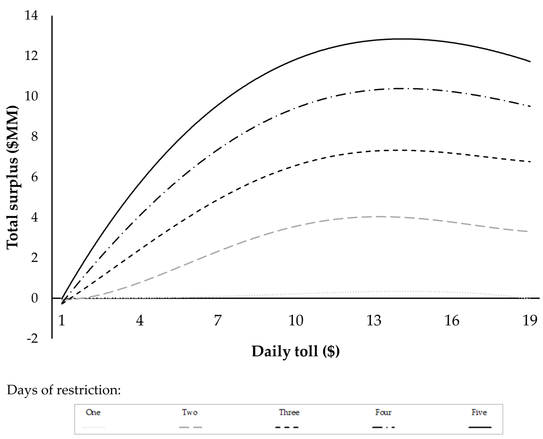
Figure 7: Total surplus change under mixed recycling



(a) Total



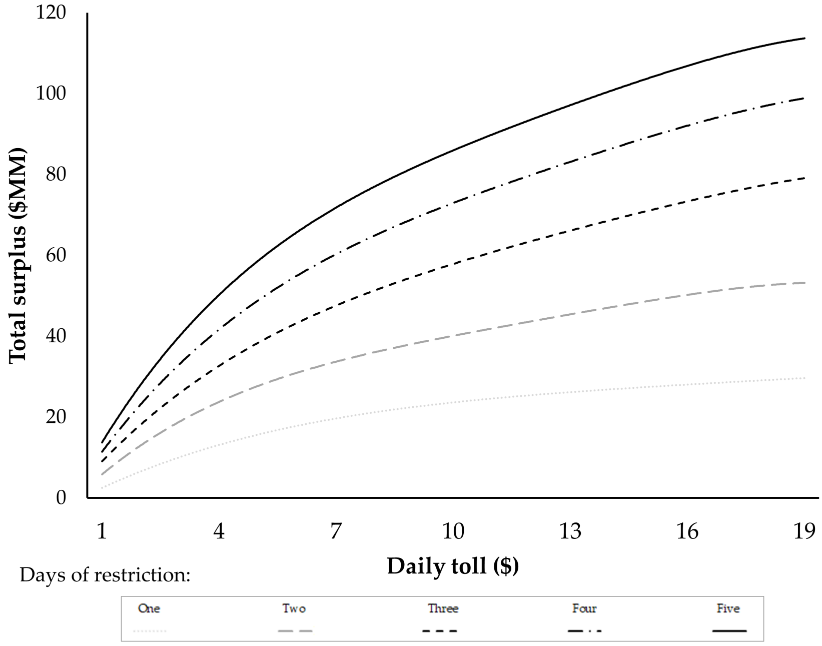
(b) Car owners



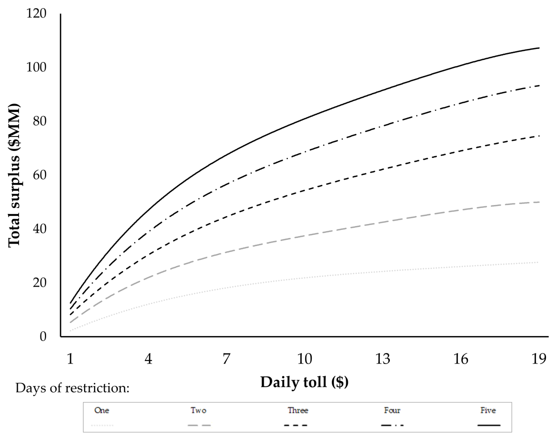
(c) Public transit users

Note: This figure depicts changes in group 1's transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels under mixed recycling, i.e., when all toll revenues are used to both: (i) reduce the transit fare, and (ii) increase the transit frequency.

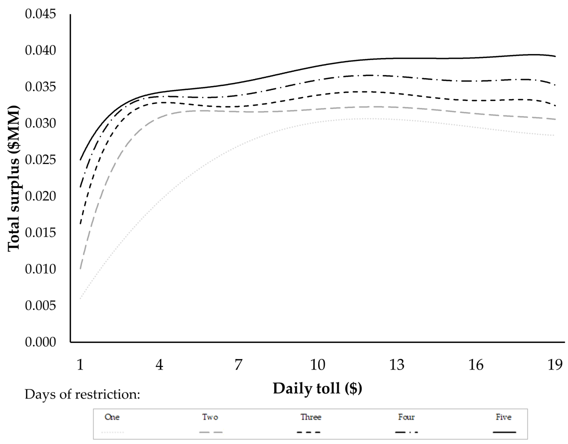
Figure 8: Group 1's surplus change under mixed recycling



(a) Total



(b) Car owners



(c) Public transit users

Note: This figure depicts changes in group 5's transport surplus for different restriction formats, from 1 to 5 days of restriction, and toll levels under mixed recycling, i.e., when all toll revenues are used to both: (i) reduce the transit fare, and (ii) increase the transit frequency.

Figure 9: Group 5's surplus change under mixed recycling

Since a similar non-monotonic pattern is observed under the previous recycling option, two important observations arise. The first is that heterogeneity across individuals (in terms of time, monetary preferences, and crowding costs) is so great that how toll collection is recycled can have quite different impacts on some individuals. As higher-income individuals prefer to spend more on increasing f_b while lower-income individual on reducing p_p, either recycling options necessarily involve some trade offs. The second observation is that no recycling scheme (other than direct transfers) leaves all groups and subgroups better off (this is no longer true, as we add pollution benefits; see next section). Car owners in group 2 end up worse off no matter what recycling scheme is used. For the remaining individuals the best course of action is to implement the most ambitious restriction (5-day restriction), with the toll set at its optimal level and much of its collection spent to reduce p_p as opposed to increase f_b. There is however, an element absent in the analysis so far that has the potential to leave all individuals better off, namely, the gains from cleaner air. We turn to this next.

6 Adding vintage exemptions

6.1 Coping with the pollution externality

So far we have assumed that vehicle emissions raise no concern insofar no car owner is prevented from paying the toll in days of restriction. In an effort to capture Santiago's pollution reality in fall and winter, we now turn to a situation where vehicle emissions do call for additional measures. To see why and how, consider the most ambitious design of Section 4: a 5-day-a-week restriction with a daily toll set at its optimal level, \$12.9 (recall that in this benchmark design, toll revenues play a neutral distributional role). While not directly designed to reduce emissions, this congestion-only policy already does that: HC and NO_x emissions drop by 27.7 and 26.4%, respectively, relative to the no-intervention scenario.

Despite this noticeable reduction, there are still too many high emitting vehicles circulating around from a social-welfare standpoint. In terms of equation (7), we are saying that there are net social gains of reducing ΔH at the expense of reducing $\sum_{i=1}^{n} \Delta S_i$. Following London, the best option to capture these extra gains would be to charge vehicles not only for their external congestion costs, as the toll already does under a 5-day-a-week-restriction scheme, but also for their external pollution costs. Here, however, we adopt a different approach following recent driving restriction programs (see BGM). We prevent owners of high-emitting cars from having the option to pay the toll in days of restriction.

We will go a step further, for reasons we elaborate in the concluding section. Instead of separating cars according to an emissions rate $\bar{e}$, we separate them according to vintage and fuel type, which happen to proxy emission rates reasonably well. Thus, owners of any class of models are entitled to pay the toll p_c as long as their cars are of certain vintage, say, τ^g and younger for gasoline cars and τ^d and younger for diesel cars. Optimal values of τ^g, τ^d and p_c are to be found by simultaneously solving the following system of first-order conditions: (i) $\sum_{i=1}^{n} \partial S_i(\tau^g, \tau^d, p_c)/\partial\tau^g - \partial H(\tau^g, \tau^d, p_c)/\partial\tau^g = \epsilon^g$, (ii) $\sum_{i=1}^{n} \partial S_i(\cdot)/\partial\tau^d - \partial H(\cdot)/\partial\tau^d = \epsilon^d$, and (iii) $\sum_{i=1}^{n} \partial S_i(\cdot)/\partial p_c - \partial H(\cdot)/\partial p_c = 0$, where ϵ^g and ϵ^d are the smallest possible errors that respect the fact that vintage is an integer variable.

Based on the average external pollution cost figures of Rizzi and De la Maza (2017) for the city of Santiago, BGM disentangle the external (local) pollution costs of different vintage models during fall and winter, which range from \$0.03 per kilometer for models less than 4 years old to \$2.85 for models more than 24 years old.[45] According to these numbers, they find that it is socially optimal to fully ban from circulation in the city of Santiago any car older than 16 years, because at that threshold the external cost is exactly equal to the private benefit from driving that car (net of the benefit from using public transport). Thus, cars older than 16 years contribute with negative net social benefits, so it is socially optimal to fully restrict their use, while newer cars contribute with positive net benefits, so it is optimal not to restrict them at all.[46]

Optimality conditions (i) and (ii) above follow the exact same principle: given some p_c, optimal values of τ^g and τ^d are such that the gain from pollution reduction of extending the vintage threshold in one year is exactly equal to the loss from using a less preferred transportation option. Based on the pollution costs of BGM, we find that the optimal 5-day-a-week restriction (with "neutral" recycling) design involves: $p_c = \$11.4$, $\tau^g = 1998$, and $\tau^d = 2003$.[47] Only owners of gasoline cars built in 1998 and later and of diesel cars built in 2003 and later are entitled to pay the toll, which is cheaper now than in spring and summer. The reason is that some owners, mostly in middle-income groups, of old cars who in spring and summer were ready to pay the toll, now cannot. And with fewer drivers there is less need to introduce too high a toll.

This optimal 5-day-restriction design leads to a significant reduction of HC and NO_x

[45] Note that BGM's dynamic model considers a single vehicle class that ages overtime. Note also that their external pollution costs are in 2006 dollars per mile and for the entire year (see their Table A.4, second row). We convert them to 2015 dollars per kilometer and for fall and winter using the corresponding conversion factors and assuming, based on Gallego, Montero, and Salas (2013), that emission harm in spring and summer is only a fifth of that in fall and winter.

[46] Recall that these driving restrictions work as proportional rationing schemes (as opposed to efficient rationing schemes) in that they do not distinguish more valuable trips from less valuable ones. Note also that in a dynamic setting "middle-age" cars may still face some restrictions, but only to further accelerate the fleet turnover toward cleaner cars. For more see BGM.

[47] If toll revenues are entirely spent on reducing the transit fare, this latter falls to \$0.07 and p_c to \$8.30, thus, preserving the difference $p_c - p_p = \$8.23$.

emissions during fall and winter: 72.2% and 69.8%, respectively. Weighting the harm from these two pollutants equally,[48] these reductions report a pollution benefit (ΔH) during fall and winter of \$483 ($= 6.8 \times (72.2 + 69.8)/2$) million, which is significantly larger than the benefit from less congestion ($\sum_{i=1}^{n} \Delta S_i$) during that period, \$301 million.

To see why this is an optimal 5-day-restriction design, consider relaxing the vintage thresholds in one year. Surplus $\sum_{i=1}^{n} \Delta S_i$ goes up in \$37 million but pollution also goes up, in 6.5 percentage points. The ratio of the two however, \$5.7 million per percentage point, falls below the cost to society of increasing pollution, \$6.8 million per percentage point. Conversely, consider tightening vintage thresholds in one year. This time, the ratio of the change in surplus (\$53 million) over the change in pollution (5.1 percentage points) is above the benefit to society of decreasing pollution. It is therefore not optimal to move the thresholds in either direction.

These thresholds, together with the toll, may need to be reconsidered, however, if the authority opts for a less ambitious restriction format, say, of 4 of fewer days of restriction. As seen in Section 4, p_c is to be adjusted upwards to account for the fewer drivers facing a price every time they take their cars. In other words, marginal external congestion costs increase as we target fewer drivers, which, in turn, call for higher tolls upon those fewer drivers.

Vintage thresholds may also need of some adjustment, but in this case it is not obvious in which way, if at all. According to BGM, vintage thresholds should be tightened only if the pollution gain from extending the restriction (i.e., banning the possibility to pay the toll) to slightly newer and cleaner cars is higher than the surplus loss (i.e., fall in $\sum_{i=1}^{n} \Delta S_i$) from doing so. This would clearly not be the case if $\sum_{i=1}^{n} \Delta S_i$ remains unchanged as we move to less ambitious formats. But $\sum_{i=1}^{n} \Delta S_i$ necessarily falls as we do that because toll (upward) adjustments are never enough to keep congestion at its most ambitious level; congestion necessarily goes up as we move to less ambitious formats. The question then is whether the drop in $\sum_{i=1}^{n} \Delta S_i$ is large enough to justify extending the restriction to slightly cleaner models. We find not, at least as we move to 4 and 3-day restriction formats.

Moving to a less ambitious format raises another issue, routinely brought up in the driving restriction debate: the possibility that a driver of a restricted car (i.e., of vintage below τ^g or τ^d) buys another, possibly older and cheaper, restricted car in order to by-pass the restriction. Clearly, this second-car strategy does not work under a 5-day-restriction format, almost by construction. But in principle, it might work under a less ambitious format. Based on the empirical evidence in BGM, however, there should be no reason for

[48] This would be unnecessary had the two emission reductions been the same.

concern. They find it cheaper for a driver to by-pass the vintage restriction not with the purchase of an old, high-emitting car, but with the purchase of a slightly more expensive, newer car, which in our pollution-congestion context would allow its owner to pay the daily toll.

6.2 Discussion of results

If the policy analysis of fall and winter can be treated independently of that of spring and summer, then there is little to add to what we have learned so far. The most ambitious format is a 5-day-a-week restriction with car owners of gasoline cars built before 1999 and of diesel cars built before 2004 being fully prevented from paying the daily toll in fall and winter. This ambitious format would not only deliver significant overall benefits, reaching $1.27 billion annually, but would also leave all income groups better off if the majority of toll revenues are allocated toward reducing public-transit fares.[49,50]

In reality, however, what happens in fall and winter is likely to affect what happens in spring and summer and vice versa. For instance, no one will keep a car to be used only during spring and summer and weekends. Without entering into a dynamic model of fleet turnover, we run two additional exercises in order to approximately account for this ownership situation. One exercise is to consider a less ambitious restriction format during fall and winter, so that owners of restricted cars —not entitled to pay the toll in days of restriction— still keep their cars in anticipation of their full use during spring and summer, provided they pay the daily toll. For example, we could consider a 5-day-a-week restriction during spring and summer (with all car owners entitled to pay the toll) that during fall and winter declines to a 3-day-a-week restriction with the same prevention on toll exemptions as in the previous section, i.e., $\tau^g = 1998$ and $\tau^d = 2003$. This time-varying design reports annual benefits of $829 million, significantly less than the $1.27 billion figure above. This drop in benefits is vastly explained by the sharp increase in pollution during fall and winter.[51]

Alternatively, we can assume that cars that are fully restricted in winter and fall are sold and, hence, not longer available in spring and summer. Under this assumption and the same vintage thresholds of the previous section, i.e., $\tau^g = 1998$ and $\tau^d = 2003$, annual benefits drop only by 7%, to $1.18 billion. And if we relax these vintage thresholds in one year, i.e.,

[49] Overall benefits can, at best, increase by 3% if 17% of the toll collection is diverted to increase f_b.

[50] One may recall from Section 5 that some income groups, particularly car owners in group 2, may end up worse off. This is not longer true if we account for pollution benefits $\Delta H = \sum_{i=1}^{n} \Delta H_i$, no matter how we divide them, whether on a per capita basis, i.e., $\Delta H_i = \Delta H/n$, or inversely proportional to income, i.e., $\Delta H_i = Y_i \Delta H/n\bar{Y}$, where $\bar{Y}$ is average income.

[51] HC and NO_x emissions increase in 27.1 and 26.2 percentage points respectively.

$\tau^g = 1997$ and $\tau^d = 2002$, annual benefits go up a bit, to \$1.19 billion. Based on this second exercise, it is safe to say that the most ambitious restriction format is likely to deliver benefits anywhere between \$1.19 and \$1.27 billion annually. Of these benefits, approximately 58% comes from lighter traffic and 42% from cleaner air. Note, however, these contributions vary widely throughout the year. In spring and summer, congestion alleviation contributes with approximately 86% of total welfare and pollution reduction with 14%, while in fall and winter, these contributions reverse, 41% and 59% respectively.

7 Final remarks

The objective of the paper has been to test the Pareto-improving property of Daganzo's hybrid driving-restriction scheme using Santiago as a case study. In addition to its toll exemptions, we have extended Daganzo's scheme to include vintage exemptions that correct for the (local) pollution externality. We found the Pareto-improving property not to hold: individuals in lower-income groups are strictly worse off, and more so as we increase the number of days of restriction.

Despite the simplicity of our model, we think we contribute with two important policy messages. The first is the need to pay close attention to distributional implications of policy design. The only way to leave all individuals better off in our application is by using the majority of the toll revenues to reduce the public-transit fare. Without this transfer, low-income groups are necessarily worse off. This is another reason to aim for more ambitious formats, as a way to increase these transfers.

The second policy message is that the use of marginal external costs for estimating the relative contribution of pollution and congestion alleviation to overall welfare can be largely misleading. Contrary to what these average cost numbers may suggest, we find pollution alleviation to contribute greatly to overall welfare (e.g., 59% in fall and winter). The reason for the discrepancy is that while all cars congest the same, old cars pollute a lot more than newer cars. Therefore, targeting old cars first, as vintage exemptions do, yields greater benefits than targeting the average car, which is what looking at these average external (pollution) costs implicitly does.

Our analysis can be extended in different directions, some of which are within the reach of our model. The first is to consider alternative instruments for pollution control. Although our focus on vintage restrictions responds to implementation constraints, one can nevertheless study the potential efficiency gains from moving to alternative, more effective instruments. One of them is the use of the congestion toll in conjunction with pollution charges proportional to a vehicle's vintage or emissions rate. Even if distance traveled must be taken as exogenously given (as done when setting the congestion toll), this would be a step closer to Pigouvian taxation by charging vehicles for actual use.

Another one, closer to our proposal, is the use of pollution thresholds based on emission rates (i.e., grams of pollutants per kilometer traveled) as opposed to vintages. The problem with this latter is the possibility of rates being manipulated during smog checks, as documented by Oliva (2015), in the case of Mexico City, and Barahona et al. (2020), in the case of Santiago. One way to get around this manipulation is to use pre-established emission rates as a function of vehicle class, vintage and fuel type, as those shown in Table 6. As time passes, the authority would need to update these figures running both own emission tests and some basic statistical analysis.

A second extension would be to consider global pollutants (e.g., CO_2), in addition to local pollutants. Vintage thresholds, even if differentiated by fuel type, would not longer be a good proxy. Perhaps mild restrictions upon all gasoline and diesel cars (e.g., one-day-a-week throughout the entire year) may prove effective accelerating the introduction of more fuel-efficient vehicles, particularly electric vehicles, at a much lower cost to government than the subsidies currently being offered in the developed world. The optimal design may be a combination of subsidies on pollution-free vehicles and restrictions on polluting ones that change over time as the fraction of electric vehicles in the market evolves. In any case, such analysis would require to extend our model to a dynamic setting, much in the spirit of BGM. This poses a major challenge, since their model builds on a representative car that ages over time. Extending their model to several type of cars would certainly complicate the resolution of the sorting conditions that determine which individuals drive which cars and when. We leave this for future work.

References

[1] Barahona, N., Gallego, F., and J.-P. Montero (2020), Vintage-specific driving restrictions, *Review of Economic Studies* 84, 1646-1682.

[2] Barahona, N., Gallego, F., Montero, J.-P., and C. von Dessauer (2020), A case for eliminating smog checks: Corruption, competition and technology loopholes, mimeo.

[3] Baranzini, A., S. Carattini and L. Tesauro (2018), The Geneva congestion charge: rationale, design, and acceptability, mimeo.

[4] Basso, L., and S. Jara-Díaz (2012), Integrating congestion pricing, transit subsidies and mode choice, *Transportation Research Part A* 46, 890-900.

[5] Basso, L., and H. Silva (2014), Efficiency and substitutability of transit Subsidies and other urban transport policies, *American Economic Journal: Economic Policy* 6, 1-33.

[6] Bento, A., Goulder, L., Jacobsen, M., and R. von Haefen (2009). Distributional and efficiency impacts of increased US gasoline taxes, *American Economic Review* 99, 667-699.

[7] Braid, R. (2018), Partial peak-load pricing of a transportation bottleneck with homogeneous and heterogeneous values of time, *Economics of Transportation* 16, 29-41.

[8] Cantillo, V., and J.D. Ortúzar (2014). Restricting the use of cars by license plate numbers: A misguided urban transport policy, *Dyna-Colombia* 81, 75-82.

[9] Cifuentes, L. (2018), Análisis exploratorio de los factores de emisión de fuentes móviles en Chile, mimeo.

[10] Currie, J., and M. Neidell (2005), Air pollution and infant health: What can we learn from California's recent experience? *Quarterly Journal of Economics* 120, 1003-1030.

[11] Daganzo, C.F. (2000). Taking turns: Rx for Congestion, *Access* 17, 14-19.

[12] Daganzo, C.F., R. Garcia (2000): A Pareto improving strategy for the time-dependent morning commute problem, *Transportation Science* 34, 303-311.

[13] David, Q., and R. Foucart (2014), Modal choice and optimal congestion, *Regional Science and Urban Economics* 48, 12-20.

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

[15] De Palma, A. and R. Lindsey (2011), Traffic congestion pricing methodologies and technologies, *Transportation Research Part C* 19, 1377-1399.

[16] 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.

[17] Fageda, X., Flores-Fillol, R., and B. Theilen (2020), Price versus quantity measures to deal with pollution and congestion in urban areas: A political economy approach, unpublished

[18] Fullerton, D. and S. West (2010), Tax and subsidy combinations for the control of car pollution, *The BE Journal of Economic Analysis & Policy* 10, 1-31.

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

[20] Harsman, B., and J. Quigley (2010), Political and public acceptability of congestion pricing, *Journal of Policy Analysis and Management* 29, 854-874.

[21] Heckmann, J., and E. Vytlacil (2005), Structural equations, treatment effects, and econometric policy evaluation, *Econometrica* 73, 669-738.

[22] INRIX (2019), INRIX Global Traffic Scorecard, Kirkland WA, U.S.

[23] Knittel, C. R. and R. Sandler, (2018), The Welfare Impact of Indirect Pigouvian Taxation: Evidence from Transportation, *American Economic Journal: Economic Policy* 10, 211-242.

[24] Leape, J. (2006). The London Congestion Charge, *Journal of Economic Perspectives* 20, 157-176.

[25] Litman, T. (2017), Understanding Transport Demands and Elasticities: How Prices and Other Factors Affect Travel Behavior, Victoria Transport Policy Institute, February.

[26] Mohring, H. (1972), Optimization and Scale Economies in Urban Bus Transportation, *American Economic Review* 62, 591-604.

[27] Molina, L. and M. Molina (eds) (2002), Air Quality in the Mexico Megacity: An Integrated Assessment, Kluwer Academic Publishers, Dordrecht

[28] Morrow, S., and K. Runkle (2005), April 2004 Evaluation of the California Enhanced Vehicle Inspection and Maintenance (Smog Check) Program, Report to the legislature, Air Resources Board, California.

[29] Nie, Y. (2017), Why is license plate rationing not a good transport policy? *Transportmetrica A: Transport Science* 13, 1-23.

[30] Oliva, P. (2015), Environmental regulations and corruption: Automobile emissions in Mexico City, *Journal of Political Economy* 123, 686-724.

[31] Parry, I., Strand, J., (2012), International fuel tax assessment: An application to Chile, *Environment and Development Economics* 17, 127-144.

[32] Ramos, R., Cantillo, V., Arellana, J., and I. Sarmiento (2017), From restricting the use of cars by license plate numbers to congestion charging: Analysis for Medellín, Colombia, *Transport Policy* 60, 119-130.

[33] Rizzi, L., de la Maza, C., (2017), The external costs of private versus public road transport in the Metropolitan Area of Santiago, Chile, *Transportation Research Part A* 98,123-140.

[34] Secretaría Interministerial de Planificación de Transporte (SECTRA) (2013), Plan Maestro Transportes de Santiago 2025, Santiago.

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

[36] Tirachini, A., Hurtubia, R., Dekker, T., Daziano, R., (2017), Estimation of crowding, discomfort in public transport: Results from Santiago de Chile, *Transportation Research Part A* 103, 311-323.

[37] Viard, V. B., and S. Fu (2015), The effect of Beijing's driving restrictions on pollution and economic activity, *Journal of Public Economics* 125, 98-115.

[38] Wang, L., Xu, J., and P. Qin (2014), Will a driving restriction policy reduce car trips? The case study of Beijing, China, *Transportation Research Part A* 67, 279-290.

[39] Wolff, H. (2014), Keep your clunker in the suburb: Low-emission zones and adoption of green vehicles, *The Economic Journal* 124, F481-F512.