MIT CEEPR

MIT Center for Energy and Environmental Policy Research

Working Paper Series

# Vintage-specific Driving Restrictions

NANO BARAHONA, FRANCISCO GALLEGO, AND JUAN-PABLO MONTERO

OCTOBER 2017

CEEPR WP 2017-018

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

# Vintage-specific driving restrictions

Nano Barahona Francisco Gallego Juan-Pablo Montero*

March 7, 2017

**Abstract**

Local air pollution problems have led authorities in many cities around the world to impose limits on car use, increasingly through driving restrictions or license-plate bans. With few exceptions, these restrictions tend to be poorly designed creating incentives for drivers to buy additional, more polluting cars. We study vintage-specific designs that place heavy restrictions on older, polluting vehicles and none on newer, clean ones. A novel model of the car market and evidence from Santiago's 1992 program, the earliest attempt to use vintage-specific restrictions, are used to show that these restrictions can be welfare enhancing by accelerating fleet turnover towards cleaner cars. These policies can be particularly effective in fighting local air pollution when alternative instruments such as scrappage subsidies and pollution-based taxes are not available.

## 1 Introduction

Local air pollution remains a serious problem in many cities around the world partly because of the steady increase in car use.[1] In an effort to contain such trend and persuade

*Barahona is with the Economics Department of Stanford University and Gallego and Montero with the Economics Department of the Pontificia Universidad Católica de Chile (PUC-Chile). Emails: hbaraho@stanford.edu; fgallego@uc.cl; and jmontero@uc.cl. We have benefited from the comments of Lucas Davis, Juan Escudero, Claudio Ferraz, Marco González-Navarro, Larry Goulder, Enrique Ide, Gastón Illanes, Kelsey Jack, Shaun McRae, Mar Reguant, James Sallee, Brandon Shaufele and audiences at Berkeley, Chicago, EAERE, Ecole Polytechnique, Harvard, LSE, Michigan, PUC-Chile, PUC-Rio, Sao Paulo School of Economics, Toronto, Toulouse School of Economics, Stanford, Universidad de Chile, Universidad de Montevideo and Universidade de Vigo. Gallego also acknowledges financial support from CONICYT/Programa de Investigación Asociativa (Project SOC 1102) and Montero likewise from Fondecyt (Grant No. 1130998) and the ISCI Institute (Basal FBO-16). An earlier version of this paper circulated under the title "Adopting a cleaner technology: The effects of driving restrictions on fleet turnover".

[1]Cars are major contributors of carbon monoxide (CO) and nitrogen oxide ($NO_x$) emissions, and to a lesser extent of particulates (PM10 and PM2.5). These local pollutants, unlike global pollutants such

drivers to give up their cars in favor of public transport, authorities in many of these cities have imposed limits on car use, based on some combination of the last digit of a vehicle's license plate and colored stickers displayed on its windshield. Good examples of these so-called driving restrictions—that, with some adjustments, remain in operation today—include Athens (restrictions introduced in 1982), Santiago-Chile (1986), Mexico City (1989), São Paulo (1996), Manila (1996), Bogotá (1998), Medellín (2005), San José de Costa Rica (2005), Beijing (2008), Tianjin (2008), several German cities (2008), Quito (2010), Hangzhou (2011), Chengdu (2012), and Paris (2016).[2]

According to the existing literature the popularity of these restrictions is problematic. As noted in a recent issue of *The Economist* ("Traffic in megacities," February 27, 2016), the take-away message from this literature is that driving restrictions create perverse incentives for drivers to buy additional vehicles, not only increasing fleet size but also moving its composition toward higher-emitting vehicles. The best documented evidence supporting this claim comes from Mexico City's *Hoy No Circula* (HNC). Some believe that HNC, as it was implemented in 1989, had a good start (e.g., Onursal and Gautam 1997, Gallego et al. 2013a), but most agree that over the longer term, it lead to an increase in the number of vehicles on the road and in pollution levels (e.g., Eskeland and Feyzioglu 1997, Onursal and Gautam 1997, Molina and Molina 2002, Davis 2008, Gallego et al. 2013a). In fact, Davis (2008) documents an increase of 20 percent in the car fleet due to HNC, an increase that took place over the course of a year (Gallego et al. 2013a & 2013b).[3]

In this paper we study an aspect of driving restrictions that has been much overlooked in the literature, yet is present in some later reforms, including Santiago and Mexico City:

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 (see, for instance, Figure 4 in Gallego et al 2013a). The adverse health effects of these local pollutants are well documented. Currie and Neidell (2005), for example, found a significant effect of CO on infant mortality.

[2]Authorities in Santiago, Brussels, London, Madrid, Milano, and Paris, to name a few, have also turned, on occasion, to one-day restrictions (in conjunction with any existing permanent programs) to combat daily episodes of critical air pollution. In December 2016, for example, Paris banned cars from circulation for three consecutive days, based on whether their license plates ended with an odd or even number. New Delhi also tried a two-week experiment in January 2016. This paper, however, focuses on restriction programs that run on a permanent basis since they have the potential to alter a city's fleet composition.

[3]Zangh et al. (2017) also failed to find air quality improvements from restrictions elsewhere, namely, Bogotá, São Paulo and Tianjin. They did find some effects from the restriction program implemented in Beijing during the Olympic games, which was later decided to be extended indefinitely. The initial gain in air quality has been confirmed by the recent work of Viard and Fu (2015) and Liu et al. (2015), but these latter also show that the gain rapidly disappeared within a year, consistent with the pattern found by Gallego et al. (2013a) for HNC.

namely, vintage-specific restrictions, or more precisely, restrictions that differentiate cars by their pollution rates. In 1992, for example, Santiago reformed its restriction program to allow all cars equipped with a catalytic converter (a device that transforms toxic pollutants into less toxic gases, which became mandatory in all post-1992 models), to be exempted from the one-day-a-week restriction, which remained active on all 1992 and older models. This exemption ended in December 2016 with a new reform to the restriction program that now extends the one-day-a-week restriction to all cars built before 2012, the year in which all cars imported into the country had to comply with the Euro 4 emission standard.

Mexico City has also reformed its HNC program to allow for vintage-specific restrictions. The reforms started in 1997 by exempting from the restriction all post-1992 vehicles equipped with a converter to evolve into HNC's current format of a three-tier restriction schedule with moving thresholds. Electric, hybrid, and gasoline/diesel-powered vehicles less than 8 years old face no restrictions. Older gasoline/diesel-powered vehicles face restriction schedules that vary depending on their age and results of smog checks. Vehicles that are between 8 and 12 years old are banned from circulation one weekday per week and every other Saturday from 5 am to 10 pm. These circulation bans are extended to vehicles that are 13 or more years old to include all Saturdays and all weekdays from 5 am to 11 am. A vehicle can always upgrade to a less restricted schedule, including full exemption, upon showing a smog-check certificate with a sufficiently low level of emissions, but this upgrade lasts only 6 months, when a new smog check is required. What is most interesting about HNC's current structure is that it not only accepts the notion that vehicles' emission factors deteriorate with age, but also recognizes that these factors vary among cars of similar vintage, which can be particularly significant in older models (Kahn 1996, Jacobsen et al. 2016).

Vintage-specific restrictions are also present in recent restriction programs in Europe. Authorities in Germany, for instance, have been adopting low-emission zones (LEZs) in several cities since 2008. Unlike the partial circulation bans in Santiago and Mexico City, LEZs impose a total circulation ban on higher-emitting vehicles in the city center. The types of vehicles restricted by the LEZ and its implementation schedule vary widely across cities (Wolff 2014). In Berlin, for example, diesel vehicles that did not comply with the Euro 2 standard (or Euro 1 plus particle filter) were not allowed into the LEZ, starting in January 2008. But starting in January 2010, access was further restricted to the Euro 4 standard (or Euro 3 plus particle filter). A similar all-or-nothing design was recently introduced in Paris. Since July 2016, all cars built before 1997 have been banned permanently from circulation within the city limits during weekdays from 8 am to 8 pm.

Of all the possible variations on a driving restriction policy one might think of, vintage differentiation represents a radical departure from early designs.[4] By allowing drivers to bypass the restriction not by purchasing a second (and possibly older, more polluting) car but by switching to a cleaner car facing lighter or no restrictions, vintage-specific restrictions have the potential to significantly alter the fleet composition towards cleaner vehicles in those places where pollution is a concern. The objective of this paper is to study such potential.

Our study begins with an illustration of the basic mechanism behind a vintage-specific driving restriction using Santiago's 1992 reform as evidence. Given the sharp discontinuity created by the reform between restricted and nonrestricted vintages, we are able to test for policy effects on fleet composition in restricted and nonrestricted areas by focusing on their fleet differences around the 92-93 vintage discontinuity. Consistent with the results of Wolff (2014) for the LEZ programs in Germany, we found that households living in areas subject to the restriction (i.e., any municipality in the city of Santiago) own a much smaller fraction of 1992 (and older) models than their counterparts living in non-restricted areas.[5]

While the evidence from these vintage-specific restriction programs, whether the 1992 reform in Santiago or the LEZs in Germany, is useful to illustrate the fleet-composition effect that vintage differentiation can produce, it still leaves many questions unanswered. For instance, it does not say much about the welfare implications of these policies and nothing about a socially optimal vintage-specific design and how that compares to optimal designs of alternative policy instruments such as scrappage subsidies, gasoline taxes and pollution-based registration/circulation fees. With the help of a novel model of the car market, the rest of the paper seeks to provide answer to all these questions.

Our model of the car market differs from existing models (e.g., Adda and Cooper 2000, Bento et al. 2009, Gavazza et al. 2014) in important ways. Like Gavazza et al. (2014), we treat cars as vertically differentiated products. This vertical differentiation is what explains trade among drivers who differ in their willingness to pay for quality. High-willingness-to-pay drivers upgrade to a new car when they decide to sell their used units to medium-willingness-to-pay drivers, who in turn sell their used units to lower-willingness-to-pay drivers, and so on. This trading process over the lifetime of a unit ends when a low-willingness-to-pay driver decides to scrap the unit. Given the average lifespan of a car (10 years in the EU, for example, according to the European Automobile

[4]Other variations might include restricting cars more than once a week, or setting restrictions only during peak hours, as opposed to all day, and in only some areas of the city, or extending the restriction to include weekends. Regarding this latter, Davis (2017) finds no evidence of air quality improvements when HNC was extended to Saturdays in July 2008.

[5]Besides covering a different restriction program, we separate from Wolff (2014) in that we develop a model of the car market to study a wide range of vintage-specific restrictions and how they compare to alternative policy instruments.

Manufacturers Association), this trading process can take a long time. Inevitably, this forces us to pay close attention to the market (equilibrium) transition when evaluating policy interventions that can have profound effects on fleet composition. So, unlike Gavazza et al.'s (2014) steady-state focus, our analysis is by essence dynamic: agents in our model are forward looking, i.e., they make investment decisions in long-lived goods based on expectations about future relevant variables, which in equilibrium must turn out to be correct.[6]

Our model is also unique in its attention to a variety of policy interventions to curb local air pollution, most notably driving restrictions in a wide range of formats, from the uniform restriction introduced in Mexico City in 1989 to the all-or-nothing design introduced in Paris in 2016. Since in all these programs the car market affected by the policy intervention extends well beyond the geographic area directly targeted by the policy, our model also considers households living in less polluted or unpolluted zones that are only affected by the policy through its effect on the car market. This is an important mechanism that can affect the optimality of these restrictions by allowing the flow of older cars to these zones where they still have value to some drivers. Our model is also flexible to allow for temporal variation in pollution harm, which is prevalent in many cities that suffer from local air pollution. When this is the case, it may be optimal to place restrictions only during those hours of the day, days of the week, and months of the year when ambient concentrations are likely to exceed safety standards. Unlike some other instruments, driving restrictions can be easily adjusted to cope with this temporal variation, as currently done in some of the programs described above.

The main message that arises from our model is that vintage-specific restrictions can be an effective pollution-control tool by helping to accelerate the fleet turnover towards lower-emitting vehicles. The optimal vintage-specific design takes the all-or-nothing form seen in Paris and Germany's LEZs: there is an optimal (moving) vintage threshold that separates cars between full exemption and (nearly) complete restriction. In the case of Santiago, and after calibrating the model with data from the 1992 reform, we find this optimal threshold to be 16 years of age.[7]

[6]The forward-looking assumption is also absent in Bento et al. (2009). Although it is present in Adda and Cooper (2000), they eliminate any differentiation among agents which implies that at all times, agents must be indifferent to keeping their used cars or scrapping them and having them replaced with new ones. This imposes, among other things, an immediate adjustment of the equilibrium scrappage age following a policy intervention, which remains constant thereafter during the entire market transition. This is not what we observe when one allows for heterogeneity in agents' preferences for quality, which is what we do.

[7]Strictly speaking, the threshold is an emissions rate (i.e., amount of pollution per distance traveled). Our model does not needto make the distinction because of the one-to-one mapping between emission rates and vintage. In reality, and notwithstanding enforcement concerns (see, for example, Oliva 2015),

The reason for this all-or-nothing design is because the first best (i.e., Pigouvian pollution taxation) does not require to persuade households to drive a lot less but to drive cleaner cars. A vintage-specific restriction can achieve this reasonably well by placing heavy restrictions on models that pollute the most. This has two important implications. The first is that a driving restriction's relevant margin of action is ownership, not use. In fact, a driving restriction that places a uniform restriction upon all cars can result in a significant welfare loss, even without accounting for the "second-car effect" documented for these types of restrictions (for instance, a one-day-a-week restriction in Santiago leads to a welfare *loss* that is 53 percent larger than the welfare gain from implementing the first-best). Related to this latter, the second implication is that the optimal vintage-specific design eliminates by construction any concern that some drivers may bypass the restriction by purchasing an additional, higher-emitting vehicle.

We also extend the model to study alternative instruments with potential to affect the fleet turnover. Much has been written on the use of scrappage subsidies, also known as cash-for-clunker programs, to help accelerate the retirement of old vehicles and stimulate the purchase of new ones (e.g., Hahn 1995, Adda and Cooper 2000, Mian and Sufi 2012, and Hoekstra et al. 2017). Our model shows that these subsidies, even when optimally designed, do not present any efficiency advantage over equally well-designed vintage-specific restrictions. The reason is simple: both instruments seek to affect the fleet turnover by aiming at the removal of the most polluting cars; one with a prohibition, the other with a reward. If anything, it appears that implementation constraints should favor the use of vintage-specific restrictions. For a scrappage subsidy to replicate the work of a vintage-specific restriction, the regulator must prevent old cars from outside the restricted area to be entitled to the subsidy. This can be done, although at the cost of introducing friction in the car market, requiring the vehicle to have a number of years of registration history in the restricted area. More importantly, even when these subsidies have been used, whether in the US or Europe, they have tended to be short-lived, lasting only a few months. This is partly explained by the high fiscal cost incurred by the government, but also because in most cases these programs were conceived as a temporary stimulus to boost the local auto industry, not as a permanent environmental policy.

A growing literature has also been examining the effect of registration fees/subsidies on new car purchase decisions (e.g., d'Haultfoeuilley et al. 2014, Adamou et al. 2014, and Rivers and Sahufele 2015).[8] We extend the model to consider not only (pollution-based)

---

one could follow the recent HNC reform and use smog checks to extend exemptions only to those cars with emission rates below the threshold.

[8]So far, these fees/subsidies cover only $CO_2$ emissions. See Drummond and Ekins (2016) for a proposal to extend them in the UK to also cover $NO_x$ emissions from new diesel cars.

registration fees on new units but also annual circulation fees on used units.[9] Because of the temporal variation in pollution harm, the optimal design, which comes remarkably close to implementing the first best, is to offer each year a menu of circulation fees that vary by vintage: drivers have the option to pay either a positive fee for unlimited use of the car or no fee for its use only during hours of less pollution, say, during weekends and late at night. In equilibrium, there is a cutoff vintage above which all car owners opt for paying the fee and below which none does. Not surprisingly, this cutoff is very similar, and sometimes equal, to the threshold in the optimal vintage-specific restriction design. Whether these registration/circulation fees can be used in practice, and in the menu format that we propose, is an open question, particularly in developing and emerging economies, which, according to Posada et al. (2015), tend to favor quantity instruments (e.g., restrictions) over price instruments (e.g., taxes and subsidies). But if not, the results of this paper show that well-designed, vintage-specific restrictions are a good alternative.

The rest of the paper is organized as follows. Using Santiago's 1992 reform as motivating evidence, Section 2 illustrates how vintage-specific restrictions can have a significant effect on fleet composition. The model of the car market is developed in Section 3 and calibrated using data from the 1992 reform in Section 4. Policy exercises for different driving restriction formats and alternative instruments are in Section 5. Distributional implications are also discussed. We conclude with Section 6.

# 2 Motivating evidence: Santiago's 1992 driving restriction

The city of Santiago, Chile's capital and home to 40% of the country's 17.5 million people, exhibits one of the worst air pollution problems of any urban center in Latin America, due partly to its geography but also to a steady increase in car use. Efforts to control vehicle emissions date back at least to the mid 1980's, first in 1985 with a total prohibition on the import of used cars and then in the winter of 1986 with the introduction of a driving restriction program. At the time, the restriction was intended to operate as an exceptional measure by banning the circulation of 20 percent of the vehicle fleet only on those days when air pollution was expected to reach critical levels. Over time these restriction episodes were called upon more often, and by 1990 the restriction program applied every weekday from 6.30 am to 8.30 pm from March through September, when

[9]We also consider gasoline taxes despite the common knowledge they are not a good instrument for handling local pollutants (e.g., Kahn 1996), which is the focus of this paper and of all driving restrictions currently in place. Our simulations for Santiago indicate that the optimal gasoline tax delivers 47% of the efficiency gains delivered by the optimal vintage-specific restriction.

thermal inversions and lack of wind prevent the rapid dispersion of pollutants.

The restriction program experimented an important change in 1992, when the government established by executive order that, starting in 1993, any new vehicle must be equipped with a catalytic converter to circulate in Santiago. And to accelerate the turnover toward these cleaner vehicles, the government also decided to exempt from the existing driving restriction all cars equipped with a converter. Given the absence of converters in 1992 and older models,[10] the 1992 reform introduced a sharp discontinuity between the 1992 and 1993 vintages that we exploit here as motivating evidence to illustrate the potential of vintage-specific restrictions for affecting fleet turnover. We do so by studying its effects on both quantities (i.e., fleet composition) and prices, using different estimation strategies and datasets for each dimension.

## 2.1 Effects on fleet composition

### 2.1.1 Data

The main database to study changes in fleet composition comes from vehicle circulation permits at the municipality level collected by the National Statistics Bureau. In March every year, each car owner is required to obtain a circulation permit upon payment of an annual fee to her home municipaliticy. We use data for 323 municipalities for the 2006-2012 period.[11] The data includes the number of cars of each vintage by municipality for each year, thus capturing the age profile of the fleet of cars in municipalities around March of each year. The data is available only since 2006, 13 years after the implementation of the 1992 reform.[12]

We also use information for a vector of controls at the municipality level, which include, among others, total population and urbanization rate, mean and coefficient of variation of income per capita at the municipality level, and some additional geographic controls related to location.

### 2.1.2 Results

We start by discussing some stylized facts that motivate the econometric analyses below. Figure 1 presents fleet composition by vintage for Santiago (the area affected by the

[10] There was a negligible number of pre-1993 Honda Accord models already equipped with converters at the time of the 1992 reform. We exploit this for a robustness check in Section 2.2.

[11] Our dataset includes 93 percent of the 346 municipalities in the country. The municipalities with missing information are in remote areas with low population density.

[12] We use 2006 data in the main estimates of the paper because that is the year closer to the implementation of the policy. This is not a problem for our empirical strategy, as argued below. We actually exploit the more recent data in Section 4 to validate the model calibration.

driving restriction) and for the rest of the country. Darker bars correspond to pre-1993 models (i.e., 1992 and older), the ones subject to the restriction, and lighter ones to post-1992 models. Several interesting facts emerge. First, it is noted that Santiago represents an important share of the national fleet, which has implications for the equilibrium effect of the policy, as discussed below in the empirical and theoretical analyses. Second, the fleet in Santiago is cleaner (i.e., has a larger fraction of post-1992 cars) than is the fleet in the rest of the country. However, it is not obvious a priori how much of this is due to the 1992 reform and how much to characteristics specific to Santiago that can be affecting car-purchasing decisions (e.g., higher average income in Santiago). The third stylized fact sheds some light on this. While most jumps in the number of cars by vintage are positively correlated between Santiago and the rest of the country, the jump for the 1992 and 1993 vintages (and vintages around them) are negatively correlated between Santiago and the rest of the country, suggesting that there may be something special related to those vintages that is different in Santiago than in the rest of the country.

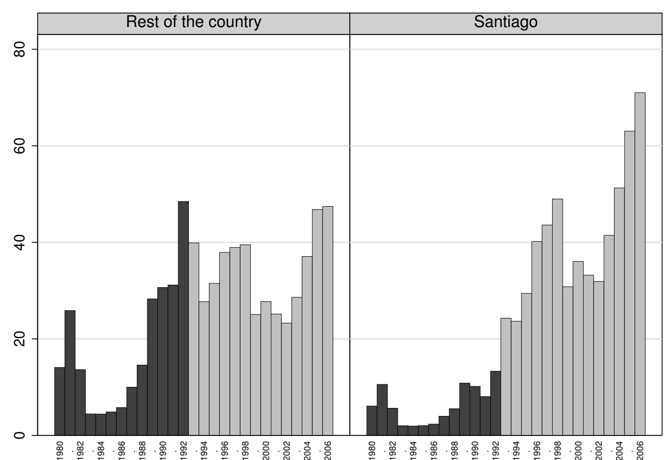



Figure 1: Santiago's fleet vs the rest of country's in 2006

Notes: Each bar represents number of cars (in thousands) of each vintage.
Pre-1993 vintages are highlighted in darker bars.

We want to estimate the effects of the policy, which affects only some municipalities, on $q_i^\tau$, the number of cars of vintage-$\tau$ in municipality $i$ in a specific year. Therefore, we exploit the geographic distribution of cars of different vintages in a particular year. We use three estimation approaches, which complement each other, as they rely on different identification assumptions.

First, we estimate the effect of being subject to the driving restriction on the ratio $y_i^{92,93} = q_i^{92}/(q_i^{92}+q_i^{93})$ using data for the cross-section of municipalities in 2006. Focusing on vintages just around the 92-93 discontinuity appears to be a clean way to test for policy effects.[13] In particular, our estimating equation is

$$y_i^{92,93} = \beta^{92,93} DR_i + x_i'\gamma + \varepsilon_i \qquad (1)$$

where $DR_i$ is a dummy that takes the value of 1 if municipality $i$ is affected by the driving restriction (i.e., if it is located in Santiago) and $x_i$ is a vector with the municipality's characteristics such as income per capita, population, distance to Santiago, income dispersion, level of urbanization, and two dummies indicating whether the municipality is located north of Santiago and in any of the country's furthest north and south regions (regions I, XI, XII and XV),[14] as purchases of new cars in these extreme regions are entitled to tax breaks. The identification in this exercise assumes that, controlling for the variables included in the $x$ vector, $\beta$ captures the causal effect of driving restrictions on $y_i^{92,93}$. A simple check of this assumption is to run placebo regressions of the ratio for other pairs of vintages different than the 92-93 dyad. Under our assumptions, we should find a zero effect for other dyads as for them the policy should not create incentives for big jumps in car ownership in Santiago relative to the rest of the country.

Table 1 presents results of estimating equation (1) for different ratios of cars of adjacent vintages. The results imply that the coefficient of being affected by the policy is negative and significant for the 92-93 ratio (column 1). Adding a vector of control variables at the municipality level does not change the results in any significant way (in column 2). The effects we find are not only statistically significant but also economically significant. In fact, if in a given municipality not affected by the restriction we observe one 93 model for each 92 model (i.e., $y_i^{92,93} = 0.5$), in a similar municipality in Santiago that ratio would be 3.44 ($= [0.5 - 0.275]^{-1} - 1$).

We also run regressions for the $y_i^{91,92}$ and $y_i^{93,94}$ ratios as placebo exercises to check the validity of the identification assumptions of this cross-section empirical strategy (in columns 3 and 4). We find that the point estimates are not only statistically insignificant but also economically irrelevant, in contrast to our results in columns 1 and 2. This is reassuring as it is expected if we were really capturing effects of the driving restrictions and not the effects of other confounding factors.[15]

[13]Notice that some of the post-1992 cars that circulated outside Santiago were not equipped with a converter. As we do not have information on their exact locations, we correct our estimates using the information from Onursal and Gautam (1997, p. 177), which reports that only 79, 87.6, and 94.8% of all new models registered in 1993, 1994, and 1995, respectively, came with a catalytic converter. If we run the regressions with the raw data (i.e. without applying this correction), our results are qualitatively similar. See the online Appendix (Section C).

[14]The country is organized in 16 regions, numbered I through XV in addition to Santiago's Metropolitan Region.

[15]We do not report results for all combinations of $\tau, \tau+1$ to save space in the table. We present results

Table 1: The effects of the driving restriction on the share of cars for contiguous vintages

| | 92-93 | 92-93 | 91-92 | 93-94 |
|---|---|---|---|---|
| Driving restriction | -0.304*** | -0.275*** | 0.0023 | 0.0142 |
| | (0.014) | (0.042) | (0.013) | (0.016) |
| Controls | No | Yes | Yes | Yes |
| Observations | 268 | 268 | 268 | 268 |
| $R^2$ | 0.480 | 0.527 | 0.090 | 0.320 |

Notes: OLS regressions with one observation per municipality. Municipalities with less than 300 cars were dropped from the sample. The dependant variable corresponds to the ratio given by $q_i^\tau/(q_i^\tau + q_i^{\tau+1})$, where $q_i^\tau$ is the total number of cars from vintage $\tau$ found in municipality $i$ in 2006. The first two columns correspond the the regression where $\tau = 1992$, while column 3 and 4 correspond to $\tau = 1991$ and $\tau = 1993$ respectively. Standard errors are calculated via block bootstrap at the province level (53 provinces in total). Municipality controls include: Income per capita (in linear and quadratic form), population, coefficient of variation of income per capita, urbanization ratio, distance to Santiago (in linear and quadratic form), and dummies for municipalities in northern and far away regions. $^*$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Second, we implement a regression discontinuity design for cars in Santiago (the treated region) with vintage ($\tau$) as the running variable, where we consider $\tau \geq 1993$ as the treated vintages. It is worth noting that implementing an RDD is challenging in our context. As shown in Figure 1, there is a jump in the stock of cars in Santiago between vintages 1992 and 1993, but similar jumps can be found in other pair of vintages (e.g., for the 1998 and 1999 vintages). These jumps are driven by national level shocks affecting the total number of cars in the country in specific years. To account for this, we first run a regression of the number of cars in each municipality on vintage fixed effects using information for all the municipalities in 2006. By construction, the average across municipalities of the residuals of this regression is 0 for each vintage and therefore we will have a normalized version of the size of each vintage (in terms of the number of cars). Then, we keep the municipalities located in Santiago and run a regression discontinuity design on the residuals of the former regression. Following the usual assumptions of RDDs, this estimator allows us to identify the *local* effect of the treatment on the discontinuity (i.e., the difference in the stock of cars between the 1993 and 1992 vintages) for municipalities in Santiago. We run a local linear regression using a uniform Kernel and a bandwidth of three vintages at each side of the discontinuity to get an estimate of 1.222 and a standard error of 0.201. The graphic implementation can be found in the online Appendix (Section

for all the combinations in the online Appendix (Section B, Table B.1). There is a couple of significant effects for other pairs of vintages but none of them are economically significant and are much smaller than the estimates for the 92-93 ratio.

A, Figure A.1).[16]

Third, while the previous two identification strategies focus on the estimation of the effect of the policy on the discontinuity between the 1992 and 1993 vintages in treated (i.e., Santiago) versus non-treated (i.e., the rest of the country) municipalities, we also estimate a more general model in which we study the effect of the driving restriction on other vintages. This is important because, based on what we know from other restriction programs, it could well be that the exodus of 1992 models was completely undone if a good fraction of drivers were by-passing the restriction not with the purchase of a post-1992 model but with the purchase of a second and possibly much older pre-1993 model. Thus, we estimate the following regression

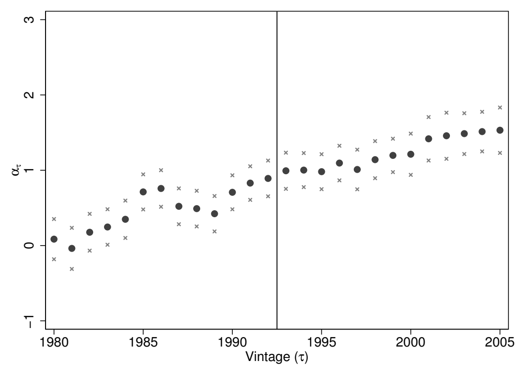
$$\log(q_{i\tau}) = \beta_\tau DR_i + \alpha_\tau \log(INCOME_i) + \gamma_\tau \log(POP_i) + x_i'\zeta + \delta_\tau + \varepsilon_{i\tau} \qquad (2)$$

where $INCOME_i$ is municipality $i$'s income per capita, $POP_i$ is the municipality's total population, $x_i$ is a vector that includes the same controls as in (1), and $\delta_\tau$ is a vintage fixed effect. We expect to obtain a vector of $\beta_\tau$'s for the different vintages $\tau$ with a discontinuous jump around the 92 and 93 vintages. In contrast, we expect a smooth evolution of $\alpha_\tau$ and $\gamma_\tau$ across vintages. This is a simple test of the identification strategy around the discontinuity produced by the policy. In this approach, we also estimate the effect of the policy outside the neighborhood of the discontinuity. To do so, we assume that conditional on other determinants, the driving restriction ($DR_i$) dummy captures only the effects of driving restrictions on the stock of cars of each vintage. This is similar in spirit to the estimator suggested by Angrist and Rokkanen (2015) to identify effects outside the discontinuity, which relies on the identification of the effect of the policy conditional on the included covariates.

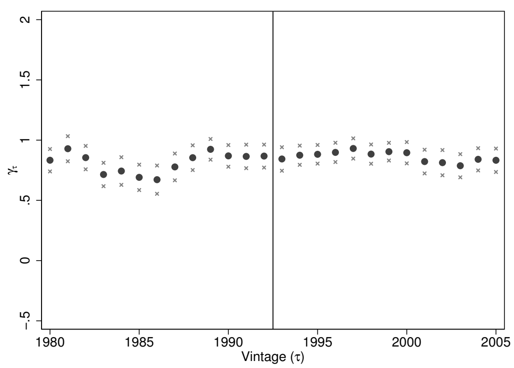
Figure 2 presents regression results equation (2) with estimates of the effects of municipality income, municipality population, and of being a municipality affected by the policy on different vintages. Results in Panel (a) are consistent with the idea that income is a main factor behind purchasing decisions and, therefore, newer models are indeed concentrated in richer municipalities. More importantly for our identification strategy, the relationship is smooth with no jump around the 92 and 93 vintages. The same happens with vintage effects for population (Panel (b)), as we do not observe any discontinuity across the 92-93 vintages.

In contrast, the 92-93 discontinuity is clearly present in Panel (c) of Figure 2, which plots the evolution of the estimated effect of being affected by the policy. The point

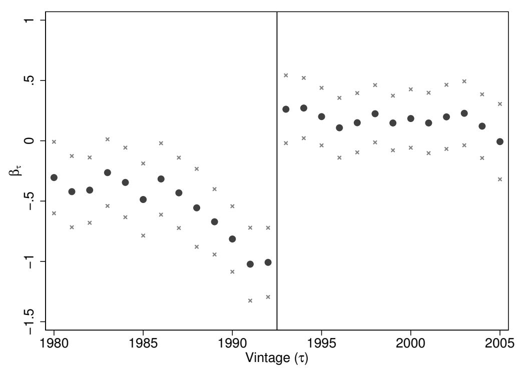
[16] Notice that since our running variable is discrete in nature (i.e. vintage, measured in years), we do not follow the existing literature in calculating the optimal bandwidth, as those methods are developed for assignment variables with density $g(x)$, where $g(.)$ is continuous and bounded away from zero (Calonico et al 2014). Our results are robust to different bandwidths and Kernel choices.




(a) $INCOME$ coefficients




(b) $POP$ coefficients




(c) $DR$ coefficients

Figure 2: Vintage effects of driving restrictions, income, and population

Notes: This figure presents the estimated vintage effects after estimating equation (2) using data at the municipality level for 2006. Panel (a) presents the coefficients for municipality income, Panel (b) for municipality population, and Panel (c) for the effect of the driving restriction. Dark dots represent the point estimates for each coefficient and light gray dots correspond to the 95% confidence intervals using robust standard errors. The vertical line marks the division between 1992 and 1993 vintages.

estimate for vintage 92 is −1.008, which is statistically significant at the 1 percent level. This implies that for each 1992 model circulating in a given municipality in Santiago, there will be 2.74 such models in a similar municipality not subject to the restriction. Conversely, the point estimate for vintage 93 (0.239 and statistically significant at the 10 percent level) indicates that for each 1993 model circulating in a given municipality in Santiago, there will be only 0.79 such models in a similar municipality not subject to the restriction.

It is interesting to note that if we are really identifying causal effects of the driving restriction program, the difference between the estimates of policy effects for the 1992 and 1993 vintages (i.e., $\beta_{93} - \beta_{92}$) should be consistent with the estimates we obtained with our RDD. In the two cases, we aim to estimate the effect of the policy on the 1993-1992 margin but we use different assumptions (in the RDD we rely on a non-parametric approach to identify effects on the discontinuity; in estimating equation (2), we take a parametric approach, which depends on the vector of included controls). It is easy to see that if $\beta_{RDD} = \beta_{93} - \beta_{92}$ both approaches produce the same estimates with different assumptions. Reassuringly, this is the case, as our estimate of $\beta_{93} - \beta_{92}$ equals 1.247, which is very close to our estimate of 1.222 for $\beta_{RDD}$. Thus, our results for the effects of the policy on the 92-93 discontinuity are roughly consistent using two identification strategies, which we interpret as a robust result. Moreover, given that our third approach relies on stronger (parametric) assumptions, we argue that this result suggests the estimates away from the 92-93 discontinuity may be reasonable.

The evolution of the estimated effects of the driving restriction as we move away from the 92-93 discontinuity in either direction is also informative, as noticeable in Figure 2c. In a market for products that are vertically differentiated and where consumers differ in their willingness to pay for higher quality (i.e., newer models), the null effects of the driving restriction program for the newest models should come as no surprise. Regardless of location, a driver's alternative to, say, a 2004 model is not a model that is ten or more years older but a model closer to 2004. In other words, ownership decisions concerning models away from the discontinuity should be independent of the policy. The same logic applies to the fact that the policy (i.e., $DR$) coefficients reverts toward zero for the very old models, so it would be wrong to interpret this reversal as an indication that some drivers who own a pre-1993 model are bypassing the restriction by purchasing an additional 1980-86 model rather than a 1993 or newer one.[17]

## 2.2 Effects on car prices

In addition to fleet composition effects, documenting price effects is important for several reasons. First, the effect on prices provides an indirect check of whether the policy was enforced or not. If the driving restriction were actually binding on drivers, one would expect to find a large impact not only on the allocation of pre and post-1992 models,

[17]This "second-car" effect is, in any case, highly unlikely to be of first order importance in our context. First, data from CASEN surveys (the main household survey applied in Chile) indicate that in 1998 (2006) about 4.7% (4.5%) of the households that owned at least one vehicle in Santiago owned more than one. Second, while this fraction was much lower outside Santiago (about 2.5% for 1998 and 2.0% for 2006), when one controls for a vector of socioeconomic characteristics, the difference between Santiago and the rest of the country drops and is not statistically significant. We implement these exercises using probit and Hurdle poisson logit. Results are in the online Appendix (Section B, Table B.2).

but also on market prices given Santiago's large market share (41.8% of the national fleet in 2006). Second, since we do not have data on fleet composition before 2006, price effects give a sense of the policy's effects in years closer to its implementation. And third, estimating the effect on prices is also important, as that provides an estimate of the cost of the restriction to individuals, and in particular, of the (lower) cost of bypassing the restriction not by purchasing a second old, polluting car but rather by upgrading to a newer, exempt car, preventing the "second-car effect" well documented for restriction programs that make no vintage distinctions (e.g., Davis, 2008).

### 2.2.1 Data

We assembled a new dataset of newspaper ads with car offers published in "El Mercurio" —Chile's main newspaper— during 1988-2000. Our sample considers price offers for both used and new models for a set of the most traded models in the market covering a wide price range (i.e., Fiat Uno, Honda Accord, Honda Civic, Mazda 323, Peugeot 205, Peugeot 505, and Toyota Corolla).[18]

### 2.2.2 Results

We use three empirical strategies to identify policy effects on car prices. We start with the following estimating equation for each of the seven models mentioned above

$$\log(P^m_{i\tau t}) = \beta^m DR_\tau + g(\tau;\sigma) + \delta^m_a + \delta^m_t + \varepsilon^m_{i\tau t} \quad (3)$$

where $P^m_{i\tau t}$ is the price offer of ad $i$ placed at time $t$ for a vehicle model $m$ of vintage $\tau$, $DR_\tau$ is a dummy equal to one for cars equipped with a catalytic converter, i.e., for all $\tau \geq 1993$,[19] $g(\tau;\sigma)$ is a parametric function of $\tau$ (more of which below), $\delta^m_a$ and $\delta^m_t$ are age of the car (where age $a = t - \tau$) and date of the ad fixed effects, respectively, and $\varepsilon_{i\tau t}$ is the error term. Notice that we cannot control for vintage fixed effects, as the $DR_\tau$ variable is collinear with them. The identification in this case relies on the assumption that all price differences across vintages unrelated to the driving restriction are captured by age and time fixed effects and by $g(\tau;\sigma)$. We implement this approach for each model for which we have data and for the pooled observations of all the models (in the latter case, we add model fixed effects to equation 3).

[18] In the online Appendix (Section A, Figure A.2) we present offers of Toyota Corollas as anecdotal evidence. Our empirical strategy is motivated by the evident price discontinuity between vintages 1992 and 1993.

[19] As mentioned before, this also includes a handful of pre-1993 Honda Accords that had a catalytic converter reported in the ads.

We experiment with three parametrizations of $g(\tau;\sigma)$. First, we control for a proxy of car quality of different vintages. We assume that prices of new cars are a good approximation of the cars' intrinsic qualities, as they do not depend on internal factors since cars are imported to Chile from international markets (in contrast with what happens in the second-hand market, which operates more as a close economy).[20] Using this assumption, we construct our proxy for the quality of the model of a specific vintage. We start running a regression of log prices of all the ads in our sample on the date of the offer fixed effects, model fixed effects and age of the car fixed effects. Then we average the residuals of this regression by vintage and model using only new car offers and use this as a proxy for the vintage-model specific quality of a car. Second, we add a linear function of vintage that can take a different slope for before and after 1993. This specification allows the cars' depreciation to differ between affected and non-affected vintages, trying to capture any technological change that is correlated with vintages that are affected by the policy. Third, we add interactions of age with linear trends in time, to allow for different depreciation rates.

Table 2 presents the results, first, for the seven models for which we have data available and, next, for all the models, with each row representing a different estimator of $\beta^m$. Column 1 presents the results of estimating equation (3), controlling only for age and date fixed effects (and model fixed effects in the case of the pooled estimation at the bottom of the table). The coefficients of having a catalytic converter suggest, after controlling for these fixed effects, that drivers are willing to pay a premium for having a catalytic converter installed in their cars (and to not be subject to the restriction). The fact that the premium tends to be higher in the more expensive models (e.g., 10 log points for a Peugeot 505 vs. 3 log points for a Peugeot 205) is consistent with a situation in which individuals who own more expensive cars have a greater opportunity cost of not driving every day and, therefore, are willing to pay more for cars exempted from the restriction.

Columns 2 to 4 present the estimates after controlling for different specifications of $g(\tau, \sigma)$. Most of the estimates do not change in any significant way, suggesting that vintage effects do not affect results (if anything, the catalytic converter estimates increase once we control for them). In all, the estimates for the pooled data (i.e., for all models) presented at the bottom of the table imply a 6.5 log point premium for having a catalytic converter installed in the car.[21]

[20] We see this assumption as particularly relevant for Chile, where there are no car manufacturers and the import of used cars has been forbidden since 1985.

[21] Notice that the cost of \$265 for replacing a catalytic reported in Onursal and Gautam (1997) can not explain the effects we report in Table 2 because this is a small share of the total price of a new car (i.e., 1.8% for a Toyota Corolla in 1995). Moreover, this difference in cost should also be captured by our control for the differences in prices of new cars reported in column 2. Besides, if differences

Table 2: The effects of having a catalytic converter on the price of used cars

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| FIAT | 0.031*** | 0.027*** | 0.034*** | 0.027*** |
| UNO | (0.006) [5220] | (0.006) [4705] | (0.007) [4705] | (0.006) [4705] |
| HONDA | 0.127*** | 0.105*** | 0.121*** | 0.122*** |
| ACCORD | (0.008) [10583] | (0.011) [3978] | (0.011) [3978] | (0.011) [3978] |
| HONDA | 0.031*** | 0.069*** | 0.054*** | 0.05*** |
| CIVIC | (0.007) [7281] | (0.007) [5655] | (0.007) [5655] | (0.007) [5655] |
| MAZDA | 0.031*** | 0.054*** | 0.055*** | 0.052*** |
| 323 | (0.006) [8377] | (0.005) [5576] | (0.005) [5576] | (0.005) [5576] |
| PEUGEOT | 0.033*** | 0.025*** | 0.024*** | 0.021*** |
| 205 | (0.007) [4285] | (0.008) [3716] | (0.008) [3716] | (0.008) [3716] |
| PEUGEOT | 0.103*** | 0.138*** | 0.116*** | 0.114*** |
| 505 | (0.008) [11665] | (0.009) [5115] | (0.01) [5115] | (0.01) [5115] |
| TOYOTA | 0.094*** | 0.17*** | 0.174*** | 0.175*** |
| COROLLA | (0.011) [9385] | (0.01) [6564] | (0.012) [6564] | (0.012) [6564] |
| ALL | 0.048*** | 0.065*** | 0.065*** | 0.064*** |
| MODELS | (0.006) [56796] | (0.004) [35309] | (0.004) [35309] | (0.003) [35309] |
| Controls: | | | | |
| Age, Model and Date f.e. | YES | YES | YES | YES |
| Vintage qual. | NO | YES | YES | YES |
| f(Vintage) | NO | NO | YES | YES |
| flexible age f.e. | NO | NO | NO | YES |

Notes: The unit of observation corresponds to a car offer published in the newspaper the first Sunday of every month between 1988 and 2000. Each row corresponds to estimates of the effect of having a Catalytic Converter in the context of equation (3) using different specifications for different models. Standard errors clustered by ad date are presented in parentheses. The number of observations in each specification are presented in squared brackets. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

were explained by this fixed cost of installing a catalytic converter, we should expect greater percentage differences in prices for less expensive cars, which is exactly the opposite of what we observe. In addition, notice that converters can only be installed in vehicles with spark-ignition engines (Onursal and Gautam, 1997), which explains why, at least in Santiago, we did not observe pre-1993 vehicles being retrofitted with converters.

Our second empirical strategy is a regression discontinuity design with $\tau$ as the running variable and, as before, we consider $\tau \geq 1993$ for all models as the treated vintages. Following the same approach used for the RDD for quantities, we start calculating residuals from a regression of log prices on date of the offer fixed effects, model fixed effects, age of the car fixed effects, and our proxy for car quality. We do this for the pooled sample and for each model. Thus, under the usual assumptions, this estimator identifies the *local* effect of the treatment on the 92-93 discontinuity. Let $\beta^{m}_{RDD}$ be this estimator. As in the case of our estimates for the effects on quantities, we run a local linear regression using a uniform kernel and a bandwidth of 3 vintages at each side of the discontinuity. We obtain a point estimate of 0.061 with a standard deviation of 0.004 (the graphic implementation of this estimator is reported in the online Appendix, Section A, Figure A.3). This number is remarkably similar to the estimate we find using our parametric approach in equation (3), i.e., 0.065.

Third, we run price regressions using newspaper ads for Honda Accords exploiting the fact that some pre-1993 models were already equipped with catalytic converters, and therefore, exempted from the restriction. This exercise is important as one may argue that 1993 models could be more expensive than 1992 models not because of the driving restriction, but because of a discrete jump in quality or costs between these two vintages (note that this concern is not relevant as long as our proxy for the quality of cars of different vintages is related to the true variable). This is unlikely to be the case for models of the same make of the same year. We exploit the fact that in many instances this feature of the car (i.e. having a catalytic converter) was explicitly reported in the ads along with the price quote. So we run the following cross-section regression

$$\log(P_i^{HA,\tau}) = \beta^{HA}_{\tau} CONVERTER_i + \varepsilon_i \tag{4}$$

where $P_i^{HA,\tau}$ refers to the price of a Honda Accord of vintage $\tau$, and $CONVERTER_i$ is a dummy that takes a value of one if the ad reports that the car has a catalytic converter. We test for the effect of reporting a catalytic converter on the price offer by running four separate OLS regressions for vintages 1991 through 1994 using ads published in October, November, and December of 1995. This provides an additional test that exploits the fact that since converters were required by law in all post-1992 models, reporting its existence in ads for these models should make no difference.

This is precisely what we see in the last two columns of Table 3, where having a catalytic converter is not statistically different from zero. This contrasts with the catalytic premiums observed in the first two columns of the table. Moreover, we can compare the estimate of $\beta^{HA}_{\tau}$ with the estimate provided for the same model using our previous empirical strategies, as a robustness check. Albeit somewhat larger, the 1991 and 1992 premiums are not that different from the 12 log point premium reported in Table 2 for the same model.

Table 3: The effects of reporting a catalytic converter on the price of used Honda Accords

| | (1991) | (1992) | (1993) | (1994) |
|---|---|---|---|---|
| CONVERTER | 0.223*** | 0.189*** | 0.0206 | -0.00487 |
| | (0.054) | (0.035) | (0.035) | (0.010) |
| Constant | 15.60*** | 15.68*** | 15.96*** | 16.40*** |
| | (0.032) | (0.034) | (0.025) | (0.010) |
| Observations | 47 | 53 | 58 | 49 |
| $R^2$ | 0.245 | 0.309 | 0.006 | 0.001 |

Notes: Each observation corresponds to a Honda Accord ad published in the newspapers in October, November and December of 1995. The dependent variable is posted price of the car (in logs). We present the estimates of a dummy that takes a value of 1 if the ad reports a catalytic converter is installed in the car in the context of equation (4). Each column presents the results for the different car vintages. Robust standard errors are presented in parenthesis. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

# 3 A model of the car market

The key message that emerges from the Santiago-1992 reform is that vintage-specific restrictions are a potentially useful tool to fight air pollution by helping accelerate the fleet composition towards lower-emitting vehicles. Yet, there are many policy-relevant questions that the empirical analysis cannot answer. What are the welfare implications of these restrictions? What is the socially optimal restriction design? How does this latter compare to alternative policy instruments? We address these questions by developing a model of the car market that is then calibrated and applied to Santiago and the rest of the country based on data generated from the 1992 reform. Although the numbers that emerge from the model are specific to Santiago, their qualitative implications apply more broadly, since there is nothing specific in the model that prevents its application to other cities.

We present the model in this section and leave the calibration and application for the following two sections. Our model shares the vertical-differentiation structure of some existing models (e.g., Gavazza et al. 2014) but differs from them on its emphasis on a wide range of pollution-control policies and their forward-looking implications on fleet evolution. Car vintage is the only attribute that creates product differentiation. New models enter the market at perfectly competitive prices and used models are traded in a frictionless secondary market by drivers with different willingness-to-pay for quality (i.e., vintage). We abstract from multiple ownership and, hence, from the possibility of a ”second-car” effect. We do this not only because the empirical evidence from the Santiago-1992 reform finds no support for this possibility but, more importantly, because the optimal vintage-specific design eliminates this possibility by construction, as we shall see.

## 3.1 Notation

There are three agents in the economy: car producers, car dealers and drivers or households. They all discount the future at $\delta \in (0,1)$. The cost of producing a new car is $c$, which is also the price at which perfectly competitive producers sell new cars to car dealers.[22] A large number of car dealers buy new cars from car producers and rent them together with second-hand cars to drivers.[23] The (annual) rental price for a car of age $a = \{0,1,2,...\}$ at date $t$ is denoted by $p_{a,t}$ ($a = 0$ corresponds to a new car). Note the change of notation from vintage $\tau$ to age $a = t - \tau$. Our model makes no distinction between the two because the car technology is invariant to time (i.e., there is no technological progress), so age is used only to facilitate the exposition.

Cars exit the market at some exogenous rate due to crashes, fatal malfunctioning, etc. This rate may vary with age, so the probability that at age $a$ car is still in the market next period is $\gamma_a \in (0,1)$, with $\gamma_a \geq \gamma_{a+1}$ (to simplify notation we assume throughout this section, but not in the calibration and simulations, that $\gamma_a = \gamma$ for all $a$). All surviving cars at time $t$ are (endogenously) scrapped at age $T_t$ for a value of $v$, which can be the price a dealer gets for an old vehicle exported to another country, for example.

There is a continuum of households/drivers of mass 1 that vary in their willingness to pay for quality and also in how much they drive. A driver of type $\theta$ who benefits from driving a car of quality $s_a$ for $x$ miles has to pay a variable cost of $\psi x$ and a rental price $p_a$ every period. Every period she obtains a utility of (to save on notation henceforth, in many places we will omit the subscript “$t$” unless it is strictly necessary)

$$u(\theta, a, x) = \frac{\alpha}{\alpha - 1}\theta s_a x^{\frac{\alpha-1}{\alpha}} - \psi x - p_a \tag{5}$$

where $\theta$ is the consumer’s type, $s_a > 0$ is the quality of a car that is $a$ years old, $x$ corresponds to miles traveled during the period, $\alpha > 1$ is a parameter that captures

[22]We could change the interpretation of $c$ to represent marginal cost plus a mark up in non competitive markets and the conclusions from the model would remain the same. The main mechanism of the model is driven by the relationship between car dealers and drivers rather than car producers and car dealers.

[23]Note that the renting assumption, which is also in Bento et al. (2009), is equivalent to assuming a frictionless secondary market that clears once per period. Evidence provided in the online Appendix (Section B, Table B.3) suggests that markets are fairly integrated across the country. Cars in Santiago tend to be 2-3% cheaper, consistent with the costs of moving them from one city to another.

decreasing returns on car use, $\psi$ is the per-mile cost of using the car (including parking, gasoline, maintenance, insurance, inspections, circulation fees, etc.),[24] and $p_a$ is the rental price. The quality of a car falls with age according to $s_{a+1} = \varsigma s_a$ with $\varsigma \in (0,1)$, either because older cars are more likely to break down or because they lack the latest technological advances.[25]

A consumer $\theta$ who rents an $a$-year-old car anticipates that she will drive

$$x(\theta) = \left(\frac{\theta s_a}{\psi}\right)^{\alpha} \tag{6}$$

miles. Her utility from renting an $a$-year-old car reduces then to

$$u(\theta, a) = \kappa \left(\theta s_a\right)^{\alpha} - p_a \tag{7}$$

where $\kappa = [(\alpha - 1)\psi^{\alpha-1}]^{-1}$.

Our formulation captures with a single parameter two empirical regularities: that households which value quality more tend to drive newer cars and that newer cars are, on average, run more often (e.g., Lu 2006). Households are distributed according to the cummulative distribution function $F(\theta)$ over the interval $[0, \bar{\theta}]$. A driver $\theta$ who does not rent a car obtains an outside utility from riding pollution-free public transport, which is assumed constant across households and normalized to zero.

## 3.2 The market equilibrium

At the beginning of some given period, say, year $t$, there will be a stock of used cars $\mathbf{S}_t = (q_{1t}, q_{2t}, .....)$. As a function of that stock, the market equilibrium for the year must satisfy several conditions. First, it must be true that in equilibrium, drivers of higher-valuation types rent newer cars. There will be a series of cutoff levels $\{\theta_{0t}, \theta_{1t}, ...\}$ that precisely determine the prices at which certain drivers are renting certain cars. Denote by $\theta_{at}$ the driver who, at time $t$, is indifferent to renting a car of age $a$ at price $p_{at}$ and one of age $a+1$ at a lower price $p_{a+1,t}$, that is

$$\kappa \left(\theta_a s_a\right)^{\alpha} - p_a = \kappa \left(\theta_a s_{a+1}\right)^{\alpha} - p_{a+1} \tag{8}$$

for all $a = 0, 1, ..., T-1$, where $T$ is the age of the oldest car rented. Consumers of type $\theta \geq \theta_a$ rent age-$a$ vehicles or newer while consumers of type $\theta < \theta_a$ rent older vehicles (or not at all for $\theta$'s sufficiently low). As in any vertical differentiation model, an obvious corollary from (8) is that a higher valuation consumer obtains strictly more surplus than a lower valuation consumer.

[24] If congestion is a problem, $\psi$ may also include (socially optimal) congestion charges, which we do not model explicitly.

[25] A linear quality decay rate is also in Gavazza et al. (2014).

In equilibrium, the series of cutoff levels $\{\theta_0, \theta_1, ...\}$ must be consistent with the population of drivers and the existing stock of used cars $\mathbf{S}$ and the new cars coming to the market ($q_0$) in period $t$. Hence, it must also hold that

$$q_0 = 1 - F(\theta_0) \text{ and } q_a = F(\theta_a) - F(\theta_{a+1}) \tag{9}$$

for all age-$a$ vehicles that are rented in equilibrium.

Since car dealers always have the option to scrap an old car and receive $v$, in equilibrium they must also be indifferent to renting an age $T$ vehicle today (and scrap it tomorrow, if the vehicle still exists) and scrapping it today, i.e.,

$$p_T + \delta\gamma v = v \tag{10}$$

In general, only a fraction of $T$-year-old vehicles will be scrapped in equilibrium (while all older vehicles will be), so

$$F(\theta_{T-1}) - F(\theta_T) \leq \gamma q_{T-1} \tag{11}$$

where $\gamma q_{T-1}$ is the number of age $T$ vehicles that survived from the last period.[26]

In addition, in equilibrium (price-taking) car dealers must break even, so the evolution of rental prices must satisfy the following zero-profit condition

$$c = \sum_{i=0}^{\Gamma} (\gamma\delta)^i p_i + (\gamma\delta)^{\Gamma+1} v \tag{13}$$

where $\Gamma$ is the age at which a car bought today, i.e., at date $t$, is expected to be retired (or rented for the last time). Note that both $\Gamma$ and $T$ depend on the existing stock $\mathbf{S}_t$ and in steady state, $\Gamma = T$.

One last condition that must hold in equilibrium is that the lowest-valuation household to rent a car today, $\theta_T$, obtains its outside utility

$$\kappa \left(\theta_T s_T\right)^{\alpha} - p_T = 0 \tag{14}$$

If (14) does not hold, a dealer would be strictly better off renting a $T$-year-old vehicle at a price slightly above $p_T$ instead of scrapping it.[27]

[26]Note that, because quality drops discretely with age, it can happen that in equilibrium all age $T-1$ vehicles are rented but all age-$T$ vehicles are scrapped; then the relevant scrapping condition is not (10) but

$$p_{T-1} + \delta\gamma v > v > p_T + \delta\gamma v \tag{12}$$

where $p_T$ is the hypothetical rental price for an age-$T$ vehicle.

[27]The same logic applies if we are in the corner (12) of the previous footnote: Given the fixed supply of vintage $T-1$ vehicles, a dealer owning a $T-1$ vehicle could slightly raise its rental price above $p_{T-1}$ and still find demand for it.

Conditions (8)–(14) determine the unique equilibrium for any given stock of used cars $\mathbf{S}_t$, that is, rental prices of new and used cars and sales of new cars. Unlike other papers, we are not only interested in the steady-state equilibrium, but also in the equilibrium during the transition phase after a policy shock. Transitions can be particularly long in car markets, so despite the great computational demands, they cannot be neglected in policy evaluation and design.

### 3.3 The social optimum

When pollution is not a problem, the market equilibrium described above delivers the socially optimal outcome. We are interested, however, in a situation where vehicle emissions create a local air pollution problem for households in a particular area. Given that the pollution is generated by households in that area, we divide the total mass of households into those within a restricted ($r$) area subject to pollution-control policies and those within a non-restricted ($nr$) area but still subject to changes in the car market. Thus, we let households in a $k \in \{r, nr\}$ area be distributed according to the cdf $F_k(\theta)$ over the interval $[0, \bar{\theta}]$, where $F_r(\theta) + F_{nr}(\theta) = F(\theta)$.

Suppose that cars emit local pollutants at a rate of $e$ per mile, which increases with cars' age, that is, $e_{a+1} > e_a$,[28] and that pollution in the restricted area is equally harmful, no matter the day of the week or hour of the day in which pollutants are emitted (later in the simulations, we relax this temporal assumption). Denoting by $h$ the harm from a unit of pollution, the cost to society of an age-$a$ vehicle running $x$ miles in the restricted area is $e_a x h$. If the social planner can monitor emissions, $e_a x$, he can restore the social optimum by levying a Pigouvian tax equal to $h$ on each unit of pollution generated in the restricted area. This will affect decisions on car use and ownership, i.e., it will affect (6) and (7) for drivers living in the restricted area in the following way

$$x_r^*(\theta, a) = \left(\frac{\theta s_a}{\psi + e_a h}\right)^\alpha \tag{15}$$

and

$$u_r^*(\theta, a) = \kappa_a \left(\theta s_a\right)^\alpha - p_a \tag{16}$$

where $\kappa_a = [(\alpha - 1)(\psi + e_a h)^{\alpha-1}]^{-1}$. Notice that expressions (6) and (7) are still valid for drivers living in non-restricted areas.

[28]In our model, emission rates decay with time equally across cars. In reality, however, there is an important dispersion in emission rates among cars of similar vintage, particularly among older models (Kahn 1996, Jacobsen et al. 2016). Our model predictions would still be valid if the restriction were established in terms of emission rates, not vintage, and enforced with smog checks. See Oliva (2015) for some concerns on this latter.

As will be illustrated, in a market for a durable good, a policy intervention can affect the market in subtle ways. One may argue that if the government levies a tax on emissions, cars will become relatively more expensive than (pollution-free) public transport, reducing the demand for cars in the market, which in turn should reduce the number of cars that enter the market each year. This intuition is only partially correct because a car is actually not a single product, but rather a collection of different products providing different services over the car's lifetime. Newer (and cleaner) cars have become relatively cheaper than older cars, so the demand for them has increased. The overall effect is that more new cars will be coming to the market each year, but each will last fewer periods.

## 3.4 Real-world policy interventions

Since Pigouvian taxation for dealing with local air pollution is technically unfeasible,[29] policy makers must rely on alternative instruments. In addition to different forms of driving restrictions, we consider scrappage subsidies, annual circulation fees, and gasoline taxes. The way these price-based instruments enter into the model is relatively simple. A scrappage subsidy enters by increasing the scrappage value from its baseline value of $v$ to some $v^r > v$. Since the price differential $v^r - v$ creates incentives for drivers in non-restricted areas to scrap their vehicles in the restricted zone and collect $v^r$, as opposed to $v$, the regulator may try to prevent this arbitrage by requiring the scrapping vehicle to have a number of years of registration history in the restricted area. We consider two scenarios: one in which arbitrage cannot be prevented , so any scrapping car receives $v^r$, and another in which cars are required to have a full registration history in the restricted area.

Unlike scrappage subsidies, annual circulation/registration fees are vintage specific. They enter into the model by increasing the rental price of all polluting vehicles in the restricted area from its equilibrium value of $p_a$ to $p_a^r$. Under the optimal circulation-fee design, the price difference $p_a^r - p_a$ is approximately equal to the total bill of Pigouvian taxes that an $a$-year-old vehicle would be expected to pay during a year. A gasoline tax may also enter as vintage specific if one accepts that a car's fuel efficiency deteriorates with age. If so, it would enter into the model by increasing the variable cost of using an $a$-year-old car in the restricted area from the baseline value of $\psi$ to $\psi_a^r$, with $\psi_a^r$ increasing with age (in the simulations we also consider the case of a constant $\psi^r$ across vintages). We omit from the model the possibility that households may drive outside the restricted

[29] $CO_2$ emissions are a different matter: A gasoline tax can do the work of a Pigouvian tax to deal with global warming, since it is irrelevant when and where the carbon content in the gasoline is released into the atmosphere.

area only to fill their tanks with cheaper gasoline.[30]

The way a driving restriction enters into the model is more involved because it depends on the specific design, which must specify the extent of the restriction and the car vintages affected. The extent of the restriction is captured by the parameter $R_a < 1$, which says that an $a$-year-old car can only be used a fraction of the time, such as 4 of 5 weekdays each week. We understand that drivers can move some trips from one weekday to another at a low cost, so $R$ should not be read as 4/5 in this example, but probably more. It is less obvious whether the trips that can be moved are the most valuable to the driver or not. We adopt the conservative assumption —less favorable to the driving restriction option— that the driving restriction destroys an equal fraction of car trips of different values during the day of the restriction. Some of these car trips may certainly be replaced by trips in public transport, which we do not need to explicitly model.

Since we have not yet made any temporal distinction as to when pollution is emitted, the latter assumption implies that the driving restriction reduces the number of car trips a driver would otherwise make uniformly over the week, or the year, for that matter. Formally, a driver $\theta$ who owns an $a$-year-old car that faces an effective restriction of $R_a < 1$ will now drive

$$x^r(\theta, a) = R_a \left(\frac{\theta s_a}{\psi}\right)^\alpha \tag{17}$$

miles and obtain a utility of

$$u^r(\theta, a) = R_a \kappa \left(\theta s_a\right)^\alpha - p_a \tag{18}$$

per period (recall that the utility from using public transport, whether alone or in combination with private transport, is normalized to zero).

Expressions (17) and (18) raise two observations that are important for the analysis that follows. The first is that by comparing (17) and (15), it appears that the driving restriction could replicate the first-best amount of driving $x^*$ by simply setting vintage-specific restriction levels of the form $R_a = [\psi/(\psi + e_a h)]^\alpha$. While this is true if households cannot adjust their renting decisions, it is not enough to restore the first-best, not only because trips are not equally valuable, but more importantly, because a policy affects not only the amount of driving but also the cars that households rent in equilibrium.

The second, which is closely related to the previous one, is that the optimal driving restriction takes an extreme form; that is, $R_a$ should be either 0 or 1. This is because $R_a$ enters linearly in the social welfare estimate of a $\theta$-type household driving an $a$-year-old car with restriction $R_a$, $u^r(\theta, a) - he_a x^r(\theta, a)$. One immediate implication of this corner solution is that an optimal driving restriction eliminates by construction the possibility

[30] At least in Santiago, any price dispersion generated by the gasoline tax between restricted and non-restricted areas is as large as the existing dispersion observed in the city.

of observing a second-car effect. A more practical implication is that this all-or-nothing structure is already present in existing vintage-specific restrictions, namely, in Paris and in Germany's LEZs.

# 4 Calibration

We use different data sources and methodologies to obtain numerical values for the different parameters that enter into the model. Some parameters are taken directly from the data, while others are calibrated to match the distribution of cars across municipalities that we observe in the 2006 sample of circulation permits. The order in which we discuss our choices follows roughly the order in which parameters appear in the presentation of the model.

## 4.1 Car-related parameters

While the 92-93 discontinuity introduces a clear partition in car quality for those two adjacent vintages, drivers more generally tend to regard cars of slightly different vintages to be of similar quality. We address this quality overlap in a very simple way by clustering car vintages in six vintage/quality groups centered around the 92-93 discontinuity: 1981-84, 85-88, 89-92, 93-96, 97-2000, and 2001-04.[31] This grouping is equivalent to assuming that people trade their cars not every year, but every four years.

The rental price of an $a$-year-old car in period $t$ is obtained from car prices according to the no-arbitrage condition

$$p_{at} = P_{at} - \delta P_{a+1,t} \tag{19}$$

where $P_a$ is the price of an $a$-year-old car and $\delta$ is the discount factor, which we set at 0.9 per year. Car prices are generated from the dataset described in Section 2. If $P_{imat}$ is the price offer in newspaper ad $i$ published in year $t$ for model $m$ that is $a$ years old, we run an OLS regression of $\ln(P_{imat})$ on a constant and year, model and age fixed effects to predict $\hat{P}_{mat}$. With these predictions and (19), we obtain (weighted average) rental prices for each of the six vintage groups identified above. These predictions are also used to obtain the (weighted average) price of a new car (i.e., $a = 0$), which we set at $c = \$16,000$ (all our numbers are in 2006 US dollars).

Since the import of used cars is forbidden, we estimate survival rates, $\gamma_a$, directly from stock changes observed in the car registration samples from 2006 through 2012. By comparing stock changes across two consecutive samples, we obtain six data points with survival rates for each car age. Imposing $\gamma_a \leq 1$ and $\gamma_{a+1} \leq \gamma_a$, an OLS fit to these data

[31] Model years 1980 and earlier, which in any case are very few, are grouped together with 1981 models.

points delivers average survival rates for cars with ages ranging from 0 to 36 years old. Averaging these numbers at our vintage-group level leads to the survival numbers $\gamma_a$ in the online Appendix (Section B, Table B.4).

The four remaining parameters related to vehicles are the scrappage value $v$, the initial quality $s_0$, the quality decay rate $\varsigma$, and the per-unit cost $\psi$. None of these parameters can be obtained directly from the data. Based on informal conversations with car dealers we set $v = \$600$, the lowest trade-in value some of them recall to have seen in recent years (we do not see prices this low in our sample of newspaper ads). On the other hand, since $s_a = \varsigma^a s_0$ enters multiplicatively in (5), we cannot separately obtain estimates for $s_0$ and the equilibrium cutoffs $\theta_a$'s. Hence, we normalize $s_0 = 10$. A summary with the calibrated parameters is displayed in Table 4.

Table 4: Calibrated parameters

| Parameter | Value | Parameter | Value | Parameter | Value |
|---|---|---|---|---|---|
| Panel (a): Car-related parameters | | | | | |
| $c$ | 16000 | $v$ | 600 | $s_0$ | 10 |
| Panel (b): Households characteristics | | | | | |
| $\alpha$ | 2.111 | $\psi$ | 0.3732 | $\varsigma$ | 0.892 |
| Panel (c): Pollution data | | | | | |
| $\omega$ | 1.993 | $h_r$ | 0.0041 | $h_{nr}$ | 0.00048 |

Notes: Calibrated parameters of the model using data from Chile in 2006. See Section 3 of the text for more details.

Following Gavazza et al. (2014), the per-unit cost $\psi$ is assumed to be invariant to location and age. From (6) and (8), $\psi$ can be expressed as

$$\psi = \frac{1}{x(\theta_0)} \frac{(p_0 - p_1)(\alpha - 1)}{1 - \varsigma^\alpha}$$

where $x(\theta_0)$ is the average travel of a car during the first period ($a = 0$), which in our calibration corresponds to the first four years of the car's life. With values of $\varsigma$ and $\alpha$, which are obtained from the calibration to be explained next, we estimate $\psi$ to match the figure of $x(\theta_0) = 50,952$ miles that is reported in Lu (2006) for average travel during the first four years of a car's life.[32]

[32]We use Lu's (2006) numbers because we do not have a comparable study for the local fleet. According to domestic car dealers, numbers for the local fleet are very similar.

## 4.2 Households' characteristics and policy response

Two important inputs for our analysis are the extent to which households exhibit decreasing returns on driving ($\alpha > 1$) and the cdf $F(\theta)$ of their marginal valuation for quality $\theta \geq 0$. Neither input is directly observable, similar to the actual response to the policy, which is assumed to be the same over all pre-1993 models, that is, $R_\tau = R < 1$ for all vintage $\tau \leq 1992$. These three inputs, together with $\varsigma$, are calibrated to match the car-holding predictions of the model with the actual holdings in the 2006 sample for each location and vintage.

We start the calibration by reducing the dimensionality of the problem from more than 300 municipalities to 60 electoral districts. Electoral districts group together municipalities that are located in the same geographic areas and therefore share similar characteristics, most importantly, income. Since the country's population is normalized to 1, our relevant car-holding variable $q_{i\tau}$ becomes the fraction of cars of vintage-group $\tau = 0, ...., 5$ in District $i = 1, ...., 60$ relative to the district's number of households.

We need to define a functional form for $F_i(\cdot)$ and how it varies from district to district. We let $F_i$ be a cubic in $\theta$ with each coefficient in the cubic ($b^1$, $b^2$, and $b^3$) varying across Districts $i = 1, ..., 60$ according to the linear function $b_i^j = \zeta_0^j + z_i'\zeta^j$, where $j = 1, 2, 3$ denotes the coefficient in the cubic, $\zeta_0^j$ is a constant and $z_i$ is a vector that includes the following district characteristics: income per capita ($INCOME$), distance to Santiago ($DISTANCE$), and level of urbanization ($URBANIZATION$). Thus, the distribution $F_i$'s are characterized by 12 $\zeta^j$ parameters (four for each of the three coefficients in the cubic) that need to be calibrated along with $\alpha$, $\varsigma$, and $R$.

Some of the calibrated $F_i$'s are plotted in Figure 3. The lower $F$ corresponds to the country's highest-income district (District 23), and the upper $F$ corresponds to the country's lowest-income district (District 46). The plot also includes two intermediate distributions that will be used extensively in the policy simulations: the distribution for Santiago, which aggregates all 15 districts affected by the driving restriction, and the distribution for the rest of the country, which aggregates the remaining 45 districts not affected by the restriction.

According to our model, we should observed exactly the same cutoff levels $\theta_\tau$ for all districts in the restricted area, $\theta_\tau^r$, and likewise for all districts in the non-restricted area, $\theta_\tau^{nr}$. These cutoffs can be obtained from (18) and the equilibrium condition (8) as follows

$$\theta_\tau^k = \left( \frac{p_{\tau+1} - p_\tau}{R_{\tau+1}^k \kappa s_{\tau+1}^\alpha - R_\tau^k \kappa s_\tau^\alpha} \right)^{\frac{1}{\alpha}} \tag{20}$$

where $k \in \{r, nr\}$, $R_\tau^{nr} = 1$ for all $\tau$, $R_\tau^r = R < 1$ for $\tau \leq 1992$ , and $R_\tau^r = 1$ for $\tau \geq 1993$. The two rows of dots at the bottom of Figure 3 depict the cutoff predictions for each of

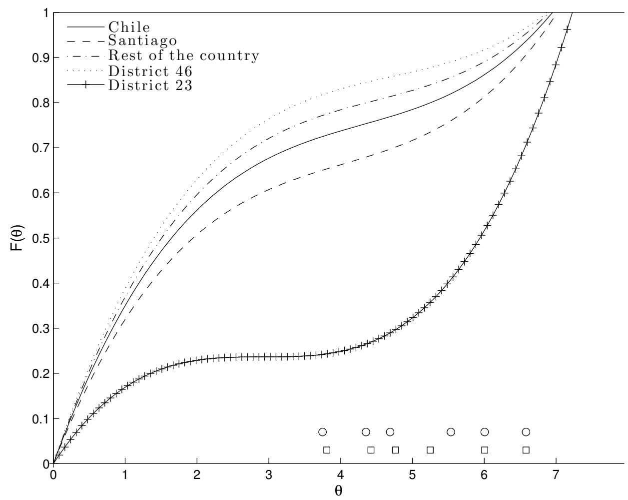



Figure 3: Estimated CDF for different groups

Notes: The figure presents $\theta$'s cummulative distributions functions for selected districts. District 23 corresponds to the richest district (located in Santiago) and 46 to the poorest (located outside Santiago). The other two curves correspond to the aggregate distribution for Santiago and the rest of the country. The bottom of the graph include square and circle dots with the estimated cutoff $\theta_a$ for Santiago and the rest of the country, respectively.

the six vintage groups for districts in Santiago (lower row of square dots) and in the rest of the country (upper row of circle dots), respectively.

Obviously, if we plug into (9) the cutoffs and distribution $F_i$'s that are in Figure 3 to obtain predictions $\hat{q}_{i\tau}$, these latter will not precisely match the shares $q_{i\tau}$ that are observed in the data. Thus, our calibration procedure looks for parameter values that minimize those mismatches. In particular, we introduce vintage-district shocks so that the cutoff in District $i = 1, ..., 60$ for vintage-group $\tau = 0, ..., 5$ is given by

$$\theta_{i\tau} = \theta_\tau^k + \varepsilon_{i\tau}$$

where $\theta_\tau^k$ is given by (20) and $\varepsilon_{i\tau}$ are error terms that make the model's prediction $\hat{q}_{i\tau} = F_i(\theta_{i\tau}) - F_i(\theta_{i,\tau+1})$ exactly match the actual share $q_{i\tau}$.[33] Since there is no reason for the error terms to be correlated with the district characteristics that define the $F_i$ distributions, and whether the district falls in the driving-restriction zone or not (i.e., $DR = 1$ or 0), our calibration minimizes the following five moments for each vintage-group $\tau$: $\sum_{i=1}^{60} \varepsilon_{i\tau} = 0$, $\sum_{i=1}^{60} INCOME_i \times \varepsilon_{i\tau} = 0$, $\sum_{i=1}^{60} DISTANCE_i \times \varepsilon_{i\tau} = 0$, $\sum_{i=1}^{60} URBANIZATION_i \times \varepsilon_{i\tau} = 0$, $\sum_{i=1}^{60} DR_i \times \varepsilon_{i\tau} = 0$.

[33] More precisely, the $60 \times 6$ $\varepsilon_{i\tau}$'s unknowns are found by solving a system of $60 \times 6$ $\hat{q}_{i\tau} = q_{i\tau}$ equations.

The resulting calibrated parameters are in Panel B of Table 4. Note that the actual policy intensity, $R = 0.967$, does not appear to be particularly large. Remember, however, that the policy was enacted 13 years prior to the 2006 data used in the calibration. Also, some values obtained are similar to those used in previous studies. The value we obtain for $\alpha$ leads to a concave utility function $u(\theta, \tau, x)$ —note that $(\alpha-1)/\alpha = 0.53$— not that different from the logarithmic utility in Gavazza et al. (2014). The value for the decay rate $\varsigma$ is also very close to the value they use.[34] Our calibrated unit-cost $\psi$, however, is twice as large as their number, partly explained by our higher gas prices.

## 4.3 Pollution parameters

The pollution-related parameters in the model are the social harm from local pollution, $h$, and the emissions rate of a $\tau$-vintage vehicle, $e_\tau$. We exploit different sources in the public domain to estimate both. The first source is Parry and Strand (2012), which contains specific estimates of vehicle external costs associated with the emissions of local pollutants for various cities in Chile (they also provide estimates of other external costs associated with car use such as accidents and noise, but these are not included in our exercises). Their damage estimate for an average vehicle in Santiago is ¢6 per mile while that for an average vehicle in the rest of the country is ¢0.7 per mile. With this information and the travel figures in Lu (2006), we estimate the annual pollution damage per car to be $720 in Santiago and $84 in the rest of the country.

In order to disaggregate average damages at the vintage level, we use information from Molina and Molina (2002, pp. 236 and 255) for Mexico City on the contribution of cars of different vintages to local emissions (i.e., CO, $NO_x$, HC, $SO_2$ and particulates). The vintage-pollution relationship they obtained is summarized in the first two columns of Table 5, which, for example, shows that the newest 60 percent of the fleet contributes only 15 percent of total emissions. The informal documentation that we have for Chile is consistent with these numbers,[35] which is not surprising, given the similarities in fleet compositions (see column 3 in Table 5).

According to our theoretical model, the damage contribution of a $\tau$-vintage car is $xe_\tau h$, so the contribution of all $\tau$ cars in area $k$ would be equal to

$$\int_{\theta^k_\tau}^{\theta^k_{\tau-1}} \left(\frac{\theta s_\tau}{\psi}\right)^\alpha e_\tau h_k dF_k(\theta)$$

where $k \in \{r, nr\}$ and $\tau = 0, ..., 5$.[36] Assuming that $e_\tau$ evolves according to the linear function $e_\tau = (1+\omega)e_{\tau-1} + \omega$, with $e_0 = 0$, we obtain values of $\omega$, $h_r$, and $h_{nr}$ that provide

[34] Their annual decay rate is 0.976 while ours is $0.892^{1/4} = 0.972$.
[35] Luis Cifuentes of the Industrial Engineering Department of PUC-Chile, personal comunications.
[36] Note that $\theta^k_{-1} = \bar{\theta}^k$.

Table 5: Fleet composition and pollution contribution

| Car vintage | Fleet Percent Share (Mexico) | Emissions Contribution (Mexico) | Fleet Percent Share (Chile) |
|---|---|---|---|
| 1993-2001 | 60% | 15% | 63.3% |
| 1985-1992 | 28% | 30% | 24.1% |
| 1980-1985 | 7% | 25% | 9.3% |
| 1979 & older | 5% | 30% | 3.3% |

Notes: Information for Mexico presented in columns 1 and 2 is borrowed from Molina and Molina (2002). Information in column 3 is based on our own calculations.

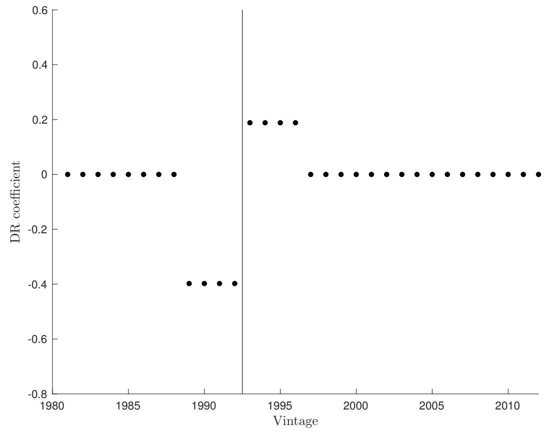
the best fit to the damage estimates of Parry and Strand (2012) and the vintage-pollution relationship of Molina and Molina (2002). The values are shown in Panel (c) of Table 4.

## 4.4 Validation checks

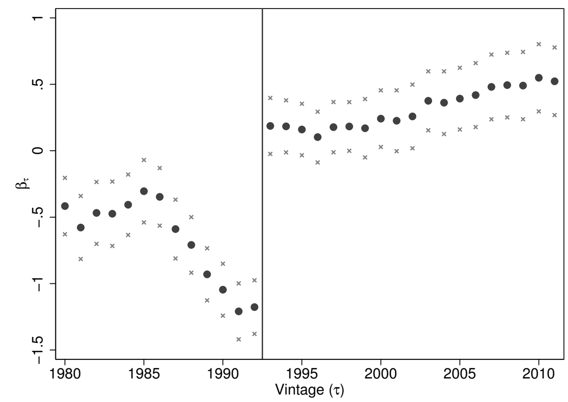
As a final step in our calibration exercise before moving to the policy simulations, we check the validity of our model and calibrated parameters by contrasting some of its predictions and assumptions to the data. We start with an out-of-sample validation that contrasts the predictions of the model for 2012 with the empirical estimation for the 2012 sample. As shown in Figure 4, the model captures reasonably well the policy effects on fleet composition both around the 92-93 discontinuity and before that. It fails, however, to capture the larger fraction of newer cars in Santiago relative to the rest of the country.[37]

According to the model predictions contained in Figure 4(a), a driving restriction like Santiago-1992 should produce only relative changes in car holdings for models just on either side of the 92-93 discontinuity. This is not only fairly consistent with what we see in Figure 4(b) but more so with what we see in Figure 2(c), which shows that the $DR$ coefficients in equation (2) for vintages away from the discontinuity are either not statistically different from zero or barely so. The fact that the $DR$ coefficients for the 94 and 95 vintages, and not just 93, are positive is not surprising because there is always noise in car quality due to individual preferences, heterogeneity in cars' aging, etc. The same is true on the other side of the discontinuity: $DR$ coefficients for 90 and 91 are comparable to that for 92. For this reason we adopted in the calibration clusters

[37]As documented in Gallego et al. (2013a), one reason for this latter that is not captured in our model is the substantive shift to private transport caused by the poorly implemented public transport reform in Santiago in February 2007, Transantiago. Unfortunately, we cannot control for this in our empirical estimation separate from the driving restriction.




(a) Model prediction for DR coefficients after 20 years of the 1992 reform.




(b) Empirical estimation of DR coefficients after 20 years of the 1992 reform.

Figure 4: Out of sample validation

Notes: Panel (a) contains model predictions for 20 years after policy implementation. The prediction is for a policy equivalent to the Santiago-1992 reform using the parameters calibrated with the 2006 data and controlling for income an population. Panel (b) shows the estimated coefficients using data from 2012, i.e., 20 years after policy implementation.

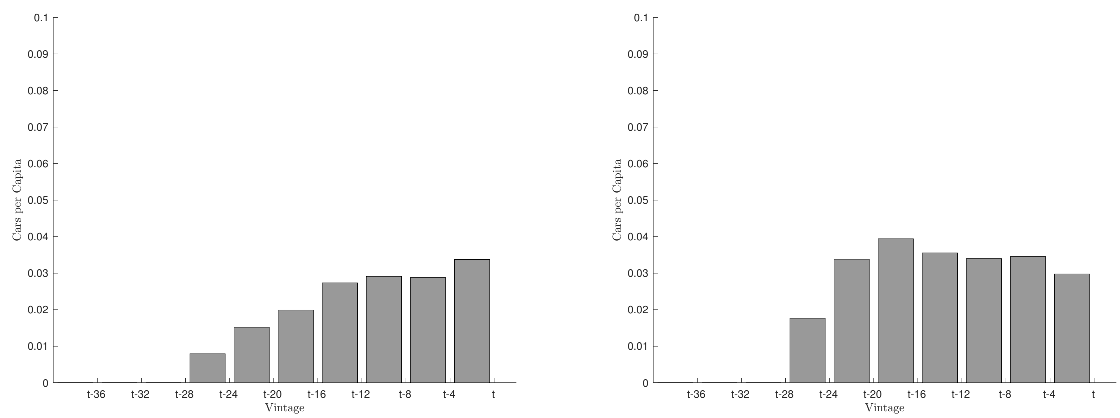
of 4-vintage groups around the 92-93 discontinuity.

Figures 2(c) and 4(b) also serve to discuss the validity of our single-ownership assumption. Some readers may question this assumption on the basis that the $DR$ coefficients for older models return to zero as we move away from the 92-93 discontinuity. This interpretation is incorrect. Using (20), our model predicts that $\theta^{nr}_\tau = (1/R)^{1/\alpha}\theta^r_\tau$ for both $\tau$ and $\tau+1$. So, if $F$ is linear in the relevant range, which is a good approximation given the large number of vintages considered, we have $q^{nr}_\tau = F(\theta^{nr}_\tau) - F(\theta^{nr}_{\tau+1}) \approx q^r_\tau = F(\theta^r_\tau) - F(\theta^r_{\tau+1})$.[38] Therefore, any evidence of a second-car effect should have been reflected in strictly positive $DR$ coefficients for the older models. Failing to find this latter in a less than optimal vintage-specific design, together with the fact that optimal vintage-specific design eliminates the second-car effect by construction, validates the use of the single-ownership assumption in our model given our focus on vintage-specific designs.

[38] The linearity in $F(\cdot)$ is not even necessary given our calibration results below: $(1/R)^{1/\alpha} = 1.02$. Notice also that we use the same $F$ for restricted and non-restricted districts because in the regression we control for other factors that explain differences across districts.

# 5 Policy simulations

With our model and the (calibrated) parameter values, we are now ready to perform some policy simulations to answer the questions posed in the introduction to Section 3. In all the simulations that follow parameter values are kept constant overtime including the speed at which a car's emissions rate deteriorates with age, population, and household characteristics. We start by constructing the no-intervention scenario. Figure 5 shows that, in the absence of any government intervention, the city of Santiago (the restricted area) already exhibits a relatively newer fleet than that in the rest of the country (the non-restricted area). While Santiago's smaller population (36 percent of the country's total in 2006) explains its smaller overall fleet, its higher income per capita explains why it nevertheless has 14 percent more of the newest models (0 to 4 years, or vintage-group 0 in our simulations) than the rest of the country. Santiago also has fewer of any of the older models.[39] Notice that in both locations, we find cars running until they are scrapped, somewhere between 24 and 28 years old.[40] This will change as the government intervenes in the market.




(a) Santiago

(b) Rest of the country

Figure 5: Steady-state fleet composition under no intervention

Notes: The figure shows the steady-state profile of car fleets in both Santiago and the rest of the country under the no-intervention scenario. Since total population has been normalized to the unity, each bar indicates the number of cars per capita of a particular vintage group.

The next set of exercises estimates the welfare gains of moving away from the no-intervention scenario. We are particularly interested in the welfare gap between the

[39]If, in comparing fleets across regions, one were to eliminate any size effect and leave only income effects, one would need to multiply the height of each bar in Figure 5a by $1.78 = 0.64/0.36$.

[40]Not surprisingly, this scrappage age differs from that in the calibration (i.e., 20-24 years old). The equilibrium in the calibration is subject to an intervention and is not a steady state.

first-best outcome and the outcome of alternative policy interventions. Since there is perfect competition in the car market, welfare in any given period $t$ is equal to the following (time index $t$ has been omitted)[41]

$$-cq_0 + vq_T + \sum_{k=r,nr} \int_{\theta_T^k}^{\bar{\theta}^k} [u^k(\theta) - h_k e^k(\theta) x^k(\theta)] dF_k(\theta)$$

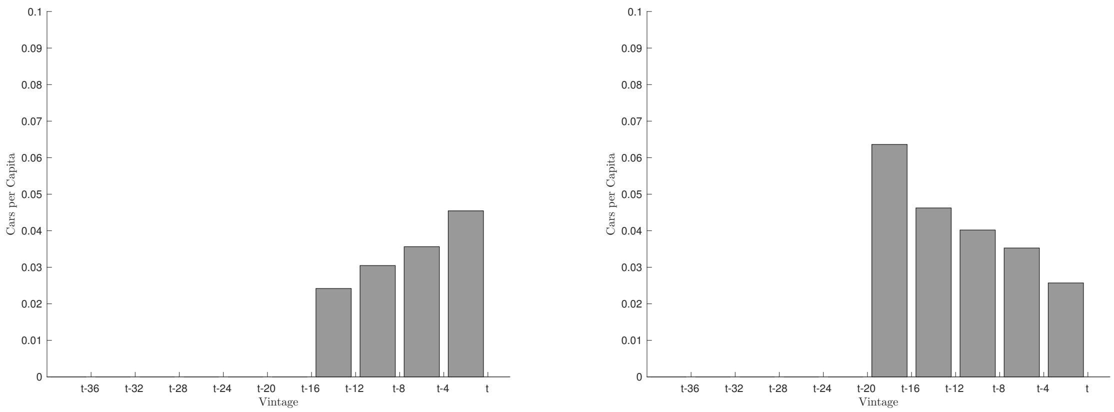
where $q_0$ is the total number of new models entering the market in period $t$, $q_T$ is the total number of cars exiting the market, $\theta_T^k$ is the last household to rent a car in region $k$ in period $t$ (see (14)), and $u^k(\theta)$, $e^k(\theta)$ and $x^k(\theta)$ are, respectively, a household $\theta$'s utility, emissions per mile, and miles traveled in region $k$ during period $t$. Notice that the specific forms of $x^k(\theta)$ and $u^k(\theta)$ vary with the policy scenario; compare, for instance, eqs. (15) and (16) with (17) and (18). We assume that all agents (including the regulator) discount the future at $\delta = 0.9$ per year, as assumed in the calibration.

## 5.1 The first-best benchmark

As observed in the steady-state outcome of Figure 6, the effect on fleet composition of levying a Pigouvian tax equal to $h_k$ upon cars circulating in region $k$ is quite dramatic. Over the long run, households in Santiago have no incentive to hold cars older than 16 years, a 12-year reduction compared to the no-intervention case. While sales of new cars in Santiago increase by 36 percent, fewer households drive cars there, but those that do, drive cleaner cars. This dramatic adjustment also has large impacts outside Santiago. The scrappage age of a car in the rest of the country is reduced in 8 years. This may look surprising at first because there is no direct intervention in the non-restricted area, but everything works through the second-hand market, as it does in all existing restriction programs. Instead of scrapping cars, Santiago is now exporting a large fraction of 16-year-old cars to the rest of the country. This increase in supply reduces the rental price of all 20-year and older models on the market to the point that is optimal to scrap them much sooner.

This adjustment has profound welfare implications. Estimating them is far from trivial because the transition from one steady state to the other is not only very long, so it cannot be omitted from any welfare estimation, but also non-monotonic (this applies to any policy intervention). Figure 7, for example, illustrates the dynamics of new car sales ($q_0$) for both Santiago and the rest of the country. In the case of Santiago, an initial jump in sales of 53 percent (from 0.034 to 0.052) is followed by a sharp drop to the steady-state level of 0.046, but a few years later, there is a new jump to finally approximate the

[41] Since a policy intervention affects the value of the existing stock of used vehicles, by altering future rental prices, there may be (unanticipated) changes in dealers' revenues that our estimation is missing.




(a) Santiago

(b) Rest of the country

Figure 6: Steady-state fleet composition under the first best

Notes: The figure shows the steady-state profile of car fleets in both Santiago and the rest of the country under Pigouvian taxion. Since total population has been normalized to the unity, each bar indicates the number of cars per capita of a particular vintage group.

steady-state level 25 years after implementation. This non-monotonicity introduces some computational challenges for determining the equilibrium dynamics, particularly when we need to find the optimal policy intervention.

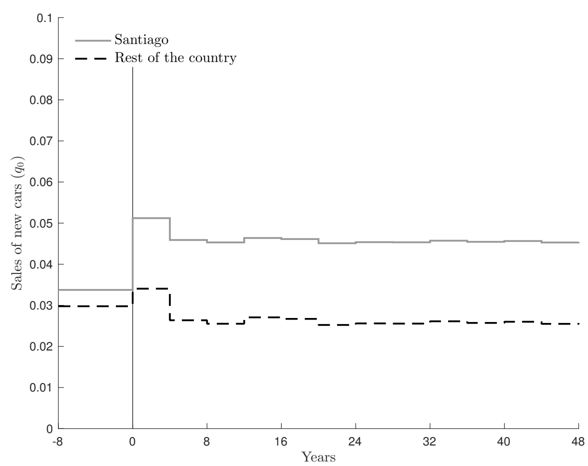



Figure 7: Transition phase: Sales of new cars

Notes: This figure shows the number of new cars ($q_0$) that are sold in every period in Santiago and the rest of the country after the implementation of a regime of Pigouvian taxes. $t = 0$ corresponds to the time when the tax policy is implemented. Values for $t < 0$ correspond to the steady state under no intervention.

In present-value terms, the welfare gain of moving from no intervention to the first-best amounts to $323.8 per household, or a 5.9 percent gain from the no-intervention baseline of $5528.5 (see the first two rows of Table 6). At the country level, this adds to a total of $1.3 billion; comparable, for example, to the gain of Germany's LEZs (Wolff 2014). In any case, we do not want to push these welfare numbers too much. Other than being rough approximation of the potential gains from curbing vehicle emissions, these numbers serve our purpose here as a benchmark for evaluating the relative performance of real-world policies like driving restrictions, circulation fees, scrappage subsidies, and gasoline taxes.

Table 6: Welfare calculations

| # | Counterfactual (in 2006 dollars) | Welfare per capita (relative to first-best) | Welfare gain/loss |
|---|---|---|---|
| 1. | No intervention | 5528.5 | 0% |
| 2. | First best | 5852.3 | 100% |
| 3. | Driving restriction with no exemptions ($R = 0.9\ \forall\tau$) | 5032.6 | -153% |
| 4. | Driving restriction with some exemptions ($R = 0.9$, $\tau > 3$) | 5606.5 | 24% |
| 5. | Optimal driving restriction ($R = 0$, $\tau > 4$) | 5815.5 | 89% |
| 6. | Scrappage subsidy ($1575), full arbitrage | 5765.1 | 73% |
| 7. | Scrappage subsidy ($2500), no arbitrage | 5781.4 | 78% |
| 8. | Circulation fees | 5836.7 | 95% |
| 9. | Gasoline tax (¢46 per gallon), no fuel-economy correction | 5642.0 | 35% |
| 10. | Gasoline tax (¢34 per gallon), with fuel-economy correction | 5664.2 | 42% |

At least two elements separate real-world policies from first-best instruments. The first is that, for either political or technical reasons, the instruments involved are never first-best. The second is that their use is restricted to geographic areas that have been declared in non-attainment with existing air quality standards. Consequently, the regulator cannot introduce policies in attainment areas only to contain any eventual pollution leakage from regulations imposed elsewhere. We adopt this geographic limitation in the simulations that follow so that policy measures are exclusive to Santiago.

Furthermore, given how long it takes to move from one steady state to another, it is natural to think that the optimal policy, whether quantity or price based, may vary over time. Given the dynamics of the first-best outcome, it appears that the regulator would like to start with a tougher policy to be gradually relaxed to its steady-state level. For simplicity, however, we focus on time-invariant policies in what follows.[42]

## 5.2 Driving restrictions

Designing a driving restriction requires both that the intensity of the restriction ($R$) and the vintages affected be defined. Before presenting the optimal (vintage-specific) design, it is instructive to go over less than optimal designs, as they help to clarify why these policies can sometimes inflict so much harm —if poorly designed— but can nevertheless be improved by introducing vintage differentiation. Take, for instance, the HNC design, as implemented in Mexico City in 1989, and, following Gallego et al. (2013a), assume a uniform restriction intensity of $R = 0.9$ upon all cars. Even neglecting the second-car effect, such a uniform design leads to a welfare *loss* that is 53 percent higher than the welfare gain from implementing the first-best (see the last column of Table 6).

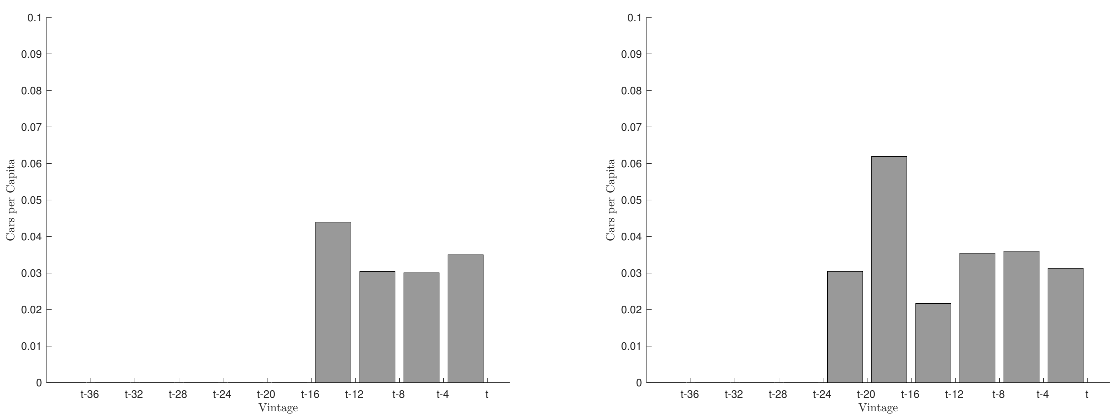
An HNC-1989 design not only fails to remove old cars from the road (actually, it extends their lives by reducing their rental prices), but it also reduces sales of new cars in Santiago (a figure with the steady-state fleet composition can be found in the online Appendix, Section D, Figure D.1). Since new cars are driven by households that most value quality, a uniform reduction in quality is felt more heavily on these new cars. As a result, demand for them falls, and with that, their rental prices and sales. As the demand for cars gets shifted towards older models, the life of the existing stock gets extended, as well. Because of this life extension, pollution may end up higher than in the no-intervention baseline unless, of course, the reduction in circulation prompted by the restriction is enough to offset the fleet's aging.

A way for the regulator to reverse the perverse outcome of a uniform restriction is to follow the reforms introduced in Santiago and Mexico City and allow some of the cleaner cars to be exempt from the restriction. In fact, if we let $R = 0.9$ be imposed only upon cars that are 12 or more years old, then this vintage-specific design results in a welfare gain, although that is only 24 percent of the welfare gain under the first-best, as indicated in the fourth row of Table 6. The exemption extended to cleaner cars solves one part of the problem: it boosts sales of new cars in Santiago (a figure with the steady-state fleet composition of this particular vintage-specific design can also be found in the online Appendix, Section D, Figure D.2). The second part of the problem, however, which is the removal of high-emitting vehicles from the road, requires a tougher restriction upon

[42]The non-monotonic dynamics described above and illustrated in Figure 7 makes the computation of time-varying (optimal) policies quite demanding. We nevertheless attempted some departures from and around the time-invariant (optimal) design. The additional gains do not qualitatively change any of the results that follow.

these cars.

Following the all-or-nothing restriction schemes seen in Paris and Germany's LEZs, it turns out that the optimal vintage-specific design is to impose a total circulation ban on vehicles that are 16 or more years of age and full exemption on newer vehicles.[43] As indicated in row 5 of Table 6, the welfare gain of doing so is significant. This latter can be explained with the aid of Figure 8. An optimal driving restriction works at both ends of the fleet spectrum, directing the removal of old cars and boosting sales of new ones. Because scrapping 16-20 year-old cars in areas where local pollution is not a problem is socially inefficient, a driving restriction works its way through the second-hand market to reallocate these cars from pollution-affected areas to pollution-free areas. But there is more. The export of these older cars to the rest of the country does not result in a sharp increase of high-emitting vehicles in this region; quite the opposite. Similar to what is behind the first-best profile of Figure 6b, the export of 16-24 year-old models to the rest of the country puts downward pressure on the price of very old cars (24-28 years old), ultimately, inducing car dealers to retire them from the market. This market dynamics may help to explain why Wolff (2014) fails to find pollution leakage from LEZs to non-affected areas.




(a) Santiago

(b) Rest of the country

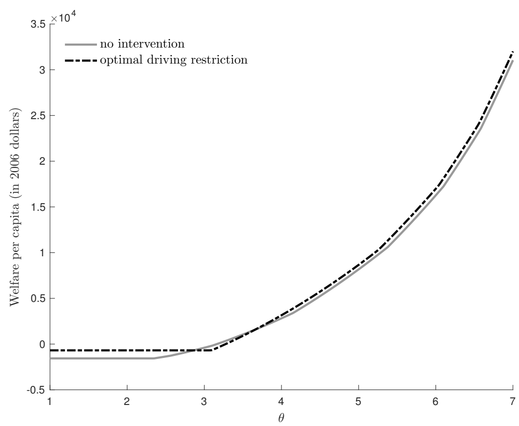
Figure 8: Steady-state fleet composition under the optimal driving restriction

Notes: The figure shows the steady-state profile of car fleets in both Santiago and the rest of the country under the optimal driving restriction. Since total population has been normalized to the unity, each bar indicates the number of cars per capita of a particular vintage group.

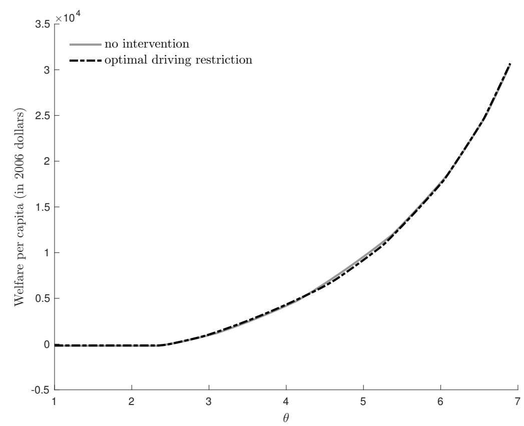
The optimal vintage-specific design raises a couple of issues that deserve further dis-

[43]Recall that strictly speaking the threshold is not a particular age but an emissions rate, say, $\bar{e}$. In our model, the distinction is irrelevant, but in reality it may not be, as cars of similar vintages may have aged differently. In that case, one could follow the recent HNC reform and use smog checks to extend exemptions only to those cars with emission rates below $\bar{e}$.

cussion (the comparison to alternative policy instruments is left for the next section). The first is that, by imposing a complete ban on old vehicles, the optimal design closes any possibility for a second-car effect. The second issue deals with distributional implications: By placing a total ban on old cars, which are mostly owned by lower-income drivers, the optimal design may raise serious distributional concerns. As Figure 9a shows, almost all drivers and public-transport users in Santiago are better off under the optimal driving restriction vis-a-vis the no-intervention scenario (households in the rest of the country are not much affected, as seen in Figure 9b). There is, however, a relatively small group of households driving cars soon to be retired that are strictly worse off. The gain in air quality, which is valued equally by all households in the economy, is not enough to compensate these drivers for the loss that implies moving to either public transport or newer but more expensive cars. In the absence of transfers, the government can still prevent this outcome, at the cost of some efficiency loss, by slightly relaxing the complete ban on old vehicles.




(a) Santiago




(b) Rest of the country

Figure 9: Distributional implications of the optimal driving restriction

Notes: This figure shows consumer surplus under two different regimes: no interventions and the optimal driving restriction.

## 5.3 Alternative policy instruments

We now take the model to study the relative performance of alternative policy instruments, namely scrappage subsidies, registration/circulation fees, and gasoline taxes. Although used on a few occasions in the US, Europe and Japan, scrappage subsidies or cash-for-clunker programs are known to have helped accelerate the retirement of old vehicles and hence, stimulated the purchase of new ones (e.g., Hahn 1995, Adda and Cooper 2000, Mian and Sufi 2012, and Hoekstra et al. 2017). According to our simulations the optimal scrappage subsidy varies between \$1575 and \$2500 depending on how much the authority can prevent car owners outside the restricted area from arbitraging the price gap in scrap values created by the subsidy. This can be done, although at the cost of introducing some friction in the car market, by requiring any given vehicle to have a number of years of registration history in the restricted area. The two simulations reported in rows 6 and 7 of Table 6 correspond, respectively, to the extreme cases of requiring no or full registration history. In any case, the numbers in the table indicate that preventing arbitrage does not make much of a difference for welfare (although it surely does for the government budget).

More importantly for the purpose of our study, scrappage subsidies provide no efficiency advantage over optimal (vintage-specific) driving restrictions. The reason is simple: both instruments seek to affect fleet turnover by aiming at the removal of the most polluting cars; one with a prohibition, the other with a reward.[44] But if anything, it appears that implementation constraints should favor the use of vintage-specific restrictions over scrappage subsidies not only because of the arbitrage problem just described, but also because restrictions are much cheaper to implement for the government under any reasonable estimate of the shadow cost of public funds (see, for example, Laffont 2005). Ultimately, this fiscal cost may explain why these subsidies are used only rarely, and when they are, for a very short time.

This fiscal cost could be avoided if, instead of paying for the removal of these old cars, the government could increase their annual circulation fees to reflect their (expected) external pollution costs during the year. Moreover, if the government not only levies pollution-based circulation fees on the oldest, most polluting cars but on all cars, including new ones, the outcome is very close to the first best (see row 8 in Table 6), provided these fees are set at their optimal levels, which roughly equal the expected Pigouvian bills. Introducing these circulation fees is, in any case, a major policy challenge for any authority as they imply a complete reversal of existing circulation-fee profiles, where older cars pay much less than newer ones. This is in sharp contrast to existing driving restrictions where older cars are already subject to much tougher restrictions than newer ones.

Another policy alternative is for the authority to increase gasoline taxes. As shown in row 9 of Table 6, by making no distinction between high- and low-emitting vehicles a gasoline tax is too bad a proxy for handling local pollution, even if optimally set at ¢46 per gallon (equivalent to a cost increase of ¢2.5 per mile). In fact, gasoline taxes impose a heavier burden on newer vehicles because those are run more intensely than

[44] As the restricted area becomes much larger than the non-restricted area so the "subsidy arbitrage" becomes less of a problem, welfare under the optimal scrappage subsidy approaches that under the optimal driving restriction.

older vehicles, doing little to move the fleet composition towards cleaner cars (see the online Appendix (Section D, Figure D.6) for the steady-state fleet composition). Since one may argue that newer cars are not only cleaner but also more fuel efficient, in the last row of the table we attempt a correction in this direction with a modest change in the welfare numbers.[45]

## 5.4 Extension: Temporal variation in pollution harm

So far, we have assumed that vehicle emissions of local pollutants are equally harmful regardless of the hour of the day, day of the week, or month of the year. In reality, however, emissions impact differently at different points in time, which has been somewhat recognized in existing restriction programs. São Paulo, for instance, places restrictions only during peak hours, from 7 to 10 am and 5 to 8 pm; Paris restricts pre-1997 models only during weekdays; and Santiago extends restrictions only from March through September, when thermal inversions and lack of wind prevent the rapid dispersion of pollutants.

Extending our model to this particular case is relatively straightforward. For simplicity let us assume that the harm caused by a unit of pollution is $h$ during a fraction $\lambda$ of the time and 0 otherwise. In such a setting, driving restrictions appear particularly flexible.[46] The owner of an $a$-year-old car subject to a restriction of intensity $R_a$ applied over a fraction $\lambda$ of the time will drive $x^{r,\lambda}(\theta, a) = \lambda R_a \left(\frac{\theta s_a}{\psi}\right)^{\alpha}$ miles during that time and $x^{r,1-\lambda}(\theta, a) = (1-\lambda)\left(\frac{\theta s_a}{\psi}\right)^{\alpha}$ during the remaining time, resulting in an overall utility of

$$u^r(\theta, a; \lambda) = (1 - \lambda + \lambda R_a)\, \kappa\, (\theta s_a)^{\alpha} - p_a \tag{21}$$

per period. Following the reasoning of Section 3.4, the social planner will choose $R_a$ to maximize $u^r(\theta, a; \lambda) - \lambda h e_a x^{r,\lambda}(\theta, a)$, which gives the same "all-or-nothing" solution obtained in that section. The reason the optimal $R_a$ is independent of $\lambda$ is because of the separability of driving between times in which pollution is of concern (e.g., peak hours, weekdays, winter and fall months) and times in which is not.[47]

[45]Since our model abstracts from technical progress, the way we introduce this fuel-efficiency consideration is by artificially allowing the fuel economy of a car to deteriorate over time, enough for our 2006 fleet to reflect the fuel economy differences observed in the US fleet based on the 2009 US Department of Transportation's National Household Travel Survey (www.nhts.ornl.gov).

[46]The model easily extends to an even higher number of partitions, as in today's Mexico City HNC, which differentiates by weekdays, Saturdays, and Sundays. Damage $h$ will vary accross these partitions, and so will the optimal vintage threshold in each of them. The procedure to obtain these thresholds is the same as in the case of two partitions.

[47]While this separability assumption is reasonable for divisions that concern months and weekdays from weekends, it may appear less reasonable when it concerns hours of the day, as drivers may substitute peak for off-peak driving. This substitutition can still be handled by the model at the cost of additional notation without any fundamental change of our results.

Of the alternative instruments considered above, it is evident that by construction scrappage subsidies cannot cope with this temporal variation as they require the scrapping car to be removed permanently from the market. Gasoline taxes face a similar problem because drivers will arbitrate any price difference created by these taxes unless, of course, they are in place for long time. In contrast, circulation/registration fees have the potential to cope with temporal variation. The authority must offer each year a menu of circulation fees that vary by vintage: drivers have the option to pay either a positive fee for unlimited use of the car or no fee for its use only during the $1-\lambda$ hours in which pollution is not a problem. In equilibrium, there is a cutoff age below which all car owners opt for the fee and above which none does. Not surprisingly, this cutoff is very similar, and sometimes equal, to the threshold in the optimal vintage-specific restriction design. If the option of offering these circulation menus is not available, however, the advantage of circulation fees over (optimal) vintage-specific restrictions rapidly vanishes as $\lambda$ drops. This is illustrated in Figure 10 with simulations that assume pollution harm in Santiago to be $h_r = 0.0041$ (see Table 4) a fraction $\lambda$ of the time and 0 otherwise (notice that $\lambda = 0$ corresponds to the case of no pollution, so the optimal policy is no intervention).

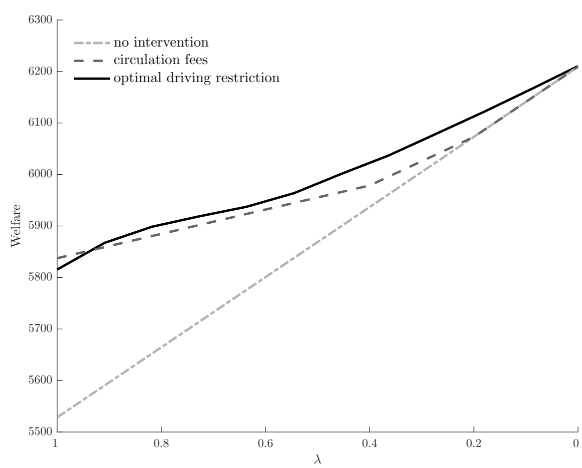



Figure 10: Welfare under temporal variation in pollution harm

Notes: This figure shows welfare estimations when extending the model to allow for temporal variation in pollution harm (i.e., $\lambda$) under various scenarios: optimal dirving restrictions, (no-menu) optimal circulation fees, and no intervention.

# 6 Conclusions

Evidence from many cities around the world experiencing local air pollution problems suggests that driving restrictions are becoming increasingly popular tools to control vehicle pollution. Previous literature (e.g., Eskeland and Feyzioglu 1997, Davis 2008, Gallego et al. 2013a), as well as this paper, show that these policies perform particularly bad when they are designed to affect a driver's *intensive* margin (i.e., amount of travel) with restrictions that treat all cars equally regardless of how much they pollute. In this paper, we have instead focused on the potential of these policies to affect a driver's *extensive* margin (i.e., type of car driven). By introducing "vintage-specific" restrictions, or more precisely, restrictions that differentiate cars by their pollution rates, the empirical and numerical results of the paper show that these policies can be quite effective in helping to accelerate the fleet turnover towards lower-emitting vehicles. The optimal vintage-specific design takes the all-or-nothing form already seen in Paris and Germany's low emission zones: there is an optimal (moving) vintage threshold that separate cars between full exemption and (nearly) complete restriction.

As these vintage-specific restrictions prove to be an effective and practical instrument for fighting local air pollution, particularly when alternative instruments such as scrappage subsidies and pollution-based circulation/registration fees are not available for either fiscal or political reasons, it would be interesting to extend our model to also explore their potential for dealing with global air pollution (unless, of course, a gasoline tax is readily available). If properly designed, these vintage-specific restrictions can also help accelerate the transition towards low- or free-carbon emitting vehicles at a lower cost for the government. An important design issue is to make sure the price of the carbon-free (and fully exempt) option is not much higher than the price of existing carbon-emitting alternatives—not enough that drivers might opt to buy a second, carbon-emitting car instead of the carbon-free option, to bypass the restriction.[48] It may turn out that the optimal design is a combination of subsidies on carbon-free cars and restrictions on carbon-emitting ones.

# References

[1] Adamou, A., S. Clerides and T. Zachariadis (2014), Welfare Implications of Automobile Feebates: A Simulation Analysis, *The Economic Journal* 124, F420-F443.

[48] One important reason the Santiago-1992 restriction worked reasonably well was precisely because the clean option at that time, which was to switch to a car with a catalytic converter sooner than otherwise, was affordable for many households.

[2] Adda, J. and R. Cooper (2000), Balladurette and Juppette: A Discrete Analysis of Scrapping Subsidies, *Journal of Political Economy* 108, 778-806.

[3] Angrist, J. and M. Rokkanen (2015), Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away from the Cutoff, *Journal of the American Statistical Association* 110, 1331-1344.

[4] Bento, A., L. Goulder, M. Jacobsen and R. von Haefen (2009), Distributional and Efficiency Impacts of Increased US Gasoline Taxes, *American Economic Review* 99, 667-99.

[5] Calonico, S., M. Cattaneo and R. Tittiunik (2014), Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs, *Econometrica* 82, 2295-2326.

[6] 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.

[7] Davis, L. (2008), The Effect of Driving Restrictions on Air Quality in Mexico City, *Journal of Political Economy* 116, 38-81.

[8] Davis, L. (2017), Saturday Driving Restrictions Fail to Improve Air Quality in Mexico City, *Scientific Reports* 7, 41652.

[9] d'Haultfœuille, X., P. Givord and X. Boutin (2014), The Environmental Effect of Green Taxation: The Case of the French Bonus/Malus, *The Economic Journal* 124, F444-F480.

[10] Drummond, P. and P. Ekins (2016), Tackling Air Pollution from Diesel Cars through Tax: Options for the UK, ISR report, UCL.

[11] Eskeland, G., and T. Feyzioglu (1997), Rationing Can Backfire: The "Day without a Car" in Mexico City, *World Bank Economic Review* 11, 383-408.

[12] Gallego, F., J.-P. Montero and C. Salas (2013a), The Effect of Transport Policies on Car Use: Evidence from Latin American Cities, *Journal of Public Economics* 107, 47-62.

[13] Gallego, F., J.-P. Montero, and C. Salas (2013b), The Effect of Transport Policies on Car Use: A Bundling Model with Applications, *Energy Economics* 40, S85-S97.

[14] Gavazza, A., A. Lizzeri and N. Roketskiy (2014), A Quantitative Analysis of the Used-Car Market, *American Economic Review* 104, 3668-3700.

[15] Hahn, R. (1995), An Economic Analysis of Scrappage, *RAND Journal of Economics* 26, 222-242.

[16] Hoekstra, M., S. Puller and J. West (2017), Cash for Corollas: When Stimulus Reduces Spending, *American Economic Journal: Applied Economics*, forthcoming.

[17] Jacobsen, M., C. Knittel, J. Sallee, and A. van Benthem (2016), Sufficient Statistics for Imperfect Externality-Correcting Policies, mimeo, UC Berkeley.

[18] Kahn, (1996), New Evidence on Trends in Vehicle Emissions, *The RAND Journal of Economics* 27, 183-196.

[19] Laffont, J.-J. (2005), *Regulation and Development*, Cambridge University Press, UK.

[20] Liu, Y., Z. Yan, S. Liu, Y. Wu, and Q. Gan (2015), The Effect of the Driving Restriction Policy on Public Health in Beijing, Working paper, Beijing Institute of Technology.

[21] Lu, S. (2006), Vehicle Survivability and Travel Mileage Schedules, NHTSA Technical Report, US DOT, Springfield, VA.

[22] Mian, A. and A. Sufi (2012), The Effects of Fiscal Stimulus: Evidence from the 2009 Cash for Clunkers Program, *The Quarterly Journal of Economics* 127, 1107-1142.

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

[24] Oliva, P. (2015), Environmental Regulations and Corruption: Automobile Emissions in Mexico City, *Journal of Political Economy* 123, 686-724.

[25] Onursal, B. and S.P. Gautam (1997), Vehicular Air Pollution: Experiences from Seven Latin American Urban Centers, World Bank Technical Paper, Washington, DC.

[26] Parry, I. and J. Strand (2012), International Fuel Tax Assessment: An Application to Chile, *Environment and Development Economics* 17, 127-144.

[27] Posada, F., D. Wagner, G. Bansal, and R. Fernandez (2015), Survey of Best Practices in Reducing Emissions through Vehicle Replacement Programs, White paper, The International Council on Clean Transportation.

[28] Rivers, N., and B. Shaufele (2015), New Vehicle Feebates, mimeo, University of Western Ontario.

[29] Viard, B., and S. Fu (2015). The Effect of Beijing's Driving Restrictions on Pollution and Economic Activity, *Journal of Public Economics* 125, 98-115.

[30] Wolff, H. (2014), Keep Your Clunker in the Suburb: Low-Emission Zones and Adoption of Green Vehicles, *The Economic Journal* 124, F481-F512.

[31] Zhang, W., C.-Y.C. Lawell, V.I. Umanskaya (2017), The Effects of License Plate-based Driving Restrictions on Air Quality: Theory and Empirical Evidence, *Journal of Environmental Economics and Management* 82, 181-220.

# Online Appendix for: Vintage-specific driving restrictions (not for publication)

Nano Barahona, Francisco Gallego and Juan-Pablo Montero*

March 1, 2017

*Emails: hbaraho@stanford.edu; fgallego@uc.cl; and jmontero@uc.cl.

# Appendix A: Additional Figures

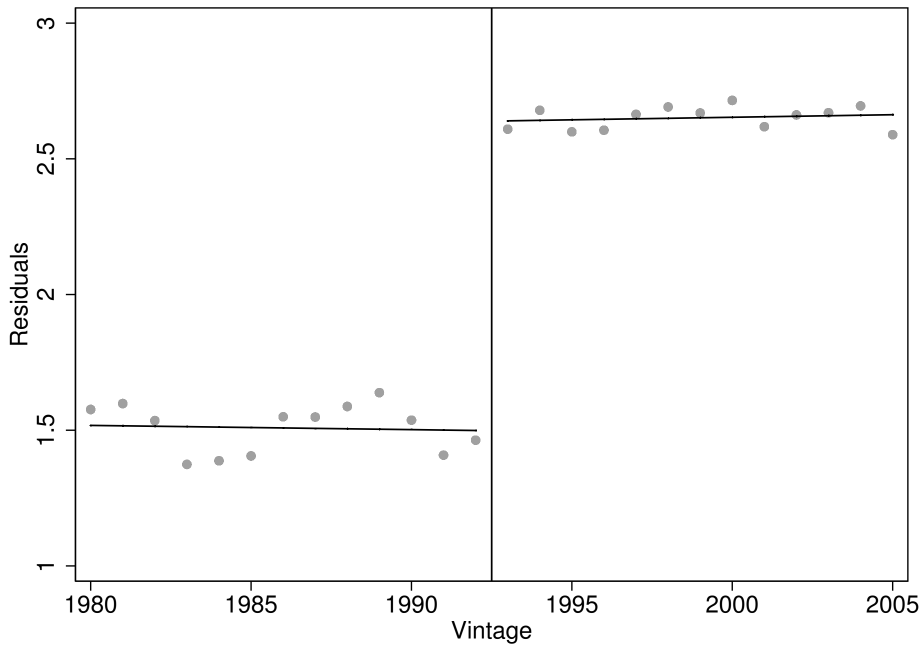



Figure A.1: Regression discontinuity design estimates for car fleet in treated municipalities (Santiago)

Notes: The units of observation are municipalities affected by the driving restriction program. We follow the procedure explained in section 2.1 of the paper. The point estimate of the RDD estimate is 1.222 and the standard error is 0.201.

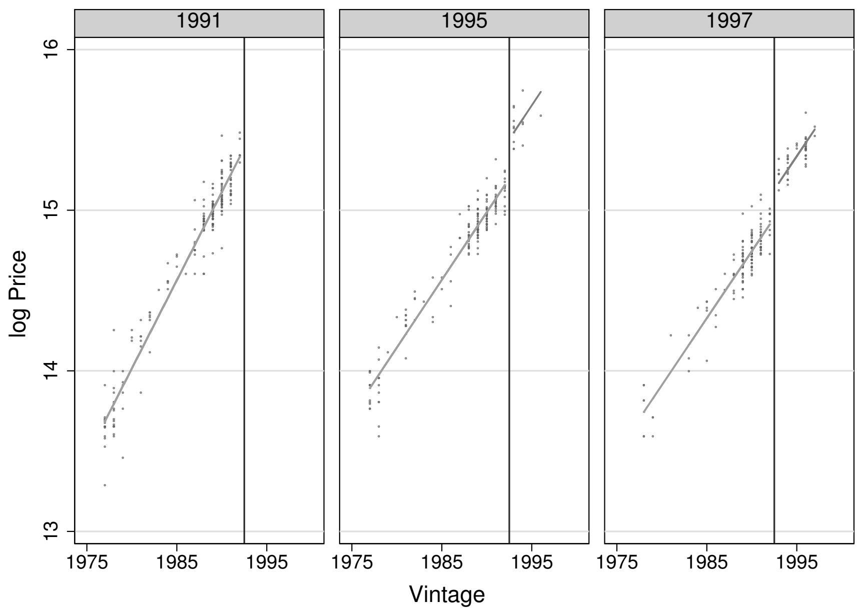



Figure A.2: Price of a second-hand Toyota Corolla

Notes: The unit of observation in this figure corresponds to a Toyota Corolla sell offer published in the newspaper in October, November and December of 1991, 1995 and 1997 respectively. The vertical line separates offers of cars vintage pre and post-1992. The lines are best linear predictor functions of price (in logs) and vintage at each side of the discontinuity.

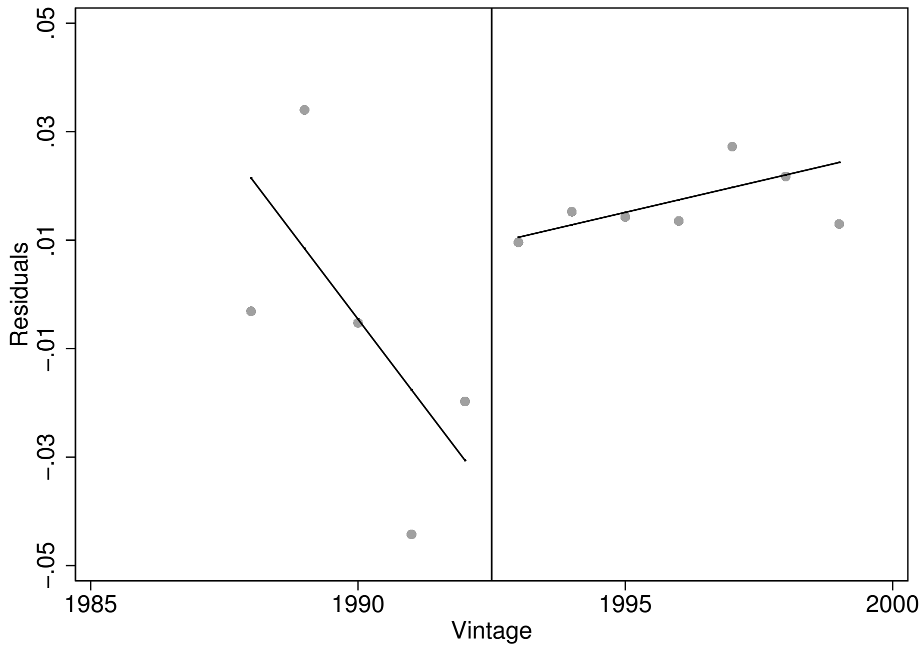



Figure A.3: Regression discontinuity design estimates for the effect of a catalytic converter on prices of used cars

Notes: The units of observation are car ads published in the newspaper the first Sunday of every month between 1988 and 2000. We follow the procedure explained in section 2.2 of the paper. The estimates (standard errors) for each individual model are as follows: FIAT UNO: 0.005 (0.007), HONDA ACCORD: 0.089 (0.012), HONDA CIVIC: 0.062 (0.008), MAZDA 323: 0.057 (0.005), PEUGEOT 205: 0.035 (0.009), PEUGEOT 505: 0.073 (0.011), TOYOTA COROLLA: 0.211 (0.024).

# Appendix B: Additional Tables

Table B.1: The effects of the driving restriction on the share of cars for contiguous vintages

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | 88-89 | 89-90 | 90-91 | 91-92 | 92-93 | 93-94 | 94-95 | 95-96 |
| DR | 0.0178 | 0.0319* | 0.0226 | 0.00231 | -0.275*** | 0.0142 | 0.0513*** | 0.0173 |
| | (0.016) | (0.018) | (0.023) | (0.014) | (0.043) | (0.015) | (0.018) | (0.014) |
| Obs | 268 | 268 | 268 | 268 | 268 | 268 | 268 | 267 |
| $R^2$ | 0.152 | 0.120 | 0.189 | 0.090 | 0.527 | 0.320 | 0.151 | 0.193 |

Notes: OLS regressions with one observation per municipality. Municipalities with less than 300 cars were dropped from the sample. Standard errors, which are in parenthesis, are calculated via block bootstrap at the province level (53 provinces in total). All columns include municipality controls: Income per capita (in linear and quadratic form), population, coefficient of variation of income per capita, urbanization ratio, distance to Santiago (in linear and quadratic form), and dummies for municipalities in northern and far away regions. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.2: Effect of living in Santiago on having more than one car

| | 1998 survey | 2006 survey |
|---|---|---|
| OLS | 0.0159 | 0.00999 |
| | (0.015) | (0.014) |
| Probit | 0.0103 | 0.00310 |
| | (0.014) | (0.011) |
| Hurdle poisson-logit | 0.062 | 0.0136 |
| | (0.081) | (0.101) |

Notes: The unit of observation is the household and the dependant variable is the number of cars in a given household. We present only the marginal effects of living in Santiago but the models also include the following variables: household characteristics related to income, assets, age, gender and employment status of the head of the household, the composition of the household (in terms of number of members and also number of employed members), and the size of the county in which the household is located. OLS and probit estimations are on households with at least one car. The Hurdle poisson-logit model uses all the observations. Observations are weighted using expansion factors. Standard errors, which are clustered at the municipality level, are in parenthesis. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Table B.3: Prices in Santiago and the rest of the country (2013)

| | (1) | (2) |
|---|---|---|
| Santiago | -0.0263$^{**}$ | -0.0198$^{*}$ |
| | (0.009) | (0.009) |
| Santiago × pre-1993 | | -0.0277$^{***}$ |
| | | (0.005) |
| Vintage f.e. | yes | yes |
| Date f.e. | yes | yes |
| Match-Model f.e. | yes | yes |
| Observations | 53915 | 53915 |
| $R^2$ | 0.717 | 0.717 |

Notes: The unit of observation is a car sell offer posted in a online platform during October of 2013 of vintages 1990 to 1995. The dependent variable is posted price of the car (in logs). $Santiago_i$ is a dummy that takes the value of 1 is the offer is the car is being sold in Santiago. $pre\text{-}1993_i$ takes the value of 1 if the car was built before 1993. We control by vintage, date of the offer and match-model fixed effects. Standard errors, which are calculated via block bootstrap clustering at the week-region level, are presented in parenthesis.
$^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$

Table B.4: Survival rates

| Age | 1-4 | 5-8 | 9-12 | 13-16 | 17-20 | 21-24 | 25-28 | 29-32 | 33-36 |
|---|---|---|---|---|---|---|---|---|---|
| $\gamma_a$ | 0.9966 | 0.9966 | 0.9966 | 0.9434 | 0.8267 | 0.7226 | 0.5828 | 0.5242 | 0.5242 |

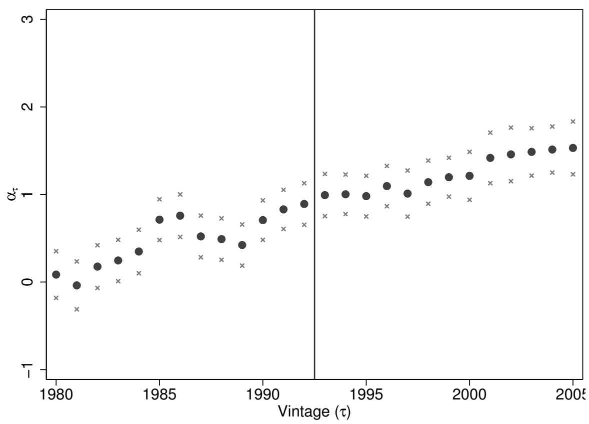
Notes: Survival rates were calculated using constrained OLS. The unit of observation is the total number of cars of a given vintage $\tau$ found in the country on a given year $t$ ($y_a^t$), which we use in the regression $y_a^t = \gamma_a y_{a+1}^{t+1} + \varepsilon_a^t$, along with imposing $\gamma_a \leq 1$ and $\gamma_{a+1} \leq \gamma_a$.

# Appendix C: Effects on Fleet Composition Without Correction

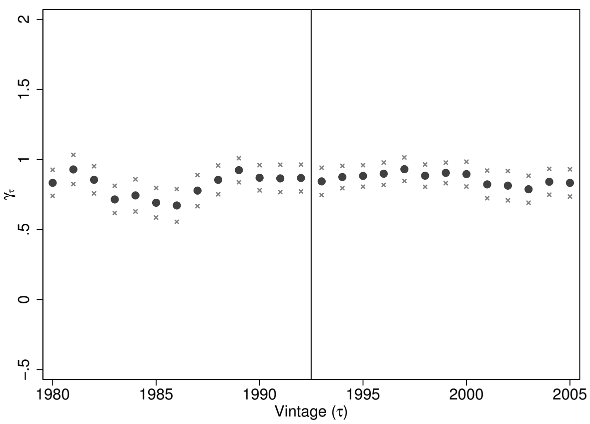
Table C.1: The effects of the driving restriction on the share of cars for contiguous vintages

| | 92-93 | 92-93 | 91-92 | 93-94 |
|---|---|---|---|---|
| Driving restriction | -0.215*** | -0.182*** | 0.00231 | -0.0223 |
| | (0.014) | (0.046) | (0.014) | (0.016) |
| Controls | No | Yes | Yes | Yes |
| Observations | 268 | 268 | 268 | 268 |
| $R^2$ | 0.480 | 0.527 | 0.090 | 0.320 |

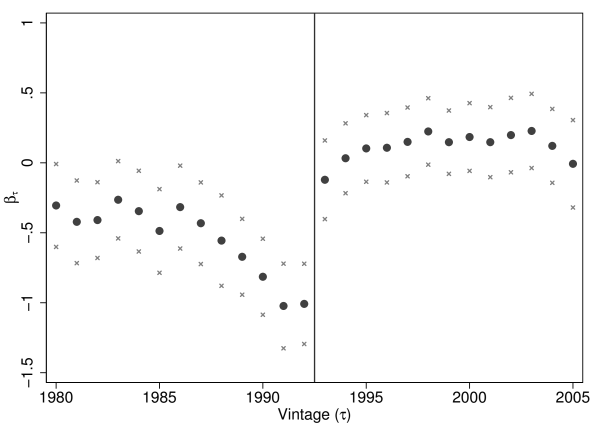
Notes: OLS regressions with one observation per municipality. Municipalities with less than 300 cars were dropped from the sample. The dependant variable corresponds to the ratio given by $q_i^\tau/(q_i^\tau + q_i^{\tau+1})$, where $q_i^\tau$ is the total number of cars from vintage $\tau$ found in municipality $i$ in 2006. The first two columns correspond to the regression where $\tau = 1992$, while column 3 and 4 correspond to $\tau = 1991$ and $\tau = 1993$ respectively. , which are in parenthesis, are calculated via block bootstrap at the province level (53 provinces in total). Municipality controls include: Income per capita (in linear and quadratic form), population, coefficient of variation of income per capita, urbanization ratio, distance to Santiago (in linear and quadratic form), and dummies for municipalities in northern and far away regions. This table differs from Table 1 in the paper as here we do not impose the correction for vintages 1993, 1994 and 1995. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$




(a) $INCOME$ coefficients




(b) $POP$ coefficients

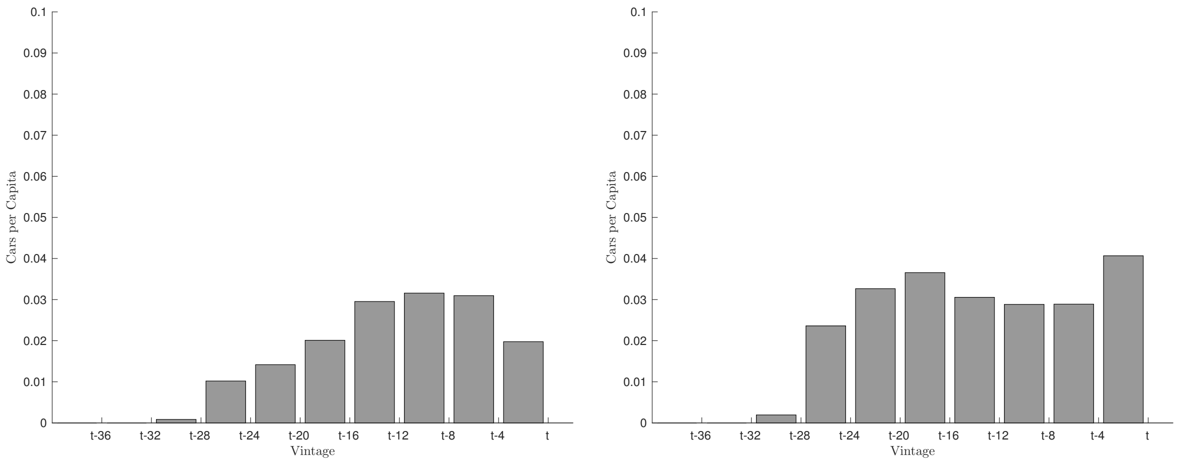



(c) $DR$ coefficients

Figure C.1: Vintage effects of driving restrictions, income, and population

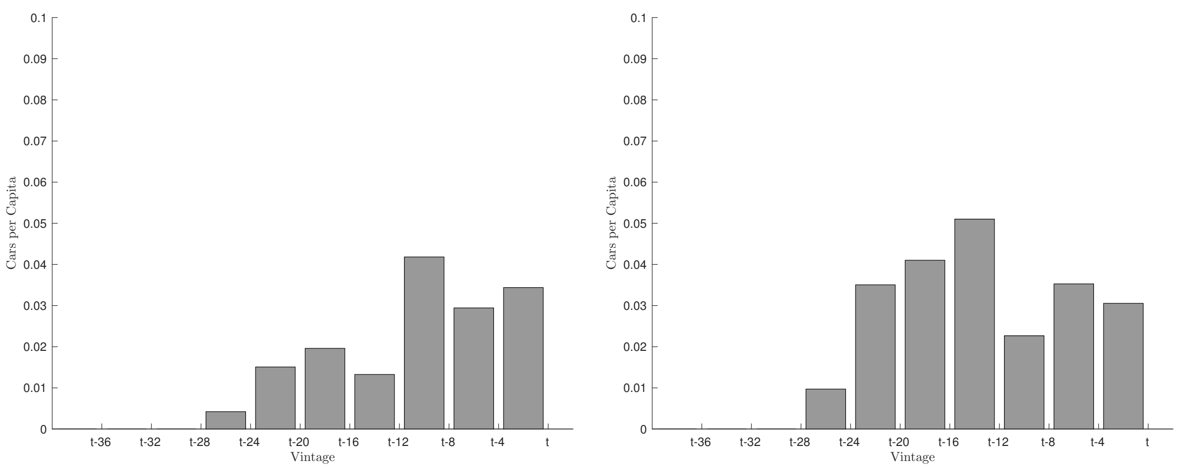
Notes: This figure presents the estimated vintage effects after estimating equation (2) using data at the municipality level for 2006. Panel (a) presents the coefficients for municipality income, Panel (b) for municipality population, and Panel (c) for the effect of being affected to the driving restriction program. Darker dots represent the point estimates for each coefficient and lighter dots correspond to the 95% confidence intervals using robust standard errors. The vertical line splits vintage between 1992 and older and 1993 and younger vintages. This figure differs from Figure 2 in the paper as here we don't impose the correction for vintage 1993, 1994 and 1995.

# Appendix D: Simulation Results



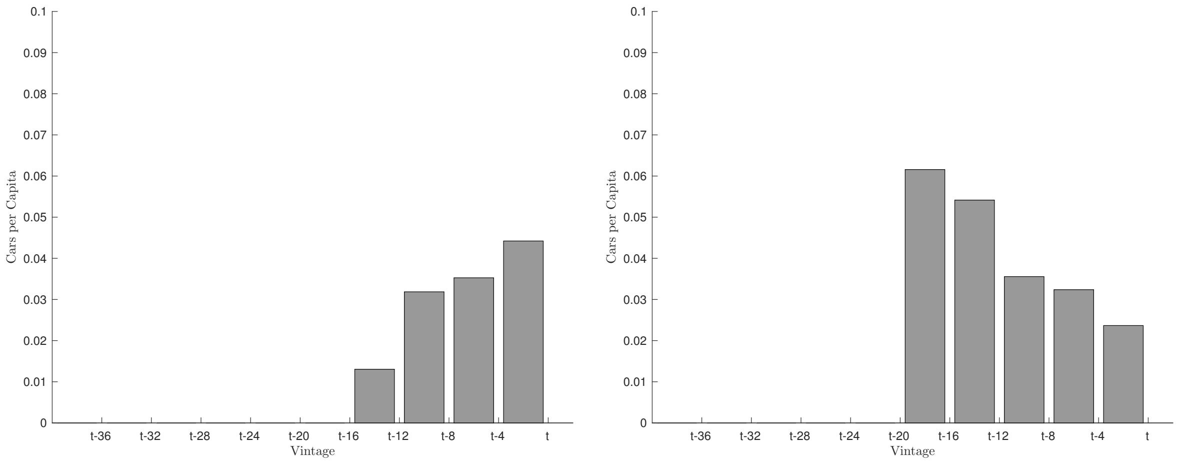

(a) Santiago (b) Rest of the country

Figure D.1: Steady-state fleet composition under a uniform driving restriction




(a) Santiago (b) Rest of the country

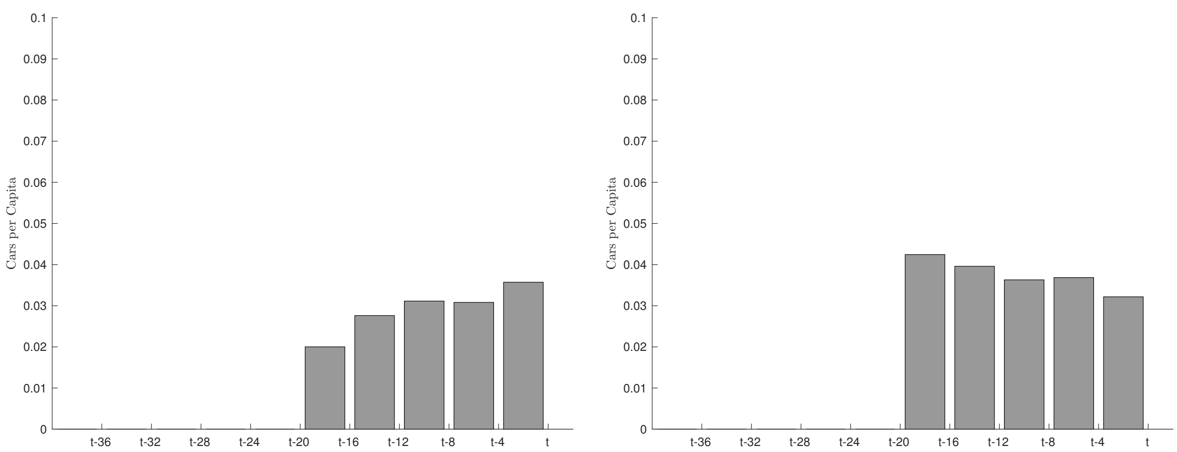
Figure D.2: Steady-state fleet composition under a driving restriction that exempts cars 12 years old and younger




(a) Santiago

(b) Rest of the country

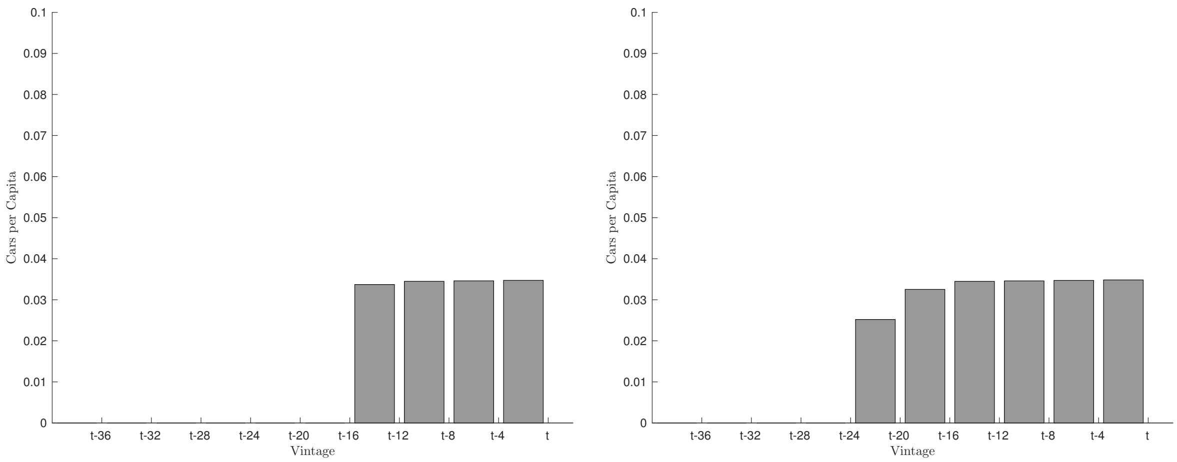
Figure D.3: Steady-state fleet composition under optimal circulation fees




(a) Santiago

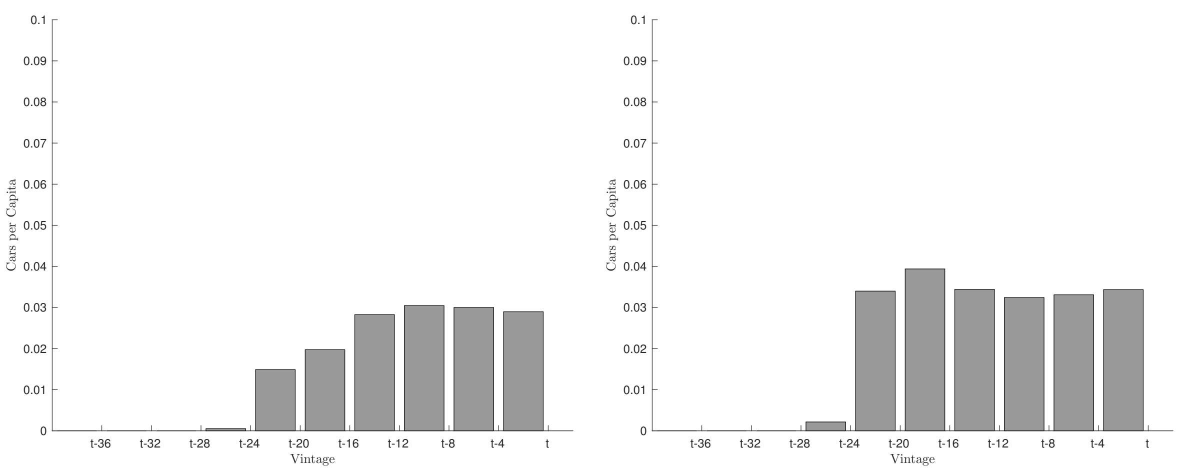
(b) Rest of the country

Figure D.4: Steady-state fleet composition under the optimal scrappage subsidy with full arbitrage between restricted and non-restricted areas ($1575)




(a) Santiago (b) Rest of the country

Figure D.5: Steady-state fleet composition under the optimal scrappage subsidy with no arbitrage between restricted and non-restricted areas ($2500)




(a) Santiago (b) Rest of the country

Figure D.6: Steady-state fleet composition under an optimal gasoline tax in Santiago with correction for fuel-economy (¢34 per gallon)

**MIT Center for Energy and Environmental Policy Research**
77 Massachusetts Avenue, E19-411
Cambridge, MA 02139
USA

Website: ceepr.mit.edu

**MIT CEEPR Working Paper Series** is published by the MIT Center for Energy and Environmental Policy Research from submissions by affiliated researchers.



For inquiries and/or for permission to reproduce material in this working paper, please contact:

Email ceepr@mit.edu
Phone (617) 253-3551
Fax (617) 253-9845