Environmental Regulations and Corruption: Automobile Emissions in Mexico City

Paulina Oliva*

University of California Santa Barbara and NBER

*This paper benefitted from multiple discussions with Bryan Graham, and helpful comments and suggestions from David Card, Michael Greenstone, Eva Arceo, Edward Miguel, John Bellows, Rema Hanna, anonymous referees and seminar participants at UC Berkeley, UCSB and other institutions. I am thankful to the staff of *Instituto Nacional de Ecología* and *Secretaría de Medio Ambiente del Distrito Federal,* who provided the data for this study.

Abstract

Emission regulations become more prevalent in developing countries; but they may be compromised by corruption. In this paper I document the prevalence of corruption and the effectiveness of vehicle emission regulations in Mexico City. I develop a statistical test for identifying a specific type of cheating that involves bribing center technicians. I also estimate a structural model of car owner retesting and cheating decisions. Results suggest that 9.6 percent of car owners paid 20 U.S. dollars to circumvent the regulation. Eliminating cheating and increasing the cost of retests would reduce emissions by 3,708 tons at a high cost for vehicle owners.

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

1 Introduction

Automobile emissions are a larger contributor to air pollution and greenhouse gases in developing countries than in developed ones. For example, in Mexico City, they are responsible for 45 percent of volatile organic compounds and 81 percent of total nitrogen oxides (Molina and Molina 2002), while these figures are 29 and 34 percent, respectively, in the United States. Three factors may potentially explain these differences: an older car fleet, less stringent manufacturer controls, and extensive cheating on emission tests in Mexico.

In this paper, I focus on the prevalence of cheating on emission tests (known as "smog-checks") in Mexico City and the extent to which it may undermine regulatory efforts to reduce vehicle emissions. Motorists in Mexico City must conduct smog-checks on their vehicles twice a year. There are no restrictions on retests; hence, emission tests can be performed as many times as needed as within the two-month window assigned to each vehicle based on their license plate termination. The regulation aims to encourage maintenance of vehicle emission control mechanisms. However, cheating, which usually involves bribing the smog-check center technician, allows motorists to circumvent the regulation without meeting the emission standards.

Documenting and studying the consequences of cheating is complicated by the lack of information on cheating decisions and true emissions from cheaters, as people do not like to reveal engaging in illegal or unseemly behaviors. 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 here because vehicle emissions may vary from test to test (Wenzel et al. 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.

I overcome the information problem by combining a non-parametric test for cheating with a structural model of the emission testing process. The first step consists of a statistical test for identifying a specific type of cheating whose patterns can be observed in the data. Interviews with mechanics and newspaper articles reveal that the most common way to cheat is to substitute clean emissions from "donor" cars for the high emissions of the actual vehicles being tested.[1] This type of cheating requires bribing a center's technician to select a suitable donor car from among other customers and record a donor's second emission reading under the cheater's registration information, yielding two consecutive readings from the same car but under two different registrations. Since emission readings from the same vehicle have a lower variance than those from different ones, test readings will appear serially correlated whenever donor cars are used. Thus, one can identify which centers are engaging in cheating based on the serial correlation patterns in the data. Results from this test suggest that 63 out of 80 centers use donor cars.[2]

The second step uses data from centers with little evidence of cheating (identified in step 1) to estimate a mapping between car attributes and the "fair" probability of passing the test. The estimated mapping is then used to predict the fair probability of passing for the entire car fleet.

The third and last step consists of estimating a structural model of car owner testing decisions that allows for both retesting and cheating. The instantaneous utility functions associated with each decision in this model can be recovered from observed testing outcomes and the variation across car types in the probability of passing the test, which is estimated in step 2. Importantly, the identification of instant utility functions associated with all decisions is possible even without explicit information about the decision to bribe. The separate identification of monetary and time cost parameters in the utility functions is possible due to a policy-induced cost difference between tests and retests. The identification of value

[1] Humberto Padgett, *Reforma*, October 4, 2012.

[2] One concern is that other unobservable characteristics could induce serial correlation in the absence of cheating. However, I control for an array of unobservable emission determinants using fixed effects and emissions of vehicles that are tested simultaneously in the same center. The results are also robust to an alternative non-parametric test that relies on permutations of the test orderings, which rules out alternative stories. I present additional evidence that implies that the observed serial correlation is consistent with cheating, such as a higher correlation when there is a shorter time between tests and a weak statistical relationship between car characteristics and emissions for centers with high levels of corruption.

functions in an infinite horizon dynamic setting follows the identification results of Magnac and Thesmar (2002). The scale of the cost parameters is identified since some of the monetary cost components, namely the testing fee and the fine for missing the testing requirement, are observable and vary between tests and retests (Heckman and Navarro 2007). Pinning down the scale of private costs, including the bribe and the cost of time, allows for welfare calculations. An extension to the model yields results of the car owner's willingness to pay for car maintenance and the social benefits from increasing enforcement.

I find that bribing to circumvent the regulation is prevalent and that the combination of retesting and bribing compromises the program's cost-effectiveness. Bribing occurs in more than 9.6 percent of the tests among owners of vehicles that are least 10 years old. The structural model allows us to simulate alternative policies, such as increasing the bribe, changing testing-fee costs and increasing the fine for non-compliance. These simulations suggest that eliminating cheating and increasing the cost of retests would eliminate 3,708 tons of emissions per year, but would do so at a high cost to vehicle owners.

The contributions of this paper extend beyond that of smog-check policy. First, this paper illustrates how a statistical test to detect unobservable behavior can be used to identify controlled environments where the counterfactual behavior is observable. Although statistical tests of this sort tend to be context-specific, there are other important examples where abnormal data patterns can be translated into statistical tests (e.g. Jacob and Levitt 2003, Della Vigna and La Ferrara 2010, Ghanem and Zhang 2013).

The second contribution is the use of the first stage reduced form results to identify a dynamic structural model of decision-making, in which some decisions are unobserved. Identification of discrete dynamic models generally requires that all decisions made by agents are observed. This allows the econometrician to identify the decision-specific transition probabilities across states (Magnac and Thesmar 2002), which determine the expected utility from each decision in each period. In many contexts, however, we might encounter incentives to keep decisions private, making these decisions inherently unobservable. In this model, the unobserved decision is whether or not to bribe; other examples include other instances of cheating, like plant inspections or standardized tests (Jacob and Levitt 2005, Duflo et al. 2012); fraud in insurance claims (Caudill et al. 2005); insider trading; program evaluation with imperfect monitoring of take-up (Kremer et al. 2011); natural resource extraction; and fertility decisions (Ebenstein 2011). Decisions might also be confounded in the presence of recall bias and misclassification of decisions (Poterba and Summers 1995, Magnac and Visser 1999). The method employed in this paper illustrates how dynamic models may still be identifiable, using out-of-sample information on the transition probabilities, and how this method could be extended to other contexts. Moreover, repeated decision-making under different prices or incentives can help uncover the prevalence and instantaneous utilities of unobserved decisions even when only a subset of the outcomes is observed (Heckman and Navarro 2007).

The paper proceeds as follows: Section 2 provides a brief description of vehicle regulations in Mexico and the process through which cheating may occur. Section 3 describes the data. Section 4 proposes a statistical test for detecting cheating. Section 5 develops a model for a 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. Section 7 concludes.

2 Vehicle Emission Testing

Compulsory vehicle emission inspections are the most common means of enforcing emission standards on vehicles throughout the world. However, the sizable gaps between emission levels from official tests measured at smog-check centers and emission levels measured on-the-road or in off-cycle tests have casted doubts on their effectiveness. 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 pass an emission test on the first attempt will fail an immediate off-cycle retest. They also find that 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.[3] 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, citing discrepancies between passing rates across private and government-owned centers and persistent differences in passing rates across smog-check facilities as evidence of cheating (Hubbard 1998, Zhang et al. 1996, Glazer et al. 1993, Wenzel 2000, Snyder and Pierce 2008, Schifter et al. 2003, and the Ministry of Environment in the Federal District of Mexico 2004). Importantly, in all these studies the empirical evidence of cheating is also consistent with alternative explanations.

In 1990, the Mexico City Metropolitan Area (MCMA) introduced compulsory, twice-a-year smog-checks for all vehicles. Every six months, vehicle owners have a two-month window to visit any registered smog-check center. A smog test is a fairly standardized process: when the vehicle owner arrives, she must pay the emissions test fee. The fee is waived on "even numbered" retests (i.e. retests 2, 4, 6, etc.). Next, a center employee electronically enters the vehicle's information, including the plate number, model-year, make, the number of cylinders, the owner's address and mileage (all of which can be read from the registration card and the odometer). The next step is a visual test of the vehicle. Upon passing this, 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, a sticker with the vehicle car plate number is pasted on the vehicle's windshield. If the vehicle fails, the owner may retest indefinitely

[3] Wenzel et al. (2004) use a sample of cars that are submitted to an off-cycle test as a requirement for a change of ownership in Phoenix and California.

upon paying the corresponding fees. About 76 percent of vehicles pass the test after the first attempt and 15 percent pass it in the first retest. A total of 2.5 percent miss the requirement each cycle and about 6.2 percent fail the retest. Non-complying vehicles are easy to spot on the road by police officers because of the windshield sticker. Fines for non-compliance range from 850 to 3,500 pesos (79 to 325 2003 U.S. dollars).[4]

In 2003, 80 licensed centers operated in the Federal District. The licenses were tendered in 1997, with very few new licenses subsequently granted. All smog-check centers are privately owned, except for three institutional centers.[5] However, they are all subject to tight government regulation. 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 local environmental authority conducts unannounced inspections of smog-check center facilities on a regular basis to check all mechanical and electronic equipment. In addition, all facilities are required to have a camera surveillance system and live publicly available video transmissions.

The emission standards are below EPA Tier 1 standard in the U.S. Columns 1 to 4 of Table 1 show the emission limits in 2003. Column 5 of Table 1 shows the emission requirements on less-than-10 year old vehicles for an exemption to the "No Driving Today" program, which restricted vehicle driving days.[6]

Finally, vehicles can be retested as often as needed within their 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 is the same at every testing center.

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

[5]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).*

[6]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 that a vehicle is restricted varies by its license plate number (Davis, 2008). Since 1997, vehicles are exempt from this program if they are model-year 1993 or newer and they meet a stricter standard (see column 5 of Table 1). In 2004, the age requirement changed to 10 years or newer, regardless of the model-year.

3 Data

This study uses data on vehicle information and test outcomes from emission tests conducted in the Federal District in 2003. The smog-check center computers are connected to a common network run by the local Ministry of Environment, which pools each center's information into one data set.

The 2003 data include information for all tests and retests for 1.6 million vehicles, with the exact measurement of the four relevant gases. Three out of the four gases are harmful pollutants: hydrocarbons (HC), nitrogen oxides (NO), and carbon monoxide (CO). 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.

The data set also contains additional information for each test. This 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 test in seconds; and smog-check center's information (center's identification number, and lane or testing equipment where the test was performed).

In addition to the 2003 data, I use data from emission tests and vehicle plate numbers for the years 2001, 2002, 2004 and 2005 to identify vehicle owners that missed the 2003 emission test. This information is used in the structural model to identify individuals who choose to postpone a test with the risk of paying a fine.[7] Table 2 presents descriptive statistics of

[7] To observe the decision to postpone, I use three pieces of data. First, I observe individuals that were ever fined in the 4 smog-check cycles that follow the cycle that I focus on (2003-2, or second semester of 2003), although I do not observe whether the fine corresponds to the 2003-2 cycle or to a different one. Second, I can see whether there is a record of this vehicle in the previous 2.5 years. Third, I observe the vehicle's absence in the 2003-2 smog-check cycle or whether the last test in the cycle was failed. Thus, the decision to postpone is coded as 1 in attempt t for those vehicles that (a) had a previous record and no change of ownership recorded, (b) have a fine in some cycle after 2003-2, and (c) had no test record, or failed the last

vehicle characteristics and emissions tests outcomes.[8]

4 Extent of Cheating

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 out of 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 assured the reporters the cheating procedure would not cause damage to the car's engine (Padgett 2002).

Newspaper articles and interviews with mechanics suggest that most cheating occurs in the form of emission substitution.[9] When bribed by a customer, technicians use a clean testing car, commonly called donor car (*auto madrina),* to provide the emission readings for the customer's dirty car. The donor car can be any vehicle that passed the emission test at the center. A donor car is needed because emissions cannot be entered manually into the computer. The car's information, on the other hand, is 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 since a single car is tested repeatedly.

---

attempt in cycle 2003-2. Note that this methodology might underestimate the number of car owners who postpone, since some of them may manage to do so for longer than 2 years.

[8] All information used in this paper, except for vehicle plate numbers, is publicly available at http://www.sma.df.gob.mx/verificentros/.

[9] The local environmental authority has expressed concern about cheating and has responded by addressing tampering with the engine. However, this form of cheating is not likely a concern by 2003, since tampering can be controlled through limits on oxygen in emissions.

4.1 Testing for Cheating

In this section, I develop a statistical test for detecting cheating that relies on measuring the serial correlation between tests of different vehicles. For motivation, it is instructive to observe a sequence of actual emissions readings. Table 3 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 pollutant readings measured at 24 kph and 40 kph. Notice that the readings that appear inside the black squares show striking similarities for all pollutants across tests despite substantial differences in car characteristics. Thus, all of the readings 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

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

where $\mathbf{x}_{it}\beta$ is the best linear prediction of emissions given observable car, test and center characteristics. The specific set of controls included in $\mathbf{x}_{it}$ is detailed below.

In the current set-up, the index i will be the true car specific identifier, as opposed to the one reported in the records. This identifier 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 immediately 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". 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 the emission readings of the donor car and the emissions readings for all of the cheaters that the donor car provides . Specifically, it assumes that

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

From now on, I will use the terms "testing equipment" and "lane" interchangeably, since there is one set of testing equipment per lane. Assumption (A1) excludes the possibility of putting aside the donor car from the testing equipment while other non-cheating vehicles get tested and then bringing it back in for a cheater's test slot. It also excludes recording the donor car's "own" emission reading using different equipment than the readings used for cheaters' certificates. However, it does allow 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 3.[10]

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

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

(A3) $$\mathbb{E}\left(\mathbf{x}_{i-1t-1}{}^{T}\, u_{it}\right) = \mathbf{0}$$

Notice that, by construction, u_{it} includes only unobservable emission determinants that are uncorrelated with observable emission determinants. Thus, assumption (A2) would be violated if, in the absence of cheating, car emission determinants not controlled for in $\mathbf{x}_{it}$ are serially correlated across cars in the sequence.

Observable determinants of emissions include brand, service and size fixed effects, a flexible function of the age of the car, age-size and age-service interactions, and flexible functions of mileage. Outside of these determinants, there are many unobservable characteristics that could induce serial correlation in the absence of cheating. Some of these unobservable characteristics can generate serial correlation through selection of a particular class of vehicles into specific centers. Serial correlation can also be generated by equipment-specific factors such as calibration. To control for additional observable and unobservable determinants of serial correlation, I exploit additional data features. First, I include center and testing equipment fixed effects, which control for center and equipment-specific features, such as selection into centers and calibration of the equipment, which could induce serial correlation. Second, I

[10] Anecdotal evidence suggests that most donor cars are also being tested themselves. If centers use 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, I 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 3, occur frequently.

use data on the time of the day and date to control for broad temporal patterns in attendance. Third, I include weather controls, which could induce correlation patterns through temperature's effect on pollutant density.

Even after controlling for car characteristics and the above controls, complex selection patterns could generate clusters of vehicles that share unobserved determinants of emissions and result in serial correlation. For example, two cars that belong to the same owner, thereby having similar maintenance histories, could arrive at the center simultaneously. To address this issue, I exploit the fact that I observe the timing of all smog checks at a given center and can therefore deduce which vehicles were being tested simultaneously (within a small time window) in adjacent lanes within the same center. Since there is a single entry lane to the center, vehicles that are lined up next to each other when entering the center wind up being tested simultaneously in adjacent lanes. If the source of serial correlation comes from adjacent vehicles having common unobservable emission determinants, controlling for the emissions of these adjacent vehicles effectively controls for any determinants that could have also generated serial correlation in the absence of cheating.

Thus, for (A2) to be violated in the presence of the controls described above, one would need to observe correlation patterns that applied to vehicles that were close, but not immediately adjacent, in the entry line for the center and that were independent of testing equipment features, date and time of the day, weather, and observed car characteristics. Since these patterns seem unlikely, (A2) is likely to hold.

Assumption (A3) is likely met given the empirical set-up. 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.

Therefore, observed emission readings under the alternative hypothesis can 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), the estimation of the following OLS regression delivers a test for cheating:

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

The t-statistic for the OLS coefficient, γ_c, can be used to test for the null hypothesis of no cheating. Assumptions (A2) and (A3) imply that in the absence of cheating $\mathbb{E}^*(r_{it}|r_{i-1t-1,}\mathbf{x}_{it}) = \mathbf{x}_{it}\beta$, where $\mathbb{E}^*$ denotes the best linear projection operator; i.e., the serial correlation coefficient in emissions, γ_c, is zero under the null of no cheating.[11]

I perform the test for cheating at the center level. Since centers differ in the number of tests performed, I draw a random same-size sample of tests from each center to ensure the power of the test is constant across centers. For each center, I estimate equation (3) 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 for each center. I provide the Holm-cutoff values that control for family-wise error rate at 0.05. These cutoff values

[11] 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), $\mathbb{E}^*(u_{it}|\mathbf{x}_{i-1,}u_{i-1t-1}) = 0$. Online Appendix 1 provides details of the proof.

assume the tests are independent, which is conservative, but reasonable for this application.[12] Under these cutoff values, I reject the null hypothesis of no corruption for 62 out of 80 centers. Panel A of Table 4 shows the test results for the 10 centers that have the strongest evidence of donor car cheating and Panel B shows results for the 10 centers that have the weakest evidence under a joint test of significance for all four specifications.[13]

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 suggest that the observed serial correlation is driven by cheating, and not other factors. This section summarizes these results (see Online Appendix 2 for details and accompanying figures).

The first robustness check consists of an alternative test for cheating that relaxes the parametric assumption on the relationship between car characteristics and emissions. The test consists of comparing the observed sequence of emission readings to the sequence that results after randomly changing the order of vehicle arrivals to the center. Therefore, I call this test the permutations test. The generated counterfactuals and observed distributions of test results are shown in Online Appendix Figure 1. The results 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 concur on which centers have insufficient evidence for corruption in 4 out of 5 cases.

The second robustness check consists of verifying that time between tests is shorter whenever the correlation between subsequent tests is stronger. This should be the case if the correlation is emerging from repeatedly testing the same car, as testing the same car consecutively is faster than testing different cars. I find that the serial correlation between

[12] See Romano et al. 2010 and Heckman et al. 2010.

[13] A technician may use a donor car, even if the owner does not request it, out of convenience. However, I find that the predicted probability of passing based on records of non-cheating centers falls below the observed probability of passing for vehicles that attend cheating centers (see section 6.1), which is not consistent with this being the predominant motivation for using donor cars.

emission readings is more important when there is a short time interval between tests (0-5 min) than when there is a long time interval (15-20 min). This relationship is shown in Online Appendix Figure 2.

The third robustness check consists of checking whether car characteristics are less predictive of emissions in centers where cheating is rampant. The logic behind this test is that emissions from cheating vehicles are provided by donor cars; hence, the car characteristics of cheating vehicles should be less correlated with the recorded emissions. I verify that this is indeed the case. The details of this test, as well as variations under weaker assumptions, are provided in Online Appendix 2 and Online Appendix Figures 3, 4 and 5. This last robustness check also suggests that the use of donor cars is the predominant way of cheating, since other types of cheating would also result in weak correlations between car characteristics and emissions but would not be captured by the correlation test.

5 A Model of Bribing Behavior

The evidence thus far suggests that cheating is a major concern for emission control policy in Mexico City. This invites questions of whether economic incentives can explain this behavior and whether bribing is as cheap as anecdotal evidence has suggested. Although the evidence provided suggests a large amount of corruption in several smog check centers, the statistical evidence provided above is not specific enough on its own yield information about the number or proportion of vehicles that rely on corruption for overcoming the smog check requirement.

In this section, I model the key factors behind the car owner's decision 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 of the implied opportunity cost of bribing and the time cost incurred by vehicle owners when complying with the regulation. Given that the most polluting vehicles are older models that are not exempt from the smog-check requirement, the model proposed in this section focuses non-exempt vehicles. As described below, the estimates found are consistent with the cheating results presented in Section 4 and with anecdotal evidence of both 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 smog-check price is constant across centers and equal to 16 U.S. dollars; retests are unlimited and "even-numbered" 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 for doing so; and, finally, centers and car owners are "price takers" with respect to the bribe.[14]

The notation in this section is entirely independent from the notation in Sections 4.1 and 4.2, and so I begin with defining the notation. The beginning of each test or retest in a single smog-check cycle will correspond to a decision round. In each decision round a car owner is assumed to choose among three different actions or choices: $i \in \{X, B, A\}$. These actions are cheating and paying a bribe to the technician, which is denoted B; lawfully carrying out the emissions test, denoted as A; and missing the smog-check requirement by postponing the check beyond the deadline in a given smog-check period, denoted by X.

There are two possible states characterized by whether the testing round is odd or even. I denote the state variables in odd states as $s = (s_X, s_B, s_A)$ and in even states as $s' = (s'_X, s'_B, s'_A)$. The state variables, s_i and s'_i, are the monetary and time costs associated with each choice and reflect the monetary cost differences between odd and even testing rounds: odd testing rounds carry a testing fee of c (equal to 16 U.S. dollars), while even testing round

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

carry no testing fee.[15][16]

The instantaneous utility function, u_i (u'_i), associated with choice i in an odd (even) round is assumed to be linear in monetary and time costs, s_i (s'_i), and a random utility shock, ε_i (ε'_i): $u_i = u(s_i, \varepsilon_i) = s_i + \varepsilon_i$ and $u'_i = u(s'_i, \varepsilon'_i) = s'_i + \varepsilon'_i$. The 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 model assumes that the car owner has full knowledge on the deterministic portions of the instantaneous utility functions in all testing rounds. However, the car owner only observes the random utility shocks of the current testing round.

If a car owner misses the smog-check requirement (i.e., $i = X$), 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. Hence, we parametrize the instantaneous utilities from choosing X in odd and even periods as $u_X = -\delta f + \varepsilon_X$ and $u'_X = -\delta f + \varepsilon'_X$, respectively, where δ is the probability of getting caught times the six-month discount rate, and f is the fine for missing the smog-check requirement, which is equal to 79 U.S. dollars.[17]

If a car owner cheats and pays a bribe to the technician (i.e. $i = B$), she is guaranteed to pass the test. Since there are a large number of centers and the services they provide are homogeneous, I assume that centers are perfectly competitive on bribes (Shleifer and Vishny 1993); hence, car owners and centers are assumed to be "bribe takers" and the market bribe is denoted by β. We parametrize the instantaneous utilities from choosing B in odd and even periods as $u_B = -\beta - \omega - c + \varepsilon_B$ and $u'_B = -\beta - \omega + \varepsilon'_B$, where ω is the monetized cost of the time it takes to go to the smog-check center.

If a car owner tests lawfully (i.e. $i = A$), she avoids the cost of the bribe. Hence, her

[15] c should not be confused with c_i from Sections 4.1 and 4.2. In this section, c without a subscript denotes a known constant.

[16] Smog-check periods are different than decision rounds. A smog-check period occurs every six months and lasts two months. 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.

[17] The parameter δ is a reduced form penalty parameter in that it may also capture the following, all as a proportion of the fine, f: the probability of bribing upon getting caught without a current smog-check certificate, the corresponding bribe, and the disutility of missing the smog-check requirement.

instantaneous utilities from this choice in odd and even periods are given by $u_A = -\omega - c + \varepsilon_A$ and $u'_A = -\omega + \varepsilon'_A$. However, if $i = A$, she also faces the risk of failing. If failing occurs, the car owner will find herself at the next decision round or retest within the same smog-check period. Note that an individual can take as many retests as she wants. The infinite horizon induced by unlimited retests combined with the two-state characterization of the test costs result in a stationary dynamic model. Hence, her present discounted utility when $i = A$ is the same in all odd and all even periods and is given by Bellman equations,

$$v_A = -\omega - c + \varepsilon_A + \mathbb{E}\left[\mathbb{E}_{\tilde{\varepsilon}'}\left(\max_j(v'_j)\right)\right] \text{ and } v'_A = -\omega + \varepsilon'_A + \mathbb{E}\left[\mathbb{E}_{\tilde{\varepsilon}}\left(\max_j(v_j)\right)\right], \qquad (4)$$

respectively. In Equations (4) the inner expectation is taken with respect to the joint distribution of random utility components, $(\tilde{\varepsilon}, \tilde{\varepsilon}') = (\varepsilon_X, \varepsilon_B, \varepsilon_A, \varepsilon'_X, \varepsilon'_B, \varepsilon'_A)$, and the outer expectation is taken with respect to the transition probabilities from one testing round to the other: the probability of failing the test $1 - P$. The probability of passing the test, P, is assumed to be car specific and known by the car owner. Hence, we can rewrite the expectations in (4) as $(1 - P)\mathbb{E}_{\tilde{\varepsilon}'}\left(\max_j(v'_j)\right)$ and $(1 - P)\mathbb{E}_{\tilde{\varepsilon}}\left(\max_j(v_j)\right)$, respectively. Note that because bribing (B) or missing the test (X) does not result in further testing rounds, $v_B = u_B$, $v_X = u_X$, $v'_B = u'_B$, and $v'_X = u'_X$; i.e., there is no continuation value when choosing to bribe or to miss the test.

In this model c and f are known constants, while ω, β and δ are parameters to be estimated along with the distribution of the unobserved multivariate vector $(\tilde{\varepsilon}, \tilde{\varepsilon}')$.

A note on maintenance

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 widespread corruption, car maintenance seems unlikely to be a common response to the smog-check requirement. The model estimation 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 assess whether maintenance is likely. Specifically, I compare the probability of passing a retest between vehicles that conduct a first test and a retest on the same day, and vehicles that conduct a first test and a retest on different days. It seems unlikely that a first test, vehicle repairs, and a retest can occur all in the same day. Hence, I can safely assume that the car owners that return for a retest on the same day, which comprise 80 percent of all retests, will not have performed any repairs to their vehicles since the first failed test. If the share of passing vehicles among same-day retesters is similar to the share of passing vehicles among different-day retesters, then I can conclude that maintenance is absent in different-day retesters. Moreover, I can calculate the share of vehicles that perform maintenance using a simple equation. Assuming that (a) there are the same number of cheaters among the two groups, and (b) the repairs guarantee passing the test, then the number of vehicles that have performed maintenance is given by M_d in the following equation:

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

where $\hat{Pr}(\text{pass}|\text{different day})$ is the share of vehicles that passed the retest conditional on having it on a different day than the first test, M_d is the share of vehicles that performed maintenance among all vehicles that went for a retest on a different day, and $\hat{Pr}(\text{pass}|\text{same day})$ is the share of vehicles that passed the retest among those who came back on the same day. The above calculation yields $\hat{M}_d = 0.074$. Given that different-day retesters are 20 percent 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. In sum, this calculation suggests that omitting maintenance as a relevant alternative is not likely to have a large impact on the structural estimates.[18]

5.2 Identification

The goal of the model estimation is to recover the deterministic components of the instant utility functions associated with each choice, as well as the scale parameter of the distribution of the random utility shocks. Identifying these structural components allows us to use the model to estimate the unobservable cheating rate and the private costs associated with the program; it also allows us to predict outcomes from changes in policy. First, I introduce the notation for the probability of choosing an action, i, $i = (A, B, X)$. This will be denoted $\Pr(i|s)$ when the choice is made in odd periods and $\Pr(i|s')$ when the choice is made in even periods. Note that the probability of making a choice in any given period depends on the probability of passing the test, P. Hence, $\Pr(i|s)$ and $\Pr(i|s')$ are functions of P. To save notation, I will not make this explicit unless it is necessary (like in Section 5.3).

The identification of the model parameters is complicated by the fact that not all decisions are observable from the data: it is impossible to distinguish someone who tests lawfully and passes from someone who bribes. A consequence of not observing all decisions made by the individuals is that the transition probabilities, $(1 - P)$, are not empirically observable from the test-outcome data. To see this, note that the observed probability of a retest is $\Pr(A|s)(1 - P)$, but $\Pr(A|s)$ is not empirically observable.

[18]The calculation is relatively robust to omitting assumption (a). To see this, assume that the share of bribers in same day retests, b_s, and the share of bribers in different day retests, b_d, were different. It can be shown that the estimated maintenance rate is given by $\hat{M}_d = M_d + \frac{(b_d - b_s)(1-M_d)}{(1-b_s)}$, where M_d is the true maintenance rate in different day retests and $\hat{M}_d$ is the calculated maintenance rate. This equation shows that the absolute value of the bias falls with M_d, and increases with b_s. For instance, assume that $b_s = 0.15$, $b_d = 0.10$ and $\hat{M}_d = 0.07$. Then, $M_d \approx 0.12$. Given that different day retesters are 20 percent of all retesters, this would mean that the true overall maintenance rate is 2.4 percent instead of 1.5 percent from our basic calculation. Note that the bias falls with the maintenance rate, so the larger maintenance rates , the more robust the calculation. Also, note that if $b_d > b_s$, the calculation would be biased upward instead of downward.

Overcoming this identification issue requires external information on P. I obtain an estimate of the probability of passing for each vehicle by estimating a reduced form model that maps testing outcomes in non-corrupt smog-check centers to car characteristics. The estimates from this model can then be used to predict the probability of passing for the rest of the car fleet. The first stage estimation of P is further discussed in Sections 5.3 and 6.1.

The model assumptions used to show identification are

(AS) Additive separability between deterministic and random components in instantaneous utility functions.

(i) Missing the test is not time consuming and, by definition, will result in no additional testing rounds. In addition, there is a known fee for odd attempts at passing the test, while even attempts are free.

(ii) Bribes are competitive and they guarantee passing the test, therefore leading to no additional testing rounds. Bribes are constant across rounds.

(iii) The probability of passing the test is a function of observable car-specific characteristics. In addition, agents have full knowledge of their probability of passing the test.

(iv) P can be identified from a reduced form model of car characteristics using data from the non-cheating centers.

(v) The random utility components have an extreme value distribution with mean zero and scale parameter τ, i.e., $F_{\tilde{\varepsilon},\tilde{\varepsilon}'}$ is known up to a single parameter.

(LU) Instant utility is linear in time and money.

Identification relies on empirically observing the following four functions of P:

The probability of failing the first test,

$$g_1(P) = \Pr(A|s)(1 - P), \tag{5}$$

the probability of failing the first retest conditional on failing the first test,

$$g_2(P) = \frac{\Pr(A|s')(1-P)^2 \Pr(A|s)}{\Pr(A|s)(1-P)} = \Pr(A|s')(1-P), \tag{6}$$

the probability of missing the test requirement without ever showing up at a smog check center during the smog-check period,

$$g_3(P) = \Pr(X|s), \tag{7}$$

and the probability of missing the test requirement conditional on failing the first test,

$$g_4(P) = \frac{\Pr(X|s')(1-P) \Pr(A|s)}{\Pr(A|s)(1-P)} = \Pr(X|s') \tag{8}$$

The identification proof, outlined in this section and detailed in Online Appendix 4 consists of using assumptions (AS)-(LU) to solve for all of the parameters in the model as a function of the probabilities given by (5)-(8).

Assumptions (i) and (ii) are structural restrictions from the economic model. Assumption (i) guarantees that (5) and (6), as well as (7) and (8), are different functions of P. Assumptions (iii) and (iv) imply that P is empirically observable for every vehicle. Finally, assumption (v) allows for a closed form solution of model parameters in terms of the empirically observable functions (5) through (8).

Note that in the case of dynamic models, the extreme value distributional assumption, (v), does not result in independence of irrelevant alternatives. For example, the attractiveness of missing the test ($i = X$) relative to taking the test lawfully ($i = A$) depends on the bribe, β, through the continuation value. Specifically, the more expensive it is to bribe, the more attractive it will be to miss the test relative to taking it lawfully. This stems from the fact that the bribe increases the expected cost of failing, which is included in the cost of choosing A. This assumption, however, does imply that unobservable determinants of choices are uncorrelated across periods and across alternatives.

Online Appendix 4 shows that under assumptions (AS)-(v), the deterministic components of the instantaneous utility functions and the expected cost of the test in odd and even rounds are identified up to the scale parameter, τ. Parametrizing instantaneous utility functions as linear functions of the costs (i.e. adding assumption (LU)) allows us to separately identify τ as well as linear parameters ω, β and δ.

5.3 Estimation

Estimation of the structural model is carried out through Maximum Likelihood. The likelihood is constructed using the model's probability for each observed outcome. Since the decision to bribe is unobserved, there are five observable outcomes from two consecutive testing rounds. These outcomes are

Outcome 1 ($O1$) Missing the test before ever showing up at the smog-check center.

Outcome 2 ($O2$) Passing the test, either through bribing or testing lawfully.

Outcome 3 ($O3$) Missing the test after the testing and failing once.

Outcome 4 ($O4$) Passing the retest, either through bribing or testing lawfully, after failing the first test.

Outcome 5 ($O5$) Failing the retest.

Since one can take the test an unlimited number of times, outcomes of further retests are also observable. However, these outcomes are very rare and therefore not used for the estimation. Moreover, the importance of learning increases for further retests and therefore assumption (iii) becomes less likely to hold. The data used to identify the model is made of observations on the five binary variables corresponding to these outcomes, $D1, ..., D5$, and vehicle-specific characteristics including model-year, service, manufacturer, engine capacity, size, mileage and vehicle's market price from blue books.[19]

The first stage of estimation consists of identifying the smog-check centers that have no evidence of corruption and using $D2$ and the observable emission determinants from this sample to estimate the mapping between car characteristics and the probability of passing. This first stage allows us to compute the predicted probability of passing, $\hat{P}$, for the entire car fleet.

The second stage consists of the maximum-likelihood estimation of the structural model. The probabilities for each observed outcome according to the model are given by:

$$\Pr(O1) = \Pr(X|\hat{P}, s)$$
$$\Pr(O2) = \Pr(A|\hat{P}, s)P + \Pr(B|\hat{P}, s)$$
$$\Pr(O3) = \Pr(X|\hat{P}, s')\Pr(A|\hat{P}, s)(1-\hat{P})$$
$$\Pr(O4) = \left(\Pr(A|\hat{P}, s')P + \Pr(B|\hat{P}, s')\right)\Pr(A|\hat{P}, s)(1-\hat{P})$$
$$\Pr(O5) = \Pr(A|\hat{P}, s')(1-\hat{P})^2\Pr(A|\hat{P}, s),$$

which are functions of model parameters and the predicted probability of passing, $\hat{P}$. In order to allow for heterogeneity in time costs across individuals, I model time cost as a function of the vehicle's market price from the corresponding year blue-books: $\omega = \omega_0 + \omega_1 CV$, where CV is the log of the vehicle price. Therefore, the unobserved parameters to be estimated are $\omega_0, \omega_1, \beta, \delta$ and τ.

The probabilities listed above are a function of the continuation values in odd and even rounds, $\mathbb{E}_{\tilde{\varepsilon}}(\max_j(v_j))$ and $\mathbb{E}_{\tilde{\varepsilon}'}(\max_j({v'}_j))$. Given observed variables $\hat{P}$ and CV and values of the parameters $\omega_0, \omega_1, \beta, \delta$ and τ, these continuation values can be computed using a fixed point algorithm (Rust 1989) (see Online Appendix 3). Each iteration of the numerical likelihood maximization computes $\mathbb{E}_{\tilde{\varepsilon}}(\max_j(v_j))$ and $\mathbb{E}_{\tilde{\varepsilon}'}(\max_j({v'}_j))$ for each individual in the sample.

[19]The share of vehicles that are characterized by each of the observable outcomes, $D1, ..., D5$, is provided in column 4 of Table 5. Summary statistics of most covariates are reported in Table 2. Average log-car value is 10.05 with a standard deviation of 0.42.

The likelihood function is given by

$$L(\omega_0, \omega_1, \beta, \delta, \tau | D1, ..., D5, \hat{P}, CV) \quad = \quad \prod_{i=1}^{N} \left[\Pr(O1)_i^{D1_i} \cdot \Pr(O2)_i^{D2_i} \cdot \Pr(O3)_i^{D3_i} \cdot \Pr(O4)_i^{D4_i} \cdot \Pr(O5)_i^{D5_i} \right]$$

where i indexes the individual observation.

The two-stage estimation may lead to an underestimate of the standard errors in the second stage if I do not account for the estimation error in the first stage. Online Appendix Table 3 (OA3) reports the estimates of a two-period model with bootstrapped errors that accounts for the first stage estimation error. Unlike the infinite horizon model, the two-period model does not require a nested fixed point algorithm, therefore the bootstrap procedure is feasible. The results of this alternative variance estimation are further discussed in Section 6.2.

6 Estimation Results

6.1 Predicting the Probability of Passing

As explained in Section 5.2, the identification of the structural model relies on observing the probability of passing for each vehicle in the car fleet. To obtain this, we estimate a mapping between car characteristics and the probability of passing the emissions tests using *only* tests from centers that were identified as low-cheating in Section 4. More specifically, after sorting the centers into 10 groups based on their cheating probabilities, I estimate the mapping using all centers in the lowest cheating group. The parameters estimated from this subsample are then used to predict the probability of passing for cars in the remaining centers.[20] The predicted probability of passing the test identified from the lowest cheating group is close or slightly above observed passing rates for the centers with low cheating evidence (first and

[20]These estimates are available in the Online Appendix Table 1 (OA1).

second lowest cheating groups), 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.[21]

If car owners self-select into cheating centers based on unobserved determinants of emissions, omitted variable bias may occur. This type of self-selection would cause one to overstate the predicted probability of pass, which would in turn generate downward bias in the simulated prevalence of cheating.[22] However, the available evidence suggests that this bias is small. First, there is substantial overlap in observable emission determinants across cheating and non-cheating centers.[23] Second, while there is some evidence of selection of older, large engine, and highly driven vehicles into low-cheating centers, these differences are small.

Although the test for cheating discussed in Section 5.1 is aimed at detecting "donor car" cheating, other forms of cheating are likely absent in the non-cheating sample. This can be inferred from the third robustness check on the test for cheating shown in Online Appendix Figure 3 (and Online Appendix Figures 4 and 5), 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 for the use of donor cars. In addition, low discrepancies between the predicted probability of passing rates and the 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.

[21] Observed passing rates and predicted probability of passing by cheating test decile are available in the Online Appendix Table 5 (OA5).

[22] Selection is stronger for the first and second deciles of the cheating test distribution. Hence, an alternative methodology uses these 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. This alternative methodology yields a lower average predicted probability of passing (0.74 instead of 0.75) and a higher predicted cheating rate (14 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.

[23] Descriptive statistics by cheating status are shown in Online Appendix Table 2 (OA2).

6.2 Parameter Estimates and the Prevalence of Cheating

In Table 5, I present the results from estimating the infinite horizon model for the second half of 2003. Since the numerical maximization of this model is computationally demanding, the model was estimated with a five percent random sample of all non-exempt vehicles (those 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-exempt vehicles are older and have higher emissions than those that are exempt.[24] Moreover, the sample was also restricted to vehicles that do not switch centers between the first test and the retest. The identification of the model relies on the difference in costs between the first and second tests. If a vehicle owner visits 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 no price difference between the two tests. About one percent of vehicles changed centers after failing the first test. The final sample used for the estimation has 17,659 vehicles.

The parameter estimates are presented in Panel A (column 1) of Table 5. 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.[25] The time cost is positively correlated with the price of the car. Column 3 of Table 5 computes the minimum, mean and maximum values of the time cost implied by the model parameters. The time cost varies between 81.26 and 150.61 Mexican pesos (MXN) across motorists. However, notice that the point estimates for time cost parameters are not very precisely estimated.[26]

The equilibrium bribe is 193 pesos (18 U.S. dollars), which is a plausible value according

[24] A decision model for exempt 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.

[25] These standard errors are not corrected for the first stage estimation error in the probability of passing. Online Appendix Table 3 (OA3) provides the results of a 2-period model with bootstrapped standard errors that account for the estimation error in this first stage.

[26] I also modeled the time cost as a function of census-level socioeconomic characteristics, however, these indicators provide very noisy measures of socioeconomic characteristics at the individual level.

to newspaper articles that have reported rates ranging from 50 to 400 pesos. The estimate for the parameter δ is 0.61. This parameter is the combination of a time discount rate, given that the fine is usually paid several months after the deadline, and the probability of paying the fine, f, for missing the test on one smog check cycle. Officially, missing a test is difficult since the car owner has to demonstrate that she paid the fine before she can get her vehicle smog-checked in the next period, and further avoiding the smog-check may result in multiple traffic tickets. However, a large percentage of car owners miss a test, which results in a relatively small δ, and suggests that alternative methods can be used to avoid paying the fine.[27]

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 could result from the simplicity of the model, given that I do not account for much of the heterogeneity across individuals. Alternatively, it could reflect an overly optimistic option value of future choices: a larger variance of the error term in (4) leads to a higher probability that a large and positive shock will appear in any of the future choices.

Panel B of Table 5 shows that the model fit is 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, O1 and O3, and worse for confounded outcomes, O2 and O4. Online Appendix Table 4 shows estimated and observed probabilities for these same outcomes when conditioning to low, medium and high probabilities of passing and for high and low values of the vehicle. The fit of the model is remarkably good when conditioning on the probability of passing. Conditioning on the value of the car gives mixed results in terms of fit, which is consistent with the low precision of the time cost parameters.

Panel C of Table 5 shows the predicted shares in each of the seven decision paths. The total cheating prevalence rate estimates from this method can be approximated by the sum of predicted second and fifth decision paths, which yields 9.6 percent.

[27]Given that the environmental authority reports that the probability of getting caught is close to one, a small δ could also reflect inconsistent time preferences.

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 know 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. This yields that the $\Pr(cheating) \approx 0.109$. 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.

Online Appendix Table 3 (OA3) shows the results of a two-period model and reports bootstrapped standard errors to account for the first-stage estimation error. This model is more restrictive: it assumes that anyone that attempts and fails the test twice chooses to cheat and bribe in the third round. Not surprisingly, the fit of this model is worse 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, making it feasible to show consistent standard errors from bootstrapping that account for the estimation error in the passing probabilities.[28] The standard errors that incorporate the estimation error in the pre-

[28] 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.

dicted probability of passing are larger than the errors from the inverted numerical Hessian. The bribe and the discount factor parameters, β and δ, are more precisely estimated than the time and scale parameters ω_0, ω_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 a cleaner vehicle. The regulation will not be binding if the non-compliance cost is low. Other aspects of the regulation may also reduce the incentives for owners to either perform car maintenance or buy new cars, such as the possibility of retesting indefinitely at a low cost (Wenzel 2004), having a low fine, or equivalently, having a low probability of paying a fine if one has not passed the emissions test.

In this section, I extend 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 by the extended corruption evidence found in Section 4 and by 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 provides 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 allows the evaluation of policy changes that are aimed at increasing the number of vehicles that get repaired. The policy simulations will be explained at length in Section 6.4.

This extension is feasible because 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 given by $\mathbb{E}_{\tilde{\varepsilon}'}\left(\max_j(v_j')\right)$ and

$\mathbb{E}_{\tilde{\varepsilon}}(\max_j(v_j))$. The expected cost of the testing requirement, $\mathbb{E}_{\tilde{\varepsilon}}(\max_j(v_j))$, is increasing in the bribe price, 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 maintenance only if it saves them money. When computing the expected cost without car maintenance, the owner takes into account all costs modeled in Section 5. Second, I specify assumptions about the costs and benefits of car maintenance using documented values. I assume that owners who perform maintenance on their vehicles will pass the emissions test with 90 percent probability in the two smog-check cycles following the maintenance (a full year). I also assume that performing maintenance improves the vehicle's gas mileage, and that this is improved relatively more for high-emitting cars than for low-emitting cars.[29] 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 gaining a probability of passing of 90 percent in one smog-check cycle, plus the improvements in gas mileage:

$$\begin{aligned} WTP_i = \; & \mathbb{E}\left(\max_j(v_j|\hat{\omega}_0, \hat{\omega}_1, \hat{\beta}, \hat{\delta}, \hat{\tau}, 0.9, CV_i)\right) + h(\hat{P}, \mathbf{x}_i) \\ & -\mathbb{E}\left(\max_j(v_j|\hat{\omega}_0, \hat{\omega}_1, \hat{\beta}, \hat{\delta}, \hat{\tau}, \hat{P}, CV_i)\right) \end{aligned}$$

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

[29] 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 (derived from www.fuel economy.gov), and that the reduction falls linearly with the probability of passing until reaching zero for those that pass with certainty.

maintenance on her vehicle.

The first column of Table 6 shows that the model estimated in Section 5 is internally consistent. The fraction of people that are predicted to perform car maintenance is 0.6 percent, according to the model extension and the parameters estimated in Section 5. In the next section, I adjust the test cost, the equilibrium bribe, and the fine in order to evaluate the extent to which these changes increase the number of owners that find maintenance preferable, and I 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. Presumably, a high enough price of the bribe can eliminate corruption entirely. However, emissions would only be reduced if those that cheat under a low-bribe price would 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. Specifically, 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 what follows, I will assume that the equilibrium bribe can be raised by increasing enforcement.[30]

Table 6 summarizes the performance results of the five policies that I consider. The outcomes for each policy in Table 5 include the minimum bribe, fine and retest fee levels to achieve each policy's goal (rows 1-3); the share of vehicle owners bribing, postponing, testing lawfully, and performing maintenance (rows 4-7); and the gains of each policy with respect to the status quo in terms of emissions avoided, lives saved, economic benefits, and

[30] The link between enforcement costs and the equilibrium bribe is unknown. Hence, in the cost benefit analysis, some costs of these alternative policies, including increased enforcement, are missing.

additional costs to car owners.[31]

Next, I describe the 5 different alternative policies explored and summarize the corresponding outcomes. Policy 1 asks what would happen if anti-corruption enforcement increased to the point that the equilibrium bribe would deter cheating. The minimum bribe that would result in no cheating is 300 MXN. This policy, however, would result in a very small increase in maintenance and a corresponding low reduction in emissions (191 tons) because rather than performing maintenance, car owners may choose to perform repeated retests until the vehicle passes the test or forgo the smog-check certificate and risk a fine. Policy 2 asks what would happen if we increased both enforcement and fines to deter cheating and postponing. The bribe and fine levels that would result in minimal cheating and postponing (less than 0.2 percent) are 475 and 1850 MXN respectively. This policy would increase the share of individuals that perform maintenance to 1.6 percent. This small increase in maintenance delivers almost eight times more emission reductions than Policy 1 and results in 1.3 more lives saved per year. Policy 3 doubles the fine. Without an increase in enforcement, this policy results in a small increase in the share that perform maintenance and over half the economic benefits in terms of lives saved compared to Policy 2. Policy 4 explores what would occur if tests were less costly to vehicle owners by eliminating the fee. This policy does little to encourage maintenance and yields modest economic benefits. However, it results in a substantial reduction of private costs (4.4 million). Finally, Policy 5 substantially increases the costs of all alternatives relative to maintenance (equilibrium bribe, fines and retesting fees). This policy results in a sizable increase in the share who perform maintenance compared to the other policies. Note, however, that even under this relatively extreme policy the share of vehicles that undergo maintenance is still relatively modest: 4.2 percent. This policy results in about 3.57 lives saved per year.

The cost side of the above analysis focuses on the costs paid by car owners. Private costs come directly from the model, corresponding to the expected cost of facing the regulation

[31] The remaining assumptions behind the calculations can be found in the notes of Table 6 and are detailed in Online Appendix 5 and Online Appendix Table 6.

under each scenario minus the current expected cost. Although the comparison of private costs to the benefits of each policy does not constitute a proper welfare calculation, this comparison might be of interest to policy-makers. Conducting a full welfare calculation would require accounting for the share of costs that are transfers. The fines, which are part of the expected cost, can be thought of as a transfer to the government. In all of the alternative policy cases considered, except for Policy 1 and Policy 4, the fine revenue goes down because of the deterrence effect created by raising the fine. The fall in revenue would need to be subtracted from the cost calculation, leading to a higher social cost of the program. It is somewhat less clear what percentage of the bribes and test fee revenues are transfers and what share are administrative costs. The fees are supposedly calculated by the government to match the administrative costs faced by the centers. Hence, the additional fees for the retest in Policy 5 could be thought of as transfers. The bribes can be thought of as the technician's marginal cost of being caught (and fired) or the center's marginal cost of being shutdown. If the marginal cost equals the average cost, then the costs of being caught are just being passed to the car owners in the form of bribes and should be incorporated in a welfare calculation. Finally, one also needs to add the change in surveillance costs that would ultimately yield an increase in equilibrium bribes. The government spends about one million pesos per month (93 thousand dollars) in the wages of the 76 inspectors that supervise the 81 smog-check centers. This adds 1.1 million dollars per year to the cost of the test. Changing the enforcement levels (increasing the equilibrium bribe) would likely increase these costs. However, it is hard to know exactly by what proportion. All together, the policies considered in Table 6 are likely more expensive to society that what is accounted for simply by the change in private costs.

The set of policies considered has at best a modest impact onair pollution, especially when compared to the additional costs to car owners generated by them. 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 is equivalent to 1.5 million dollars in savings. In contrast, the estimated cost to 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. First, 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 that some vehicles with high maintenance costs (and potentially high emissions) are misallocated into performing car maintenance and some additional vehicles with low maintenance costs (but potentially relatively low emissions) may opt for car maintenance.[32]

Second, the 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 mayalso be present for predicted emissions. More specifically, self-selection into non-cheating centersmay result in an underprediction of vehicle emissions. This

[32]If maintenance costs are higher for most vehicles, I overestimate the share that would perform maintenance in response to the regulatory changes. Since the program is far from cost-effective under this conservative scenario, higher maintenance costs would make it even less so.

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 6, the downward bias would have to be of more than fifty percent.

7 Conclusions

Researchers and governments have questioned the effectiveness of smog checks in reducing vehicle emissions, citing repeated testing and cheating as reasons why delinquent cars still pass these tests. This paper uses indirect evidence to show that cheating is wide spread in Mexico City. It develops a test for cheating that relies on detecting serial correlation patterns in consecutive emissions generated by the use of donor cars, a cheating practice that involves using emissions from a clean car to substitute for that of a cheater. This test predicts that 79 percent of centers have engaged in this behavior.

I use the results from the "cheating test" as an input into the estimation of a structural model of car owner decisions which recovers the underlying parameters of the cheating decision and uses the parameters 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 have 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 the estimation of 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.

The above conclusion begs the question of which alternative policies are most effective in reducing vehicular emissions. Technological mandates in new vehicles, such as the use of catalytic converters and the tightening of fuel economy standards, are easier to enforce, given that a small number of firms, rather than individual car owners, are required to comply with the regulation. This type of regulation is already in place in Mexico City. Although these measures have not been evaluated in the context of Mexico City, evidence from other contexts suggest they are effective. For example, Greenstone and Hanna (2012) showed that this type of policy was one of the few policies that successfully reduced air pollution in India. Regulating the older segment of the car fleet remains a challenge. The government has implemented car fleet renewal incentives, including a temporary cash-for-clunkers program and a new version of the driving restrictions program that exempts new vehicles. However, there is no evidence that these regulations have had a substantial impact in the car fleet age.

The empirical methods used in this paper illustrate how auxiliary information can be used to identify dynamic discrete choice models even when some decisions, as well as the conditional transition probabilities, are unobserved. It also shows how a statistical test for the presence of the unobservable behavior can be used to identify controlled environments where counterfactual behavior is observable.

References

[1] Arceo, Eva, Rema Hanna and Paulina Oliva. 2013. “Does the Effect of Pollution on Infant Mortality Differ Between Developing and Developed Countries? Evidence from Mexico City” *mimeo.*

[2] Bell, MichelleL., Aidan McDermott, Scott L. Zeger, Jonathan M. Samet, and Francesca 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.

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

[4] Casella, George and Roger L. Berger. 2002. *Statistical Inference* Second Edition, Duxbury.

[5] Caudill, Steven B., Mercedes Ayuso, and Montserrat Guillén. 2005. “Fraud Detection Using Multinomial Logit with Missing Information” *Journal of Risk and Insurance* 72: 539-550.

[6] Charlson, R. J. and T. M. L. Wigley. 1994. "Sulfate Aerosol and Climatic Change." *Scientific American* 270 (February): 48-57.

[7] 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).

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

[9] Della Vigna, Stefano and Eliana La Ferrara, 2010, “Detecting Illegal Arms Trade” *American Economic Journal: Economic Policy* 2: 16-57.

[10] Duggan, Mark, and Steven Levitt. 2002. “Winning Isn’t everything: Corruption in Sumo Wrestling,” *American Economic Review*. 92: 1594-1605.

[11] Duflo, Esther, Michael Greenstone, Rohini Pande and Nicholas Ryan. 2012. “Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India”. *mimeo*.

[12] Ebenstein, Avraham. 2011. “Estimating a Dynamic Model of Sex Selection in China” *Demography* 48: 783-811.

[13] Fisman. Ray. 2001. “Estimating the Value of Political Connections” *American Economic Review* 91(4): 1095-1102.

[14] Ghanem, Dalia, and Junjie Zhang. 2013. “Effortless perfection: Do Chinese Cities Manipulate “Blue Skies?” *mimeo.*

[15] Glazer, Amihai, Daniel Klein, and Charles Lave. 1995. "Clean on Paper, Dirty on the Road: Troubles with California’s Smog Check" The University of California Transportation Center, reprint No. 275.

[16] Glazer, Amihai, Daniel Klein, and Charles Lave, 1993, “Clean for a A Day: Troubles with California’s Smog Check” Report to the California State Senate Committee on Transportation, August.

[17] Hanna, Rema and Michael Greenstone. 2013. “Environmental Regulations, Air and Water Pollution, and Infant Mortality in India” Working Paper no. 17210, NBER, Cambridge, MA.

[18] Heckman, James, Seong H. Moon, Rodfigo Pinto, Peter Savelyer, and Adam Yavitz. 2010. “Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the High Scope Perry Preschool Program” *Quantitative Economics* 1(1): 1-46.

[19] Heckman, James J. and Salvador Navarro, 2007, “Dynamic discrete choice and dynamic treatment effects” *Journal of Econometrics* 136: 341-396.

[20] Hubbard, Thomas A. 1998. “An Empirical Examination of Moral Hazard in the Vehicle Inspections Market” *RAND Journal of Economics* 29(2): 406-426.

[21] Kremer, Michael, Edward Miguel, Sendhil Mullainathan, Clair Null, and Alix Peterson Zwane, 2011. “Social Engineering: Evidence from a Suite of Take-up Experiments in Kenya” *mimeo*

[22] Jacob, Brian A. and Steven Levitt. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating” *Quarterly Journal of Economics* 118: 843-877.

[23] Magnac, Thierry and David Thesmar. 2002. “Identifying Discrete Decision Processes”. *Econometrica* 70: 801-816.

[24] Magnac, Thierry and Michael Visser. 1999. “Transition Models with Measurement Errors” *The Review of Economics and Statistics* 81: 466-474.

[25] Molina, Mario and Luisa Molina, eds. 2002. *Air Quality in the Mexico Megacity*, Dordrecht: Kluwer Academic Publishers.

[26] McFadden, Daniel and Thomas Domencich. 1996. *Urban travel demand. A behavioral analysis*, North-Holland Publishing Co., 1975, Reprinted.

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

[28] Olken, Benjamin. 2007. “Monitoring Corruption: Evidence from a Field Experiment in Indonesia” *Journal of Political Economy* 115: 200-249.

[29] Olken, Benjamin. 2006. “Corruption and the Costs of Redistribution: Micro Evidence from Indonesia” *Journal of Public Economics* 90: 853-870.

[30] Padgett, Humberto. 2002. “Con propina todos pasan”, REFORMA (Newspaper Article), October 01, 2002.

[31] Poterba, James M. and Lawrence H. Summers. 1995. “Unemployment Benefits and Labor Market Transitions: A Multinomial Logit Model with Errors in Classification” *The Review of Economics and Statistics* 77: 207-216.

[32] Romano, Joseph P., Azeem M. Saikh, and Michael Wolf. 2010. “Hypothesis Testing in Econometrics” *Annual Review of Economics* 2: 75-104.

[33] Rust, John. 19890 “A Dynamic Programming Model of Retirement Behavior” *The Economics of Aging,* ed. by D.Wise. Chicago: University of Chicago Press, 359-398.

[34] Singer, Brett C., and Robert A. Harley.. 1996, “Fuel-based motor vehicle emission inventory”, *Journal of the Air and Waste Management Association* 46:6.

[35] 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.

[36] Small, Kenneth A. and Camilla Kazimi. 1995. "On the Costs of Air Pollution from Motor Vehicles" *Journal of Transport Economics and Policy*, 29.

[37] 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.

[38] Schifter, I., L. Díaz, V. Múgica, and E. López-Salinas. 2005. "Fuel-Based Motor Vehicle Emission Inventory for the Metropolitan Area of Mexico City" *Atmospheric Environment,* 39: 931–940.

[39] Shleifer, Andrei and Robert Vishny. 1993. "Corruption", *Quarterly Journal of Economics*, 108 (3): 599-617.

[40] Snyder, Jason and Lamar Pierce, 2008, "Ethical Spillovers in Firms: Evidence from the Vehicle Emissions Testing Market" *Management Science* 54: 1891-1903.

[41] 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.

[42] Schwartz, Joel. 1993. "Air Pollution and Daily Mortality: A Review and Meta Analysis" *Environmental Research* 64: 36-52.

[43] 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.

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

[45] 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* 9: 107-124.

[46] 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* 30: 1445-1450.

Table 1: 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]
	Private vehicles, model > 1990	Private vehicles, model < 1991	Corporate vehicles, model < 1994	Corporate vehicles, model > 1993	Only private and corporate vehicles, model > 1992
HC	200	300	200	350	100
NO	2500	2500	2500	2500	1200
CO	2	3	2	3	1
O2	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 2: Descriptive Statistics

Panel A: Vehicle characteristics			Panel B: Test outcomes (first test)		
	First test	All tests		Mean / SD	95th %ile / 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
Compact-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 (°C)	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 2,126,781. First time checks (or number of vehicles checked in 2003) is 1,513,111.

3. Panel A indicates the variable for the proportion of vehicles in each category, 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 3: 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, Distrito Federal.

Table 4: Test Coefficients and P-Values Compared to Holm's Cutoff Values

Center Order by χ^2	Lagged HC (Lhc)	Lagged NO (Lno)	Lagged CO (Lco)	Lagged O2 (Lo2)	Obs.	χ^2 stat, Lhc = Lco = Lno = Lo2 = 0	Prob > χ^2	Cutoff value under Holm stepdown method of FWER control	Test under Holm
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Centers with lowest p-value for the joint significance test									
1	0.0173	0.0065	0.0033	0.2248	4,500	297.789	0.000	0.0006	reject
	[0.011]	[0.012]	[0.011]	[0.013]					
2	0.0294	0.1238	0.0273	0.1795	4,500	275.616	0.000	0.0006	reject
	[0.013]	[0.014]	[0.013]	[0.012]					
3	0.0331	0.0954	0.1800	0.0585	4,500	262.939	0.000	0.0006	reject
	[0.009]	[0.012]	[0.013]	[0.014]					
4	0.0156	0.1975	0.0199	0.0129	4,500	240.057	0.000	0.0006	reject
	[0.016]	[0.013]	[0.013]	[0.013]					

5	0.0123	0.0478	0.0025	0.1949	4,500	236.284	0.000	0.0007	reject
	[0.009]	[0.013]	[0.012]	[0.013]					
6	0.0235	0.0889	0.0669	0.1670	4,500	213.567	0.000	0.0007	reject
	[0.012]	[0.014]	[0.012]	[0.014]					
7	0.0148	0.0218	-0.0082	0.1955	4,500	210.071	0.000	0.0007	reject
	[0.016]	[0.012]	[0.013]	[0.014]					
8	0.0643	0.1754	0.0163	0.0531	4,500	200.390	0.000	0.0007	reject
	[0.013]	[0.014]	[0.013]	[0.013]					
9	0.0055	0.2011	-0.0019	0.0311	4,500	182.237	0.000	0.0007	reject
	[0.013]	[0.015]	[0.012]	[0.013]					
10	0.0215	0.1387	0.0311	0.0065	4,500	166.351	0.000	0.0007	reject
	[0.012]	[0.011]	[0.012]	[0.011]					
Panel B: Centers with the highest p-value for the joint significance test									
70	-0.0049	0.0311	0.0142	0.0140	4,500	8.875	0.064	0.0045	cannot reject
	[0.014]	[0.013]	[0.013]	[0.012]					
71	0.0193	0.0167	0.0162	0.0173	4,500	6.764	0.149	0.0050	cannot reject
	[0.011]	[0.013]	[0.013]	[0.013]					
72	0.0018	0.0166	0.0001	0.0206	4,500	5.055	0.282	0.0056	cannot reject
	[0.013]	[0.012]	[0.012]	[0.012]					

73	0.0027	0.0001	-0.0207	0.0165	4,500	4.944	0.293	0.0063	cannot reject
	[0.015]	[0.013]	[0.012]	[0.013]					
74	0.0015	0.0140	0.0203	0.0085	4,500	4.317	0.365	0.0071	cannot reject
	[0.008]	[0.012]	[0.013]	[0.012]					
75	-0.0153	-0.0090	-0.0130	0.0012	4,500	3.989	0.408	0.0083	cannot reject
	[0.009]	[0.012]	[0.010]	[0.012]					
76	-0.0128	0.0045	0.0121	0.0122	4,500	3.289	0.511	0.0100	cannot reject
	[0.020]	[0.012]	[0.010]	[0.013]					
77	0.0444	0.0077	0.0342	0.0059	1,070	3.208	0.524	0.0125	cannot reject
	[0.037]	[0.025]	[0.024]	[0.024]					
78	-0.0022	0.0081	0.0088	0.0144	4,500	2.301	0.681	0.0167	cannot reject
	[0.011]	[0.012]	[0.011]	[0.014]					
79	-0.0304	0.0078	0.0345	0.0235	356	1.548	0.818	0.0250	cannot reject
	[0.049]	[0.049]	[0.045]	[0.040]					
80	0.0024	-0.0045	0.0111	-0.0038	4,500	1.280	0.865	0.0500	cannot reject
	[0.010]	[0.011]	[0.011]	[0.014]					

Notes:

1. Source: Seemingly Unrelated Regressions estimation by center (2003). Numbers in main rows are coefficients on lagged emissions. Numbers in brackets are standard errors.

2. Controls include brand categories, number of cylinders categories, 3 service categories, cylinder 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 6 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 7 indicates the chi-squared statistic for the null hypothesis that all coefficients on lagged emissions are 0. Column 8 indicates the cumulative probability of the chi-squared distribution with four degrees of freedom evaluated at the chi-squared statistic. Column 9 provides the Holm-cutoff values that control for family-wise error rate at 0.05 for the top ten and bottom ten cheating centers, Column 10 indicates whether the hypothesis is rejected or not once I account for multiple tests.

Table 5: 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
d	0.61	0.1122	SE of random shock	97.52
t	60.11	9.1528		
Panel B: Fitted probabilities for each observed outcome				
Observed outcome			Actual	Fitted
			(4)	(5)
O1: Postpone			0.025	0.025
O2: Bribe/No bribe - Pass			0.763	0.747
O3: No bribe-Fail-Postpone			0.000	0.001
O4: No bribe-Fail-Bribe/No bribe-Fail-No bribe-Pass			0.150	0.175
O5: No bribe-Fail-No bribe-Fail			0.062	0.052
Panel C: Predicted probabilities for each decision sequence				
Decision sequence				Predicted

	(6)
Postpone	0.025
Bribe	0.067
No bribe-Pass	0.680
No bribe-Fail-Postpone	0.001
No bribe-Fail-Bribe	0.029
No bribe-Fail-No bribe-Pass	0.146
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 t.

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 O5. The percentage of tests with cheating can be computed by adding up the shares of vehicles that bribe in first and second attempts, 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 the last outcome of Panel C. Additional individuals that may bribe in subsequent periods are not included in the bribing estimation. However, because the probability of continuing to a third attempt is small (0.05), the bulk of bribers is still given by the sum of bribers in the first and second attempts.

Table 6: Maintenance Decisions and Summary of Benefits for Different Policy Regimes

	Actual	**Policy 1**: Less than 0.2% of Bribing	**Policy 2**: Less than 0.2% of Bribing, and less than 0.2% of Postponing	**Policy 3**: Doubling the fine for Postponing. (Results in 14% of Bribing and less than 0.1% Postponing.)	**Policy 4**: Testing for Free. (Results in 8% of Bribing, 5% Post-poning and 5% going to 3rd Retest.)	**Policy 5**: Less than 0.2% of Bribing and less than 0.2% of Postponing when retests are costly.
	(1)	(2)	(3)	(4)	(5)	(6)
1. Bribe (MXN)	193	300	475	193	193	400
2. Fine (MXN)	875	875	1850	1750	875	2000
3. Retest (MXN)	0	0	0	0	0	175
4. Percent bribing in 1st and 2nd attempts	9.6	<0.2	<0.2	14.1	7.9	<0.2
5. Percent postponing in 1st and 2nd attempts	2.6	59.7	<0.2	<0.2	5.4	<0.2

6. Percent going to a third attempt	2.1	2.1	6.4	4.7	5.1	6.1
7. Percentage with maintenance	0.1	0.1	1.6	0.9	0.3	4.2
8. Sum of emission differences (total)		191	1,595	902	384	3,708
9. Change in number of lives saved		0.18	1.55	0.87	0.37	3.57
10. Change in benefits from reduced emissions (1000 US$)		119	1,005	568	243	2,321
11. Change in private costs (1000 US $)	$10,600^{+}$	91.6	2,091	1,752	-4,447	3,593

Notes:

1. The sum of total emissions corresponds to the sum of changes in emissions of all vehicles that would be submitted to car maintenance 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. 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). The deaths from CO emissions are exclusively infant deaths. For NOx (PM10) and HC (PM10), these numbers are calculated using Schwartz (1994) estimate of 0.5222 deaths per 100,000 people per unit of TSP. For CO, I use infant mortality estimates from Arceo, Hanna and Oliva (2013) of(166.4 per 100,000 births). The Value of Statistical Life used is 650 thousand USD and is taken from Molina and Molina (2002). 2. Rows 1-3 in this table show the amounts corresponding to the policy described in each column title. Rows 4-6 show the simulated percentage of individuals that bribe, postpone, perform a third test and undergo maintenance. 3. 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).

4. The + in Column 1, Row 11 denotes level under current policy. Rows 8-11 for the remaining columns show differences between each alternative policy and the actual policy.