INSTITUTO DE ECONOMÍA

TESIS de MAGÍSTER

2018

Smog-Checks and Local Air Pollution: Evidence from Chile

José Diego Salas

www.economia.puc.cl

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA

TESIS DE GRADO
MAGISTER EN ECONOMIA

Salas, Muñoz, José Diego

Julio, 2018

PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA

"Smog-Checks and Local Air Pollution: Evidence from Chile"

José Diego Salas Muñoz

Comisión:
Constanza Fosco
Tomás Rau
Gert Wagner

Santiago, Julio de 2018

Pontificia Universidad Católica de Chile

Instituto de Economía

M.A. Thesis:

Smog Checks and Local Air Pollution: Evidence from Chile

July 22th, 2018.

José Diego Salas M [1].

Abstract

The main goal of this paper is to establish whether the Smog-Check system improves local air quality. Using panel data from 2008-2016 for eleven air pollution monitors and a comprehensive dataset of Smog-Checks, I find that an increase in the number of rejections lowers the levels of a small group of pollutants, even after controlling for confounding variables. The main result is that an increase of one standard deviation in the number of rejections decreases $[CO]$, $[NO_2]$, $[PM_{10}]$, and $[PM_{2.5}]$ by 5,1%, 6%, 1,4%, and 5,8% of a standard deviation respectively. This is about 16.7% of the effect of what establishing an environmental pre-emergency can accomplish. These results remain significant after changing the empirical specification, and several falsification exercises are conducted to strengthen the identification strategy. Additionally, I explore the potential heterogeneity underlying these results along the station-quality dimension using metrics derived from California's STAR program, which compares the expected rejection rate versus the realization of this variable, after controlling for a comprehensive set of observable characteristics. I find that rejecting a car by a low-quality station has a negligible impact on air pollution while high-quality stations are the main driver behind the aforementioned results.

[1]Thesis written as a Master student at Department of Economics, Pontificia Universidad Católica de Chile. I would like to thank my thesis advisors Constanza Fosco, Tomas Rau, and Gert Wagner for their continuous guidance and comments. I am also very grateful to professors Juan-Pablo Montero and Francisco Gallego for their suggestions and support during my thesis. In addition, I would like to thank Sebastián Poblete and Bernardo de Moura for their support, ideas, discussion, and proofreading. This project is supported by FONDECYT No.1161802 of professor Juan-Pablo Montero. Any errors or omissions are my own responsibility. Comments at: jdsalas@uc.cl

For Josette, whom does nothing but give love.

Contents

1 Introduction

There is a long-standing debate about the most cost-effective and efficient policies to curb air pollution. Automobile emissions are an important source of pollution, and this is particularly relevant in developing countries where emissions standards tend to be less stringent. Specifically, for Chile's capital city Santiago cars and mobile sources are responsible for about 94% of Carbon Monoxide (CO) Emissions, which is one of the main components of air pollution (Gallego et al., 2013). However, private cars not only contribute to local air pollution by emitting CO, but also by emitting Nitric Oxides (NO_x), Particulate Matter of different sizes ($PM_{2.5}$ and PM_{10}), and other organic volatile compounds. There is a wide consensus among the economic and medical literature that the health effects of air contamination are sizable (see below). The main instrument used to enforce emission standards for vehicles around the globe are compulsory Smog-Checks. However, there is little empirical evidence about whether this program is being effective at improving local air quality.

The link between Smog-Checks and local air pollution could be imperceptible due to several reasons.The main objective of the Smog-Check system is to detect cars that are not complying with emission standards. However, this ability has been questioned in the literature because there seems to exist a wide gap between on-the-road emissions and levels measured by official tests at Smog-Checks plants. This gap could be explained, as in (Wenzel et al., 2004), owing to the volatile nature of car emissions or, by the short-lived nature of the repairs made to approve the Smog-Checks (Glazer et al., 1993, 1995). Additional concerns were raised by Oliva (2015), who found that in Mexico around 10% of old-car owners paid bribes in order to circumvent the regulations. All of these concerns could translate into a negligible effect of the Smog-Check program in terms of local air pollution.

In this work, I will answer if the Smog-Check program is fulfilling its objective of enforcing car emission standards and whether this is being translated into lower levels of air pollution. This will be done by assessing whether the inspections that ended in a rejection are diminishing the concentration of local air pollution after controlling for several sources of possible confounding factors. From a more general perspective, Im am explaining variations in the current levels of air pollution, after taking into account several plausible sources of confounding factors, with variations in the total number of rejections conducted on days previous to a specified date. Additionally, I will classify Smog-Checks plants according to the ratio between their expected and realized rejection rates to obtain a station-quality metric like has been recently implemented in California under the STAR program (Bureau of Automotive Repairs, 2012). This is done to understand whether there is heterogeneity in our results along this dimension.

This work uses data from the air monitoring network maintained by the Chilean Environment Office (SINCA) and data from the Smog-Check plants coming from a centralized network that the Chilean Transport and Telecommunication Office control. The main results suggest that there is a small, but significant effect of rejecting an additional high-emitting car in terms of local air pollution. With this dataset, I also classified Smog-Checks stations according to a quality metric called Similar Vehicle Failure Rate (SVFR) derived from California's STAR program (Bureau of Automotive Repairs, 2012). This metric compares the expected failure rate at a determined station to the realization of that variable, after controlling for a comprehensive set of observable characteristics. A station will be classified as low-quality if its SVFR is below a threshold of 0.75, which means that is rejecting a 25% less than what it would be expected based on the observable characteristics of the car conducting their inspection at that specific station. When I explore the potential heterogeneity along this dimension I find that rejections from low-quality stations have an imperceptible effect, while high-quality stations are the main driver behind our results.

To the best of my knowledge, there is only one article linking the Smog-Check program to local air pollution and classifying plants according to their "quality" (Sanders and Sandler, 2017). However, there are several reasons to suspect why a different context could be very important. First, in developing countries, enforcement, car-manufacturers controls, and pollution standards might be less strict. Second, an older car fleet could exacerbate some of the aforementioned problems. Finally, I am able to use a station-quality metric which is observable and viable to implement from the regulator's perspective.

Several papers have documented the negative effects of environmental pollution due to several causes on outcomes such as, infant mortality (Kampa and Castanas, 2008; Chay and Greenstone, 2003; Knittel et al., 2011; Currie and Neidell, 2005), losses of worker productivity (Hanna and Oliva, 2015; Graff-Zivin and Neidell, 2012; Chang et al., 2016b,a), School absence and cognitive performance in general (Currie et al., 2009), and so on. Most of this literature takes advantage of sources of external variation of air pollution and other sources of plausibly exogenous variation in air pollution to understand how they affect this outcomes. However, there are very heterogeneous results on the efficacy of current policies against air pollution (Harrington et al., 2000; Wolff, 2013; Davis, 2008; Gallego et al., 2013). In this context, this paper contributes to the literature by assessing the efficacy of Smog-Checks in terms of local air pollution.

The article proceeds as follows. In the next section, I describe our data and discuss some relevant features for the present work. Section 3 discuss the regulatory framework that car-owners face and how this is related to the main empirical estimations. Section 4 the main equations to be estimated and the identification assumptions are discussed. This section is divided into two parts to explain how to asses the impacts of an additional rejection and how to explore the potential sources of Heterogeneity in these results. Section 5 discusses the main results of this estimations. Finally, section 6 conclude and shed some lights on policy lessons that could be derived from the empirical exercises.

2 Data

The present work uses data from several sources. First, I have a comprehensive data set of all the Smog-Checks conducted between 2008-2016 in Chile. The information includes data about each car such as, make, model, year of fabrication, odometer readings[2], and even plate number. From this panel, I get the main right-hand-side variable which is the total number

[2]While this data is supposedly reported it is very noisy and is not available for every year.

of rejections at a specific plant. Notice that with this dataset I know specifically the cause behind the rejection (i.e. I can observe whether the rejection comes from elevated tailpipe emissions or from another mechanical malfunctioning). The other variables available will be used to construct the quality metrics as specified by California's Bureau of Automotive Repair manual (Bureau of Automotive Repairs, 2012).

Second, the main outcomes come from the air-quality monitoring network maintained by the Chilean Environmental Office (*Sistema de Información de la Calidad del Aire* - SINCA). This network consists of 11 monitors throughout the Metropolitan Region and is representative of different socio-economic sectors. They provide hourly information on the levels of several pollutants. The focus of this paper is on Carbon Monoxide due to its close link to car-use, but I report coefficients for $PM_{2.5}$, PM_{10}, NO_2 and O_3 which are available from the same source. Additionally, this dataset has information on hourly temperature, humidity, wind speed, and wind direction which are key controls for the main empirical specification.

With these data-sets I construct a panel from 2008-2016 at the geographic-unit level with the daily concentrations of different air pollutants, the total number of rejections and by levels of station-quality, and all the weather conditions relevant to air pollution. However, to construct this panel is necessary to match Smog-Check stations to specific Monitors. This is done by considering all plants within a 5-kilometer radius of the monitor to be part of the same "Geographic-Unit"[3]. The datasets consists of 11 geographic-units paired with 35 Smog-Check stations, so in average there is 3.2 stations per Geographic-Unit. I then explore changes in this criteria to understand the spatial nature of our results.

Summary statistics for the number of inspections-rejections conducted daily at each geographic unit are available in Table 1 for each year. At first sight, one can notice that the number of inspections is not monotonically increasing over time. The series reached its peak

[3]Geographic-Unit is our observational level, and consists of an air-quality monitor with all the Smog-Check plants that are within a 5-kilometer radius. The radius is constructed around the air-quality monitors. There were two stations in which there was overlap of the geographic-units. I assigned those stations to the closest monitor. Additionally, I re-estimated the same empirical specifications by dropping this stations from our sample and all results remain unchanged.

in 2013. This might be due to the fact that cars from 0 up to 2 years of age are exempt from mandatory inspections and that the car fleet has been getting younger in time (Barahona et al., 2017). In Table 2, the number of observations coming from the Smog-Check database is presented. Here, it is important to highlight that the number of total inspections has been increasing in years which is consistent with data reported by the National Institute of Statistics (*Instituto Nacional de Estadísticas*) and a growing vehicular fleet. However, these numbers do not match the total amount of cars because I am reporting the total number of inspections which is larger because there are cars conducting more than one inspection and several cars conducting multiple inspections in the same year. I discuss the balance in term of the number of cars with the number of observations in our dataset in the following section. Table 3 shows the average daily reading of several pollutants for every year. In figure 1, I plot the daily average concentration for $[PM_{10}]$, $[CO]$, and $[O_3]$. The seasonality of this variables responds mainly to weather conditions. It is important to highlight that the seasonality for $[O_3]$ is the opposite as with $[CO]$ and $[PM_{10}]$. This is due to the fact that the formation of ground-level Ozone depends on the presence of sunlight. For $[CO]$ one can see that there is a downward trend in time, which is consistent with what Bharadwaj et al. (2017) describe for a previous period. $[PM_{2.5}]$, and $[NO_2]$ are not plotted here due to their similarity with $[CO]$, and $[PM_{10}]$.

Finally, I want to refer to figure 2. On panel a), I show how car emissions, specifically for CO, behave according to the age of a specific car. A “group” is defined as cars that have between 0-4 years of age for group 1, 5-8 years of age for group 2, and so on. Cars with 20 o more years of age are pooled in group 5. The different lines of this plot show the average CO emissions for each group after controlling for fixed-effects at various levels. For example, the blue line labeled “Plate No.” plots the same relationship described, but after controlling for fixed-effects at plate number. This is extremely, demanding since I am taking into account everything that is common for a given car. If the car has not changed owner I am even controlling for the level of maintenance that a given owner puts in his car or how carefully he drives, etc. The main conclusion of this graph is that polluting cars come mainly for a specific group of the total vehicular fleet (i.e. from groups 4-5). Panel b) of the same figure, plots the variance of CO emissions by doing the same exercise as explained before. With this information, I know that not only CO emissions grow almost exponentially with the age of cars, but that they become more heterogeneous along this dimension. This two facts are related at the heart of the empirical strategy that I will develop on section 4.

3 Smog-Check Regulation in Chile

The main objective of the vehicle inspection program is to determine whether the vehicles have the technical requirements to guarantee that their circulation is safe and is complying with environmental protection laws. To this end, the law requires that private car owners conduct a comprehensive examination at least once a year. There are some exceptions. For example, new cars have a two-year period in which they are not required to be inspected. A car that has approved the examination gets a visible sticker. This certification is needed in order to obtain a valid transit permit which is to be renewed in March of every year. Driving without an approved vehicle inspection can result in a fine of 1 to 1.5 UTM or about 77 to 115 USD approximately. The fine for driving without a transit permit goes from 1.5 to 3 UTM (115 to 230 USD approximately). Additionally, if the driver does not have a valid transit permit, the car is immediately seized and taken to a municipal parking lot. To retrieve the car, the owner has to pay the fine, obtain a valid transit permit, and pay for the time the car was in the parking lot.

Car owners that are meant to conduct their vehicle inspection in the current year are subject to a pre-determined monthly schedule in which the past certificate expires. This schedule depends solely on the last number of the license plate. December and March are the only two months in which no license plate will have their certificate expiring. However, many people go to inspect their vehicles in this months for two reasons: to obtain a transit permit, and to get their vehicle inspected before driving on the highway on vacation months. This is evident in our panel data set since there are inspections and rejections every working day.

One concern that can arise with the use of this data is that, since it is "voluntarily" to obtain a valid transit permit and to have all the paperwork up to date, it may be possible that few car owners actually get their vehicles inspected. This would result in a dataset with fewer inspections than the total amount of private cars in the city. In order to check that there is some kind of balance in the number of cars and inspections conducted, I use data from the National Statistics Bureau (INE) (Subdirección Técnica, INE, 2016). These reports, show the number of private-owned cars by region. For example in the year 2009, there were 1,189,014 privately-owned cars in the metropolitan region. In our data, for that year I have 965,112 cars. However, the regulation does not require cars that are 1 and 2 years old to conduct an official inspection. With data from 2010, I know that the amount of cars manufactured in 2008 and 2007 is 230,211. So, in total there were 1,195,323 cars conducting their inspections in the Metropolitan Region. The small difference may be due to the fact that some privately-owned cars might be classified differently in the two sources of data and/or people getting their inspections outside of the Metropolitan Region, etc. This kind of balance is present in all the years that cover our dataset[4].

The main focus of this work is on smog-checks, which correspond to a specific inspection station or stage inside the vehicle inspection program. There are many steps in the inspection process. The process starts by checking visual aspects of the car, such as the plate number visibility, the seat-belts are working properly, windows and rearview mirrors are in good condition, etc. Next, they inspect specific parameters for all the lights of the car, and then they proceed to check the wheel alignment, breaks, suspension system, and finally they proceed to the emissions inspection. The first part of this stage of the process consists of a visual examination of car emissions. At this level, if cars emissions are visible, the car is immediately rejected and has to conduct a reinspection in the near future. In the next step, cars are connected through the tailpipe to a machine which detects levels for different pollutants. The readings of this machine are directly transmitted to the Chilean Telecom-

[4]From looking at table 2, for 2009 the number of total inspections conducted is 1,839,362. This number includes all cars that conducted inspections at a certified station. If I adjust this for the total rejection rate for this year I have an estimated of 1,164,624 cars conducting their inspection at a certified station which closely matches the number reported by INE.

munications and Transport Office. If any of the pollutants registered is above the regulation threshold, the car is rejected and has to conduct a reinspection (Ministerio de Transportes y Telecomunicaciones, 2015). Owners of a rejected car have an average term of two weeks to conduct the reinspection at a reduced value. However, this term depends on the seriousness of the malfunction.

4 Empirical Strategy

4.1 Impact of Rejections

In light of the context described in the previous section, what is relevant for our purposes is the following hypothetical process. First, car owners have to conduct their inspection on the month assigned to them based on their plate number. Our empirical specification is partially based on the fact that the timing of this process is quasi-random in regard to local air pollution. Next, those cars that comply with the emission norms are given the approval certificate. However, those cars whose emissions are above the limits are rejected. The following part of the process is critical because what happens at this stage is crucial to our results. The owner of a car that has been rejected has to perform a reinspection within the due date if he wants to keep driving his car. However, he can still drive his car without the possession of a valid certificate, but he will be at risk to be fined according to the amounts discussed in the previous paragraphs. To obtain a valid certificate on the next reinspection there are many possibilities. Among them are: repairing the car, trying to cheat the system (paying bribes and/or performing short-lived modifications to the car), or simply trying to get a valid certificate without conducting any further repairs. If the smog-checks system is working properly, only the first option would allow a car-owner to obtain a valid passing certificate. If there are flaws to the system, it would be possible to pay a bribe and get the certificate, and/or the machines-technicians would not be able to detect short-lived modifications or cars without repairs when inspecting the car the next time. If the latter is true, then there would be no impacts of the smog-check program on the level of air pollution.

In the section "Station-Quality Heterogeneity" I will explain how these potential differences translate into different levels of quality according to our measures. As mentioned before, the only "tool" available to curtail tailpipe emissions is to remove high-emitting vehicles from the street (by rejecting them) and induce a subsequent maintenance or repair of these cars[5]. This process is at the heart of the empirical specification I would employ in the next sections.

From a more general perspective, Im am trying to explain variations in the current level of air pollution, after taking into account several plausible sources of confounding factors, with variations in the total number of rejections conducted on days previous to a specified date. Obviously, the levels of air pollution differ from day to day due to several factors. However, the number of rejections at Smog-Checks should affect the composition of the vehicular fleet regarding their levels of emissions on a daily basis. The latter is due to the fact that the system is taking high-polluting cars out from the streets and forcing them to repair their cars in order to obtain a valid certification. So, if the number of inspections/rejections varies on a daily basis, and local air pollution also varies on a daily basis I can exploit this quasi-exogenous source of variation of rejections to understand whether it is affecting the level of air contamination.

The main equation, following on Sanders and Sandler (2017), that relates the rejection of a car by the Smog-Check program with local air pollution is:

$$p_{t,m} = \beta \left[\sum_{k=0}^{K} r_{(t-k),m} \right] + \delta X_{t,m} + \phi_m + \xi_t + \varepsilon_{t,m}. \tag{1}$$

where $p_{t,c}$ denotes the concentration of pollutant p at date t on geographic-unit m. Our main focus is on CO since it is directly attributable to cars, however, I have data from several other pollutants, such as Particulate Matter and Nitrogen Dioxide. Our main independent variable is $\sum_{k=0}^{K} r_{(t-k),m}$, where $r_{(t-k),m}$ are rejections conducted at date $t-k$ the geographic area m. So, the main variable is the total number of reinspections conducted between t and

[5] This is what hypothetical owners of rejected cars would have to do in order to pass in a subsequent inspection.

K days before date t at a specific geographic region[6]. The parameter of interest is β which captures the effect of an additional reinspection between t and K days before date t in the geographic-unit m. $X_{t,m}$ corresponds to a vector of daily weather covariates for each monitor. ϕ_m and ξ_t are geographic-unit and day fixed-effects to account for everything that is common to all units in a specific date, and all things common to a specific unit in all dates respectively.

This empirical strategy exploits the fact that vehicle inspections and rejections are independent to local air pollution, conditional on certain covariates. As mentioned in the previous paragraph, I control for several confounding factors. Additionally, a key characteristic of the regulation is that the month in which people have to complete their certification is essentially random. However, for the period that our dataset covers, around 25% of inspections are conducted at a date that is later than the specified schedule. I employ time fixed-effects and control for all shocks that are common through geographic regions at a specific date. Weather covariates are key to avoid confounding factors since local air pollution responds directly to this variables. I need to highlight the fact that $r_{t,m}$ is a proxy for removing a high-emitting-car from the streets at time t in the geographic-unit m. If cars are not being removed (i.e. people are driving without their legal certification) and/or they are getting their inspections from places that they are not driving frequently, the measurement error will bias our estimates toward zero and I will not be able to find any statistically significant results[7]. Geographic-unit fixed-effects control for all characteristics of these geographic units that do not vary with time. Standard errors will be calculated by clustering at this level.

The main concern one has to be aware of when using this strategy is that $r_{t,m}$ cannot be correlated with $\varepsilon_{t,m}$. This could be problematic if there are other variables, not included in this equation, that explain $p_{t,m}$ and are correlated with $r_{t,m}$. For instance, one of the most relevant forms of endogeneity in this context could be Reverse Causality. One example of a hypothetic situation where this could be relevant is if the authorities put pressure on inspection plants specifically when pollution is high. In Chile, there are several episodes

[6] The main results are conducted by setting K to 1 or 7 days (i.e. from one day to one week) as mentioned at the footnote of each table.

[7] See proof in 7.2

of "Environmental (Pre) Emergencies" each year. It could be the case that the Chilean Environmental Office and/or the Transport and Telecommunication Office contact inspections plant administrators to ask them to be overly restrictive on the days surrounding these episodes. One way out of this kind of problems is that the panel structure of our data allows us to control for everything that is common to all plants at each relevant point in time. So, if this kind of pressure is put on all plants, these would be taken into account with day-specific fixed-effects. However, I will conduct several robustness exercises to confirm that I am identifying the causal effect of an additional rejection correctly. For example, a key and simple robustness exercise is to conduct a falsification or placebo-test in which I estimate the main specification, but using a lag of the dependent variable (i.e. try to explain past pollution with current rejections). If there are hidden trends that drive our results this would be reflected in the finding of significant effects in this kind of exercise. If there are no effects it means that there are not trends or that these trends are extremely short-lived. For instance, if one is worried that shocks that increase the mass of vehicles in the city are behind our results, then this kind of "trends" would be reflected in finding spurious correlations. The same logic stands if one thinks of specific times of the year in which the authorities are more concerned with the leniency of inspection plants standards. Additionally, I can estimate the main empirical specification, but excluding the critical winter months in terms of high levels of air pollution to discard this plausible sources of endogeneity (i.e. reverse causality as mentioned before).

One potential source of confounding factors is related to the size of the vehicular fleet that transits nearby air pollution monitors. It could be the case that all of our results can be explained by simply reducing the total number of cars circulating in the surrounding areas. Since I do not have data on transit or vehicular fleet for each geographic-unit at daily frequencies I deal with this concern by conducting an additional falsification exercise. If the main mechanism behind our results is to reduce the number of cars, I can estimate the same equation, but using rejections by causes other than Smog-Checks. I discuss these results at the end of section 5.1.

4.2 Station-Quality Heterogeneity

Even though there are several factors affecting whether a specific vehicle passes at the smog check, one key policy question is whether one can effectively classify stations according to their quality. In recent year, the state of California approved a new legislation that allows plants that are above certain quality-threshold to conduct inspections to high-polluting vehicles. For our purposes, it is interesting to understand if the effects that I am trying to estimate are affected by what I will label as “station-quality”. Before describing the specific detail of this metric, it is important to consider what quality means and how it is relevant in this context.

From my perspective what I can effectively observe is that the “quality” level of each station is the result of an economic equilibrium. On one side of this market, there are stations. There is some evidence that is consistent with inspection stations involved directly in cheating, and evidence that shows how plants change their levels of leniency to environmental standards to attract more customers (Oliva, 2015; Hubbard, 1998). On the other side, there are customers trying to get their vehicles approved. I have informal and anecdotical evidence that there is a wide number of common practices to pass the Smog-Check without taking measures to improve car emissions, such as forcefully accelerating the vehicle for about 15 minutes before the inspection, and/or washing the tailpipe with water and soap. What I expect from the “quality” metric is that it should be able to capture whether a specific plant is being involved (directly or indirectly) in this kind of practices, and this should also be reflected in the heterogeneity of the results along this dimension.

To translate the Similar Vehicle Failure Rate (SVFR), one of the STAR-quality metrics, I follow Sanders and Sandler (2017). It is important to remember that this metric is trying to capture by how much each station is deviating from their expected failure rates compared to what is seen in the Smog-Check system as a whole. First, I need to take into account that car-owners can conduct several inspections each year. I will denote vehicle type, that is make-model, with k. I will index by $n \in \{1, \dots, N\}$ the number of reinspections for each car within a year. For the following calculations, I only consider the first inspection within each cycle for each car, and each cycle is considered to last one year. For example, if a car fails its first inspection and passes the second $N = 2$. The group of all cars conducting their n-th inspection at a station s within a year is denoted by $\zeta_s(y, n)$, where y stands for the year. I will denote by l the time level of aggregation of the data for this exercise[8]. With all these ingredients, the expected rejection rate of a station is:

$$\Xi(\zeta_s(y,n), y, n) = \frac{1}{\zeta_s(y,n)} \sum_{i \in \zeta_s(y,n)} \mathcal{P}(fail_{i,y,n} = 1 | k_i, l, X_{iky}). \tag{2}$$

That is to say, that the expected rejection rate at a station s is the average of all the individual predicted probabilities of failure for each car, after considering all the available observable characteristics. Notice also that I am only considering the first inspection for every car in the construction of the predicted values, since there could be more noise in this values on the subsequent inspections. In this context, $fail_{i,y,n}$ is an indicator function that equals to one when vehicle $i \in \zeta_s(y, n)$ fails its n-th inspection in year y. X_{iky} is a vector of car characteristics, such as age, and make-model. I express $\mathcal{P}(fail_{i,y,n} = 1 | k_i, l, X_{iky})$ in the following form:

$$\mathcal{P}(fail_{i,y,1} = 1) = \alpha_k + \eta_l + \beta X_{iky} + \eta_k. \tag{3}$$

where, α_k controls for all common things of models that do not change in time, and η_l controls for things common to all models during period l. It is important to highlight the fact that the previous linear probability model only calculates the probability of failure of different models using the initial inspection data (i.e. $n = 1$).

The SVFR for a specific station s during the period of time l can be written in the following form:

$$SVFR_{sl} = \frac{\frac{1}{\zeta_s(y,1)} \sum_{i \in \zeta_s(y,1)} fail_{i,y,1}}{\Xi(\zeta_s(y,1), y, 1)}. \tag{4}$$

The interpretation of this metric is how much more(less) strict is a specific station, when

[8] The main results of this exercise were done by setting l to one month.

compared to the average station within the system, in rejecting cars of similar observable characteristics. The threshold specified by the STAR program is 0.75. This means that a station rejecting 25% less cars of similar characteristics than what it is to be expected will be qualified as a low-quality station [9]. With this information, I aggregate the total number of rejections conducted at "high" ("low")-quality stations for each geographic-unit. Estimating this specification will shed information on how does the effects of conducting the inspection at a station with different levels of "quality" compares in terms of local air pollution. In concrete terms the main equation to be estimated at this section is the following:

$$p_{t,m} = \beta^H \left[\sum_{k=0}^{K} \sum_{s^H=1}^{S_m^H} r^H_{(t-k),s,m} \right] + \beta^L \left[\sum_{k=0}^{K} \sum_{s^L=1}^{S_m^L} r^L_{(t-k),s,m} \right] + \delta X_{t,m} + \phi_m + \xi_t + \varepsilon_{t,m}. \tag{5}$$

where β^H captures the effect of an additional reinspection at a high-quality station between t and K days before date t in the geographic-unit m and β^L captures the analogous, but for low-quality stations. Notice that S_m^H (and S_m^L) depends on m because the total number of high(low) quality stations at each geographic-unit might differ. Intuitively, one can expect the coefficient β^H to be negative and statistically significant. This potential heterogeneity is based on the fact that a car being rejected at a high-quality station is probably a car that should be taken out from circulation based on its level of emissions. However, a car rejected in a low-quality station might not be a car that should be taken out from circulation, or at least the metric is capturing that this stations cannot effectively recognize which cars are more(less) polluting. In this sense, low-quality stations are not being effective at changing the composition of the car fleet in terms of emissions as it is expected and this, in turn, should be captured as β^L being close to zero and noisily estimated. Notice that the identification strategy for estimating this equation has the same assumptions as our main specification.

[9] I did the same exercise, but using rejections for causes other than Smog-Checks and the correlation between these two metrics is about 40%.

5 Results

5.1 Impact of Rejections

All coefficients are rescaled to represent the impact of one additional standard deviation of rejections on the concentration of pollutants, also in terms of standard deviations. This means that when I change rejections by one standard deviation, β represents how much of a standard deviation the specific pollutant is affected. This standardization is done because there are different measurement units across pollutants and for the ease of the interpretation and comparison across pollutants I preferred to standardized them so they are all comparable. Additionally, while I am estimating the impact of an additional rejection by the Smog-Check system, this coefficients would be very small and this would make them hard to compare between specifications and across pollutants[10].

On table 4 the results from different exercises can be found, all of them with the average daily concentration of $[CO]$ at each monitor as the left-hand-side variable. In the first column, results of the most simple regression of $[CO]$[11] on the total number of rejections without adding any controls is estimated. Notice that these two variables correlate positively, possibly this is due to the fact that both series reflect the growth of the vehicular fleet across the time-span of our panel. However, as one can see in column number 2, as soon as I control for geographic-unit fixed-effects the sign changes. This means that, after controlling for everything constant in time for each geographic-unit, one can see that the sign of the coefficients turns to negative. Notice that in this specification I am only explaining the residual variation that reinspections generate on $[CO]$ within each geographic-unit. When one looks at column 3, after adding weather controls the coefficient increases and I gain statistical significance. This is crucial since it is known that air pollution responds directly to several weather conditions. For example, an increase in wind speed will move pollution usually to the coast, therefore, reducing measured air pollution by the monitors in our sample. With low temper-

[10]Remember that different pollutants are reported with different measurement units.
[11]Recall that I am using daily averages for each pollutant at the Geographic-Unit level.

ature, the cloud of gas contracts and therefore augments its relative concentrations. Also, I know that local air pollution is directly affected by humidity and the direction of the wind. All of this facts explain why after controlling for weather covariates one can see an increase the statistical significance. On the last column, I add day fixed-effects that capture all the variation that is common across plants for a specific day. For example, if there are forest fires and/or if police increase their monitoring, and so on, this would be captured by these fixed-effects. Notice that the magnitude of the coefficient, these remains fairly unchanged after adding geographic-unit fixed effects, but I am only improving the statistical significance of it. The coefficient in the last column, which is our preferred specification, means that an increase in one standard deviation in the number of rejections within a specific geographic unit decreases the concentrations of Carbon Monoxide by around 4.8% of a standard deviation.

In table 5 I show the results of the same exercises as described above, but with $[NO_2]$ as our main left-hand-side variable. The results are more or less unchanged in terms of the evolution of the sign and magnitude of the coefficients. However, I am not able to find any statistically significant effects for this pollutant. I do not find this to be problematic since this is the most demanding specification, because I am letting the effects to accumulate for only one day. In addition, it is important to highlight that $[CO]$ is our main dependent variable since there is evidence that links its pollution directly to car emissions, while for $[NO_2]$ and the other studied pollutants this is not the case. The coefficient in column 4, means that an increase of a one standard deviation in the number of rejections for a specific monitor translates into a decrease of about 3.4% of a standard deviation in the concentration of Nitrous Dioxide. From now on, all estimations are conducted with our full empirical specification except when explicitly mentioned.

Table 6 shows the results of our preferred specification for all pollutants available in our database. The first two columns show the impact of rejections on particulate matter of different sizes. While car emissions are certainly one of the main contributors to this contaminants by emitting Volatile Organic Compounds, there are many other sources, such as smoke and resuspension of organic matter. However, it is highly reassuring that I find statistically significant and economically relevant estimates for these two pollutants ($[PM_{10}]$, and $[PM]_{2.5}$) since they are one of the most problematic sources of smog (Enviromental Protection Agency, U.S., 2006). While for $PM_{2.5}$, the smallest measured size of particulate matter our coefficient is statistically significant and about 5% of a standard deviation for PM_{10} is significant at 10% and about less than 1% of a standard deviation. This reflects the fact that probably the production function for $PM_{2.5}$ has cars emissions as one of the main sources and that is not the case for PM_{10}. Also, it is important to notice that PM_{10} might be subject to potential sources of confounding variables that I am not able to control for such as a barbecues near the monitors or increased vehicular flow. This is particularly important with these two pollutants due to the fact that the size of this particulate matter is related to ground particles that are resuspended and may be harder to link to car emissions. Most notably, the coefficient for O_3 is around 1% of a standard deviation, positive and not even nearly statistically significant. This might be due to the fact that the production function for O_3 is highly complex and its formation probably takes more time to form since it depends on the concentration of Nitric Oxides $[NO_x]$, the presence of sunlight, etc.

When I move to a less demanding specification in which I allow the effects of rejections to take up to one week to translate into local air pollution. The results presented in table 7 are for this specification, but the details of timing seem not to be extremely important (from 7 to 5 or 8 days the results practically remain unchanged). Now one can see that the magnitudes for all pollutants are a little bit higher, except for $[O_3]$. Notice that I now find statistically significant results for PM_{10} and for NO_2 the effects remain marginally significant which is reassuring. For $[CO]$ the results are virtually unchanged.

In Table 8, I conduct the same exercise as described in the previous paragraph, but the level of Geographical Aggregation correspond to the Metropolitan Region. This is done to understand to what extent our results depend on the geographical level of aggregation. Remember that for the estimation of 1 I was using a 5-kilometer radius to define our geographic-units. However, this is arbitrary and, here I run the same specification, but collapsing our panel dataset into a times series in order to increase the radius of the geographic-units. It is important to highlight that, when I move to this scenario there are many things changing, for instance, at this level of aggregation I lose the ability to control for day and geographic-unit fixed effects. One of the upsides of this specification is that at this level I do not need to make assumptions about whether car-owners that conduct their inspections at determined stations do or do not circulate in the proximity of those areas. The magnitude of the coefficients reported here is a little lower for every pollutant[12], which could be a result of the movement of the smog cloud, the geographical heterogeneity within the city, and that between geographical-units variation is lost at this level of aggregation[13]. Calculating the impacts of rejecting an additional car on $[CO]$, when I move from a local context to a more aggregated one our magnitudes drop by about 6 times. This is consistent with the fact that air quality monitors capture only local pollution and that aggregating could hinder the relevant local variation.

It is very important to remember, as said in section 2, that our results are driven by high-emitting or high-polluting cars. To be more emphatic about this, I took the gas readings reported in our database and compared the average $[CO]$ emissions of cars that passed and from those that were rejected at the Smog-Check. An average car that passed the inspection has a $[CO]$ reading of 0.155 (%v/v), while a car that was rejected has a $[CO]$ reading of 1.45 (%v/v). This means that, on average, rejected cars emit a little more than 9 times the $[CO]$ of what non-rejected cars emit. Notice that the two distributions do not share anything in common because rejected cars are always above the maximum pollution levels of the pollution distribution of non-rejected cars. This is true both at 24(km/h) and at 50(km/h). Additionally, the pollution "production function" does not take as an input average emissions, but the sum of emissions. This is crucial because the aforementioned differences are exacerbated by this fact. If I want to convert this measurements unit to (*ppb*) I have to

[12] A standard deviation for all pollutants is more or less the same than the magnitudes reported in 2. However, a standard deviation in the number of rejections is about 74 cars, whereas before was about 25 cars.

[13] At this level of aggregation I have a time-series dataset and I can not control for geographic-unit fixed-effects and/or day fixed-effects.

multiply by a factor of 10000000[14]. So again, these differences become wider when I look at the appropriate measurement units for this context. However, it is highly complex to take these measurements and extrapolate them to local air pollution levels. The main issue relies on the fact that all these measurement units are concentration units and are calculated over different volumes[15]. However, this analysis help us to understand that what it is being removed are extremely high-polluting cars and that in average cars that are rejected are not rejected by small margins.

To understand the economic magnitude and significance of this estimations I will follow Bharadwaj et al. (2017), where they estimate that exposing children while in utero to an increase in the $[CO]$ concentration of one standard deviation[16] decreases standardized test scores of affected children in the long run by about 6% of a standard deviation [17]. The main mechanism they explore is that increased local air pollution affects both birthweight and gestational age negatively. These coefficients are obtained for the same city which I analyze in this work, thus their coefficients are easier to extrapolate in this context. However, it is worth noticing that they are evaluating the long-run effects of the exposure to air pollution[18]. Additionally, their sample covers from 1990 to 2005. Applying this results one would be looking at an increase in standardized tests of about 3% of a standard deviation if exposed to an environment in which the number of rejections was increased by one standard deviation (i.e. about 25 cars for an average geographic-unit).

Another way to put our results into perspective is to consider alternative policies implemented and their effects. For example, Troncoso et al. (2012) find that during pre-emergencies in Santiago (the same city as in this work), $[PM_{10}]$ decreases by about 12% of its mean. Translating our results to make them comparable to this context, I find that one

[14] from %v/v to (ppm) I have to multiply by 10000. From (ppm) to (ppb) I have to multiply by 1000.

[15] To do this, I would have to know the relevant air volume of each geographic-unit to calculate how these concentrations alter the concentration for each pollutant.

[16] For the period they cover the standard deviation for $[CO]$ in *ppb* was 0.95, whereas in our sample is about 0.51.

[17] The standardized test scores they use (SIMCE) is standardized to have mean 200 and a standard deviation of 50.

[18] Their main outcomes are measured around 8-9 years after being exposed to air pollution.

would have to increase rejections by about 6 standard deviations to reduce $[PM_{10}]$ by the same amount. Repeating the same exercise, but with $[PM_{2.5}]$ one would have to increase the number of rejections 6 times (i.e. almost the same as with $[PM_{10}]$). This helps to put into perspective our results and their magnitude. Comparing this results, I can conclude that our estimates suggest a small, but significant impact on local air pollution.

In a context more similar to ours in terms of levels of air pollution, Viard and Fu (2015) finds that one-digit driving restrictions decrease PM_{10} levels by about 21%. Our estimates suggest that a one standard deviation increase in the number of rejections translates into a reduction of about 3% of PM_{10} with respect to its mean. The big difference in magnitude is due to the fact that they are removing about 11% of the vehicular fleet from circulation, but they are doing so in an indiscriminate fashion. The more comparable results in terms of context are those of Sanders and Sandler (2017). These authors find that reinspecting a 1000-high-emitting cars decreases $[CO]$ and $[NO_2]$ by about 7% of a standard deviation. However, I am considering much smaller boundaries for our geographic units (i.e. 5-kilometer radius versus 15-kilometer radius) and this could explain the difference because air pollution concentration might probably wash-out at larger distances as I discussed previously.

To end this section, I want to refer to our main falsification exercises. In table 9, I present the coefficients associated with the estimation of our main specification, but now with the left-hand-side variable lagged. In other words, I am estimating past pollution with the current number of reinspections. Needless to say that it would be alarming if I find any statistically significant result and/or big magnitudes. The latter, because it is not physically plausible that rejections affects past levels of pollution. Nonetheless, if there are hidden trends not captured by our fixed-effects or by weather covariates that affect the way in which rejections and local air pollution relate this would be evident in this specification. From the evidence displayed by this coefficients, one can say that these trends are not present, or that if there is any trend it is extremely short lived and it does not persist for a long time (i.e. it has to be present for less than one week). Even though this is not definitive evidence it is a small guarantee for the absence of reverse causality in the most literal way. The magnitude of the coefficients is severely reduced from the previous estimations for every pollutant, except for O_3 and I do not find any statistically meaningful results which are coherent with our main hypothesis.

Table 10 estimates equation 1, but excluding the critical winter months. For this estimation I dropped from the estimation the months in which the average pollution (for all pollutants) was higher: May, June, and July[19]. For example, for $[CO]$ the mean during these three months is almost twice the sample mean. This creates additional concerns regarding identification since these months are particularly sensitive to the authorities regarding air pollution and acute respiratory diseases. The magnitudes of the results are a little bit higher for the case of PM_{10}. For $[CO]$, and $[PM_{2.5}]$ the magnitudes are somewhat smaller. For the case of $[NO_2]$, I find a coefficient near zero and with the wrong sign. Notice that up to this point our results for $[NO_2]$ do not seem to be very robust. For $[O_3]$ the story is more or less the same as with $[NO_2]$, but with the opposite sign. In terms of statistical significance, I lose the ability to find significant results for $[PM_{2.5}]$. This could be explained by the fact that I am dropping around 25% of our dataset. The conclusion of this exercise is that the main results are not primarily driven by the different nature of winter months regarding air pollution and rejections.

In table 12, I try to discard the fact that our outcomes could be the result of reducing the number of cars circulating near the air quality monitor at each geographic-unit (i.e. that reducing the total number of circulating cars is the mechanism by which rejections affect pollution). For this reason, I estimate the main empirical equation, but using as the main explanatory variable the total number of rejections by causes other than Smog-Checks (i.e. malfunctioning lights, tire problems, etc. as described in section 3). Here, one can see that all columns have a negative sign which is to be expected since removing cars from the streets mechanically lowers environmental pollution. However, all coefficients are noisily estimated and their magnitude is about half or less than the main results. This is reassuring since I can state that simply removing cars from the street is not the main explanation for our results

[19] Recall from figure 1 that all pollutants, but O_3 reached its peak during the winter.

and that one need to detect and remove high-emitting cars to significantly reduce local air pollution. Additionally, it is possible that there exists a positive correlation between being rejected from, say having a bad tail light and having emissions above the norm since both cases can be interpreted as signs of bad maintenance.

5.2 Station-Quality Heterogeneity

In this section, I study the potential heterogeneity behind our main results regarding station-quality by estimating 5. As I discussed in section 4.2 I expect the effects of additional rejections conducted on low-quality stations to be noisily estimated and of a smaller magnitude than those conducted at high-quality stations. It is important to remember that in this context "quality" is an equilibrium result and even though I can hypothesize about the reasons behind this, I am not able to unravel the underlying mechanisms behind this. As discussed in section 4.2, the quality of plants is calculated based on one-month periods. In table 13, I estimated the main equation, but separating those rejections conducted at high(low)-quality stations according to the STAR program. I repeated this exercise for all the pollutants available in our dataset. In line with our main hypothesis, I do not find any statistically significant result for rejections certified at low-quality stations. Additionally, the magnitude of this coefficients is never above 1% of a standard deviation and even has the opposite sign as one would have expected. However, when I refer to the coefficients of high-quality plants, one can see that the sign is agreement with our main results of the previous section and the magnitudes are always higher. The latter is consistent with the fact that I was averaging among qualities and now I can divide the contribution of quality separately. Even more impressive is the fact that I am able to find a negative and statistically significant coefficient for $[O_3]$. This probably is reflecting the fact that $[O_3]$ formation depends on $[NO_2]$ among other things and that now I can find statistically significant effects for it, but I do not have a solid explanation to support this specific result.

In table 14, I repeat the same exercise but allowing for a one-week period so the impacts can take place. Again, the magnitude of the results remains fairly unchanged. However, a little of significance on $[NO_2]$, and $[O_3]$ is lost. However, they are still marginally significant. This reinforces our prior that the contribution of car-pollution to the production function for $[O_3]$ seems to empirically depend on the concentration of $[NO_2]$. Once more, for simplicity, I only show the results for this specification, but the specific details of the lag-period seem to have little implications.

One additional conclusion of the two previous exercises is related to the falsification exercise in which I use the total number of rejections by causes different than Smog-Checks to explain local air pollution. Since I concluded that it is not enough to simply remove cars from the street to reduce local air pollution and that one needs to remove high-emitting cars, the results of this section can be interpreted in this fashion. The quality-measure I am using is intended to capture specifically this effect: how good are plants in recognizing, in average, which cars should be removed because of their elevated tailpipe emissions. The evidence that low-quality plants do not help in reducing local air pollution reinforces the fact that removing cars from the street is not the main driver of our results. In fact, only by removing cars because of their Smog-Checks can explain the main results. Additionally, these results are strengthened when one looks at Smog-Checks rejections from stations which are good at detecting high-emitting cars.

Even though the STAR program grades stations as high-quality if they are above the 0.75 SVFR threshold, on table 15 I present the same exercise as described in the previous paragraph, but changing the quality-threshold. Notice that the results remain fairly unchanged in terms of their magnitude and statistical significance for both β^H and β^L. I am not able to keep moving the threshold to a more "relaxed" scenario (i.e. a lower SVFR-threshold), because I am left with very few observations and start estimating this equation with several zeros in it, and if I drop the zeros I am left with a small number of observations so the exercise would not make much sense. When I move to a stricter scenario (i.e. I move to a higher SVFR-threshold) the results remain more or less the same in terms of magnitudes and statistical significance. This means that even though this metric is capturing the heterogeneity in station quality the specific level of the threshold does not seem to be of much relevance. What seems to be crucial is to exclude from the estimations the rejections coming from plants situated at the lower end along the quality distribution.

6 Conclusions

In this work, I estimate the effects of removing an additional high-emitting car by the Smog-Checks system in terms of local air pollution. I used data from the air quality monitoring network (SINCA) and data from the Smog-Checks system maintained by the Office of Transport and Telecommunications. I constructed a panel data from 2008-2016 associating each monitor with the smog-checks stations situated nearby. The main results indicate that removing an additional standard deviation of cars due to their emissions decreases $[CO]$, $[NO_2]$, $[PM_{10}]$, and $[PM_{2.5}]$ by 5,1%, 6%, 1,4%, and 5,8% of a standard deviation respectively. As discussed earlier, this is about 16.7% of the decreased accomplished by establishing and environmental pre-emergency. Not only these results are statistically significant, they also seem to be economically significant according to the related literature of the economic costs of air pollution. These results remain significant after a series of robustness checks and are subject to a couple of falsification exercises.

Additionally, I explored the potential heterogeneity behind our results due to the differences in station quality. By using a metric derived from California's STAR program, which compares the expected rejection rate for each station with the realized rejection rate, after controlling for all the available car characteristics, I find that low-quality stations have virtually no impact in terms of air pollution, while high-quality station seems to drive the main results. This is relevant in terms of public policy since the regulator could use this metric to increase monitoring in low-quality stations and possibly enhance welfare through the reduction of local air pollution.

Finally, based on this work, I can conclude that the Smog-Check system is fulfilling its main objective. However, there is plenty of room for improvement. One way the regulator can further reduce local air pollution through this system is by adopting monitoring tactics based on the quality-metrics used in this work. Furthermore, the regulator could use this dataset to notify owners who are not complying with their inspection schedule[20].

[20]i.e. based on this dataset the regulator can send notifications to car-owners in case they are behind their inspection schedules.

References

Arceo, E., Hanna, R., and Oliva, P. (2015). Does the effect of pollution on infant mortality differ between developing and developed countries? evidence form mexico. *The Economic Journal*, 126((March)):257–180.

Barahona, N., Gallego, F., and Montero, J. (2017). Vintage-specific driving restrictions.

Bharadwaj, P., Gibson, M., Zivin, G., and Neilson, C. (2017). Gray matters: Fetal pollution exposure and human capital formation. *Journal of the Association of Environmental and Resource Economists*, 4(2):505–542.

Bureau of Automotive Repairs (2012). Star program questions and answers. Technical report, Department of Consumer Affairs: Bureau of Automotive Repair.

Chang, T., Graff-Zivin, J., Gross, T., and Neidell, M. (2016a). The effect of pollution on worker productivity: Evidence from call-center workers in china. Working Paper 22328.

Chang, T., Graff-Zivin, J., Gross, T., and Neidell, M. (2016b). Particulate pollution and the productivity of pear packers. *American Economic Journal: Economic Policy*, 8(3):141–169.

Chay, K. and Greenstone, M. (2003). The impact of air pollution on infant mortality: Evidence from geographic varition in pollution shocks induced by a recession. *The Quarterly Journal of Economics*, 118(3):1121–1167.

Currie, J., Hanushek, E., Kahn, M., Neidell, M., and Rivkin, S. (2009). Does pollution increase school absences? *The Review of Economics and Statistics*, 91(4):682–694.

Currie, J. and Neidell, M. (2005). Air pollution and infant health: What can we learn from california's recent experience? *The Quarterly Journal of Economics*, 120(3):1003–1030.

Davis, L. (2008). The effect of driving restrictions on air quality in mexico city. *Journal of Political Economy*, 116(1):38–81.

Enviromental Protection Agency, U.S. (2006). Air quality criteria for ozone and related photochemical oxidants. Technical report, United States Enviromental Protection Agency.

Ernst, A. and Zibrak, J. (1998). Carbon monoxide poisoning. *New England Journal of Medicine*, 339(22):1603–1608.

Gallego, F., Montero, J., and Salas, C. (2013). The effect of transport policies on car use: Evidence from latin american cities. *Journal of Public Economics*, 107:47–62.

Glazer, A., Klein, D., and Lave, C. (1993). Clean for a day: Troubles with california's smog check. University of California Transportation Center.

Glazer, A., Klein, D., and Lave, C. (1995). Clean on paper, dirty on the road: Troubles with california's smog check. *Journal of Transport Economics and Policy*, 29(1):85–92.

Graff-Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. *American Economic Review*, 102(7):3652–3673.

Hanna, R. and Oliva, P. (2015). The effect of pollution on labor supply: Evidence from a natural experiment in mexico. *Journal of Public Economics*, 122:68–79.

Harrington, W., McConnell, V., and Ando, A. (2000). Are vehicle emission inspection programs living up to expectations? *Transportation Research Part D*, 5(3):153–172.

Hubbard, T. (1998). An empirical examination of moral hazard in the vehicle inspection market. *The RAND Journal of Economics*, 29(2):406–426.

Kampa, M. and Castanas, E. (2008). Human health effects of air pollution. *Environmental Pollution*, 151:362–367.

Knittel, C., Miller, D., and Sanders, N. (2011). Caution, drivers! children present: Traffic, pollution, and infant health. NBER Working Paper 17222.

Ministerio de Transportes y Telecomunicaciones (2015). Revisiones técnicas tipo b: Manual de procedimientos e interpretación de resultados. Technical report, Ministerio de Transportes y Telecomunicaciones.

Oliva, P. (2015). Environmental regulations and corruption: Automobile emissions in mexico city. *Journal of Political Economy*, 123(3):686–724.

O'Ryan, R. and Larraguibel, L. (2000). Contaminación del aire en santiago: ¿qué es, qué se ha hecho, qué falta? *Perspectivas (Departamento de Ingeniería Industrial, Universidad de Chile).*, 4(1):153–191.

Sanders, N. and Sandler, R. (2017). Technology and the effectiveness of regulatory programs over time: Vehicle emissions and smog checks with a changing fleet. NBER Working Paper 23966.

Seaton, A., MacNee, W., Donaldson, K., and Godden, D. (1995). Particulate air pollution and acute health effects. *Lancet*, 345(8943):176–178.

State of California (2014). Laws and regulations pertaining to automotive repair dealers, smog check stations and technicians, official lamp and brake adjusting stations laws and regulations. Technical report, California Department of Consumer Affairs: Bureau of Automotive Repair.

Subdirección Técnica, INE (2016). Informe anual del medio ambiente. Technical report, Instituto Nacional de Estadísticas.

Troncoso, R., de Grange, L., and Cifuentes, L. (2012). Effects of enviromental alerts and pre-emergencies on pollutant concentrations in santiago, chile. *Atmospheric Environment*, 61:550–557.

Viard, B. and Fu, S. (2015). The effect of beijing's driving restrictions on pollution and economic activity. *Journal of Public Economics*, 125:98–115.

Wenzel, T., Singer, B., Slott, R., and Stedman, D. (2004). Short-term emissions deterioration in the california and phoenix i/m programs. *Transportation Research Part D*, 9(2):107–124.

Wolff, H. (2013). Keep your clunker in the suburb: Low-emission zones and adoption of green vehicles. *The Economic Journal*, 124:F481–F512.

7 Appendix

7.1 Appendix 1: Tables and Figures

Table 1: Descriptive Statistics of Daily Inspections and Rejections Conducted at each Geographic-Unit for Each Year.

	Inspections		Total Rejections		S.C. Rejections	
Year	Mean	S.D.	Mean	S.D.	Mean	S.D.
2008	310.40	244.68	53.96	52.67	18.26	18.29
2009	336.67	255.24	159.74	150.82	30.00	30.35
2010	395.02	317.01	176.72	181.18	29.89	33.93
2011	412.47	342.40	182.95	190.69	26.39	30.60
2012	459.15	388.62	167.15	166.16	23.21	23.46
2013	477.31	392.05	148.50	145.91	19.59	20.54
2014	469.85	370.61	140.62	134.62	16.55	16.27
2015	456.78	365.24	125.44	122.11	16.41	13.88
2016	428.22	332.25	139.59	102.35	20.74	14.46
Total	418.20	342.49	143.69	147.78	22.28	23.87

Inspections consists of all the Smog-Checks conducted daily by each geographic unit. Total Rejections are inspections that ended in a rejection by all causes. S.C. Rejections are rejections exclusively from the Smog-Checks.

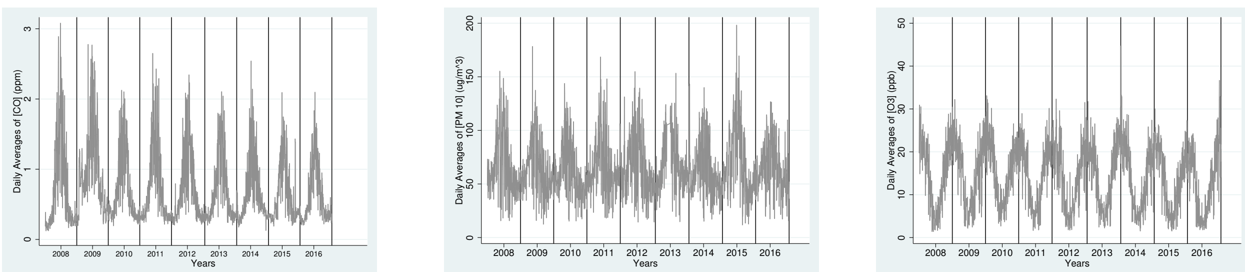
Table 2: Number of Observations (Inspections Conducted) by year in the Metropolitan Region.

Year	Observations
2008	1,533,376
2009	1,660,754
2010	1,839,362
2011	1,855,867
2012	2,190,446
2013	2,401,996
2014	2,432,208
2015	2,577,507
2016	2,809,005
Total	19,300,521

Table 3: Mean of Daily Concentrations of Several Air Pollutants by Geographic Region for Each Year.

	$[CO]$		$[NO_2]$		$[PM_{2.5}]$		$[PM_{10}]$		$[O_3]$	
Year	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**	**Mean**	**S.D.**
2008	738.73	644.23	23.74	11.30	31.49	15.35	67.09	28.10	15.95	7.89
2009	962.65	571.76	23.81	9.83	27.95	12.58	66.37	125.72	6.07	7.46
2010	760.18	500.45	22.11	9.31	25.97	12.30	63.42	125.08	6.16	6.95
2011	742.91	537.37	22.04	10.01	26.55	12.57	67.67	26.88	15.04	6.94
2012	702.95	464.51	21.97	9.93	25.00	12.21	66.18	128.53	4.46	7.17
2013	703.85	472.87	23.01	9.76	25.86	11.06	69.62	125.83	5.51	7.23
2014	644.99	406.34	20.23	8.98	27.91	17.01	62.72	24.42	14.83	7.48
2015	682.74	451.10	22.79	12.44	30.22	19.13	70.38	33.64	14.07	6.96
2016	616.97	400.32	23.10	9.35	29.74	18.20	65.25	123.06	2.92	7.14
Total	728.40	508.04	22.43	10.18	27.76	14.89	66.49	127.05	4.81	7.32

All pollutants are in ppb, except $PM2.5$ and $PM10$ which are in $\mu g/m^3$. PM_{10} and $PM_{2.5}$ corresponds to Particulate Matter of $10\mu M$ and $2.5\mu M$ respectively. NO_2 stands for Nitrous Dioxide.

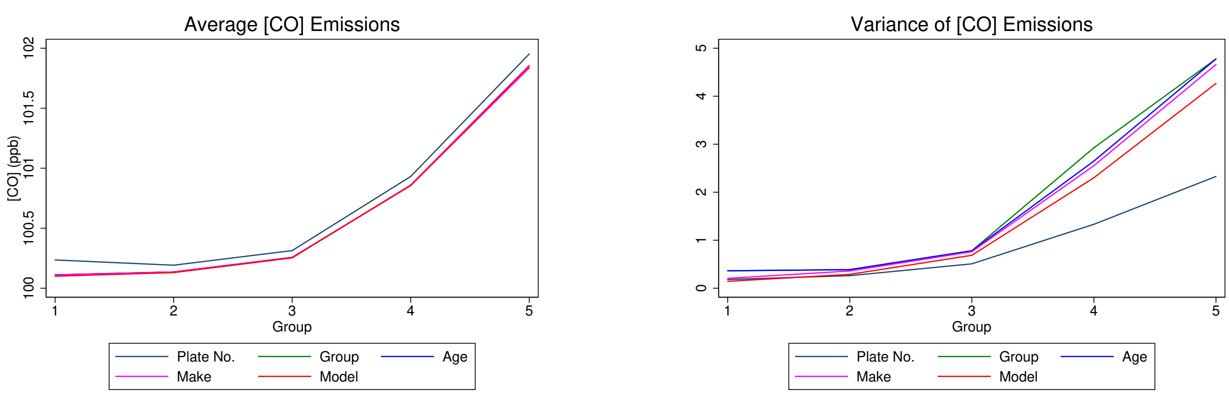


(a) Daily Average of $[CO]$ (ppm).

(b) Daily Average of $[PM_{10}]$ ($\mu g/m^3$).

(c) Daily Average of $[O_3]$ (ppb).

Figure 1: Daily Concentration Averages for Different Pollutants across our Sample Period.



(a) Average Carbon Monoxide Emissions by car (%v/v) using different controls.

(b) Variance of Carbon Monoxide Emissions by car using different controls.

Figure 2: Descriptive Statistics using data from 2008. Groups are defined as tranches of 4 years of age.

Table 4: Impact of Rejections on Daily Carbon Monoxide Concentrations

	(1) CO	(2) CO	(3) CO	(4) CO
Rejections	0.012	-0.044^{+}	-0.046*	-0.048**
	(0.0124)	(0.0202)	(0.0149)	(0.0150)
N	16484	16484	15479	15479
R^2	0.001	0.033	0.315	0.317
adj. R^2	0.001	0.032	0.315	0.316
County FE	NO	YES	YES	YES
Weather Controls	NO	NO	YES	YES
Day FE	NO	NO	NO	YES

Regressions of Daily [CO] on Rejections from the Day Before, controlling for Weather Covariates, and Day and Geographic-Unit Fixed Effects.

Clustered Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Impact of Rejections on Daily Nitrous Dioxide Concentrations

	(1) NO_2	(2) NO_2	(3) NO_2	(4) NO_2
Rejections	0.0610	-0.0124	-0.0239	-0.0343
	(0.0410)	(0.0183)	(0.0183)	(0.0210)
N	15326	15326	14329	14329
R^2	0.035	0.164	0.383	0.388
adj. R^2	0.035	0.163	0.382	0.387
County FE	NO	YES	YES	YES
Weather Controls	NO	NO	YES	YES
Day FE	NO	NO	NO	YES

Regressions of Daily [NO_2] on Rejections from the Day Before, controlling for Weather Covariates, and Day, and Geographic-Unit Fixed Effects.

Clustered Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Impact of Rejections on Daily Concentrations of Different Air Pollutants

	(1) $PM2.5$	(2) $PM10$	(3) NO_2	(4) CO	(5) O_3
Rejections	-0.0523**	-0.0074*	-0.0343	-0.048**	0.012
	(0.0161)	(0.003)	(0.0210)	(0.0150)	(0.0166)
N	15323	15846	14329	15479	15514
R^2	0.216	0.248	0.388	0.317	0.624
adj. R^2	0.215	0.247	0.387	0.316	0.624

Regressions of Daily [*Pollutants*] on Rejections from the Day Before, controlling for Weather Covariates, and Day and Geographic-Unit Fixed Effects.

Clustered Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Impact of Lagged Rejections on Different Air Pollutants

	(1) $PM2.5$	(2) $PM10$	(3) NO_2	(4) CO	(5) O_3
Rejections	-0.058**	-0.014***	-0.060*	-0.051***	0.027
	(0.0148)	(0.0019)	(0.0236)	(0.0108)	(0.2120)
N	15288	15807	14299	15440	15475
R^2	0.270	0.305	0.467	0.385	0.648
adj. R^2	0.269	0.304	0.466	0.385	0.648

Regressions of Daily [*Pollutants*] on Rejections (one-week lag), controlling for Weather Covariates, and Day and Geographic-Unit Fixed Effects.

Clustered Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Aggregated Impact of Lagged Reinspections on Different Air Pollutants

	(1) $PM2.5$	(2) $PM10$	(3) NO_2	(4) CO	(5) O_3
Rejections	-0.0379*	-0.0641**	-0.0625**	-0.0231+	0.00669
	(0.0149)	(0.0156)	(0.0132)	(0.0126)	(0.00952)
N	2704	2704	2704	2704	2704
R^2	0.381	0.365	0.554	0.581	0.770
adj. R^2	0.380	0.363	0.553	0.580	0.769

Regressions of Daily [*Pollutants*] on Rejections (one-week lag), Controlling for Weather Covariates, and Polynomial Time Trends. The Geographical Aggregation Level is Metropolitan Region.

Newey-West Standard Errors in Parentheses (one-period lag).

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Falsification Exercise: Impact of Rejections on the Lag of Different Air Pollutants

	(1) $PM2.5$	(2) $PM10$	(3) NO_2	(4) CO	(5) O_3
Rejections	-0.0055	-0.0001	0.0030	-0.0130	0.0027
	(0.0114)	(0.0024)	(0.0118)	(0.0120)	(0.0166)
N	15254	15774	14270	15418	15438
R^2	0.829	0.820	0.846	0.729	0.903
adj. R^2	0.792	0.782	0.810	0.671	0.882

Regressions of Daily [*Pollutants*] (one-week lag) on Current Rejections, Controlling for Weather Covariates, and Day and Geographic-Unit Fixed Effects.

Clustered Standard errors in parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Impact of Lagged Rejections on Different Air Pollutants Excluding Critical Winter Months

	(1) $PM2.5$	(2) $PM10$	(3) NO_2	(4) CO	(5) O_3
Rejections	-0.0318^{+}	-0.0193	0.00828	-0.0295**	-0.0152
	(0.0154)	(0.0156)	(0.0238)	(0.00780)	(0.0248)
N	15678	16326	14422	16043	16067
R^2	0.806	0.836	0.851	0.611	0.878
adj. R^2	0.777	0.813	0.827	0.555	0.861

Regressions of Different Air Pollutants on the Number of Rejections (one-week lag), Controlling for Weather Covariates with Day, and Geographic-Unit Fixed Effects. Critical Winter Months are Excluded from Estimation.

Clustered Standard Errors in Parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 11: Aggregated Impact of Lagged Rejections on Different Air Pollutants Excluding Critical Months.

	(1) $PM2.5$	(2) $PM10$	(3) NO_2	(4) CO	(5) O_3
Rejections	-0.0407*	-0.0826**	-0.0696**	-0.0295*	0.0177
	(0.0172)	(0.0177)	(0.0154)	(0.0146)	(0.0125)
N	2703	2703	2697	2703	2702
R^2	0.362	0.356	0.533	0.553	0.746
adj. R^2	0.360	0.354	0.532	0.552	0.745

Regressions of Daily [*Pollutants*] on Rejections (one-week lag), Controlling for Weather Covariates, and Polynomial Time Trends. Critical Winter Months are Excluded from Estimation. Geographical Aggregation Level is Metropolitan Region.

Newey-West Standard Errors in Parentheses (one-period lag).

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Falsification Exercise: Impact of Lagged Rejections from Other Causes on Different Air Pollutants

	(1) $PM_{2.5}$	(2) PM_{10}	(3) NO_2	(4) CO	(5) O_3
Rejections (Other-Causes)	-0.0230	-0.00643	-0.00118	-0.0294	-0.0189
	(0.0310)	(0.0168)	(0.0249)	(0.0171)	(0.0225)
N	21175	21896	19604	21455	21462
R^2	0.838	0.843	0.848	0.736	0.910
adj. R^2	0.814	0.821	0.824	0.699	0.898

Regressions of Different Air Pollutants on the Number of Rejections (one-week lag) by Causes Other Than Smog-Checks Controlling for Weather Covariates and Day, and Geographic-Unit Fixed Effects

Clustered Standard Errors in Parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Impact of Rejections Conducted on High(Low)-Quality Stations on Different Air Pollutants

	(1) $PM_{2.5}$	(2) PM_{10}	(3) NO_2	(4) CO	(5) O_3
R (q="High")	-0.0165	-0.0203	-0.0605**	-0.0655**	-0.0513***
	(0.0293)	(0.0143)	(0.0235)	(0.0254)	(0.0150)
R (q="Low")	-0.0124	0.0148	0.0112	0.0137	0.0138
	(0.0192)	(0.0173)	(0.0188)	(0.0141)	(0.0167)
N	21211	21949	19634	21508	21515
R^2	0.838	0.844	0.850	0.739	0.912
adj. R^2	0.814	0.822	0.826	0.701	0.899

Regressions of Different Air Pollutants on the Number of Rejections from the Day Before, Conducted at High (Low)-Quality Stations with Weather Controls, Day, and Geographic-Unit Fixed Effects.

Clustered Standard Errors in Parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: Impact of Lagged Rejections Conducted on High(Low)-Quality Stations on Different Air Pollutants

	(1) $PM_{2.5}$	(2) PM_{10}	(3) NO_2	(4) CO	(5) O_3
R (q="High")	-0.0160	-0.0306+	-0.0611+	-0.0733**	-0.0482**
	(0.0279)	(0.0140)	(0.0280)	(0.0242)	(0.0176)
R (q="Low")	-0.0159	0.00201	0.0104	0.00304	0.0128
	(0.0139)	(0.0143)	(0.0198)	(0.00986)	(0.0192)
N	21181	21905	19609	21464	21471
R^2	0.838	0.844	0.850	0.739	0.912
adj. R^2	0.814	0.822	0.826	0.702	0.900

Regressions of Different Air Pollutants on the Number Rejections (one-week lag) Conducted at High(Low)-Quality Stations with Weather Controls, Day, and Geographic-Unit Fixed Effects.

Clustered Standard Errors in Parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Impact of Lagged Rejections Conducted on High(Low)-Quality, Changing the Quality-Thresholds Stations on Carbon Monoxide

	(1) $[CO]$	(2) $[CO]$	(3) $[CO]$
R (q="High")	-0.0647**	-0.0638**	-0.0621**
	(0.0252)	(0.0265)	(0.0274)
R (q="Low")	0.0118	0.0138	0.0197
	(0.0108)	(0.0115)	(0.0112)
SVFR-Threshold	0.75	0.80	0.85
N	21508	21508	21508
R^2	0.739	0.739	0.739
adj. R^2	0.701	0.701	0.701

Regressions of $[CO]$ on the Number of Rejections (one-week lag) Conducted at High(Low)-Quality Stations, with different Quality-Thresholds, Controlling for Weather Covariates, Day, and Geographic-Unit Fixed Effects.

Clustered Standard Errors in Parentheses

$^{+}$ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

7.2 Appendix 2

In this section, the nature of the measurement error regarding the main independent variable is discussed. This is relevant since the number of rejections is a proxy for the total number of cars being effectively removed/repair. Suppose that the real model is:

$$\mathbf{Y} = \beta \mathbf{R}^* + \epsilon. \tag{6}$$

Where $\mathbf{Y}$ corresponds to daily pollution levels for each geographic-unit. $\mathbf{R}^*$ represents the number of cars rejected that are effectively removed/repair after being rejected by the Smog-Checks system. Notice that we are only able to observe:

$$\mathbf{R} = \mathbf{R}^* + \mathbf{Z}. \tag{7}$$

Where, $\mathbf{R}$ is the daily number of rejection at each geographic-unit and $\mathbf{Z}$ represents the daily number of car-owners that keep driving and/or do not conduct any significant repairs to their cars, even after being rejected.

This case is a little bit different from the classical measurement error since I know that $cov(Z, \epsilon) \geq 0$. In intuitive terms, this means that when there is a larger number of cars that do not stop driving and/or not repair their cars after getting a rejection, the error term is going to be larger (i.e. pollution will be higher than what our model would be able to predict or capture). I will call $\mathbb{V}[R] = \sigma^2_R$, $\mathbb{V}[Z] = \sigma^2_Z$, and $cov[Z, \epsilon] = \sigma_{Z,\epsilon} \geq 0$. Additionally, I know that if $\mathbf{Z}$, and $\mathbf{R}^*$ are correlated, then $cov(Z, R^*) = \sigma^2_{Z,R^*} > 0$. Our identification assumption holds: $cov[R^*, \epsilon] = 0$.

The OLS estimator of β is:

$$\hat{\beta} = [\mathbf{R}'\mathbf{R}]^{-1}(\mathbf{R}\mathbf{y}). \tag{8}$$

Replacing equation 7 and 6 in 8:

$$\begin{aligned}
\hat{\beta} =& [(\mathbf{R}^* + \mathbf{Z})'(\mathbf{R}^* + \mathbf{Z})]^{-1}((\mathbf{R}^* + \mathbf{Z})'\mathbf{y}) \\
=& [\mathbf{R}^*\mathbf{R}^{*'} + 2\mathbf{R}^{*'}\mathbf{Z} + \mathbf{Z}'\mathbf{Z}]^{-1}[(\mathbf{R}^* + \mathbf{Z})'(\beta\mathbf{R}^* + \epsilon)] \\
=& [\mathbf{R}^*\mathbf{R}^{*'} + 2\mathbf{R}^{*'}\mathbf{Z} + \mathbf{Z}'\mathbf{Z}]^{-1}\left\{[\mathbf{R}^{*'}\mathbf{R}^*\beta] + [\mathbf{Z}'\mathbf{R}^*\beta] + [\mathbf{R}^{*'}\epsilon] + [\mathbf{Z}'\epsilon]\right\}.
\end{aligned} \tag{9}$$

Rewriting the last expression in scalar notation I get:

$$\hat{\beta} = \frac{\left[\frac{1}{NT}\sum_i^N\sum_t^T R^*_{it}R^*_{it}\beta\right] + \left[\frac{1}{NT}\sum_i^N\sum_t^T Z_{it}R^*_{it}\beta\right] + \left[\frac{1}{NT}\sum_i^N\sum_t^T R^*_{it}\epsilon_{it}\right] + \left[\frac{1}{NT}\sum_i^N\sum_t^T Z_{it}\epsilon_{it}\right]}{\left[\frac{1}{NT}\sum_i^N\sum_t^T R^*_{it}R^*_{it} + 2\frac{1}{NT}\sum_i^N\sum_t^T R^*_{it}Z_{it} + \frac{1}{NT}\sum_i^N\sum_t^T Z_{it}Z_{it}\right]}. \tag{10}$$

Now applying limits in probability to the whole expression in 10, I get:

$$\underset{N,T\to\infty}{plim}\ [\hat{\beta}] = \underset{N,T\to\infty}{plim} \left\{ \frac{\frac{1}{NT}\sum_i^N\sum_t^T R_{it}^* R_{it}^*\beta + \frac{1}{NT}\sum_i^N\sum_t^T Z_{it}R_{it}^*\beta + \frac{1}{NT}\sum_i^N\sum_t^T R_{it}^*\epsilon_{it} + \frac{1}{NT}\sum_i^N\sum_t^T Z_{it}\epsilon_{it}}{\frac{1}{NT}\sum_i^N\sum_t^T R_{it}^* R_{it}^* + 2\frac{1}{NT}\sum_i^N\sum_t^T R_{it}^* Z_{it} + \frac{1}{NT}\sum_i^N\sum_t^T Z_{it}Z_{it}} \right\}. \tag{11}$$

Using basic limits' properties I will do this separately for the ease of presentation. I will treat the four terms in the numerator separately from the denominator. Starting with the four terms in the numerator of 11:

$$\underset{N,T\to\infty}{plim} \left[\frac{1}{NT}\sum_i^N\sum_t^T R_{it}^* R_{it}^*\beta\right] = \sigma_R^2\beta. \tag{12}$$

$$\underset{N,T\to\infty}{plim} \left[\frac{1}{NT}\sum_i^N\sum_t^T Z_{it} R_{it}^*\beta\right] = \sigma_{Z,R^*}^2. \tag{13}$$

$$\underset{N,T\to\infty}{plim} \left[\frac{1}{NT}\sum_i^N\sum_t^T R_{it}^*\epsilon_{it}\right] = 0. \tag{14}$$

$$\underset{N,T\to\infty}{plim} \left[\frac{1}{NT}\sum_i^N\sum_t^T Z_{it}\epsilon_{it}\right] = \sigma_{Z,\epsilon}. \tag{15}$$

Distributing the limits in probability to the terms in the denominator of 11 and proceed from left to right we get:

$$\underset{N,T\to\infty}{plim} \left[\frac{1}{NT}\sum_i^N\sum_t^T R_{it}^* R_{it}^*\right] = \sigma_R^2. \tag{16}$$

$$2\underset{N,T\to\infty}{plim} \left[\frac{1}{NT}\sum_i^N\sum_t^T R_{it}^* Z_{it}\right] = 2\sigma_{R^*,Z}^2. \tag{17}$$

$$\underset{N,T\to\infty}{plim} \left[\frac{1}{NT}\sum_i^N\sum_t^T Z_{it} Z_{it}\right] = \sigma_Z^2. \tag{18}$$

Gathering all those terms:

$$\underset{N,T\to\infty}{plim}\ [\hat{\beta}] = \beta\frac{\sigma_R^2 + \sigma_{R^*,Z}^2}{\sigma_R^2 + 2\sigma_{R^*,Z}^2 + \sigma_Z^2} + \frac{\sigma_{Z\epsilon}}{\sigma_R^2 + 2\sigma_{R^*,Z}^2 + \sigma_Z^2}. \tag{19}$$

Notice, that the term that accompanies β is less than one. This term is similar to the term that appears in a classical measurement error context. Additionally, since I know, from theory, that $\beta < 0$ and that $\sigma_{Z\epsilon} > 0$, I am adding a positive term to a negative one. This means that the presence of cars that are not complying with the regulations and/or owners that keep driving without conducting any meaningful repairs will bias our estimates toward zero. The conclusion of this section is that our estimates correspond to lower-bounds of the true impact of rejecting a high-emitting additional car in terms of local air pollution.