Environmental Regulations and Corruption: Automobile Emissions in Mexico City

Paulina Oliva*

March, 2012

Abstract

Emission regulations become more prevalent in developing countries as car-fleets grow; but they may be compromised by corruption. To shed light on this issue, I follow three steps. First, I develop a statistical test for identifying a specific type of cheating that involves bribing center technicians. Second, I predict fair probabilities of passing the test for the entire car-fleet by using low-cheating centers identified in step 1. Third, I estimate a structural model of car owner retesting and cheating decisions, whose parameters are recovered from observed testing outcomes and the empirical distribution of the probability of passing the test. No direct information on cheating decisions is required. I find that at least 9.6 percent of old-car owners paid bribe amounts of 20 U.S. dollars to circumvent the regulation. Simulations suggest that eliminating cheating and increasing the cost of retests would eliminate 1,443 tons of emissions, but would do so at a high cost for vehicle owners.

Keywords: environmental regulation, cheating, corruption, smog-checks

1 Introduction

Automobile emissions are an important contributor to air pollution and greenhouse gases, especially in developing countries. In Mexico City, automobile emissions are responsible for 45 percent of volatile organic compounds and 81 percent of total nitrogen oxides (Molina and Molina 2000). These gases are responsible for ozone formation, which is harmful for health at low atmospheric levels; and particulate

*University of California at Santa Barbara, oliva@econ.ucsb.edu

matter, which has been shown to be associated with severe respiratory illness. In contrast, automobile emissions in the United States are responsible for about 29 and 34 percent of these gases, respectively. There are three reasons for this difference: a younger car fleet, stricter manufacturer controls in the U.S., and extensive cheating on emission tests in Mexico City. This paper will focus on the third reason.

Compulsory vehicle emission inspections, known as smog-checks, are the most common means of enforcing emission standards on vehicles throughout the world. However, their effectiveness in reducing on-the-road emissions has been questioned widely due to sizable gaps between emission levels from official tests measured at smog-check centers and emission levels measured on-the-road or off-cycle tests (Glazer et al. 1993, Wenzel et al. 2004, *etc*). Some studies have attributed this discrepancy to a high variance in emissions or a fast deterioration of emission controls. Wenzel et al. (2004), for example, find that eight percent of the cars in Phoenix that passed an emission test on the first attempt will fail an immediate off-cycle retest. And, those cars that failed the first attempt but passed the official test on the second attempt would fail an immediate retest with a probability of 32 percent.[1] Emission testing requirements may be ineffective for reducing average emissions if emission variance is high and affordable retesting is available.

Other studies have emphasized "cheating" or smog-check center fraud as the main source of test ineffectiveness (Hubbard 1998, Zhang et al. 1996, Glazer et al. 1993, Wenzel 2000, Snyder and Pierce 2008, and the Ministry of Environment in the Federal District of Mexico 2004), but the evidence they present is often subject to confounding factors. Wenzel (2000), for example, finds that private centers in California have higher passing rates than government owned centers in Arizona, and attributes this discrepancy to fraud. Similarly, Hubbard (1998) argues that private centers in California have incentives to "help" car owners pass the test, citing differences in passing rates between private and government inspectors to support his argument. In a recent U.S. study, Snyder and Pierce (2008) blame unethical behavior for persistent differences in passing rates across smog-check facilities. In Mexico City, remote sensing studies have found that on-the-road emissions are substantially higher than same-vehicle emissions measured at testing centers.[2] The local Ministry of Environment

[1] He uses a sample of cars that are submitted to an off-cycle test as part of the requirements for change of ownership in Phoenix and California. The failure rates reported correspond to Phoenix. California failure rates on immediate off-cycle tests are 6 and 20 percent respectively.

[2] There were remote sensing studies in Mexico City in 2000 and 2003 (Schifter et al. 2003).

(Secretaría de Medio Ambiente, D.F. 2004) attributes the discrepancy to cheating.

This paper assesses the prevalence of cheating on smog checks in Mexico City and the extent to which cheating undermines regulatory efforts to reduce vehicle emissions. Documenting and studying cheating is complicated by the lack of information on cheating decisions and true emissions from cheaters. Some papers in the corruption literature have overcome similar data issues by comparing administrative records with survey information or independent assessments (Olken 2006, 2007). This approach would not be reliable in the case of smog-checks since vehicle emissions may vary from test to test (Wenzel, 2004). Other papers have relied on indirect evidence to document and study cheating and corruption (Fisman 2001, Levitt and Jacob 2006, Duggan and Levitt 2002). This paper's approach is closer to the latter literature.

The methodology used in this paper overcomes the lack of information issue by combining a non-parametric test for cheating with a structural model of the emission testing process. In a first step, I develop a statistical test for identifying a specific type of cheating that involves bribing the smog-check center personnel. Second, I estimate a mapping from car attributes into "fair" probability of passing the test by only using data from centers with little evidence of cheating (identified in step 1). The mapping is then used to predict the fair probability of passing for the rest of the car-fleet. Third, I estimate a structural model of car owner testing decisions that allows for both, retesting and cheating. This allows me to study simultaneously the two issues with smog checks that have been identified by previous literature. The parameters in this model can be recovered from observed testing outcomes and the empirical distribution of the probability of passing the test, which is estimated in step 2. No direct information on cheating is required for the estimation. The maximum likelihood estimation of the model yields estimates for the prevalence of cheating and the equilibrium bribe in the cheating market. An extension to the model yields results on the car owners' willingness to pay for car maintenance and the social benefits from increasing enforcement.

The test for cheating in step one relies on the identification of cheating-consistent patterns in smog-check center records. Interviews with mechanics and anecdotal evidence suggest that the most common way to cheat is to substitute clean emissions from "donor" cars for high emissions from high emitting vehicles. This type of cheating requires bribing a center's technician to select a suitable donor car among other customers and match a donor's second emission reading to the cheater's registration card. Cheating centers can be identified because emission readings from the same vehicle have a lower variance than readings from different vehicles; hence, test readings will appear to be serially correlated whenever donor cars are used. Results from this test suggest that 63 out of 80 centers accept bribes for cheating using donor cars. Three different robustness tests, that rely on additional evidence, support the main test's methodology.

Although this test does identify cheating centers, it cannot be used to identify individual cheaters. However, the cheating test does provide an important input to the estimation of the structural model that does estimate cheating prevalence at the individual level. Specifically, the test identifies a group of low-cheating centers and, in a second step, I use test information in low-cheating centers to estimate a mapping between "fair" pass rates and observed car characteristics. This mapping is then combined with the distribution of car characteristics and observed pass rates across the entire car-fleet to construct an estimable model for the decisions car owners make when facing the smog-check requirement.

The structural model of car owner decisions in step three assumes individuals take their expected probability of passing as given and choose between fair testing, cheating and paying a bribe or postponing the test beyond the testing deadline. The probability of cheating is expressed as a function of the expected costs of bribing and not-bribing decisions. These expected costs depend in part on the fair probability of passing the test, and the amount paid to the technician as a bribe. Because of the large number of centers and the homogeneity of the service offered, I assume that centers are perfectly competitive on bribes (Shleifer and Vishny 1993); hence, car owners and centers are assumed to be "bribe takers." The model is estimated by maximizing the likelihood of the observed sequences of retesting decisions and test outcomes, given the theoretical probabilities of these sequences from the model. Despite not observing individual cheating decisions, estimates for the parameters in the structural model are recovered from an excess of passed tests in the observed sequence of tests and retests, given the predicted fair probabilities of passing. The likelihood maximization yields estimates for the bribe amount and the time cost associated with smog checks. Most importantly, the maximization yields an estimate of cheating prevalence and the probability of cheating for each vehicle. The parameter estimates from the model suggest that the average bribe was about 20 U.S. dollars in 2003.

The results of the model estimation likely provide a lower bound of the prevalence of cheating, since the predicted "fair" probability of passing may overestimate the true probability of passing. This would be the case if potential cheaters know which centers welcome bribes and self-select into them. In the structural model estimation, the upward bias of the estimated probability of passing would lead underestimating amount of cheating. In other words, the model predicts that *at least* 9.6 percent of emissions tests appear to be fraudulent.

The structural model is used to simulate car owner decisions under increased enforcement and increased cost of retests. I combine predicted decisions on car maintenance with the information on emissions from low-cheating centers to perform a back of the envelope calculation of the environmental benefits from eliminating corruption and at the same time increasing the cost of repeated tests. Results suggest that a set of policies that address both cheating and inexpensive retesting would avoid 1,443 tons of pollution, with the estimated benefit of 1.5 million dollars in reduced mortality. However, these policies would be accompanied by an increase in costs of about 3.7 million dollars to car owners. The simulation results highlight how inducing a small percent of the car-fleet to reduce their emissions using smog-check inspections comes at a high cost for the entire car-fleet.

The rest of the paper proceeds as follows: Section 2 provides a brief description of the vehicle regulations and describes the context in which cheating occurs. Section 3 describes the data used for the cheating detection and the structural model estimation. Section 4 proposes a statistical test for detecting cheating. Section 5 develops a model for car owner's decisions with respect to cheating. Section 6 extends the model to simulate willingness to pay for car maintenance and car maintenance decisions under different policy scenarios; and Section 7 concludes.

2 Vehicle Regulations in Mexico City

Mexico City introduced twice-a-year smog checks for all vehicles in 1990. The requirement is compulsory and universal within Mexico City Metropolitan Area (MCMA), which comprises of all the Federal District and a large portion of the state of Mexico. Vehicle owners have a two-month window to attend any smog check center within their state and full fill the smog-check requirement by passing the emissions test and obtaining the corresponding certificate. The certificate is given in the form of a sticker that is pasted on the vehicle's windshield by center staff. Non-complying vehicles are easy to spot on the road by police officers, since they do not carry a valid sticker. Fines for non-compliance range from 850 to 3,500 pesos (79 to 325 2003 U.S. dollars).[3]

Smog checks consist of the following steps: upon arrival at each test attempt the vehicle's owner pays the emission test fee whenever required (even attempts at passing the test are free). A center employee enters the vehicle's information into the computer. This information includes plate number, model-year, make, number of cylinders, owner's address and mileage, all of which can be read from the registration card and the odometer. Once the vehicle's information is recorded, one or two technicians perform a visual test of the vehicle in order to identify any obvious malfunctioning of the engine or tailpipe. Upon passing the visual test, the vehicle is placed on the dynamometer and the reader is connected to the vehicle's tailpipe to perform the emission test. Emissions are read directly by the computer and cannot be entered manually. After the test is complete, the corresponding test certificate is imprinted with the vehicle's plate number. The technician does not observe the recorded emissions until the certificate is printed. If the vehicle passes the emission test, a sticker with the vehicle car plate number is pasted on the vehicle's windshield. If the vehicle fails the test, the owner may retest indefinitely upon paying the corresponding fees.

In 2003 there were a total of 80 licensed centers in the Federal District. The licenses for these centers were tendered in 1997 and very few new licenses have been granted ever since. All smog-check centers are privately owned, except for three institutional centers.[4] However, they are all subject to tight regulations from the government. The smog-check centers are obliged to purchase their testing equipment and computer software from government-approved providers. The software, which contains the current emission norms, is renewed annually. The ministry of environment conducts unannounced inspections of smog-check center facilities on a regular basis. During these inspections, all mechanical and electronic equipment is checked. In addition, all facilities are required to have a camera surveillance system and publicly available video transmissions.

[3]Fines are set in minimum wages. A fine for smog-checking after deadline is 20 daily minimum wages. A fine for circulating without a valid smog-check certificate is 40 daily minimum wages. A fine for failing to comply with the smog check requirement within 30 days after first fine is 80 daily minimum wages. (Gaceta Oficial del Distrito Federal 2004)

[4]Institutional centers belong to the Department of Defense (*SEDENA*), the National Power and Electricity Company (*Luz y Fuerza del Centro*) and the Department of Water Resources (*Sistema de Aguas).*

Emission standards are somewhat below EPA Tier 1 standard in the U.S. Columns 1 to 5 of Appendix Table 1 show the emission limits in 2003. Vehicles can be retested as many times as needed within their corresponding two-month window. In 2003 each test cost 175 pesos (16 U.S. dollars) and every second retest was (and remains) free. This pricing structure has been determined by the Ministry of the Environment. Column 6 of Appendix Table 1 shows the emission requirements on less-than-10 year old vehicles for an exemption to the "No Driving Today" program.[5]

3 Data

This study uses data on vehicle information and test outcomes from emission tests conducted in the Federal District in 2003. The computers in each smog-check center are connected to a common network run by the local Ministry of Environment, which pools the information from all centers into a single data set. The resulting data set includes information on each test and retest performed on every car that visits an authorized smog check center.

The data for 2003 include information for all tests and retests for 1.6 million vehicles. The information on each test includes the exact measurement on four relevant gases. Three out of the four gases are harmful pollutants: hydrocarbons (HC), nitrogen oxides (NO) and carbon monoxide (CO); while the fourth, oxygen (O_2), is measured to confirm the proper balance in the combustion process and avoid passing tampered vehicles. Each test consists of two different readings of each of these four gases. The first reading is taken at 24 kilometers per hour (kph), and the second one is taken at 40 kph. In order to pass the emission test, a vehicle must have emissions below the standard in both readings.

Information available for each test also includes car characteristics (plate number, model year, brand and size of the engine); test outcomes (pass/fail status, reason for failure, visual conditions of the car, and whether or not the owner paid a fine for non-compliance in the last smog-check period); beginning and ending times of the

[5]The "No Driving Today" program has been in place since 1989 and originally restricted all vehicles from circulating one day a week. The day of the week a vehicle is restricted varies with the last digit of the vehicle's plate number (see Davis 2008 for a recent evaluation of this program). Since 1997, vehicles are exempt from this program if they are model-year 1993 or newer and they meet a stricter standard (see column 6 of Table 1). This age requirement changed in 2004. Exempt vehicles should now be 10 years or newer, regardless of the model-year.

test in seconds; and smog check center's information (center's identity number, and lane or testing equipment where the test was performed). Appendix Table 1 presents descriptive statistics of emission readings and car characteristics.

4 Extent of Cheating

Cheating at smog-check centers is a major concern in Mexico City. Anecdotal evidence and newspaper articles suggest that fraud is a common practice. In 2002, undercover newspaper reporters took a car with substandard emissions to seven randomly selected smog check centers. In six of the seven, the reporters were able to obtain the emission test certificate by paying an additional "tip" that ranged from 5 to 40 U.S. dollars. The technicians at the fraudulent centers did not reveal any details about the cheating procedure. However, they assured the reporters the procedure would not cause damage to the car's engine (Padgett 2002).

The local Ministry of Environment has also expressed concern about cheating. However, they have emphasized cheating in the form of tampering with the engine, which is likely not the main form of cheating by 2003.[6]

Anecdotal evidence suggests that most cheating occurs in the form of emission substitution. When bribed by a customer, technicians use a clean testing car, commonly called donor car (*auto madrina),* to provide the emission readings for the bribing costumer's dirty car. The donor car can be any other vehicle that passed the emission test at the center. A donor car is needed because emissions cannot be entered manually into the center's computer. The car's information, on the other hand, has to be entered manually into the computer, which allows the technicians to enter the information from a dirty car and match that with emissions from a clean car. An observable consequence of this type of cheating is that consecutive emissions readings in a single lane will have strong serial correlation. The serial correlation arises from

[6]The local Ministry of Environment published a report in 2004 where they expressed concern about a cheating problem in the form of tampering with the engine. The report states that tampering is detectable through the levels of other gases such as oxygen and carbon dioxide. They also notice that after including some checks for these additional gasses in 2002, evidence of tampering was reduced in the smog check center data. In their report, they propose using additional controls of this sort to eradicate tampering. These measures, however, do not seem to have solved the cheating problem. Moreover, because of the additional emission requirements and because vehicles with computerized systems can be severely damaged when manually altering the combustion process, tampering is no longer the main cheating method at emission tests.

having a single car as a provider of emissions for more than one consecutive test. Presumably, one of these tests is assigned to the vehicle to which it belongs. The rest of the emission readings are assigned to vehicles owned by bribers.

4.1 Testing for cheating

The statistical test for detecting cheating developed in this section relies on measuring the serial correlation between tests of different vehicles. In motivating the statistical test used to detect cheating, it is instructive to observe a sequence of actual emissions readings in the order they occur. Table 2 shows a fragment of the sequence of emission readings for a single smog check center in a single lane. Column 1 shows the exact time at which the test was performed. Columns 2 to 4 show the model-year of the vehicle tested, the number of cylinders, and the volume of displacement in the engine. The remaining columns show the reading of each pollutant measured at 24 kph and 40 kph. Notice that the sequences of readings that appear inside the black squares show striking similarities in all pollutants across tests despite substantial differences in car characteristics. All readings within these sequences could correspond to the same donor car, even though they are listed under different vehicle records.

I use an autoregressive model for each measured pollutant to estimate the extent of serial correlation between consecutive readings. The model controls for flexible functions of car, car owner and test characteristics. A significant positive coefficient on the preceding test will be interpreted as evidence of cheating.

More formally, let $\tilde{r}_{jit}$ be the true emissions of pollutant j, for car i at time t. In this case, j will denote one of the four gases that are involved in the test when measured at 24 kph. To simplify the notation, I will omit the index j in the analysis that follows. However, the regression analysis proposed below will take into account all four pollutants. True emissions are given by

$$\tilde{r}_{it} = \mathbf{x}_i\beta + u_{it}, \tag{1}$$

where $\mathbf{x}_i\beta$ is the mean square error minimizing linear prediction given observable car, test and center characteristics. The specific set of controls included in $\mathbf{x}_i$ is detailed below.

In the current setup, the index i will be a car specific identifier that will also keep track of the order in which vehicles show up at a center. To ease the exposition, assume each center has a single lane. For example, the vehicle arriving after vehicle i will be denoted $i+1$. Emission components in u_{it} are indexed by time, t, since emissions of the same car may vary from test to test. The index t will then denote the order in which vehicles were actually tested or "test slot". The distinction between i and t is necessary only when cheating occurs. In what follows, I will refer to vehicles whose owner decides to cheat as "cheaters" or "cheating vehicles".[7] Under no cheating, these two indexes should be one to one. However, in the presence of cheating, the same vehicle may be tested in two or more consecutive test slots. For example, if $i-1$ is a donor car, i is a cheater, and $i+1$ is not a cheater, then the smog-check data set will have the following sequence of emissions on test slots $t-1$ through $t+1$: $\tilde{r}_{i-1t-1}, \tilde{r}_{i-1t}, \tilde{r}_{i+1t+1}$. If none of the three cars are cheaters, then the sequence of emissions in the smog-check record will be: $\tilde{r}_{i-1t-1}, \tilde{r}_{it}, \tilde{r}_{i+1t+1}$.

The test for cheating relies on the contiguity between emission readings of the donor car and all cheaters it provides emissions for. Specifically, it assumes that

(A1) all vehicles the donor car provides emissions for have the consecutive test slots in the same lane.

In the rest of the paper, I will often use the terms "testing equipment" and "lane" interchangeably, since there is one testing equipment per lane. Assumption (A1) excludes the possibility of putting away the donor car from the testing equipment while other non-cheating vehicles get tested and bringing it back in for a cheater's test slot. It also excludes recording the donor car's "own" emission reading in a different equipment than the readings used for cheaters' certificates. However, it allows for multiple vehicles to cheat using the same donor car, as long as they are all "tested" consecutively after the donor car. Assumption (A1) is supported by anecdotal evidence and the observation of sequences of emission test like the one shown in Table 2.[8]

It is plausible to assume that, under the null of no cheating, there should be no dependence between subsequent emission readings, conditional on car and center

[7] Obviously, the decision to cheat is made by the owner, not the vehicle.

[8] There is anecdotal evidence that the use of donor cars that are not being tested themselves is not common. However, the use of this type of donor cars would violate assumption (A1). If centers used a home car as a donor car, e.g. the technician's car, the donor car's emissions would not appear in the test slot preceding a cheater. In this case, the test proposed below would only detect corruption if the donor car was used consecutively for more than one briber. Groups of consecutive cheaters, like the one exemplified in Table 2, do occur frequently.

characteristics. The regression methodology outlined below requires an even weaker set of assumptions: first, the unobserved components of linearized emissions, u_{it}, are assumed to be serially uncorrelated, and second, the correlation between the unobserved component of emissions and every observed car characteristic of the preceding vehicle is assumed to be equal to zero:

$$\text{(A2)} \qquad \mathbb{E}(u_{i-1t-1}u_{it}) = 0$$

$$\text{(A3)} \qquad \mathbb{E}(\mathbf{x}_{i-1}u_{it}) = \mathbf{0}$$

Notice that, by construction, u_{it} includes only emission determinants that are uncorrelated with observed emission determinants. Thus, assumption (A2) would be violated if, in the absence of cheating, car emission determinants not controlled for in $\mathbf{x}_i$ are serially correlated. For example, consider two cars that arrive simultaneously at the center and belong to the same owner, say the same household or same company, and thus have similar maintenance histories. For these type of car pairs to violate the identification assumptions, the maintenance histories, or other unobserved car characteristics, would have to be correlated with emissions but be uncorrelated with the observed emission determinants previously described. In the estimation of (1), observed emission determinants include brand, service and size fixed effects, a flexible function of the age of the car, age-size and age-service interactions, flexible functions of mileage and time of the day, week and center-lane (testing equipment) fixed effects. Note that the testing equipment fixed effects would pick up similarities in emissions such that may come from differences in calibration or any time-fixed patterns in the type of cars that are directed to a specific lane. To further control for non directly-observable emission determinants that might induce serial correlation if similar vehicles coincide in timing and center choices, $\mathbf{x}_i$ also includes emissions recorded in preceding and subsequent tests that are performed on neighboring lanes in the same center. Notice that, since all centers have multiple lanes, vehicles that are in consecutive places in the queue will be distributed across the different lanes. Hence similarities within groups of vehicles with similar arrival times are likely picked up by emissions from vehicles in contiguous lanes that are tested simultaneously or closely in time. Controlling for these emissions, assumption (A2) would only be violated if cars that are similar in non-observable ways are directed consistently to the same lane upon arriving simultaneously at the center.

Assumption (A3) is presumably easier to meet than assumption (A2). For it to be violated, some unobserved emission determinant that is uncorrelated with any of the observed own car and test characteristics would have to be correlated with the car characteristics of the contiguous vehicle in a single testing lane.

Let r_{it} be observed smog-check center emissions. Under the null hypothesis of no cheating, r_{it}, should be equal to true emissions, $\tilde{r}_{it}$. Under the alternative hypothesis, an observed reading can be either a measure of true emissions or a measure of a donor car emissions. Hence, observed emission readings under the alternative hypothesis be expressed as:

$$r_{it} = c_i \tilde{r}_{i-kt} + (1 - c_i)\tilde{r}_{it}, \tag{2}$$

where r_{it} are observed emissions of car i in slot t, $\tilde{r}_{i-kt}$ are true emissions from the first donor car that preceded vehicle i in the same lane and c_i is a binary variable, such that $c_i = 1$ if i is a cheater.

Since k is unknown, true emissions, $\tilde{r}_{i-kt}$, are not observed; and therefore, it is not possible to test for H_0 by regressing r_{it} on $\tilde{r}_{i-kt}$. However, we do observe r_{i-1t-1}, the observed emission reading for car $i-1$ in test slot $t-1$, which will originate from the donor car, $i-k$, whenever i is a cheater under assumption (A1). Under assumptions (A1)-(A3), an unbiased test for whether a center uses donor cars for cheating is given by the estimation of the following OLS regression:

$$r_{it} = \gamma_c r_{i-1t-1} + \mathbf{x}_i \gamma_x + \nu_{it}, \tag{3}$$

The OLS estimate of coefficient γ_c can be used as a test statistic for the null hypothesis of no cheating. Assumptions (A2) and (A3) imply that, in the absence of cheating, $\mathbb{E}^*(r_{it}|r_{i-1t-1,}\mathbf{x}_i) = \mathbf{x}_i\beta$.[9]

The test for cheating outlined above is performed at the center level. In order to maintain the power of the test constant across centers, I draw a random same-size sample of tests from each center. For each center, equation (3) is estimated jointly for all four pollutants using a Seemingly Unrelated Regressions model. I test for cheating using the chi-square statistic for the joint hypothesis that the serial correlation coefficients for all four equations are zero. This produces a single p-value

[9]To see this, note that $\mathbb{E}^*(r_{it}|r_{i-1t-1,}\mathbf{x}_i) = \mathbb{E}^*(\mathbf{x}_i|r_{i-1t-1,}\mathbf{x}_i)\beta + \mathbb{E}^*(u_{it}|r_{i-1t-1,}\mathbf{x}_i) = \mathbf{x}_i\beta + \mathbb{E}^*(u_{it}|r_{i-1t-1}) = \mathbf{x}_i\beta + \mathbb{E}^*(u_{it}|\mathbf{x}_{i-1,}u_{i-1t-1})$. Under assumptions (A2) and (A3), $\mathbb{E}^*(u_{it}|\mathbf{x}_{i-1,}u_{i-1t-1}) = 0$. See Appendix 1 for the details of this proof.

for each center. I find evidence of corruption at the 5 percent level for 63 out of 80 centers. Panel A of Table 3 shows the test results for the 10 centers that have the lowest evidence of donor car cheating and Panel B shows results for the 10 centers that have the strongest evidence of donor car cheating under a joint test of significance for all four specifications.

4.2 Robustness tests and other types of cheating

Although, I cannot directly test the validity of assumptions (A2) and (A3), I perform three different robustness tests that support the cheating interpretation of the results. First, I propose an alternative test for cheating that also relies on detecting the abnormal occurrence of sequences of close emissions, but under slightly different assumptions. This alternative test allows for some cheating-unrelated serial correlation between observed and unobserved determinants of vehicle emissions to occur, but constrains this correlation to occur at the center, lane, date and day-shift level (a day is considered to have 4 shifts). The identification assumption of this test would be violated if the correlation occurred at a more specific level such as center, lane, date and hour of the day. The mechanics of this test rely on using permutations of the data to generate different draws of the empirical distribution of emission differences between consecutive pairs. Under the null hypothesis of no cheating, these random draws should not be statistically different than the observed draw. These empirical distributions generated from the permutations are compared to the observed distribution of pairwise differences. With a large number of draws, I can test for whether the the observed distribution has a significantly larger amount of small differences compared to the distributions from permuted data. The generated and observed distributions are shown in Appendix Figure A1 and the mechanics of this alternative test are explained at length in Appendix 2. The results of this alternative test suggest that corruption occurs at 75 out of 80 centers. Although this test appears to have a stricter standard for ruling out corruption (the p-values for the test in 73 centers are less than 0.001), the two tests agree on which centers have insufficient evidence for corruption in 4 out of 5 cases.

The second robustness test for the proposed methodology relies on testing for the presence of an additional implication of donor car use. Specifically, test results from the same car should be recorded close in time, in addition to being similar in measured pollutants, as two tests from the same car would not be separated by the time required to disconnect one vehicle and connect the next vehicle to the testing equipment. The hypothesis that tests that occur closer together in time also have similar recorded emissions can be tested by interacting time between tests with lagged emissions. Results confirm this prediction by showing that the serial correlation between emission readings is more important for short time intervals between tests (0-5 min) than for long time intervals between tests (15-20 min).[10]

Finally, notice that the more donor cars are used, the more often vehicle emissions will be matched with the "wrong" car characteristics. The final robustness test evaluates whether the relationship between car characteristics and emissions varies across centers by the strength of the cheating evidence in each center. I classify centers in 10 groups according to the p-value of the joint test for whether emissions are serially correlated. Centers with the lowest p-values are more likely to engage in cheating using donor cars, while centers with higher p-values are less likely to engage in cheating. I then run a simple linear regression of emissions on model-year and a constant for each group. The estimated coefficient on model-year should be more negative for the group of centers that, based on the statistical test, are unlikely to have cheating (group 1) and less negative for centers that are more likely to have cheating (group 10). Figure 1 depicts the strength of the estimated relationship between the year the car was made and the measured emissions by how likely a center is to be engaged in cheating.

Figure 1 suggests that centers that do not use donor cars are less likely to cheat in other ways. To see this, note that centers were grouped in Figure 1 according to a measure of serial correlation, which reflects the prevalence of cheating with donor cars. However, other types of cheating would also produce a discordance between car characteristics and emissions. Therefore, an additional implication of Figure 1 is that centers that are classified as having low cheating of the form of donor cars also have low prevalence of other forms of cheating.

The validity of this falsification test can be further verified by checking whether the negative slope in the graphs disappears when an environmental predictor of emissions, such as temperature, is used instead of a car attribute. The environmental determinant of emissions is always matched to the right emission readings despite the

[10] The results for this robustness test are available in as Figure WA1 of the online appendix (http://www.econ.ucsb.edu/~oliva/Research/smogchecks/Web_Appendix.pdf).

presence of corruption, hence the coefficient should not fall as corruption increases. A figure analogous to Figure 1 but using temperature instead of model year (Figure WA2 in Web Appendix) confirms this prediction.

Last, because of potential self selection into cheating and non-cheating centers, vehicle characteristics may differ across centers with different levels of cheating (see section 7 and Table WA2 in Web Appendix). If the effect of age on emissions is non-linear, we would expect the marginal effect of age to differ across centers as well. Figure WA3 in the Web Appendix replicates Figure 1 holding the age distribution of vehicles constant across centers. The positive relationship between the coefficients and the level of cheating rules out that non-linear effects of age are driving the results of the robustness test.

5 A model of bribing behavior

The results from Section 4 suggest that cheating is a major concern for emission control policy in Mexico City. The evidence of widespread corruption shown invites the questions of whether economic incentives can explain the observed behavior and whether bribing is as cheap to car owners as anecdotal evidence has suggested. In addition, although the evidence provided suggests a large amount of corruption in several smog check centers, the statistical evidence is not specific enough as to inform about the number or proportion of vehicles that rely on corruption for overcoming the smog check requirement.

This section models the key factors behind the car owner's decisions in order to validate and refine the reduced form evidence of corruption in Section 4. Using the full information available on car owner decisions as well as official test results, I estimate a dynamic model of bribing decisions that delivers estimates for the implied opportunity cost of bribing as well as the time cost incurred by vehicle owners when complying with the regulation. Given that the most polluting vehicles are older models which cannot exempt the smog-check requirement, the model proposed in this section focuses on the behavior of owners of non-exemptible vehicles, who are required to pass an emissions test in order to drive in Mexico City. The estimates found are consistent with cheating results in Section 4 and with anecdotal evidence on the prevalence of corruption and the amount of money paid in bribes.

5.1 Model set-up

The model proposed in this section incorporates the main features of the smog-check requirement in Mexico City for non-exemptible vehicles: the price of smog checks is constant across centers and equal to 16 U.S. dollars; retests are unlimited and even retests are free; the cost of not passing the test and getting caught with an expired certificate is 79 U.S. dollars; all car owners have the choice of cheating and are not punished from doing so; finally, centers and car owners are "price takers" with respect to the bribe.[11]

I preview the next paragraphs by noting that the notation in this section is entirely independent from the notation in sections 4.1 and 4.2. Specifically, the time subscript in this model will count the number of attempts to pass the test that a single vehicle undergoes within a single smog-check cycle. The beginning of each test or retest in a single smog-check cycle will correspond to a decision round, denoted t. In each decision round, indexed by t, a car owner is assumed to decide between three different actions: cheating and paying a bribe to the technician ($s_t{=}B$); submitting her car to a fair test ($s_t{=}A$); and postponing the check beyond the deadline ($s_t = X$), which entails a fine upon getting caught. If a car owner cheats and pays a bribe to the technician, she is guaranteed to pass the test. If she doesn't cheat, then she faces the risk of not passing the test, but she also avoids the cost of the bribe. If failing occurs, the car owner will situate herself at the next decision round or retest within the same smog-check cycle. The probability of passing the test is assumed to be car specific known by the car owner. If the car owner postpones the test to the next smog-check cycle, she will avoid any costs in the current smog check period, but will risk the payment of a fine in the next smog-check period.[12]

The model proposed so far does not explicitly allow for car maintenance decisions: *i.e.* owners cannot chose to perform a car tune up in order to increase the probability of passing the emissions test. An important reason for excluding maintenance from the car owner's set of options is the absence of observable data on maintenance decisions. However, given the evidence of wide spread corruption, car maintenance seems unlikely to be a common response to the smog-check requirement. The model estimation

[11] Although there is no risk of getting caught for the car owner, the equilibrium bribe may incorporate technician's risk of getting caught.

[12] Notice that smog-check periods are different than decision rounds. Each smog-check period starts with a first decision round. Subsequent decision rounds in the same smog-check period appear as the vehicle fails the tests and retests.

does not exclude maintenance from happening, but it does constrain maintenance to be unrelated to car owner's decisions regarding the smog check requirement. Indirect evidence can help us assess how unlikely is maintenance comes from the differences in probability of passing between the first test and the retest among vehicles that fail the first test. It seems unlikely that a first test, vehicle repairs, and a retest can occur all in the same day. Hence, we can safely assume that the car owners that return for a retest on the same day, which make for 80 percent of all retests, will not have performed any repairs to their vehicle since the first failed test. We can then compare the share of passing vehicles among same-day retesters with the share of passing vehicles among different-day retesters. Assuming that (a) there is the same number of cheaters among the two groups, and (b) the repairs guarantee passing the test; we can roughly calculate the number of vehicles that have performed maintenance by solving the following equation for M:

$$\hat{Pr}(pass|different\ day) \approx M + \hat{Pr}(pass|same\ day)(1 - M)$$

where $\hat{Pr}(pass|different\ day)$ is the share of vehicles that passed the rests conditional on having it on a different day than the first test, M is the share of vehicles that performed maintenance among all vehicles that went for a retest on a different day and $\hat{Pr}(pass|same\ day)$ is the share of vehicles that passed the retest among those who came back on the same day. The above calculation yields $\hat{M} = 0.074$. Given that different-day retesters are 20 per cent of all retesters, this calculation suggests that only up to 1.5 percent of all retesters may have conducted maintenance. Hence, this maintenance response to the smog-check is relatively unimportant in 2003. Alternatively, assuming that individuals that perform car maintenance pass with 0.90 probability, yields 2 percent of vehicles that might have undergone maintenance. This calculation suggests that omitting maintenance as a relevant alternative will not have a large impact in our structural estimation results.

As is standard in discrete choice models, we can represent car a owner's decisions in a random utility framework. I assume a car owner's utility is linear in money and she chooses the action $s_t \in \{X, A, B\}$ that maximizes her expected utility. Random utility depends of unknown parameters (denoted with greek letters) ω, β, δ, and τ; known constants (denoted in latin letters) c, and f; vehicle-specific probability of passing, P; and random utility components ϵ_t^s, which are decision round and choice specific.[13] Parameters ω and β as the time cost associated with each visit and the bribe paid, respectively. Both parameters are known to the car owner, but unknown to the econometrician. In this initial setup, these parameters are assumed to be constant across vehicle owners. However, the estimated version of this model allows the time cost, ω, to vary across car owners and smog-check centers. The constants c and f denote the fee charged by the center for every odd test (even tests are free) and the fine that applies to belated smog-checks, respectively. Both are measured in pesos and are known to both, car owners and the econometrician.[14] The probability of passing the test, P, is assumed to be vehicle-specific and known to both, the econometrician and vehicle owners. Random utility shocks, denoted by $\tau\epsilon_t^s$, are specific to choice s and decision round t and are assumed to be independent across choices and across decision rounds. These utility shocks can be understood as unobserved events that may change the relative cost of each choice, including for example, unforeseen time constraints that make postponing the test more attractive, a donor car constraint, *etc.* The random utility components, $\tau\epsilon_t^s$, are observed by the car owner right before making their decision at the beginning of round t, but not before that. As is usually done in discrete choice models, I assume that these shocks are extreme value distributed with mean zero and scale parameter τ. Finally, the unknown parameter, δ, can be interpreted as a discount factor on the fine that applies to belated smog-checks, f. This parameter may be interpreted as a combination of time preference and the probability of punishment.

Given the notation described above, the expected utility for each of the available choices at odd rounds is given by:

$$\begin{aligned} \mathbb{E}_t u_t^{od}(X) &= -\delta f + \tau\epsilon_t^X \\ \mathbb{E}_t u_t^{od}(B) &= -\omega - c - \beta + \tau\epsilon_t^B \\ \mathbb{E}_t u_t^{od}(A) &= -\omega - c + (1-P)\cdot\mathbb{E}_t(V_{t+1}^{ev}) + \tau\epsilon_t^A \end{aligned} \tag{4}$$

When t is an even round, the expected utilities are

$$\begin{aligned} \mathbb{E}_t u_t^{ev}(X) &= -\delta f + \tau\epsilon_t^X \\ \mathbb{E}_t u_t^{ev}(B) &= -\omega - \beta + \tau\epsilon_t^B \\ \mathbb{E}_t u_t^{ev}(A) &= -\omega + (1-P)\cdot\mathbb{E}_t(V_{t+1}^{od}) + \tau\epsilon_t^A \end{aligned} \tag{5}$$

[13] c should not be confused with with c_i from Sections 4.1 and 4.2. In this section, c without a subscript denotes a known constant equal to the cost of odd tests (16 U.S. dollars).

[14] The test fee is 16 U.S. dollars and the fine is approximately 79 U.S. dollars.

The value of decision round t is given by: $V_t^{od} = \max\left(\mathbb{E}_t u_t^{od}(X), \mathbb{E}_t u_t^{od}(B), \mathbb{E}_t u_t^{od}(B) \middle| \Omega_t\right)$ and $V_t^{ev} = \max\left(\mathbb{E}_t u_t^{ev}(X), \mathbb{E}_t u_t^{ev}(B), \mathbb{E}_t u_t^{ev}(A) | \Omega_t\right)$, where $\Omega_t = \{\epsilon_t^X, \epsilon_t^B, \epsilon_t^A\}$.

There is a trade off between the uncertainty of a fair test and the cost of bribing. The expected utility of a fair test in equations (4) includes the expected value of entering a new decision round, $\mathbb{E}_t V_{t+1}^{ev}$, multiplied by the probability of failing the test, $(1-P)$. If the car owner decides to cheat and pay a bribe, then her expected utility does not include the expected value of entering another round, because after paying the bribe she passes the test with certainty.

The value of entering another decision round differs between even and odd rounds, since the centers do not charge a testing fee for even retests. However, because there is an unlimited amount of retests, the utility associated with each of the decisions is independent of t. Hence, the value of entering another decision round is also independent of t, conditional on whether t is even or odd.

The stationarity of the expected utility and the extreme value distribution assumption for the random shocks in the model facilitate the estimation. Following McFadden and Domencich's (1996), the extreme value distribution assumption results in a closed form expression for the expected value of subsequent rounds:

$$\begin{aligned} \mathbb{E}_{t-1} V_t^{od} &= \tau\left(k + \log\left(\exp\left(\tfrac{-\delta f}{\tau}\right) + \exp\left(\tfrac{-\omega - c - \beta}{\tau}\right) + \exp\left(\tfrac{-\omega - c + (1-P)\cdot \mathbb{E}_t V_{t+1}^{ev})}{\tau}\right)\right)\right) \\ \mathbb{E}_{t-1} V_t^{ev} &= \tau\left(k + \log\left(\exp\left(\tfrac{-\delta f}{\tau}\right) + \exp\left(\tfrac{-\omega - \beta}{\tau}\right) + \exp\left(\tfrac{-\omega - (1-P)\cdot \mathbb{E}_t V_{t+1}^{od}}{\tau}\right)\right)\right) \end{aligned} \tag{6}$$

where k is the Euler constant and τ is the scale parameter of the random utility shock distribution. The stationarity of expected utility facilitates the solution to (6), since $\mathbb{E}_{t-1} V_t^{od} = \mathbb{E}_{t+1} V_{t+2}^{od}$ $\forall$ odd t and $\mathbb{E}_{t-1} V_t^{ev} = \mathbb{E}_{t+1} V_{t+2}^{ev}$ $\forall$ even t. In estimating the model, I use a nested fixed point algorithm to solve for the system of equations in (6) as a subroutine to the standard maximum likelihood problem (Rust, 1987).[15]

For computational reasons, I also estimate a two round version of the model. The two round version has the same payoffs than the infinite horizon model in the first round. However, it assumes that there is a maximum of three attempts to pass the test and all individuals choose to pay a bribe upon failing two consecutive times. Therefore, in the second round $\mathbb{E}_2 u_2(A)$ is given by:[16]

[15] For a formal proof of the existence and uniqueness of the fixed point see the Web Appendix 1.

[16] The *ev* and *od* superscripts, that denote even and odd rounds respectively, are omitted in the two round model to avoid redundancy.

$$\mathbb{E}_2 u_2(A) = -\omega + (1-P)(-\omega - c - \beta) + \tau\epsilon_2^A \tag{7}$$

The expected value of entering another decision round in $t = 1$ is given by

$$\mathbb{E}_1 V_2 = \tau\left(k + \log\left(\exp\left(\frac{-\omega-\beta}{\tau}\right) + \exp\left(\frac{-\omega+(1-P)(-\omega-c-\beta)}{\tau}\right) + \exp\left(\frac{-\delta f}{\tau}\right)\right)\right) \tag{8}$$

The two round version of the model is more efficient to compute for two reasons: first, the estimation requires computing only tone expected value of entering another round and seconds, the model has a closed form representation given by (8). Hence, estimating the two round version does not require computing the numerical solution described above.

In both models, the extreme value distribution of the random shocks implies that the probability of each choice is given by a multilogit-like expression. For example, the probability of $s_t = B$ in the infinite horizon model is given by

$$\Pr(s_t = B) = \frac{\exp\left(\frac{-\omega-c-\beta+\delta f}{\tau}\right)}{1+\exp\left(\frac{-\omega-c-\beta+\delta f}{\tau}\right)+\exp\left(\frac{-\omega-c+(1-P)\mathbb{E}_t V_t^{ev}+\delta f}{\tau}\right)}$$

for all odd t's, with $s_t = X$ as the base category. Similar probability expressions can be derived for each of the other choices for both even and odd rounds (see Web Appendix 2).

5.2 Estimation methodology

This Section describes how the structural parameters −time cost, ω, bribe, β, discount rate, δ, and random shock variance, τ− in the model of Section 5.1 can be identified from estimation. The identification of these parameters, despite the lack of data on bribing decisions, represents a methodological contribution of this paper. The identification strategy proposed relies on being able to estimate the probability of passing the test without cheating for at least two different types of vehicles. We proceed assuming this condition is met. In practice, P, is estimated in a first stage as its predicted value given the car's characteristics and the parameter estimates from a low-cheating subsample identified in Section 4.2.[17]

[17] Section 6.1 discusses in further detail how the results of Section 4 are used in the estimation procedure.

The parameters ω, β, δ and τ will be estimated by maximizing the likelihood of observed decisions and test outcomes, where the likelihood is constructed using the model's representation of the probability of each decision-outcome history, as discussed later in this section. A history of decisions and test outcomes is given by the sequence of bribe/no-bribe/postpone decisions and pass/fail test outcomes. For example, one possible history of decisions and test outcomes would be not bribing and failing in the first test, then bribing and passing in the second test.

Denote the observed test result in test round t as R_t, where $R_t = 1$ indicates a passing result, $R_t = 0$, and $R_t =$ N/A indicates that vehicle was not tested in round t. The complete set of decision and test outcome histories for the infinite horizon model is described by the sequences of test outcome vectors, (s_1, R_1), $(s_2, R_2), \ldots$, below, where the first entry in each vector corresponds to a vehicle owner decision and the second entry corresponds to a test result

$H1$: $(s_1 = X, R_1 = \text{N/A})$

$H2$: $(s_1 = B,\ R_1 = 1)$

$H3$: $(s_1 = A,\ R_1 = 1)$

$H4$: $(s_1 = A,\ R_1 = 0)$,$(s_2 = X,\ R_2 = N/A)$

$H5$:$(s_1 = A,\ R_1 = 0)$,$(s_2 = B,\ R_2 = 1))$

$H6$:$(s_1 = A,\ R_1 = 0)$,$(s_2 = A,\ R_2 = 1)$

$H7, H8, \ldots$:$(s_1 = A,\ R_1 = 0)$,$(s_2 = A,\ R_2 = 0)$,...

A car owner that does not bribe an fails in the first test and then bribes and passes in the second test would correspond to $H5$. The sequence of potential histories continues to infinity, since a vehicle can get retested as many times as the car owner chooses. However, as I explain below, the first few histories are sufficient to estimate the infinite horizon model, and I do not need to consider every potential history to estimate the parameters of interest. The possible histories in the two round model are the same as in the infinite horizon model except for the last one, which is given by

$H7$:$(s_1 = A,\ R_1 = 0)$,$(s_2 = A,\ R_2 = 0)$, $(s_3 = B,\ R_3 = 1)$

An expression for the probability of each of histories $H1$ to $H7$ for both the infinity horizon and the two round models can be derived from the probabilities for each decision and the probability for passing the test. For example, the probability of history $H4$ can be derived from the model to be:

$$\Pr(H4) = \Pr(s_1 = A) \cdot (1 - P) \cdot \Pr(s_2 = X),$$

where $\Pr(s_1 = A)$ is the theoretical probability of attempting without bribing in the first round, $1 - P$ is the probability of failing the test and $\Pr(s_2 = X)$ is the theoretical probability of postponing the test to the next smog check cycle. The model probabilities for each of remaining histories as well as the formula for the likelihood function are listed in the Web Appendix 2.

Because bribing is unobserved, it is impossible to distinguish between histories $H2$ and $H3$. In both cases the car owner will be observed to have a single try at the smog test and obtain a pass. Similarly, it is impossible to distinguish between histories $H5$ and $H6$. In what follows, I will refer to the two pairs of histories that we cannot observe separately as "confounded" histories and to the histories that are observed independently ($H1$, $H4$ and $H7$) as "unconfounded" histories. The methodological contribution of this paper is to show that all four unknown parameters are identified and, therefore, we can predict the probability of all seven histories despite observing confounded decisions.

The full proof of identification is outlined in detail in the Web Appendix 3. Here I discuss briefly the main sources of identification and provide some insights on how the model pins down the most important parameters of the model. Broadly speaking, the identification of all structural parameters comes from being able to observe official failure rates (and official missed test rates) in the first two attempts to pass the smog check, conditional on also observing the true probability of passing. Given that the cost of the test is different in odd and even attempts, the failure rates in first tests and retests provide distinct information about the four underlying parameters of the model. The next paragraph focuses on providing intuition about how the bribe and the time cost parameters are separately identified in the model. The source of identification of the two remaining parameters is more obvious, since missed tests help pin down δ (the time discount times probability of getting caught parameter) and knowledge on the constant c helps pin down the scale parameter, τ. For details, we refer the reader to the formal proof in the Web Appendix 3.

Even though the model in Section 5.1 is estimated using Maximum Likelihood, identification is more intuitive from identifying the key moments that are used to solve for the model parameters. The intuition for the identification of the bribe and time cost parameters, β and ω, is as follows: Start by assuming that both, the bribe paid, β, and the time cost, ω, are constant across car owners and that we can identify a group of vehicles that have the same known probability of passing, P'.[18] We can observe the failure rate in the first test among this subclass of vehicles, which will equal the probability that they test honestly (without cheating) times the probability of failure, $\Pr(s_1 = A|P')(1-P')$.[19] This rate is informative about the size of the bribe and the time cost: First, $\Pr(s_1 = A|P')$ is increasing in the amount of the bribe, β, since higher bribes will deter car owners from cheating. Second, $\Pr(s_1 = A|P')$ is decreasing in the time cost, ω, since a large opportunity cost of time means getting to the back of the line upon failing the test is very costly.

Hence, a low failure rate (conditional on the probability of passing) is also consistent with a high time cost. If we only observed the official rate of failure for the first test and for a single group of vehicles with common probability of passing, we would not be able to separately identify these two parameters. This identification issue is illustrated by Panel A of Figure A2 in the Appendix. This figure is a three dimensional surface that plots the theoretical probability of failing the first test, $\Pr(s_1 = A|P)(1 - P)$, for $P = 0.4$ as a function of the bribe and the time cost parameters. This figure also plots two level curves on the bribe-time cost plane, which represent bribe-time cost combinations that would result in a low probability of observed failure (dark level curve) and a high probability of observed failure (light level curve). An identification issue would arise if we only observed the failure rate in the first test for a group of vehicles with a common probability of passing the test. For instance, take the low probability of failure case. Since multiple combinations of the bribe and the time cost levels would be consistent with this low observed level, we would not be able to separately identify the bribe and the time cost parameters.

Panel B of Figure A2 in the Appendix illustrates how observing the failure rates for the retests as well as for the first test helps overcome this identification issue. Because the fee for even tests, including the first retest, is zero, the relationship between failure rates in even tests and the model parameters is different than in

[18]The estimated model allows time cost to differ across vehicle owners by having a component that is proportional to the value of their vehicle.

[19]In terms of the probabilities listed in Web Appendix 2, the fraction of failed first tests among vehicles with probability of passing P' is given by the sum of $\Pr(H4|P')$ to $\Pr(H7, H8, ...|P')$, which equals $\Pr(s_1 = A|P')(1 - P')$.

odd tests. The theoretical relationship is illustrated by Panel B of Figure A2 in the Appendix. Notice that this figure contains the same surface depicted in Panel A, but adds an additional surface which represents the probability of failure in a retest, $\Pr(s_2 = A|P)(1-P)$, for $P = 0.4$ as a function of the bribe and time cost parameters. The fact that these two surfaces intersect contributes to solving the identification issue posed in the previous paragraph. To illustrate this point, imagine the observed failure rate in the first test can be mapped to bribe-time cost combinations corresponding to the dark level curve in Panel A, and the observed failure rate in the retest can be mapped to the light level curve in Panel B that intersects it at bribe = 450 and time cost = 220. The intersection of these two level curves pins down the bribe and time cost parameters in the model.[20]

6 Results from estimation

6.1 Predicting the probability of passing

As explained in Section 5.2 the identification of the structural model relies on observing the distribution of the probability of passing for the car fleet. If the smog check center data were reliable, I could approximate the distribution by using the empirical distribution of pass/fail outcomes. However, pass rates in the data are confounded by cheating; hence, the empirical distribution strongly overstates pass rates. As an alternative, I estimate a mapping between car characteristics and the probability of passing the emissions tests using only tests from centers that have low evidence of cheating, identified in Section 4. More specifically, after sorting centers in 10 groups, where the last group has the highest evidence of donor car cheating, I estimate the mapping using all centers in the first group. The parameters estimated from this

[20]The reason why the two surfaces representing the decision to test without cheating intersect is two-fold. First, for all levels of the bribe and time cost, individuals are more likely to bribe in even tests than in odd tests. This is because risking test failure in an odd test is cheap as the retest is for free, while risking test failure in an even test is expensive because of the non-zero smog-check fee. This explains why the probability for the first round test decision, which is one of the alternatives to bribing, lies above the probability of the second round test decision for low values of the bribe and time cost parameters. Second, for high values of the bribe and time cost parameters, individuals are more likely to skip the smog check requirement all together and fall into non-compliance status. Individuals are more sensitive to these additional test costs when the smog check fee is high than when the smog check fee is zero; hence the surface representing the decision to test without cheating is steeper in the first test attempt.

subsample of tests are then used to predict the probability of passing for cars in the remaining centers.[21] Predicted probability of passing the test is close or slightly above observed passing rates for the centers with low cheating evidence (first and second deciles of the cheating test), and up to six percentage points below centers with medium to high evidence of cheating. This is consistent with cheating centers having passing rates that are abnormally high.[22]

While using tests from non-cheating centers significantly reduces the chances that the probability of passing is overstated, self-selection by car owners into cheating centers based on their own tendency to cheat may lead to omitted variable bias in the estimates. Specifically, if the owners of high emitting cars self-select into cheating centers, there will be an over-representation for low-emitting cars in non-cheating centers. This type of self-selection may cause the estimated mapping to overstate the predicted probability of passing.

The empirical evidence on self-selection into cheating and non-cheating centers is mostly as expected. There is evidence of selection of older, large engine and highly driven vehicles into low-cheating centers. However, most differences are small and there is a substantial amount of overlap in characteristics of cars at cheating centers and non-cheating centers. If unobserved determinants of the probability of passing follow a similar selection pattern, the predicted probability of passing will be biased upward, although the bias is unlikely to be very large. Upward bias in the probability of passing will generate downward bias in the simulated prevalence of cheating.[23] [24]

Although the test for cheating discussed in Section 5.1 is aimed at detecting "donor car" type of cheating, other forms of cheating in the no-cheating sample are likely absent in the non-cheating sample. This can be inferred from Figure 1 (and Web

[21] Results of the probability of passing estimation are available in the Web Appendix (Table WA1).

[22] Observed passing rates and predicted probability of passing by cheating test decile are available in the Web Appendix (Table WA3)

[23] Descriptive statistics by cheating/non cheating status according to the cheating test are shown in Web Appendix to this paper as Table WA2.

[24] Selection is stronger for the first and second deciles of the cheating test distribution (descriptive statistics by all deciles of cheating test are available upon request). Hence, an alternative methodology uses the second and third deciles of the cheating test distribution to predict the probability of passing the test. This strategy reduces the self selection problem but also results in noisier predictions of the probability of passing, since centers in these groups may have some cheating. The results from this alternative methodology (available on request) yield a lower average predicted probability of passing (0.74 instead of 0.75) and a higher predicted cheating rate (14 percent instead of 9 percent). This is consistent with self-selection leading us to overestimate the probability of passing the emission test and underestimate cheating in the structural model.

Appendix Figures WA2 and WA3) which shows that for centers with high p-values in the cheating test, car characteristics are better predictors of emissions. We would not observe this pattern if other types of cheating were substituting the use of donor cars in high p-value centers. In addition, low discrepancies between predicted probability of passing rates observed passing rates for centers with high p-values for the donor-car cheating test seem to suggest that these centers have low cheating over all.

6.2 Parameter estimates and the prevalence of cheating

The results from estimating the infinite horizon model for the second half of 2003 are presented in Table 4. Because the numerical maximization of this model is computationally intense, the model was estimated with a five percent random sample of all non-exemptible vehicles (vehicles that cannot exempt the driving restriction) with rational result histories (*e.g.* no retesting after passing), consistent car characteristics across rounds and non-missing values for car characteristics and test outcomes. Non-exemptible vehicles are older and have higher emissions than exemptible vehicles.[25] The sample was also restricted to vehicles that do not switch centers between the first test and the retest. Recall that the identification of the model relies on the difference in costs between the first and second tests. However, if a vehicle owner decides to go to a different center after the first failed test, he loses the retest discount, and therefore the model cannot be identified from these observations since there is not price difference between the two tests. About one percent of vehicles appear to have changed centers after failing the first test, which cannot be explained by the model. The final sample used for the estimation has 17,659 vehicles.

The parameter estimates are presented in Panel A (Column [1]) of Table 4. Time cost, ω, is modeled as a linear function of the approximate log-value of the car from 2003 newspapers: $\omega = \omega_0 + \omega_1 \cdot \ln(\text{car value})$. To the right of the parameter estimates, I report the standard errors from the numerical Hessian of the likelihood function. These errors are not corrected for the first stage estimation of the probability of passing (see footnote 26). Notice that time cost varies positively with the price of the car and has some variation across periods. Column [3] of Table 4 computes the minimum, mean and maximum values of the time cost implied by the model parameters. The time cost for each period varies between 106 and 151 pesos (7.5 U.S.

[25]A decision model for exemptible vehicles would have to incorporate the incentive to pass the exemption threshold: driving one more day a week. This model is left for future work.

dollars and 14 U.S. dollars respectively). However, notice that the point estimates for time cost parameters are not very precisely estimated.

The equilibrium bribe is 193 pesos (18 U.S. dollars), which is a plausible value according to newspaper articles on the matter (50 to 400 pesos). The estimate for the parameter δ is, 0.61. Recall that this parameter is the combination of a time discount rate, given that the fine is usually paid a few months after the deadline expires, and the probability of paying the fine, f, for missing the test on one smog check cycle missing a test. Officially, missing a test is not easy since the car owner has to show a proof of fine payment in order to get their vehicle smog-checked the next period; and further avoiding the smog-check may result in multiple traffic tickets. However, the large percentage of car owners that miss a test, which results in a relatively large δ, suggests that there might be a way around paying the fine.

Finally, the estimate for the standard deviation of the random shock implied by the estimate of τ is somewhat large: 60.11 (5.5 U.S. dollars). This may be a result of the simplicity of the model, since it does not account for much of the heterogeneity across individuals. It could also reflect an overly optimistic option value for the future choices: a larger variance of the error term in (4) and (5) leads to a higher probability that a large and positive shock will appear in any of the future choices.

Panel B of Table 4 shows the fit of the model as well as the predicted probabilities for each history from H1 to H7. The fit of the model is in general quite good since all outcomes are predicted within 2.5 percentage points of accuracy and most of them within one percent. The fit is best for the probability of postponing, H1 and H4, and worse for confounded histories, H5 and H6. The total cheating prevalence rate estimates from this method can be approximated by the sum of predicted histories H2 and H5, which yields 9.6 percent.

A way to assess the size of the bias in the structural model estimate for the rate of cheating is through a simple calculation: Suppose we knew that the probability of passing the test for those vehicles that decide not to cheat is equal to $\tilde{P}$. Then, the observed passing rate would be a weighted average of one and $\tilde{P}$ where the weights correspond to the percentage of cheaters and the percentage of non-cheaters. If we approximate the probability of passing the test of non-cheating vehicles, $\tilde{P}$, as the observed passing rate in non-cheating centers, 0.759, we can then approximate the rate of cheating by solving for Pr(*cheating*) in the following equation:

$$0.786 = \Pr(cheating) + (1 - \Pr(cheating)) \times 0.759$$

where the observed passing rate is equal to 0.786. According to this simple calculation, $\Pr(cheating) \approx 0.109$. Hence, the downward bias in the structural model estimate for cheating resulting from selection of clean cars into non-cheating centers appears to be small.

The results for the two period version of the model are presented in Table 5. This model is more restrictive, since it assumes that any owner that attempts and fails the test two times chooses to cheat and pay a bribe in the third round. Not surprisingly, the fit of this model is less good than the infinite horizon model; yielding differences between predicted and observed outcomes of up to 6 percentage points. However, the parameter estimates for the bribe and the predictions for cheating are similar to the infinite horizon model: 231 and 8 percent respectively. This model is less computationally intensive than the infinite horizon version of the model. This makes it feasible to show consistent standard errors from bootstrapping, which account for the estimation error in the passing probabilities.[26] The standard errors that incorporate the estimation error in the predicted probability of passing are larger than the errors from the inverted numerical hessian. The bribe and the discount factor parameters (β and δ) and are more precisely estimated than the time and scale parameters ($\omega_0,\ \omega_1$ and τ).

6.3 Implied willingness to pay for car maintenance

Cheating interferes with the objectives of the regulatory policy by lowering the willingness to pay (WTP) for having cleaner vehicles. The regulation of vehicle emissions will not be binding if the cost of not complying with the regulation is low. Other aspects of Mexico City's regulation may also reduce the incentives for owners to either

[26] In order to compute bootstrapped standard errors, I assume individuals decide on a center and on cheating behavior jointly based on imperfect information of center's cheating practices. Furthermore, I assume that some centers are inherently more corrupt than others. For each bootstrapped sample of vehicles, I repeat the cheating test at the center level and identify the centers in the first decile of the cheating test chi-squared statistic, which correspond to the centers with the lowest evidence of cheating. I use this sample to estimate the probit model for the probability of passing the test and predict the probabilities of passing for the remaining vehicles. Finally, I estimate the structural model using the observed histories of the sampled vehicles with no access to an exemption.

perform car maintenance or to buy new cars. These include the possibility of retesting indefinitely at a low cost (Wenzel 2004) and also having a low fine or equivalently, a low probability of paying a fine for not passing the emissions test.

This section extends the model in Section 5 to approximate the WTP for car maintenance implied by a given value of the parameters. The model in Section 5 does not incorporate maintenance as an option available to the car owner. This was justified on the grounds of the extended corruption evidence found in Section 4 and of indirect evidence from same-day retests. The model estimates, however, can be used to approximate the WTP for pre-smog-check car maintenance. This calculation is important for two reasons. First, it will provide a test of internal consistency for the model under the current policy parameters: if the WTP for car maintenance implied by the model is above the cost of maintenance for a substantial proportion of the sample, then the model will violate internal consistency since the maintenance option would be preferred by several car owners according to the model predictions. Second, it will allow the evaluation of changes in policy aimed at increasing the number of vehicles that get repaired. The policy simulations will be explained at length in Section 6.4.

The key feature of the model that allows for this extension is that we can calculate the estimated expected cost of facing the smog check regulation for each individual for a given value of the parameters by solving numerically for the fixed point of the system of equations in (6). Because utility is expressed in "dollar" terms (more accurately, in peso terms), the expression for $\mathbb{E}_{t-1}(V_t^{odd})$ given in (6) is the expected cost facing the smog-check requirement given the best possible response of the car owner to the available options, which include retesting an infinite amount of times, bribing and postponing (or not complying with) the requirement. This expected cost is increasing in the price of the bribe, the fine and the cost of the test, and is decreasing in the probability of passing the test.

The estimation of the WTP for car maintenance proceeds in two steps. First, I assume that the car maintenance decision is made before the set of decisions modeled in Section 5: car owners evaluate the expected cost of facing the smog-check requirement with and without maintenance, and perform car maintenance only if car maintenance saves them money. When computing the expected cost without car maintenance, the owner takes into account all costs modeled in Section 5. Second, I make a few necessary assumptions about the costs and benefits of car maintenance that are based on documented values. I assume that owners who perform maintenance on their vehicles will pass the the emissions test with 90 percent probability in the two smog check cycles following the maintenance (a full year). I assume that performing maintenance also improves the vehicle's gas mileage, and that the gas mileage is improved more for high-emitting cars than for low-emitting cars.[27] Given the above assumptions, the six-month equivalent of the willingness to pay for car maintenance is equal to the change in expected value from improving the probability of passing to 90 percent in one smog-check cycle, plus the improvements in gas mileage:

$$\begin{aligned} WTP_i = \; & \mathbb{E}_{it-1}\left(V^{odd}|\hat{\omega}_0, \hat{\omega}_1, \hat{\beta}, \hat{\delta}, \hat{\tau}, 0.9, y_i\right) + g(\hat{p}_i(\mathbf{x}_i), \mathbf{x}_i) \\ & -\mathbb{E}_{it-1}\left(V^{odd}|\omega_0, \hat{\omega}_1, \hat{\beta}, \hat{\delta}, \hat{\tau}, \hat{p}_i(\mathbf{x}_i), y_i\right) \end{aligned}$$

where $\hat{p}_i(\mathbf{x}_i)$ is the predicted probability of passing according to vehicle characteristics, y_i is the log-value of the vehicle, and $g(\hat{p}_i(\mathbf{x}_i), \mathbf{x}_i)$ are the approximate 6-month gas mileage benefits from a car tune-up. The WTP is compared by each car owner to the six-month equivalent of the tune-up cost, which I assume to be 400 pesos based on the prices for car tune-ups in Mexico City reported in the consumer survey of 2005. If the WTP is higher than the cost, the car owner performs maintenance on her vehicle.

The first column of Table 5 shows that the model estimated in Section 5 is internally consistent. The first column of Table 5 shows that the fraction of people that are predicted to perform car maintenance according to the model extension and the parameters estimated in Section 5 is 0.6 percent. In the next section, I adjust the cost of the test, the equilibrium bribe, and the fine in order to evaluate the extent to which these changes increase the number of owners that find maintenance to be preferable, and I also evaluate whether these changes are cost-effective ways to improve air quality.

6.4 Policy evaluation

A crackdown on corruption through either increased enforcement or increased fines to cheating smog-check centers should result in a higher equilibrium bribe price. A high

[27] I assume all vehicles with 50 percent or less probability of passing the test receive a 4 percent reduction in their gas mileage due to maintenance and that the reduction falls linearly with the probability of passing until reaching 0 for vehicles that pass the test with certainty. The 4 percent average reduction in gas mileage is taken from www.fuel economy.gov for vehicles that have failed a smog-check.

enough price of the bribe can presumably eliminate corruption entirely. However, emissions would only be reduced if car owners that would cheat under a low bribe price, decide to perform car maintenance under a high bribe price.

The model proposed in Section 5.1 and the extension discussed in the previous section allow us to predict the individual response to an increased bribe. The extended model allows individuals to decide whether or not they want to perform car maintenance to increase the probability of passing the test before they are faced with the postponing/bribing/testing decision.

In order to approximate the environmental benefits from eliminating corruption, I simulate owners' decisions under the minimum bribe price that reduces the number of cheaters to zero. Total emission reductions will be calculated as the sum of differences between current average emissions and after-maintenance emissions for all vehicle owners whose simulated decision is to perform car maintenance. My assumptions for after-maintenance emissions are explained below.

Notice that the number of owners that opt for actually reducing their emissions when facing the high bribe is most likely lower than the number of cheaters in the current scenario. Vehicle owners can also react to an increased bribe by accepting higher risks of failing the test and having to take a retest, or by choosing not to comply with the regulation and and risking paying the corresponding fine. The second column of Table 5 shows the number of owners that choose to perform car maintenance in response to an increased bribe price is indeed very low (0.4%). Most vehicle owners react to the increased bribe by accepting a higher risk of failing the test. This is unsurprising given that the cost of every other retest is zero. Hence, the only cost incurred by retesting once is the time cost.

The third column of Table 5 presents the results of an alternative set of policies aimed at reducing the appeal of all non-maintenance options. The policies include equating the cost of the odd and even retests (both are now 175 pesos), increasing the minimum fine for non-compliance to 1,800 pesos and further increasing enforcement such that the equilibrium bribe price is increased to 700 pesos. Under this set of policies, about four percent of owners decide to reduce their emissions by performing car maintenance. For this set of policies to be cost-effective, the additional costs faced by all car owners must be lower than the benefits of emission reductions. The rest of this section will estimate the benefits to society from reduced emissions as well as the costs of these policies.

The calculation of the benefits from reducing emissions is done in four steps. First, I estimate emission reductions due to increased maintenance by computing the difference between emissions after maintenance and predicted emissions without maintenance. Predicted emissions are based on the sample of low-cheating tests. Emissions after maintenance are assumed to be the average emissions of all vehicles whose predicted probability of passing the test is above the 90th percentile. Since true emissions are not available, I use predicted emissions from 2003 low-cheating centers for both 90th percentile averages and owners who decide to perform maintenance.

Second, I use the fuel-based methodology for emission inventories proposed by Singer and Harley (1999), and applied by Schifter et al. (2004) for Mexico City, to convert emission concentrations at vehicles' exhaust, which is recorded by the smog check readings in parts per million (ppm), to total emission tons released into the environment. This methodology uses fuel economy, miles driven and pollutant densities. The results of this calculation are given in Column 1 of Table 6.

Third, I follow Small and Kazimi (1995) in attributing ambient concentrations of air pollutants to vehicle emissions in tons. Health costs from pollution have been estimated for ambient concentrations of air pollutants such as particulate matter (PM_{10}) and ozone (O_3), but have not been estimated for NO_x and HC, which are the pollutants measured in the smog check readings. However, PM_{10} and O_3 are partially formed from NO_x and HC, so following Small and Kazimi (1995), I use the elasticities of conversion from the literature (*e.g.* Charlson and Wigley, 1994) to estimate the concentration of air pollutants due to emissions in tons of primary pollutants. Also, I follow Song et al. (2010) to link O_3 to HC emissions. They estimate that one percent increase in volatile organic compounds (VOC) emissions (HC are a sub-class of VOC) results in about a half percent increase on average O_3 ambient concentration. These calculations are in Column 2 of Table 6.[28]

Fourth, I calculate the health costs of emissions by summing up mortality costs from the different pollutants. I use Schwartz' (1994) estimates of mortality from PM_{10} concentration and Bell et al. (2004) estimates from O_3-related mortality. In the case of CO, I use the direct health impacts on infant mortality from Neidell and Currie (2005). To calculate the mortality costs in dollars, I use the value of statistical life used by Molina and Molina (2002) for Mexico City. The lives-per-ton calculations

[28]In their work within the MILAGRO project, they find that, contrary to what was previously believed (e.g. Molina and Molina, 2002), ozone formation in Mexico City is VOC-sensitive.

implied by these numbers are in Column 3 of Table 6

Table 6 summarizes the benefit calculations from the set of policies proposed in Column 3 of Table 6. Column 1 is the result of steps 1 and 2 above: it is the sum of emission differences in tons per year across all vehicles that are simulated to opt for car maintenance. Column 2 corresponds to step 3 above: it presents the factors needed to convert a ton of primary emissions into its contribution to ambient concentration of air pollutants computed using Small and Kazimi's (1995) methodology for PM_{10} and Song et al. (2010) conversion rates. Column 3 corresponds to step four: it shows the mortality estimates from the literature for each atmospheric pollutant assuming a population of 20 million is exposed to the increase in concentration. Column 4 shows the deaths per ton of vehicle emissions and is the product of columns 2 and 3. Column 5 shows the value of statistical life taken from Molina and Molina (2002). Column 6 shows the cost in dollars of a ton of pollutant and is the product of columns 4 and 5. Columns 7 and 8 are the product of columns 1 and 4, and 1 and 6 respectively: they present the benefits of the set of policies considered in number of lives per year and in 2003 U.S. dollars respectively.

Notice that, although CO emission reductions are very large, the health benefits are dominated by changes in NO_x emissions. One one hand the emissions of NO_x gas from mobile sources constitute a large percentage of total NO_x emissions, and hence of PM_{10} concentrations. On the other hand, PM_{10} has, by far, the worst documented health consequences of any of the gases considered. In contrast, large health impacts from atmospheric CO have not been documented in the literature. Hence, I restrict to Neidell and Currie's (2005) estimate for infant mortality due to CO for the cost calculations.

The set of policies considered has at best a modest impact in air pollution, especially when compared with the additional costs to car owners that they generate. The estimated emission savings from eliminating corruption, charging for retests, and doubling the fine (Column 6) are roughly equivalent to half a day of vehicle emissions in Mexico City. The effect on total emissions is small because the proportion of vehicles that respond to the policy change is only four percent of the non-exemptible car fleet. The emission reductions from this set of policies would save approximately four lives per year (Column 7), which are equivalent to 1.5 million dollar savings. In contrast, the estimated cost for car owners of this set of policies amounts to 3.7 million dollars. This cost can be estimated by adding up the increases in the expected cost of facing the test across the entire car fleet. About 2.2 million of this additional cost is a transfer to smog-check centers: retests are costly under the proposed policy. The rest of the cost is in the form of additional time costs from increased retesting among individuals that do not opt for car maintenance and unobserved costs represented by the random shocks to utility.

It is worth noting the likely consequences of some of the simplifying assumptions that were necessary for the calculation. One important assumption is that maintenance costs are equal across vehicles. However, the actual costs of car maintenance aimed at reducing emissions may be lower than the cost of tune-ups for some vehicles (*e.g.* some vehicles may substantially increase their probability of passing the test by changing the air filter, replacing the spark plug wires, etc.) and larger for others (*e.g.* if major repairs are needed to meet the emission standards and these are more costly than tune ups). The net effect of varying maintenance costs on the benefits is ambiguous, since it could be the case that some vehicles with high maintenance costs (and potentially high emissions) are miss-allocated into performing car maintenance and some additional vehicles with low maintenance costs (but potentially relatively low emissions) may opt for car maintenance.

Also, predicted emissions used in this calculation were identified by using the low-cheating centers identified in Section 4.1 and the upward bias related to the selection problem discussed in Section 6.1 might also be present for predicted emissions. More specifically, self-selection into non-cheating centers might result in an underprediction for vehicle emissions. This would cause a downward bias in the calculated benefits from the proposed set of policies. However, notice that for the set of policies to be cost-effective according to the calculations in Table 5, the downward bias would have to be of more than half.

7 Conclusions

Researchers and governments have questioned the effectiveness of smog checks in reducing vehicle emissions. They have cited repeated testing and cheating as potential explanations for why vehicles with high on-the-road emissions have been able to pass the emissions test. This paper uses indirect evidence to show that cheating is a wide spread practice in Mexico City. I develop a test for cheating relies on detecting serial correlation patterns in consecutive emissions generated by the use of donor cars.

This test predicts that 79 percent of centers have engaged in donor car use, a cheating practice that involves using emissions from a clean car to substitute for emissions of a cheater.

The test for cheating is also an input for the estimation of a structural model of car owner decisions that recovers the underlying parameters of the cheating decision and is used to simulate individual responses to the smog check requirement. Although cheating decisions are unobserved, the parameters of the model can be recovered without any explicit information on cheating decisions. The model's identification relies on the difference in costs between odd and even retests, and on observing the distribution of the probability of passing the test.

The maximum likelihood estimation of the model yields an estimate for the bribe amount of about 20 U.S. dollars. This estimate is within the range of bribes that has been reported in newspapers. The simulations of individual decisions suggest that about 9 percent of car owners choose to cheat on the smog check. Because cheating is an alternative to car maintenance, and the price of the bribe is relatively low, the model suggests that incentives for car maintenance are very low or non-existent.

An extension to the model further allows to estimate the benefits and costs from boosting incentives for car maintenance through plausible policies such as increased enforcement and higher retesting costs. These combined policies are predicted to induce car maintenance in 4 percent of the vehicles. The resulting emission reductions are equivalent to less than one day of Mexico City traffic a year. However, the emission reductions come at a high cost for the entire car fleet: smog check-costs for car owners increase by about 3.7 million per cycle. These calculations suggest that, forcing car owners to pass smog checks twice a year is not a cost effective policy for reducing vehicle emissions in Mexico City.

Table 1: Descriptive statistics

Panel A: Vehicle characteristics	First Test	All Tests	Panel B: Test outcomes (First Test)	Mean/SD	95th Pctile/Max
Model-year					
Older than 1970	0.008	0.009	HC40	57.01	191
1970-1979	0.051	0.059		[168.74]	9884
1980-1984	0.065	0.074	HC24	51.06	187
1985-1989	0.080	0.091		[147.64]	9992
1990-1994	0.233	0.258	CO40	0.45	2.12
1995-1999	0.260	0.242		[0.91]	14.66
2000-2003	0.303	0.265	CO24	0.44	2.12
				[0.90]	17.21
Size			NO40	445.50	1595
VW Sedan and Chevy	0.243	0.260		[597.40]	9964
Mini-Compact	0.461	0.446	NO24	324.57	1172
Compart-Medium	0.100	0.100		[449.34]	9958
Medium-Large	0.036	0.036	O2_24	0.63	2.50
Sport	0.038	0.035		[1.00]	28.30
Minivan	0.037	0.032	O2_40	0.54	2.20
Van	0.008	0.008		[0.87]	31.30
Pick Up	0.078	0.084	Pr(Pass)	0.90	1.00
				[0.30]	1.00
			Odometer (km)	106,381	114,653
				[148,243]	157,200
			Temperature (ºC)	21.98	22.06
				[7.95]	7.68

Notes:

1. Source Smog Check Center data for 2003.

2. The total number of tests in this period was 2,126781. First time checks (or number of vehicles checked in 2003) is 1,513,111.

3. Panel A indicates proportion of vehicles in each category variable, except for the last two variables, which indicate mean (main row) and standard deviation (below main row) of odometer reading and air temperature when the test was taken.

4. Panel B indicates mean and standard deviation (in brackets) in first column. The second column indicates the 95th percentile (main row) and maximum value (below).

Table 2: Example of a suspicious sequence of emission readings in a single lane.

Time	Model	Number of Cylinders	Engine Displacement	HC 24kph	HC 40kph	CO 24kph	CO 40kph	NO 40kph	NO 24kph	O2 40kph	O2 24kph
11:04:37	1988	4	1800	43	127	0.15	0.59	107	162	1.2	1.0
11:08:55	2000	8	4600	33	25	0.13	0.12	2	3	1.1	1.2
11:17:06	1971	4	1600	34	32	0.20	0.19	80	86	1.1	1.0
11:26:02	1993	6	3100	1	2	0.12	0.12	5	2	1.1	1.1
11:33:52	1998	4	2000	10	10	0.13	0.13	10	10	1.1	1.1
11:38:19	1997	6	3100	0	0	0.00	0.00	0	0	0.0	0.0
11:51:05	2000	4	1400	4	0	0.06	0.05	112	10	1.2	1.2
11:54:45	1980	4	1600	18	30	0.16	0.22	26	52	1.1	1.0
12:05:35	1999	4	1600	21	16	0.16	0.15	3	1	1.0	1.0
12:07:53	1981	4	1600	27	29	0.16	0.19	39	56	1.1	1.7
12:16:38	1988	4	2200	31	37	0.20	0.28	52	68	1.0	1.1
12:24:23	1987	4	1800	59	20	0.32	0.17	39	18	1.4	1.0
12:28:15	2001	4	1600	0	0	0.00	0.00	0	0	0.0	0.0
12:39:48	1975	4	1600	72	84	0.46	0.89	31	38	2.4	4.0
12:41:56	1997	6	3100	12	12	0.15	0.16	4	8	1.0	1.0
12:47:08	1984	8	5000	8	11	0.17	0.17	3	3	1.0	1.0
12:57:58	1992	4	1600	26	28	0.16	0.17	12	13	1.0	1.0
13:03:23	1998	4	1600	39	39	0.22	0.22	23	17	1.0	1.0
13:14:21	1968	4	1600	63	78	0.38	0.53	63	58	1.2	1.6
13:19:40	1991	6	3100	19	17	0.22	0.22	1	2	0.9	0.9
13:26:01	1994	4	1800	22	23	0.23	0.22	9	9	0.9	0.9
13:34:12	1992	4	1600	31	30	0.23	0.23	8	9	0.9	0.9
13:39:53	1990	4	2300	26	26	0.23	0.23	5	4	0.9	0.9
13:50:17	1977	6	3700	27	27	0.23	0.23	7	5	0.9	0.9
13:57:51	1984	4	2000	80	118	0.49	0.86	72	77	0.9	0.8
14:04:17	1985	4	2200	87	163	0.26	0.32	174	167	1.0	0.8
14:15:36	1993	4	2500	38	45	0.14	0.18	28	31	0.9	0.9
14:20:49	1987	6	2800	41	56	0.15	0.24	26	91	0.9	0.9
14:29:25	1991	4	2300	34	34	0.14	0.14	22	13	0.9	0.9
14:40:21	1985	6	3800	109	106	0.46	0.44	91	104	0.9	0.8
14:44:17	1993	4	1600	47	35	0.15	0.14	34	14	0.9	0.9
14:53:49	1978	8	5000	1061	97	0.30	0.14	141	39	3.2	0.9
14:59:16	1994	4	1600	28	23	0.14	0.13	10	14	0.9	0.9
15:33:34	1990	4	2500	137	196	0.55	0.46	1458	1000	0.9	1.2
15:50:02	1986	4	1600	358	359	3.82	3.33	1107	1128	0.5	0.5

Notes:
1. Source: Fragment of time-ordered smog-check center data from 2003, 1st half D.F.

Table 3: Testing for cheating at the center level

Center	[1] HC	[2] NO	[3] CO	[4] O2	[5] Number of tests sampled	[6] Chi-squared	[7] P-value
Panel A: Centers with highest p-value for the joint significance test							
1070	-0.0304	0.0078	0.0345	0.0235	356	1.55	0.82
	[0.049]	[0.049]	[0.045]	[0.040]			
9029	0.0040	0.0099	-0.0007	0.0126	4,500	2.00	0.74
	[0.006]	[0.012]	[0.012]	[0.012]			
9081	-0.0044	-0.0181	-0.0054	0.0010	4,500	2.67	0.61
	[0.018]	[0.012]	[0.012]	[0.013]			
9053	-0.0082	0.0179	0.0036	0.0076	4,500	3.05	0.55
	[0.012]	[0.013]	[0.010]	[0.013]			
1069	0.0444	0.0077	0.0342	0.0059	1,070	3.21	0.52
	[0.037]	[0.025]	[0.024]	[0.024]			
9018	-0.0039	0.0034	-0.0065	0.0227	4,500	3.25	0.52
	[0.013]	[0.013]	[0.011]	[0.014]			
9034	0.0150	0.0102	0.0029	0.0060	4,500	3.56	0.47
	[0.009]	[0.012]	[0.010]	[0.010]			
9059	0.0235	0.0170	-0.0051	0.0071	4,500	5.41	0.25
	[0.013]	[0.012]	[0.013]	[0.012]			
9066	-0.0031	-0.0119	-0.0032	0.0277	4,500	5.57	0.23
	[0.008]	[0.013]	[0.011]	[0.013]			
9004	-0.0062	0.0120	0.0159	0.0178	4,500	5.73	0.22
	[0.009]	[0.011]	[0.013]	[0.012]			
Panel B: Centers with the lowest p-value for the joint significance test							
9011	-0.0024	0.1247	0.0064	0.1134	4,500	174.55	0.00
	[0.015]	[0.013]	[0.013]	[0.012]			
9009	-0.0022	0.1048	-0.0004	0.1277	4,500	176.22	0.00
	[0.012]	[0.013]	[0.013]	[0.012]			
9077	0.0013	0.0228	-0.0108	0.1795	4,500	201.39	0.00
	[0.017]	[0.012]	[0.013]	[0.013]			
9028	0.0165	0.0065	-0.0210	0.1871	4,500	205.74	0.00
	[0.012]	[0.012]	[0.013]	[0.013]			
9052	0.0429	0.1052	0.0280	0.1665	4,500	218.86	0.00
	[0.016]	[0.013]	[0.012]	[0.013]			
9010	0.0039	0.0374	0.0045	0.1870	4,500	230.72	0.00
	[0.020]	[0.013]	[0.014]	[0.012]			
9035	0.0496	0.1853	-0.0041	0.1080	4,500	251.47	0.00
	[0.015]	[0.014]	[0.013]	[0.012]			
9027	0.0002	0.2301	0.0228	0.0410	4,500	281.76	0.00
	[0.017]	[0.014]	[0.012]	[0.013]			
9055	0.2312	0.0313	0.0565	0.1244	4,500	336.82	0.00
	[0.014]	[0.012]	[0.012]	[0.013]			
9038	0.0075	0.0389	0.0399	0.3097	4,500	588.57	0.00
	[0.005]	[0.011]	[0.010]	[0.013]			

Notes:

1. Source: Seemingly unrelated regressions estimation by center (2003). Numbers in main rows are coefficients on lagged emissions.

2. Controls include brand categories, 8 size categories, 3 service categories, size and service interactions with model-year, model-year polynomial, mileage polynomial, lane (equipment) and center fixed effects, test results from contemporaneous tests in neighboring lanes and time of the day polynomial.

3. Column 5 indicates the number of tests sampled from each center during period 2003. All centers except for three have more than 4,500 tests for the period. The average number of tests per center is 23,000. When the number of tests is smaller than 4,500, not enough tests were found to balance the number of tests. Column 6 indicates the chi-squared statistic for the null hypothesis that all coefficients on lagged emissions are 0. Column 7 indicates the cumulative probability of the chi-squared distribution with four degrees of freedom evaluated at the chi-squared statistic.

Table 4: Infinite horizon model parameters and demand for bribes

Panel A: Parameter Estimates

Model Parameters	Estimate	SE (see note 3)		Implied estimates
	[1]	[2]		[3]
w-intercept	-18.32	56.1038	Mean time cost	105.70
w-slope	12.34	6.1120	Minimum time cost	81.26
b	192.67	27.8798	Maximum time cost	150.61
δ	0.61	0.1122	SE of random shock	97.52
τ	60.11	9.1528		

Panel B: Fitted and predicted probabilities for each history

Observed Histories	Actual	Fitted
	[4]	[5]
H1: Postpone	0.025	0.025
H2,H3: Bribe/No bribe - Pass	0.763	0.747
H4: No bribe-Fail-Postpone	0.000	0.001
H5, H6: No bribe-Fail-Bribe/No bribe-Fail-No bribe-Pass	0.150	0.175
H7: No bribe-Fail-No bribe-Fail	0.062	0.052

All Histories	Predicted
	[6]
H1: Postpone	0.025
H2: Bribe	0.067
H3: No bribe-Pass	0.680
H4: No bribe-Fail-Pospone	0.001
H5: No bribe-Fail-Bribe	0.029
H6 No bribe-Fail-No bribe-Pass	0.146
H7: No bribe-Fail-No bribe-Fail	0.052
Total Bribing in First and Second Tests	0.096

Notes:

1. Table 5 shows the results of the infinite horizon maximum likelihood estimation of the bribing behavior model developed in Section 6 using a 5% random sample (17,365 vehicles). Panel A shows the model's parameter estimates and Panel B shows the actual and fitted probabilities as well as the simulated probabiliites for all decision histories, including the unobserved ones.

2. The time cost,w, and the bribe, b, are in Mexican pesos (2003). The first 2 model parameters correspond to constant and a slope on the log price of the vehicle. The mean log price of the vehicle is 10.5. The right-most column in Panel A shows the minimum, maximum and mean values of the time cost given the intercept and the slope shown in the first column of Panel A. The right most column also shows the standard error of the random shock in the utility functions implied by the parameter esitmate of τ.

3. The standard errors in the second column of Panel A correspond to the squared root of the inverse numerical Hessian. These standard errors are inconsistent since they do not account for the fact that the probability of passing the test is a predicted value.

4.The top of Panel B shows Actual and Fitted mean probabilities for observed histories. The fitted probabilities correspond to the simulated averages given the estimated values of the parameters.

5. The bottom of Panel B shows mean Predicted probabilities for all histories up to H7. The percentage of tests with cheating can be computed by adding up the shares of vehicles that in histories H2 and H5, given as result 14 percent. The infinite horizon model includes everyone on subsequent histories (e.g. those with three attempts, those who bribe in the third period, etc.) in H7. So additional individuals that may bribe in subsequent periods are not included in the bribing estimation. However, because the probability of H7 is small (0.05), the bulk of bribers is still given by H2 and H5.

Table 5: Maintenance decisions for different policy regimes

	Actual: Bribe=$193 Fine=$875 Retest=$0	Policy 1: Bribe=$700 Fine$875 Retest=$0	Policy 2: Bribe=$700 Fine=$1800 Retest=$175
	[1]	[2]	[3]
Fraction with maintenance	0.0061	0.0104	0.0427

Notes:
1. This table shows the simulated proportion of individuals that perform maintenance to their vehicles under the current policy (Column [1]), a policy that eliminates cheating by increasing enforcement (Column [2]) and a policy that further increases the incentives for car maintenance by increasing the fine and the cost of retests (Column [3])

2.Individuals that choose to perform car maintenance are those for whom the expected cost of the test with car maintenance is lower than the expected cost without it (see Section 6.3).

Table 6: Benefit calculations from Policy 2 in Table 5.

	Sum of emission differences between predicted and after-maintenance emissions in tons[a]	Contribution to average concentration (in mg/m^3 or ppm) per ton of emissions[b,c]	Deaths per year per unit of concentration[d,e]	Deaths per ton of emissions	Value of life in $1000[f]	Cost of a ton of pollutant in U.S. dollars	Total number of lives saved per year	Total benefits from reduced emissions in $1000
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
NOx (PM10)	864	2.18E-05	104.400	2.28E-03	650	1,480.8	1.967	1,279
HC (PM10)	292	9.14E-06	104.400	9.55E-04	650	620.5	0.278	181
HC (O3)	292	7.34E-08	10.142	7.44E-07	650	0.5	0.000	0
CO	288	1.19E-06	129.640	1.54E-04	650	53.3	0.044	29
Total	**1,443**						**2.290**	**1,489**

Notes:

[a] The sum of total emissions corresponds to the sum of changes in emissions of all vehicles that would be submitted to car maintenace according to model predictions (see Section 6). The changes in emissions are calculated as the difference between predicted emissions and average emissions of (non-exemptible) vehicles with a probability of passing the test that is 0.9 or larger.

[b]Concentration for PM10 is defined in mg/m3, and concentration of O3 and CO is defined in ppm.

[c] The contributions to concentration for NOx (PM10) and HC (PM10) were calculated using Small and Kazimi (1995) conversion rates between primary pollutants (NOx and HC) and PM10. For HC (O3), I use the estimated elasticity of 0.52 between concentration of O3 in ppm and tons of VOC emissions from Song et al. (2010).

[d] The deaths from CO emissions are exclusively infant deaths.

[e] For NOx (PM10) and HC (PM10), these numbers are calculated using using Schwartz (1994) estimate of 0.5222 deaths per 100,000 people per unit of TSP. For CO, I use infant mortality estimates from Currie and Neidell (2005) (34 per 100,000 births). Estimates assume a population of 20 million and 381,288 births per year.

[f] The Value of Statistical Life is taken from Molina and Molina (2002).

References

[1] Bell, M.L., A. McDermott, S.L. Zeger, J.M. Samet, and F. Dominici., 2004, "Ozone and short-term mortality in 95 U.S. urban communities, 1987-2000", *Journal of the American Medical Association*, 292(19): 2372-2378.

[2] CAM/IMP (Comisión Ambiental Metropolitana / Instituto Mexicano del Petróleo), 2000, Auditoria Integral del Programa de Verificación Vehicular, Report Publisher by Comisión Ambiental Metropolitana and Instituto Mexicano del Petróleo. Mexico City, November.

[3] Charlson, R. J. and T. M. L. Wigley, 1994, "Sulfate aerosol and climatic change." *Scientific American*, 270 (February): 48-57.

[4] Currie, Janet and Matthew Neidell, 2005, "Air pollution and infant health: what can we learn from california's recent experience?", *Quarterly Journal of Economics*, 120(3)

[5] Davis, Lucas, 2008, "The Effect of Driving Restrictions on Air Quality in Mexico City", *Journal of Political Economy*, 116(1)

[6] Duggan, Mark, and Steven Levitt, 2002, "Winning Isn't everything: Corruption in Sumo Wrestling," *American Economic Review*, 92:1594-1605

[7] Fisman, Ray, 2001, "Estimating the value of political connections", *American Economic Review*, 91(4):1095-1102

[8] Glazer, A., D. Klein, and C. Lave, 1995, "Clean on Paper, Dirty on the Road: Troubles with California's Smog Check", The University of California Transportation Center, reprint No. 275.

[9] Glazer, A., Klein, and D. Lave, C., 1993, "Clean for a A Day: Troubles with California's Smog Check"; Report to the California State Senate Committee on Transportation; August

[10] Hubbard, Thomas A., 1998, "An empirical examination of moral hazard in the vehicle inspections market", RAND Journal of Economics, 29(2), 406-426

[11] Jacob, Brian A. and Steven Levitt, 2003, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating, Quarterly Journal of Economics, August

[12] Molina, M. and L. Molina, eds., 2001, Air Quality in the Mexico Megacity, Dordrecht: Kluwer Academic Publishers

[13] McFadden, D. and Domencich, T., 1996, "Urban travel demand. A behavioral analysis", North-Holland Publishing Co., 1975., Reprinted

[14] National Research Council, 2001. Evaluating Vehicle Emissions Inspection and Maintenance Programs. National Academy Press, Washington, DC.

[15] Olken, Benjamin, 2007, "Monitoring Corruption: Evidence from a Field Experiment in Indonesia", *Journal of Political Economy* 115, no. 2; pp. 200-249

[16] Olken, Benjamin, 2006, "Corruption and the costs of redistribution: Micro evidence from Indonesia", *Journal of Public Economics* 90, no. 4-5, pp. 853-870

[17] Padgett, Humberto, 2002, "Con propina todos pasan", REFORMA (Newspaper Article), October 01

[18] Rust, John, 1989, "A Dynamic Programming Model of Retirement Behavior" in *The Economics of Aging,* ed. by D.Wise. Chicago: University of Chicago Press, 359-398

[19] Singer, B.C. ; Harley, R.A., 1996, "Fuel-based motor vehicle emission inventory", Journal of the Air and Waste Management Association, Volume: 46; Issue: 6; Jun

[20] Secretaría de Medio Ambiente del Distrito Federal, 2004, "Elementos para la Propuesta de Actualización del Programa Hoy No Circula de la Zona Metropolitana del Valle de México", Gobierno del Distrito Federal/Dirección General de Gestión Ambiental del Aire

[21] Small, Kenneth A. and and Camilla Kazimi, 1995, “On the Costs of Air Pollution from Motor Vehicles”, Journal of Transport Economics and Policy, Vol. 29, No. 1, Motor Vehicle Air

[22] Schifter I., L. Díaz, J. Durán, E. Guzmán, O. Chávez and E. López Salinas, 2003, “Remote Sensing Study of Emissions from Motor Vehicles in the Metropolitan Area of Mexico City”, Environ Sci. Technol, *37*, 395-401.

[23] Schifter, I.; L. Díaz, V. Múgica, E. López-Salinas, 2005, “Fuel-based motor vehicle emission inventory for the metropolitan area of Mexico city”, Atmospheric Environment 39, 931–940.

[24] Shleifer, Andrei and Robert Vishny, “Corruption”, *Quarterly Journal of Economics*, vol. 108, no. 3, pp. 599-617, 1993

[25] Snyder, J. and L. Pierce, 2008, “Ethical Spillovers in Firms: Evidence from the vehicle emissions Testing Market”, Management Science, Vol. 54, No. 11, November 2008, pp. 1891-1903.

[26] J. Song, W. Lei, N. Bei, M. Zavala, B. de Foy, R. Volkamer, B. Cardenas, J. Zheng, R. Zhang, and L. T. Molina, 2010, “Ozone response to emission changes: a modeling study during the MCMA-2006/MILAGRO Campaign”, Atmos. Chem. Phys., 10, 3827–3846

[28] Wenzel, T., Singer, B.C., Gumerman, E.Z., Slott, R.S., Sawyer, R.F., 2000. Section III: Evaluation of the enhanced smog check program: a report to the California Inspection and Maintenance Review Committee. Final report submitted to the California Inspection and Maintenance Review Committee, LBNL-46117, Berkeley

[29] Wenzel, T., 2001. Evaluating the long-term effectiveness of the Phoenix IM240 program. Environmental Science and Policy 4, 377–389.

[30] Wenzel, Thomas p., Brett C. Singer, Robert R. Slott, and Donald H. Stedman, 2004, "Short-Term Emissions Deterioration in the California Phoenix I/M Programs." Transportation Research Part D 9, 107-124

[31] Zhang, Yi., Donald H. Stedman., Gary A. Bishop., Stuart P. Beaton., and Paul L. Guenther, 1996, "On-Road Evaluation of Inspection/Maintenance Effectiveness" Environmental Science & Technology, Vol. 30, No. 5, 1445-1450.

Appendix 1: Consistency of test for cheating

The statistical test from the regression methodology has null hypothesis:

$$H_0 : r_{it} = \tilde{r}_{it} = \mathbf{x}_i\beta + u_i,$$

and, according to (A1), alternative hypothesis:

$$H_1 : r_{it} = c_i\tilde{r}_{i-kt} + (1 - c_i)\tilde{r}_{it}$$

Under assumptions (A1) - (A3), the OLS coefficient $\hat{\gamma}_c$ from regression 3 is a consistent test statistic of null hypothesis.

Proof: Under the null hypothesis:

$$\mathbb{E}^*(r_{it}|\mathbf{x}_i, r_{i-1t-1}) = \mathbf{x}_i\beta + \mathbb{E}^*(u_{it}|\mathbf{x}_i, r_{i-1t-1})$$

Define $\mathbf{w}_i = [\mathbf{x}_i, r_{i-1t-1}]$ and $\mathbf{W} = [\mathbf{w}_2, \mathbf{w}_3, ..., \mathbf{w}_{N_1}, \mathbf{w}_{N_1+2}, \mathbf{w}_{N_1+3}, ..., \mathbf{w}_{N_1+N_2}, ...]$, where N_l indexes the number of tests in lane l; and $l = 1, 2, ..., L$, where L is the total number of lanes in all 80 centers. Then

$$\mathbb{E}^*(u_{it}|\mathbf{x}_i, r_{i-1t-1}) = \mathbf{w}_i\mathbb{E}(\mathbf{w}_i\mathbf{w}_i')^{-1}\mathbb{E}(\mathbf{w}_i u_{it}) \tag{9}$$

Equation 9 will equal zero under the null hypothesis if all elements of $\mathbb{E}(\mathbf{w}_i u_{it})$, *i.e.* $\mathbb{E}(r_{i-1t-1}u_{it})$ and $\mathbb{E}(\mathbf{x}_i u_{it})$, are zero. Under the null, $\mathbb{E}(r_{i-1t-1}u_{it}) = \mathbb{E}(\tilde{r}_{i-1t-1}u_{it})$. Hence, the first term equals zero because

$$\mathbb{E}(\tilde{r}_{i-1t-1}u_{it}) = \mathbb{E}((\mathbf{x}_{i-1}\beta + u_{i-1t-1})u_{it}) = \mathbb{E}(\mathbf{x}_{i-1}\beta u_{it}) + \mathbb{E}(u_{i-1t-1}u_{it}) = \mathbf{0},$$

by assumptions (A2) and (A3). The second term equals zero, $\mathbb{E}(\mathbf{x}_i u_{it}) = \mathbf{0}$, since $u_{it} = \tilde{r}_{it} - \mathbb{E}^*(\tilde{r}_{it}|\mathbf{x}_i) = \tilde{r}_{it} - \mathbf{x}_i\beta$.

Appendix 2: Robustness check 1: Permutations Test for Cheating

As mentioned in section 4.1, under donor car cheating, some consecutive tests are closer to each other than what we would expect from two randomly arriving vehicles. This alternative test acknowledges that some cheating-unrelated serial correlation between observed and unobserved determinants of vehicle emissions may occur, but constrains this correlation to occur within certain time and location windows. The permutations test assumes that, under a no cheating regime, car arrivals to smog check centers are random conditional on the center, lane, date and day-shift level (a day is considered to have 4 shifts). More formally, this is equivalent to assuming car arrivals are independent and identically distributed random draws from some unknown distribution within center-date-time blocks. Under the independent arrival assumption, shifting the observed order of the tests should not affect the distribution of differences between consecutive tests in the absence of cheating. I will test for the null hypothesis of no cheating by comparing the distribution of distances between consecutive tests in the order they occurred with several distributions of distances when the order is randomly permuted.

In order to measure the difference between consecutive tests, I consider all readings for each car, *i.e.* each pollutant in each of the different driving conditions. This amounts to 8 readings per test: HC at 24kmh, HC at 40kmh, CO at 24kmh, CO at 40kmh, $O2$ at 24kmh $O2$ at 40kmh, NO at 24kmh and NO at 40kmh. The measure of multivariate distance I will use is

$$d_t = \left(D_t V^{-1} D_t'\right)^{\frac{1}{2}},$$

where $D_t = (r^s_{1,t}, r^s_{2,t}, ..., r^s_{8,t}) - (r^s_{1,t-1}, r^s_{2,t-1}, ..., r^s_{8,t-1})$, $r^s_{j,t}$ is a transformed measure of pollutant j in test t[29], and V is the estimated variance covariance matrix across standardized pollutants. This multivariate distance measure is commonly known as Mahalanobis distance and is similar to the Euclidean distance except that differences of each pollutant are weighted not only by the variance of each pollutant but also by the covariance across pollutants. Hence, if pollutants HC and CO are positively correlated, a large difference in HC will have a smaller weight whenever the difference in CO is also large. Notice that t indexes tests in the order in which they occurred. The solid lines in Appendix Figure 1 show the distribution of d_t for two different smog-check centers in Panels A and B.

In the absence of cheating, the distribution of d_t should not change when changing the order in which tests occurred. Hence the no cheating assumption can be tested by comparing the observed distribution of d_t with distributions simulated by randomly permuting the data. To allow for common determinants of emissions across "nearby" vehicles, I will shift the order randomly within center-date-time blocks.[30] After permuting the data within the previously defined time-location window, I generate d_{t^n}, which is analogous to d_t except that it is obtained from a new order of the data, t^n, corresponding to permutation n. Panels A and B of Figure 1 show ten different dashed lines; each of which corresponds to the distribution of the d_{t^n} associated with a different random permutation of the data ($n = 1, ...,10$). The center depicted in Panel A of Appendix Figure 1 shows little evidence of fraud: permuting the order of tests generates distributions of d_{t^n} (dashed lines) that are very similar to the observed one: d_t (solid line). In contrast, the center depicted in Panel B shows strong evidence of corruption: the observed distribution of d_t has a larger proportion of small consecutive differences in tests compared to the distributions of $d_{t^1}, d_{t^2,}...d_{t^{10}}$, in which cars are assumed to arrive randomly to the center.

A formal statistical test can be derived from this methodology. First, I compute a test statistic that describes the relevant part of the distribution. Since I'm looking for an excessive amount of

[29]The transformation used is given by $r^s_{j,t} = \sqrt{r_{j,t}}$. Other standardizations, such as $ln(r_{j,t})$, yield similar results.

[30]Blocks will be defined by the smog-check center, the date and time of the day, which can be morning (8 to 12 hrs.), afternoon (12 to 16 hrs.) or evening (16 to 20 hrs.). The average number of tests per block is 83, with less than 1 percent of blocks with fewer than 11 checks. Alternative definitions of the blocks, such as center-lane-month-day of the week and center-lane-month-half of the month yield similar results. The validation of the results with the inclusion of lane in the block definition is particularly important, since one may argue that different testing machines within the center may be calibrated differently and, hence, generate similar consecutive readings.

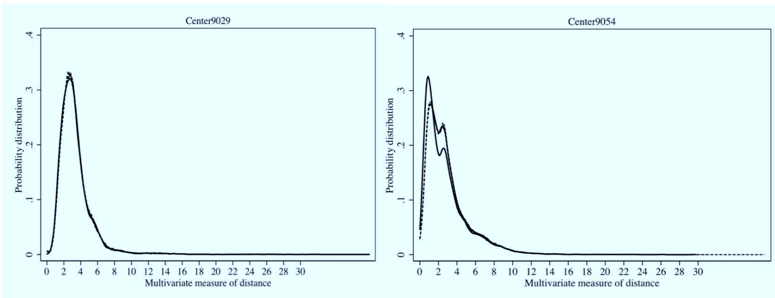
differences close to zero, or an excessive amount of small d_t draws, I chose the fifth percentile, $\hat{q}^d_{.05}$, of the distribution of d_t.[31] I will compare this statistic with the distribution of the fifth percentile of d_{t^n} for $n = 1, ..., 1000$. *I.e.*, I will generate an empirical distribution of $q^d_{.05}$ by performing 1000 different random permutations of the smog test ordering. The comparison yields p-values for the null hypothesis of no cheating. The test is performed at the center level for each of the 80 smog-check centers in the 2003-2nd-half round of D.F.'s smog-check center data. The p-values for all centers are less than 0.001, which rejects the null hypothesis of no cheating for all centers with 99.9 percent of confidence.

The permutation test generates very clear results regarding the extent of cheating across centers. According to this test, all centers participate in fraud. Results are robust to changes in the definition of the blocks within which data is reordered and to alternative test statistics (alternative results not shown).

[31] Other test statistics such as the proportion of vehicles under $d_t = 1$ and the tenth percentile of d_t yield similar results.

Figure A1: Nonparametric test for cheating

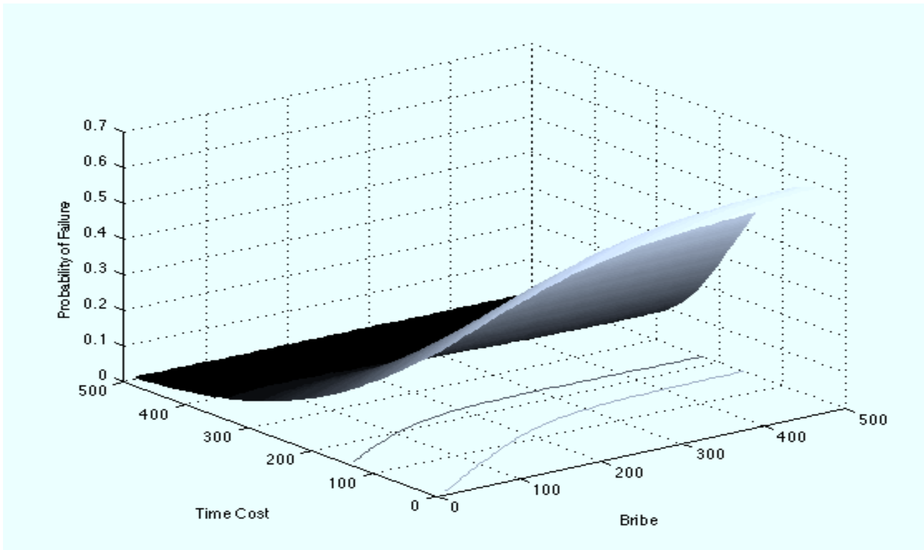
Panel A: Center with low cheating prevalence | Panel B: Center with high cheating prevalence



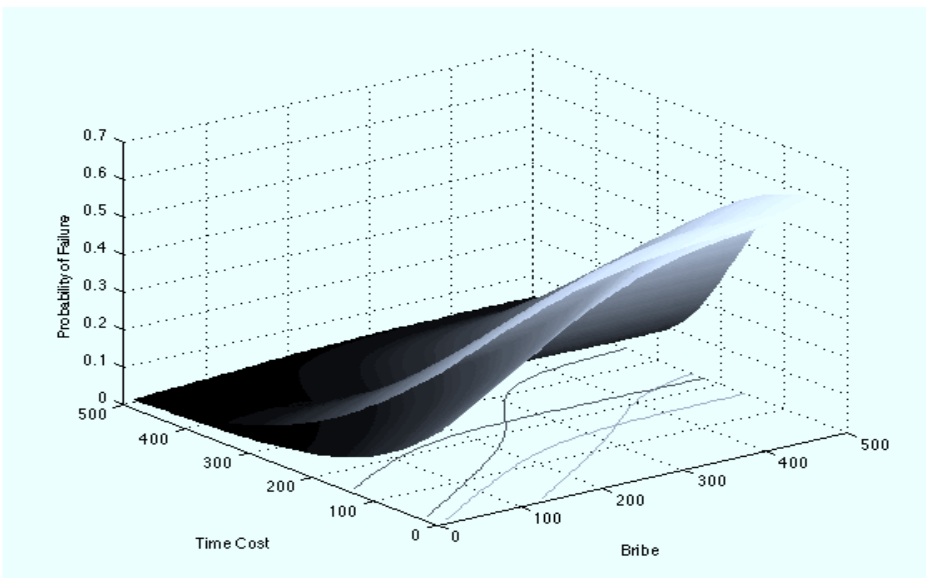
Notes:
1. Smog check center data from 2003 (2nd half), D.F.
2. The measure of distance corresponds to the square root of the sum of square differences between pollutants standardized by the variance covariance matrix
3. Dotted lines correspond to each of 10 different random permutations of the original data. The solid line correspond to the actual order of the data.

Figure A2: Identification of Structural Model

Panel A: Probability of Failure in First Test



Panel B: Probabilities of Failure in First Test and Retest



Notes: This figure illustrates identification by showing the surface that maps the bribe and the time cost parameters into the theoretical probability of failure in the first test (Panel A) and how the corresponding surface for the probability of failure in the retests intersects it (Panel B). Identification of these two parameters is achieved by finding the intersection of the level curves that correspond to the observed rates of failure in each test (test and retest).

Table A1: Emission restrictions in 2003 -2004

	Maximum level of emissions to obtain smog-check certificate					Maximum level of emissions for HNC exemption
	[1]	[2]	[3]	[4]	[5]	[6]
	Private vehicles, model>1990	Private vehicles, model<1991	Corporate vehicles, model<1994	Corporate vehicles, model>1993	Public transportation vehicles	Only private and corporate vehicles, model >1992
HC	200	300	200	350	100	100
NO	2500	2500	2500	2500	1200	1200
CO	2	3	2	3	1	1
O_2	15	15	15	15	15	15

Notes:

1. Source: Secretearía de Medio Ambiente, Distrito Federal
2. Limits are given in parts per million for HC and NO and in percent of volume for CO and O2.
3. The numbers indicate the maximum level of emissions a vehicle needs to attain in all 4 gases.

Table A2: Two period model parameters and demand for bribes

Panel A: Parameter Estimates

Model Parameters	Estimate	Bootstrapped Standard Errors		Implied estimates
	[1]	[2]		[3]
w-intercept	-34.42	80.52	Mean time cost	45.07
w-slope	7.57	12.73	Minimum time	26.52
b	230.67	81.73	Maximum time	69.21
δ	0.624	0.089	SE of random s	74.36
τ	57.98	47.14		

Panel B: Fitted and predicted probabilities for each history

Observed Histories	Actual	Fitted
	[4]	[5]
H1: Postpone	0.0832	0.0169
H3,H4: Bribe/No bribe - Pass	0.755	0.7422
H4: No bribe-Fail-Postpone	0.0037	0.0001
H5, H6: No bribe-Fail-Bribe/No bribe-Fail-No bribe-Pass	0.1142	0.1792
H7: No bribe-Fail-No bribe-Fail-bribe	0.0439	0.0516

All Histories	Predicted
	[6]
H1: Postpone	0.0169
H2: Bribe	0.0476
H3: No bribe-Pass	0.6946
H4: No bribe-Fail-Pospone	0.0001
H5: No bribe-Fail-Bribe	0.0310
H6 No bribe-Fail-No bribe-Pass	0.1482
H7: No bribe-Fail-No bribe-Fail-Bribe	0.0516
Total bribing	0.0786

Notes:

1. Table A3 shows the results of the two-period maximum likelihood estimation of the bribing behavior model developed in Section 6 using a 5% random sample (17,365 vehicles). Panel A shows the model's parameter estimates and Panel B shows the actual and fitted probabilities as well as the simulated probabiliites for all decision histories, including the unobserved ones.
2. For notation, see notes on Table 4. The last row of Panel B offers the total amount of bribing, i.e. the sum of H2, H5 and H7.
3. The standard errors for parameters and predictions correspond to bootstrapped standard errors. These errors where obtained from 100 replications of the test for cheating, estimation of the mapping between car characteristics and probability of passing, and MLE of car owner decision model.