Environmental Regulations and Corruption: Automobile Emissions in Mexico City

Paulina Oliva*

UC Berkeley

Preliminary. Do not cite.

Abstract

Abstract: This paper uses Mexico City's smog-check center data to estimate a demand model for cheating on automobile emission tests. Cheating behavior is modeled as a sequence of decisions between bribing the mechanic and retesting the car. Each decision depends on the size of the bribe and the true probability of passing the emissions test. I rely on a non-cheating subsample of centers to estimate the true probability of passing the emissions test based on car characteristics. Non-cheating centers are identified by a statistical test for cheating that is developed in the first part of the paper. The statistical test is specific to the most common method of cheating used in Mexico City; a clean car may be tested repeatedly to provide false emission measures for high polluting vehicles. This type of cheating results in serial correlation across emission readings, which is the basic idea behind the test. I find strong evidence of corruption in all but 11 out of 80 smog-check centers. The second part of the paper uses the predicted probabilities of passing for all vehicles as an input into the bribing behavior model. Model estimates find that car owners choose to bribe in 27 percent of all tests. The model is further used to evaluate policy changes such as toughening emission standards and sorting vehicles across centers according to their model-year to reduce the availability of clean-testing vehicles for potential cheaters. The model simulations show that these alternative policies could substantially reduce cheating and boost incentives for car maintenance.

*Department of Economics, University of California Berkeley. Contact: paulinaoliva@berkeley.edu. A copy of this paper can be found at http://ocf.berkeley.edu/~olivap

I am particularly thankful to Bryan Graham for all the helpful discussions, and the invaluable advice. Without his guidance and encouragement, this paper would not have been possible. I am also very thankful for the advice received from Michael Greenstone, David Card, Edward Miguel, and Steve Raphael. This work also benefitted from discussions with Eva Arceo, John Bellows, Damon Jones, Ken Chay, Steven Levitt, Kenneth Train, Chang Tai Hsieh, Kevin Stange, Avi Ebenstein, Santiago Guerrero, Gabriela Dobrescu, Mame Fatou Diagne, Lori Beaman, Jeremy Magruder, Bob Anderson, Owen Ozier, Fred Finan, Alain de Janvry, Leslie Martin, Daniel Egel and all the participants in the U.C. Berkeley Development Seminar and U.C. Berkeley Development Lunch. Finally, many thanks to Adán Martínez, José Carlos Fernández and Ana Laura Lozada from *Instituto Nacional de Ecología*. All errors are my own.

1 Introduction

In a context of growing concern about air pollution and global warming, automobile use continues to grow rapidly in less developed countries. World vehicle production reached 70 million in 2007 and the global stock of vehicles grew at the fast rate of 6.5 percent per year (McKillop, 2007). Meanwhile, in Mexico City the average rate of growth of vehicles in use has been 10 percent between 1996 and 2006 (Secretaría de Medio Ambiente y Recursos Naturales, 2007).

Automobile emissions are an important contributor to air pollution and greenhouse gases, especially in developing countries. In Mexico City, automobile emissions are responsible for 45 percent of volatile organic compounds and 81 percent of total nitrogen oxides (Molina and Molina, 2000). The two gases are responsible for ozone formation, which is harmful for health at low atmospheric levels. In contrast, automobile emissions in the United States are responsible for about 29 and 34 percent of these gases, respectively. A younger car fleet and stricter manufacturer controls are two important factors that explain this difference. A third important factor that may explain why Mexico City is lagging behind in keeping vehicle emissions under control is extensive cheating with the environmental regulations.

With economic growth extending to a broad set of nations, controlling vehicle emissions will remain an important policy concern. Compulsory vehicle emission inspections, known as smog-checks, are the most prevalent way of enforcing emission standards on vehicles throughout the world. However, their effectiveness in reducing on-the-road emissions has been questioned widely as researchers have found sizable gaps between emission levels at official tests and in off-cycle or on-the-road samples (Livernois and Moghadam, 2005; Glazer *et al.*, 1993; Green, 1997; Wenzel *et al.,* 2003). 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, finds that 8 percent of the cars in Phoenix that passed an emission test on the first attempt will fail an immediate off-cycle re-test. And, those cars that passed the official test on the second attempt will fail an immediate re-test with a probability of 32 percent.[1] Emission testing requirements may be ineffective for reducing average emissions if emission variance is high.

Other studies have emphasized "cheating" or smog-check center fraud as the main source of test inefficiency (Zhang *et al.,* 1996; Glazer *et al.*, 1993; Wenzel 2000, and the Ministry of Environment in the Federal District of Mexico, 2004). Wenzel (2000), for example, argues that private centers in California are more prone to fraud than government owned centers in Phoenix, Arizona. In Mexico City, remote sensing studies[2] have found that on-the-road emissions are substantially higher than

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

[2] Remote sensing is an on-the-road emission measuring technique that can be used with passing vehicles. The remote sensing equipment is placed on a street and the fumes are analyzed as vehicles pass by it. The equipment

same-vehicle emissions measured at testing centers. The Ministry of Environment (Secretaría de Medio Ambiente, D.F., 2004) attributes the discrepancy to cheating, and emphasizes tampering with the engine as the main method.

Interviews with mechanics and anecdotal evidence suggest that the most common way to cheat is through clean emission substitution from clean "donor" cars to high emitting vehicles. This type of cheating requires bribing the technician to perform the substitution. Emission substitution is prefered to tampering in part due to the pemanent damage that the later can cause to computerized engines.[3]

This paper provides evidence that cheating in the form of donor cars remains an important problem for smog-check efficiency, and develops a structural model for car owner behavior that can be used to evaluate policy changes. The first part of the paper proposes parametric and non-parametric methodologies for detecting cheating, both of which are based on searching for serial correlation patterns in emissions data. Most centers in Mexico City show strong evidence of corruption by both methodologies. However, there are at least 11 centers for which the stricter parametric methodology finds no evidence of cheating. The tests from these 11 centers are used to predict emissions and the probability of passing the test for the rest of the car fleet by estimating predictive functions of observed car characteristics.

The second part of the paper develops a structural model for car owner behavior that expresses the probability of cheating as a function of the expected costs of bribing and not-bribing decisions. These expected costs involve the predicted probability of passing the test without bribing, which is predicted in the first part of the paper; and the size of the bribe, which is unknown. The model is estimated by maximizing the likelihood of the observed sequences of retesting decisions and test outcomes, given the theoretical probabilities of these sequences from the model. The likelihood maximization yields estimates for the standard bribe, the time cost associated with smog checks and the intertemporal discount rate. The model also delivers an estimate of the probability of bribing at the level of individual car owners.

The model can be used to compute the expected private cost of the smog-check as a function of the probability of passing the test. The slope of this curve alows us to indirectly evaluate the willingness to pay for car maintenance. This curve is useful to evaluate the consequences of changes in model parameters, like an increase of the bribe, on the incentives for car maintenance.

A demand curve for bribing can also be simulated by varying the amount of the bribe and

includes a camera that registers the plate number of the vehicles that are tested. There were remote sensing studies in Mexico City in 2000 and 2003 (Shifter *et al.* 2003).

[3]In addition expanding emission requirements to complementary gases has substantially reduced the amount of cheating of this sort (Secretaría de Medio Ambiente, D.F., 2004). The government now requires checking for the exahustion level of additional gases, like oxygen, which should remain under certain levels under normal combustion. Also, newer generations of vehicles are less suitable for tampering since computerized systems do not easily allow for manual alterations to the combustion mechanism.

simulating the amount of bribing at each value. The elasticity of this curve will be a function of model parameters such as the time cost of a smog check, the level of uncertainty and the time discount rate. The elasticity is also a function of the overall distribution of the probability of passing the test.

An additional assumption on the link between cheating feasibility and the number of potential donor cars provides a stilized model for bribe-taking or the "supply" of cheating. This completes a model for the market of bribing, which can be further used to evaluate changes in policy. The paper uses the model to assess the impact on cheating of a contraction of donor availability that could be achieved by sorting vehicles across centers according to their model year. It also evaluates the effect of toughening of the emission standards on the amount of cheating and the incentives for car maintenance.

The cheating prevalence rate from structural estimation is 27 percent, which is consistent with a lower bound estimate of 20 percent from the reduced form methodology. The estimates from the model suggest that the average bribe is about nine US dollars and the average time cost from attending the center is about 4.5 US dollars. According to the model, restricting the supply of donor cars by 70 percent would yield a fraud reduction of about 50 percent. This supply contraction could be accomplished by sorting vehicles across centers by model-year in order to reduce the availability of donor cars for potential cheaters. The model also suggests that this strategy would produce a threefold increase in the willingness to pay for car maintenance. In other words, relatively inexpensive policy changes could obstruct fraud and encourage standard compliance.

The results of the structural model should be interpreted carefully, since the predicted probability of passing is likely an overestimate of the true probability. This is because the sample of tests from non-cheating centers that is used for the estimation is unlikely to be fully representative of the overall car fleet. Presumably, vehicles with lower emissions, and hence in less need of cheating, will attend centers in which bribery is not available. Nevertheless, the predicted probability of passing is still lower than the observed probability of passing by 3 to 5 percent.

Using the sample on non-cheating centers, I can also predict true emissions for the overall car-fleet. I find that average hydrocarbon and nitrogen oxide emissions are, respectively, 20 and 10 percent above the level observed from smog-check data.

The rest of the paper proceeds as follows: Section 2 provides a brief description of the vehicle regulations that constitute the context in which cheating occurs. Section 3 describes the data used for the cheating detection and the structural model estimation. Section 4 proposes two different methodologies for detecting cheating. Section 5 presents results of overall cheating tests and tests at the center level. Section 6 develops a structural model for car owner's cheating behavior. Section 7 proposes a simple model for cheating availability in the centers that completes a market equilibrium model for bribes. Section 8 discusses the estimates from the structural model and policy simulations. Section 9 concludes and proposes directions for future work.

2 Vehicle Regulations in Mexico City and Smog Check Centers

2.1 Current Vehicle Regulations

Mexico City introduced biennial smog checks for all vehicles in 1990. The requirement is compulsory and universal within Mexico City Metropolitan Area (MCMA), which comprises all of the Federal District and a large proportion of the state of Mexico. Vehicle owners have a two-month window to attend any smog check center within their state and obtain an emission test certificate upon passing the test. The certificate is given in the form of a sticker that is pasted on the windshield by center staff. Non-complying vehicles are easy to spot on the road by police officers, since they do not carry a valid sticker. Fines for non-compliance range from 850 to 3500 pesos (85 to 350 dollars)[4].

Emission standards are comparable to those in the U.S. Columns 2 to 6 of Table 1 shows the emission limits in 2003. Vehicles can be retested as many times as needed within their corresponding two-month window. Each test costs 175 pesos[5] (17 dollars) and every second retest is free. The price structure of retests is fixed by the Ministry of Environment and *de facto* prices were confirmed by centers' personel during interviews.[6]

In addition to the smog check requirement, MCMA vehicle regulations include a driving restrictions program: No Driving Today (*Hoy no Circula*). This program has been in place since 1989 and originally restricted all vehicles from circulating one day a week. The day of the week a vehicle is restricted varies with the last digit of the car plate number.

Shortly after its implementation, the driving restrictions program received numerous criticisms regarding its environmental impact. In a fast growing city with limited public transportation, a common reaction to the program was the purchase of a second vehicle. In 1994, a study from UAM (Universidad Autónoma Metropolitana) found that 20 percent of the families with a second car mentioned the program as the main reason for its purchase. Out of these families only 5.4 percent used the car exclusively during the day their main car was restricted.

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

[5]Cost in 2003.

[6]A total of 5 centers were interviewed.

In the light of those criticisms, the government introduced exemptions from the driving restrictions program for cleaner vehicles. An optional second emission standard was added to the smog check requirement. Those vehicles meeting the tougher standard could be exempt from the driving restrictions program. The exemption standard and model year requirements for 2003 are presented in Column 1 of Table 1. The additional cost of the exemption is of of 4.3.

Vehicles that may qualify, here forth exemptible, are those that meet model year and vehicle use requirements as described in Table 1. Notice that in order to be exempted from the driving restriction, a vehicle must meet tougher a emission standard in addition to falling into the exemptible category. The non-exemptible category of vehicles includes model-year 1993 or older and transportation vehicles. The structural model proposed in Section 6 will focus exclusively on non-exemptible vehicles. The classification of vehicles between exemptible and non-exemptible will also be important when analyzing changes in policy.

2.2 Smog check centers and procedures

There are a total of 80 licensed centers (2003) in the Federal District and 103 in the state of Mexico[7]. These licenses were tendered in 1997 and very few new licenses have been granted ever since. All smog check centers are privately owned in MCMA, except for three institutional centers.[8] However, they are all subject to tight regulations from the government. The smog check centers are obliged to purchase their testing equipment and computer software from government-approved providers. The software, which contains the current emission norms, is renewed periodically. Centers are responsible for acquiring the newest version every year. The ministry of environment conducts unannounced inspections of the smog check center facilities. During these inspections, all mechanical and electronic equipment is checked. In addition, all facilities are required to have a camera surveillance system and publicly available video transmissions.

Smog checks consist of the following steps: At her arrival, the vehicle's owner chooses the type of certificate she wants, exemption or no-exemption, and pays the corresponding amount. Upon paying, a center employee types the vehicle's information into the computer. The information is read from the vehicle's registration card and includes plate number, model-year, make, number of cylinders, and the owner's address. In addition, the employee reads and registers the mileage from the odometer. Once the vehicle information is recorded, one or two technicians perform a visual test of the vehicle. The vehicle can fail the smog check during the visual test if there is an obvious malfunctioning of the engine or tailpipe. The information on the visual test is also typed directly into the computer. Upon passing the visual test, the vehicle is placed on the dynamometer and

[7]Numbers correspond to 2003, 1st half. This paper will focus on data from the Federal District

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

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. In the case of exemptible vehicles the vehicle owner receives the non-exemption certificate and a refund for the difference in prices if the vehicle fails the exemption standard but meets the universal standard. Finally, the corresponding test certificate is imprinted with the vehicle's plate number and the sticker with a serial number is pasted on the windshield.

The extent of cheating at smog-check centers is a major concern in Mexico City. There exist anecdotal evidence and newspaper articles that suggest that fraud is a common practice. In 2002, under-covered newspaper reporters took a car with substandard emissions to seven randomly selected smog check centers. In six of them, they were able to obtain the emission test certificate by paying an additional "tip" that ranged from five to 29 US dollars. The technicians at the fraudulent centers did not reveal any details about the cheating procedure. However, they assured the reporters the procedure would not cause damage to the car's engine (Padgett, 2002).

2.3 Cheating

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

In discussions with mechanics that have repeated interactions with smog-check center technicians, I have learned that most cheating occurs in the form of emission substitution. When bribed by a costumer, technicians at the center use a clean testing car to provide the emission readings for the bribing costumer. The second car or the clean emission-provider can be any other vehicle that passed the emission test at the center and is commonly called *auto madrina* (donor car). A second car is needed because emissions cannot be entered manually to the test record. Nevertheless, the car's information has to be typed into the computer. This allows the technicians to match the briber's car information to the emissions of a cleaner vehicle. Upon interviewing car mechanics that have repeated interactions with center technicians, they confirmed this type of cheating is the most prevalent.

An observable consequence of the cheating in the form donor car is serial correlation between consecutive emission readings. The serial correlation arises from having a single car providing consecutive emissions for more than one test. Presumably one of these tests is assigned to the vehicle it belongs by typing the correct information into the computer. The rest of the emission takes are assigned to vehicles of bribers by linking their information to the emission test record.

3 Data

This study makes use of smog check center data, which originates from each emissions test and includes vehicle information and test outcomes. The smog check computers are connected to a common network run by the local Ministry of Environment, which pool the information from all centers. The resulting data set includes information on each test and retest performed to every car that visits an authorized smog check center since. This amounts to information on at least one test for about 1,200,000 cars for Mexico State and 1,600,000 for the Federal District (D.F.) in every smog check cycle. The analysis of this draft is based on D.F. data for the second half of 2003.[9]

The information on each test includes the exact measurement on each of the relevant gases. Three out of the four gases are harmful pollutants: hydrocarbons (HC), nitrogen oxides (NO) and carbon monoxide (CO); while the fourth one is oxygen ($O2$) and it is measured to confirm the proper balance in the combustion process and avoid passing tampered vehicles. Throughout the paper I will refer to "the four pollutants" as the set of these four gases, even though one of them is not harmful for the environment.

Each test consists of two different readings of each of these four gases. The first reading is taken while the vehicle is being run at 24 kph and 50 rpm, and the second one is taken with the vehicle running at 40 kph and 25 rpm. In practice both sets of measures are very similar, especially among the three polluting gases. This similarity will be useful for improving identification in the estimation of the cheating prevalence rate. The role of the second measure will be thoroughly explained in Appendix 2.

Information available for each test also includes car characteristics (plate number, model year, brand, size of the engine, etc), test outcomes (pass/fail status, reason for failure, visual conditions of the car), beginning and ending times of the test in seconds, smog check center's information (center's identity number, and lane where the test was performed) and, for some rounds, car owner's zip code.

The main features of the data that will be used below are the full set of pollutant readings, the type of certificate requested, the result of the test, date, time and location and emission recordings

[9] There are two reasons for the choice of this particular round of data. First, neighboring rounds are needed for the structural model estimation and the rounds of data for 2000-1 and 2002-2 were not available due to format issues that may be resolved in the future. Second, later rounds have many missing values in mileage, which is an important covariate for estimations.

and vehicle characteristics. The timing information will be the basis of the cheating detection methodology, which is based on serial correlation patterns across different cars that were tested subsequently. The exact start times of the tests will also allow me to compute a time-between-tests measure, which will be used in some of the robustness checks. Also, the number of observations will allow several cuts of the data, which will be useful to test the reliability of the methodology. Finally, the availability of car identifiers allows me to put together detailed decision histories for each vehicle owner. The availability of several rounds of data are promising features that will add new dimensions to this study in later revisions.

4 Detecting cheating

As mentioned above, the strategy for detecting cheating will rely on serial correlation between consecutive tests of different vehicles. This strategy is best conveyed by observing the sequence of emission readings in the order they appear. Table 2 shows a fragment of this sequence. Column one shows the exact time at which the test was performed. Columns two to four show the model-year of the vehicle tested, the number of cylinders and the volume of displacement in the engine. The remaining columns show the reading of each pollutant measured at 24 kph and 40 kph. Notice that the sequence of readings that appears inside the black square shows striking similarities in all pollutants across tests despite the wide differences in car characteristics. All readings of this sequence could in fact correspond to the same donor car, even though they are listed under different vehicle records.

The first test, here forth permutation test, will formalize the intuitive idea that observing too many close consecutive readings may be evidence of corruption. In order to implement this test, I first construct a measure of multidimensional distance that incorporates the closeness between consecutive readings of all pollutants tested. Second, I assess the magnitude of this measure, assuming that in a no-corruption regime, changing the order of the tests should not affect the observed distance between consecutive readings. More specifically, I compare the distribution of the observed distance measure with the distribution of the same measure after randomly altering the order of the observed tests. A large number of test order permutations will generate a distribution of distance distributions. This empirical distribution of distributions under the null of no cheating can then be used to perform a statistical test on the likelihood of the observed distance distribution.

The permutation test relies on the assumption that the ordering of vehicles arriving to the testing center should be random. This assumption can be relaxed by allowing for unintended vehicle sorting across some dimensions, such as center, date, and time of the day. I do so by permutating the test order within groups defined by these sorting variables. This allows me to control for a number of dimensions in a non-parametric way. However, the number of dimensions on which unintended sorting is allowed cannot be too large, since additional controls reduce the size of the groups. The second test for cheating will propose a way to control for a broader set of sorting characteristics in a more parsimonious way.

The second test, henceforth the regression test, will test for serial correlation with an autoregressive model for each measured pollutant. A significant positive correlation between a car's test result and the test that preceded it will be interpreted as evidence of corruption. The advantage of the regression test *versus* the permutation test is that the former can control for a broader set of observed test and car characteristics. However, in contrast with the permutation test, the regression test does impose a functional form on the relationship between emissions and controls. This last caveat can be ameliorated by controlling for flexible functions of the covariates.

Both tests are conducted at the smog check center level. The results from the first methodology suggest that cheating occurs in every center. However, the figures of the non-parametric distributions of consecutive test differences also suggest that there is wide variation across centers in the amount of corruption that takes place. The regression test finds evidence of cheating in 86 percent of the centers.

Appendix 1 shows that the serial correlation coefficients from the regression methodology are also informative about the overall cheating prevalence. However, OLS estimation of these coefficients does not provide a consistent prevalence estimate. Appendix 1 shows that the prevalence estimates from this methodology can be interpreted as a lower bound on the true percentage of corrupted tests.

The rest of Section 4 will develop in detail each of the methodologies outlined above. Section 5 will compare the cheating test results from each methodology and will discuss prevalence rate estimation results.

4.1 Permutations test

Under donor car cheating, some consecutive tests are closer to each other than what we would expect from two randomly arriving vehicles. The permutations test assumes that, under a no cheating regime, car arrivals to smog check centers are random conditional on the smog check center, month, day of the month and day shift. More formally, this is equivalent to assuming car arrivals are independent and identically distributed random draws from some unknown distribution within center-date-time blocks. Under the independent arrival assumption, shifting the observed order of the tests should not affect the distribution of differences between consecutive tests in the absence of cheating. I will test for the assumption of no-cheating by comparing the distribution of distances between consecutive tests in the order they occurred with several distributions of distances when the order is randomly permutated.

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

$$d_t = \left(\sum_{j=1}^{8} (r^s_{j,t} - r^s_{j,t-1})^2 \right)^{\frac{1}{2}},$$

where $r^s_{j,t}$ is a standardized measure of pollutant j in test t, where $j = 1, ..., 8$.[10] Notice that t indexes tests in the order in which they occurred. The solid lines in Figure 1 show the distribution of d for two different smog check centers in Panels A and B.

As mentioned in the above paragraph, in the absence of cheating, the distribution of d should not change when changing the order in which tests occurred. Hence the no-cheating assumption can be tested by comparing the observed distribution of d with distributions simulated by randomly permutating the data. In order to respect potential similarities across vehicles that attend the center on a certain day or during a certain time of the day, I will shift the order randomly within center-date-time blocks[11]. Panels A and B of Figure 1 show ten different dashed lines; each of which corresponds to the distribution of d associated with a different random permutation of the data. The center depicted in Panel A shows little evidence of fraud: permutating the order of tests generates distributions of d (dashed lines) that are very similar to the observed one (solid line). In contrast, the center depicted in Panel B shows strong evidence of corruption: the observed distribution of d (solid line) has a larger proportion of small consecutive differences in tests compared to 10 different distributions of d where cars are assumed to arrive randomly to the center (dashed lines).

A formal statistical test can be derived from this methodology. First, I compute a test statistic that describes the relevant part of the distribution: the fraction of test differences close to zero. The chosen statistic is the fifth percentile, $\hat{q}^d_{.05}$, of the empirical distribution of d.[12] I will compare this statistic with the distribution of the fifth percentile of d over 1000 different permutations of

[10]The standardization is necessary to give equal weight to each of the four measures since pollutants are measured in different scales. Define $r_{j,t}$ as the observed measure of pollutant j for test t. Define also μ_j as the sample mean of $\sqrt{r_{j,t}}$. Also, define σ_j as the sample standard deviation of $\sqrt{r_{j,t}}$. The standardization used is given by $r^s_{j,t} = \frac{\sqrt{r_{j,t}} - \mu_j}{\sigma_j}$. Other standardizations, such as $ln(r_{j,t})$, yield similar results.

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

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

the smog test ordering. The comparison yields p-values for the null hypothesis of no cheating. The test is performed at the center level for each of the 80 smog check centers in the 2003-2nd-half round of D.F.'s smog check center data. The p-values for all centers are less than 0.001, which rejects the null hypothesis of no-cheating for all centers with 99.9 percent of confidence.

The permutation test generates very clear results regarding the extent of cheating across centers. According to this test, all centers participate in fraud. Results are robust to changes in the definition of the blocks within which data is reordered and to alternative test statistics (alternative results not shown). However, the test relies on the assumption of random arrival of vehicles to the center. One may argue that, even after controlling for some of the time and spatial variables that may be correlated with car arrival at the centers, similar vehicles may tend to group together generating the observed serial correlation. The following subsection proposes an alternative methodology for testing for corruption that controls for vehicle characteristics in addition to center, lane and time of the day. This alternative methodology offers the additional advantage of providing estimates that are informative about corruption prevalence rates.

4.2 Regression methodology

The regression methodology consists of a regression model of smog check center emission measures on lagged emissions, or emissions of the previous vehicle and control variables. A significant coefficient on lagged emissions can be interpreted as evidence of corruption under the assumptions outlined in this section.

Denote $\tilde{r}_{jit}$ the true emissions of pollutant j,[13] for car i at time t. To simplify the notation, I will omit the index j in the analysis that follows. However, keep in mind that the regression analysis proposed below applies to a single pollutant at a time.[14]

Denote the linear prediction of true emissions given car characteristics as: $\mathbb{E}^*(\tilde{r}_{it}|\mathbf{x}_i) = \mathbf{x}_i\beta$, where $\mathbf{x}_i$ is a vector of observable car, test and center characteristics[15] and $\mathbb{E}^*$ is the linear predictor operator. Hence, true emissions are given by the following equation:

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

where $u_{it} = \tilde{r}_{it} - \mathbf{x}_i\beta$. In the current setup, the index i will be a car specific identifier that will also keep track of the order in which vehicles show up at the center. For example, the vehicle that was tested after vehicle i will be denoted $i+1$. Emission components in u_{it} are indexed by time, t,

[13]In this case, j will denote one of the four gases that are involved in the test when measured at 24 kph.

[14]The availability of four different pollutants will play an important role in the analysis of consistency of the coefficients on lagged emissions when interpreted as cheating prevalence estimates (Appendix 2).

[15]These include brand, service and size fixed effects, a flexible function of the age of the car, age-size and age-service interactions, flexible functions of mileage and time of the day and testing equipment fixed effects

since emissions of the same car may vary from test to test. The index t will then denote the actual, but unobserved, order in which vehicles were checked or "test slot". The distinction between i and t is necessary only when cheating occurs. Under no cheating, these two indexes should be one to one. However, in the presence of cheating, a vehicle may be tested in two or more consecutive time slots. For example, if i is a cheater, $i-1$ is a donor car and $i+1$ is not a cheater, then the smog check data set will have the following sequence of emissions on time slots $t-1$ to $t+1$: $\tilde{r}_{i-1t-1}, \tilde{r}_{i-1t}, \tilde{r}_{i+1t+1}$. If none of the three cars are cheaters, then the sequence of emissions in the smog check record will be: $\tilde{r}_{i-1t-1}, \tilde{r}_{it}, \tilde{r}_{i+1t+1}$.

It is plausible to assume that, under the null of no cheating, there should be no correlation between subsequent emission readings, conditional on car and center characteristics. The regression methodology outlined below requires an even weaker set of assumptions: that the unobserved components of linearized emissions, u_{it}, are serially uncorrelated and that the correlation between unobserved component of emissions and every observed car characteristic of the preceding vehicle is equal to zero:

(A1) $\mathbb{E}(u_{i-1t-1}u_{it}) = 0$

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

Notice that, by construction, u_{it} includes only emission components that are uncorrelated with observed emission components. Hence, assumption (A1) and (A2) would only be violated if, in the absence of cheating, vehicles are sorted based on unobserved emission components that are uncorrelated with $\mathbf{x}_i$.

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

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

where r_{it} are observed emissions of car i in slot t, $\tilde{r}_{i-k,t}$ is a true measure of emissions from a vehicle that was tested k slots before and c_i is a binary variable, such that $c_i = 1$ if i is a cheater. Equation (2) assumes that the donor car is always tested before the cheating car (in test slot $t-k$) and on the same lane as the cheating car. However, it allows for the donor car to be an emission provider for other cheating cars in test slots $t-k+1$ to $t-1$.

Consider the regression of r_{it} on r_{i-1t-1} and $\mathbf{x}_i$:

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

where r_{i-1t-1} is the observed emission reading for car $i-1$ in slot $t-1$. The OLS estimate of coefficient γ_c can be used as a test statistic for the null hypothesis of no cheating . To see this, notice that under assumptions (A1) and (A2) a non-cheating regime implies $\mathbb{E}^*(r_{it}|r_{i-1t-1,}\mathbf{x}_i) = \mathbf{x}_i\beta$.

Second, notice that if vehicle i is a cheater, vehicle $i-1$ will be either the donor car for i or a cheater using the same donor as i.[16] Hence, its observed emissions, r_{i-1t-1}, will be equal to $\tilde{r}_{i-k,t-1}$, where $i-k$ is the donor car. Observed lagged emissions from the donor car, $\tilde{r}_{i-k,t-1}$, differ from unobserved donor car emissions, $\tilde{r}_{i-k,t}$, only by the time varying component of emissions in u_{it}. Assume u_{it} can be decomposed as $u_{it} = a_i + e_{it}$, where a_i is an unobserved but fixed determinant of emissions and e_{it} is an unobserved and time varying component of emission measures. This time varying component may arise from machine measuring error or from variations in emissions within the same vehicle. Then $\tilde{r}_{i-k,t-1} = \tilde{r}_{i-k,t-k} + e_{i-k,t-1} - e_{i-k,t-k}$. The difference between $\tilde{r}_{i-k,t-1}$ and $\tilde{r}_{i-k,t}$ may lead to underestimating the serial correlation between emissions when cheating occurs. However, if $c_i = 0$ for all i , γ_c should still be zero. Hence, a serial correlation coefficient significantly larger than zero will lead to rejecting H_0.

4.3 Cheating prevalence rate interpretation

The serial correlation coefficient on equation (3) can also be interpreted as the mean prevalence of cheating. To see this, consider the equation

$$r_{it} = p_0 r_{i-1t-1} + \mathbf{x}_i\gamma_x + \nu_{it}, \tag{4}$$

and compare them with equation (2). Given this comparison, p_0 can be interpreted as $\Pr(c_i = 1)$, or the cheating prevalence rate, and $\gamma_x = (1-p_0)\beta$, while $\nu_{it} = r_{it} - p_0 r_{i-1t-1} - \mathbf{x}_i\gamma$. Parameters p_0 and γ_x can be estimated as the coefficients on r_{i-1t-1} and $\mathbf{x}_i$, respectively and the parameter of interest in the above equation will be p_0. There are two issues with consistent estimation of the cheating prevalence rate using the regression framework. The first one is measurement error in the regressor r_{i-1t-1}, which was mentioned briefly in the section above. Although measurement error does not compromise the consistency of the test for cheating, since under the null $p_0 = 0$, it will result in a downward bias for $\hat{p}_0$ from OLS estimation. The second issue with consistency is correlated random coefficient bias. Notice that p_0 is in fact the average value of a random

[16] If centers used a single car, *e.g.* the technician's car, as emission donor instead of incoming vehicles, the donor car wouldn't need to be tested in advance. Observed emissions would then be given by $r_{it} = c_i\tilde{r}_{kt} + (1-c_i)\tilde{r}_{it}$, where $\tilde{r}_{kt}$ is the emission mesasure of the fixed donor car at center k in time t. In this case an OLS regression of r_{it} on r_{i-1t-1} would not detect corruption unless the donor car was used consecutively more than once. As long as some cheating occurs with incoming donor cars as opposed to fixed donor cars, or fixed donor cars are used more than once consecutively, cheating will be detected when testing for the significance of γ_c in equation (3). Upon interviewing a mechanic that is aware of the cheating procedures, I confirmed that donor cars are usually incoming vehicles as opposed to a fixed vehicle in the garage.

coefficient, c_i, since the true model is given by equation (2). In model (4), the random part of c_i, $c_i - p_0$, is left in the error term ν_{it}. If c_i is uncorrelated with r_{i-1t-1}, the randomness in c_i would not affect the consistent estimation of p_0. However, c_i and r_{i-1t-1} are likely negatively correlated since for cheating vehicles $c_i = 1$ and the emission of the previous car in line (the donor car) are presumably lower than the emissions for an average vehicle. The negative correlation between c_i and r_{i-1t-1} will cause downward bias in the OLS estimate of p_0 from regression (4). In the light of these sources of bias, the coefficient on r_{i-1t-1} can be regarded as a lower bound on the cheating prevalence rate. Appendix 2 expands on these sources of bias and studies different specifications that can ameliorate the downward bias. The most reliable estimates suggest that cheating is present in at least 20 percent of the tests.

5 Empirical results from regression methodology

Estimates of cheating prevalence rates can be obtained from the smog-check center data by using the methodology described in the previous section. Specification (4) can be estimated for all four gases observed in the data. Since the data contains the exact time of the test, the tests can be sorted in order to construct a lagged emission variable for each gas. Lagged emissions, r_{i-1t-1}, will be the right-hand-side variable of interest throughout this section.

5.1 Test results and cheating prevalence

Table 3 presents OLS estimation results for specification (3) on all tests for the second smog check cycle of 2003. Each row corresponds to each of the four gases measured during the test at 24 kph.[17] The first column presents the serial correlation coefficients without controls for car characteristics. These results can be interpreted as a parametric version of the permutation test for cheating, since unintentional sorting of vehicles across observed car and test characteristics is not controlled for. The test for corruption rejects the null of no cheating for all 4 pollutants. The second column controls for car charactistics such as brand, use of the vehicle (private, transportation), size of the engine, size and service specific age trends and flexible functions of vehicle's age, mileage and time of the day. The serial correlation estimates fall from previous specification, suggesting there may be some sorting of vehicles across centers or by the time of the day. Specifications reported on the third and fourth columns control for center and testing equipment fixed effects respectively.[18] Although the coefficients fall slightly from previous specifications, the evidence of corruption remains significant for all pollutants at the 1 percent level.

[17]Results for gas measures at 40 kph (not shown) are very similar.
[18]Testing equipment indicators are center-specific.

As mentioned in the previous section, the serial correlation coefficients have a prevalence of corruption interpretation from equation (4). However, the serial correlation coefficient under the alternative hypothesis is inconsistent, hence this coefficient is only informative as a lower bound on corruption prevalence. Measurement error and correlated random coefficients produce a downward bias in the prevalence rate estimate. Notice that the actual prevalence rate parameter should be the same in all four specifications. However, the size of the bias in the prevalence rate estimate varies across pollutants. The size of the measurement error depends on the variance of the time varying component of the gas; and the correlation between the random coefficient and the regressor depends on how stringent is the selection of donor cars on that particular pollutant.[19] This explains the variation in coefficient estimates across pollutants.

The case of the oxygen specification is particularly interesting. Oxygen is not a pollutant *per se*; however, centers test for the percentage of oxygen in the emissions to avoid cheating in the form of tampering. This type of cheating sometimes involves piercing the tail pipe or affecting the amount of oxygen that enters the combustion process. Both of these strategies result in oxygen levels above the normal standards. Vehicles owners know this requirement exists, hence tampering in these forms has been reduced. As a result, 99.7 percent of vehicles have oxygen levels under the threshold. Since meeting the oxygen threshold is not challenging, donor car selection is unlikely to be based on this pollutant. Hence, correlated random coefficient bias should be smaller in the oxygen specification. This is consistent with results of Table 3, which show the largest prevalence rate estimates for oxygen (fourth row).

Appendix 2 discusses this interpretation of the serial correlated coefficient in more detail and proposes an IV strategy that eliminates the measurement error bias under plausible assumptions. Appendix Table A1 compares the OLS and IV results. The best estimates for corruption prevalence rate correspond to the IV-oxygen specification, since measurement error is controlled for and correlated random coefficient bias is minimal. These estimates provide a lower bound for the corruption prevalence rate of 20 percent.

5.2 Robustness tests

In order to confirm that serial correlation originates from cheating, I conduct two different robustness tests. The first one relies on the test length difference between a normal test and a corrupt test. Corrupt tests may take less time to be completed, since the donor vehicle remains on the dynamometer instead of placing a different car. This can be tested by interacting time between tests with lagged emissions. Appendix Table A2 shows the results of a regression of vehicle emissions on a set of indicator variables for various time lengths between tests and their interaction with lagged

[19] See Appendix 2 for a thorough explanation of these biases.

emissions. The table shows that the serial correlation is more important for short time intervals between tests (0-5 min) than for long time intervals between tests.

The second robustness tests confirms that serial correlation is related to cheating by comparing serial correlation between vehicle pairs where the first vehicle just passed the test and vehicle pairs where the first vehicle just failed the test. If donor cars are selected based on whether they passed the test or not, pairs with a failing first car should exhibit no serial correlation, while pairs with a passing first car should.

I restrict the analysis to vehicles around the emission threshold (those who barely passed and those who barely failed), since failing to do so could lead to spurious correlations between subsequent tests. Since most vehicles pass the emission test, restricting the analysis to pairs with one failing vehicle could induce a negative correlation between the pair. However, I avoid this problem by focusing on pairs with similar emissions for the first vehicle.

The results of this analyisis are shown in Appendix Table A3. Since the large majority of vehicles pass the oxygen test, there is no substantial number of tests around the threshold. Hence, Table A3 shows results for the remaining 3 pollutants. Columns 1 and 2 show specifications for vehicles that just passed while columns 3 and 4 show similar specifications for vehicles that just failed. The serial correlation estimate is very different between the two groups. In the case of HC and NO, there is no significant serial correlation for vehicles that just failed. This is consistent with serial correlation being a product of corruption in the form of donor cars.

5.3 Detecting cheating at the center level and predicted emissions

Because of the large number of test observations per center (usually between 10 and 30 thousand), prevalence rates can be estimated at the center level, yielding results on the extent of cheating across centers. I estimate equation (3) at the center level controlling for all car and test characteristics. I find evidence of corruption at the 5 percent level for 69 out of 80 centers for at least one pollutant. Panel A of Table 4 shows the test results for the 11 centers that have no evidence of cheating and Panel B show results for the 6 centers that have evidence of cheating under all four specifications.[20]. P-values are in brackets below each coefficient. The remaining centers (not shown) have evidence of corruption for one to three pollutant specifications.

Notice that the first two centers in Panel A have very few tests throughout the smog check cycle. These two centers are called "institutional" centers, since they are operated by a government institution as opposed to a private entity with a government license. The first center is operated by the Department of Defense (*SEDENA*) while the second one is operated by the National Power and Electricity Company (*Luz y Fuerza del Centro*). There is only one other institutional center,

[20] Standard errors are clustered at the testing equipment level.

which shows evidence of cheating in the nitrogen oxide specification (result not shown). The fact that two of these government operated centers have no evidence of cheating is an interesting finding given Wenzel's (2000b) observation that government centralized centers in the U.S. are less prone to fraud. However, the institutional centers in Mexico seem to differ considerably from private centers since the number of tests they conduct per cycle is very small.

The emission test outcomes from centers in Panel A are used to estimate true emissions for the rest of the car-fleet. It is important to keep in mind that vehicle selection into centers is not random. High polluting vehicles may avoid centers where bribing is not available. If there are unobserved emission components that affect self selection into centers, then an emission prediction based on observed car cahracteristics may underestimate actual emissions. Panel A of Table 5 analyzes self selection into non-cheating centers by comparing observed car characteristics across cheating and non-cheating centers. Most differences in car characteristics are consistent with high polluting vehicles selecting out of centers with no cheating. Non-cheating centers are avoided by VW and Chevy economy[21] vehicles and larger vehicles, and more heavily attended by small and medium sized vehicles. Cars with bigger engines (higher displacement volume) attend cheating centers more often. However, vehicles do not differ in age across the two types of centers, and vehicles that attend non-cheating centers seem to have traveled more kilometers.[22]

Panel B of Table 5 shows that, despite selection, all measured emissions are significantly higher in non-cheating centers. However, the probability of passing the emission threshold for non-exemptible vehicles is only lower by 3 percent in the non-cheating sample.

Observed and predicted average emissions for the whole car fleet are shown in Table 6. Separate averages are shown for exemptible and non exemptible vehicles.[23] Observed emissions from non-exemptible vehicles are 5 times as large as emissions from exemptible vehicles for some pollutants. Predicted emission are 16 percent higher than observed HC emissions, 17 percent higher than observed NO, and 9 percent higher for CO. The differences in NO are more important for non-exemptible vehicles (25 percent) while the differences in HC are more relevant for exemptible vehicles (up to 22 percent).

5.4 Predicting the probability of passing

Section 6 will develop a structural model for bribing behavior that will translate the car-specific probability of passing the emissions test into the probability of bribing in each smog-check center's visit. The true probability of passing the test for each car type cannot be computed directly

[21]These category includes old VW Beatles, which tend to be high polluting vehicles because of their age and heavy use.

[22]Cumultive kilometers are somewhat unreliable as a measure of vehicle use since reseting the odometer is a common practice in Mexico.

[23]See Section 2.1 for a characterization of the two groups.

from the smog check center data, since cheating will lead to overestimating the probability of passing. However, the tests from non-cheating centers can be used to estimate a probit model for the probability of passing the test as a function of car characteristics . The estimated coefficients on car characteristics can then be used to make an out-of-sample prediction of the probability of passing the test for the rest of the car fleet.[24]

The last column of Table 6 compares the observed and the predicted probabilities of passing the test for vehicles with no access to the exemption. Notice that mean observed and predicted probabilities of passing differ by only 3 percentage points. It's important to point out that the predicted probability of passing may be biased upwards because of sample selection into low-cheating centers (see Table 5). The predicted probability of passing the test and the observed test outcomes will be used for estimating the structural model for bribing behavior.

Since they are used to predict the probability of passing the test for the rest of the car fleet, the correct identification of the non-cheating centers is key for the identification of the structural model parameters. Hence, I test for whether car characteristics are indeed better predictors of vehicle emissions in the sample of non-cheating centers than for centers with substantial evidence of corruption. In order to do so, I classify centers in five groups (numbered 0 to 4) according to the number of different specifications (HC, NO, CO and $O2$) for which the null hypothesis of no corruption cannot be rejected. Centers with four failures to reject correspond to the sample I used for predicting emissions and the probability of passing, while centers with zero failures to reject are the ones listed in Panel B of Table 4. I then run a simple linear regression of emissions on model-year and a constant for each subsample. The slope coefficient of this regression should be more negative for the group of centers in which evidence of cheating does not exist (group 4) and less negative for centers with high evidence of corruption (group 0). Appendix Figure A1 depicts the coefficients and their confidence intervals for each group. Groups are numbered on the horizontal axis of each panel. The figure has a panel for each pollutant. Notice that the coefficients generally become more negative as the evidence of cheating becomes less persuasive. Model-year has the largest effect in absolute value on all pollutants when the sample is constrained to centers with no evidence of corruption.[25]

The next three sections of this paper will develop a structural model for the market of bribes that will provide estimates of the prevalence of cheating and will be helpful for analyzing policy changes that can improve the emission testing program's effectiveness.

[24] Appendix Table A5 shows the probit model coefficients on the non-cheating sample of centers.

[25] I am thankful to David Card for suggesting this falsification test.

6 A model of bribing behavior

The above results show that cheating is a major concern for emission control policy in Mexico City. A better understanding of the key factors behind the decisions to cheat will help redesign policy to decrease the amount of cheating and create incentives for lowering emissions. This section proposes a structural model for car owner behavior that will translate information on the probability of passing the test (predicted for each vehicle in the previous section) and the information on the costs of the test into a probability of bribing for each vehicle. In Section 8, the model proposed will also allow me to estimate the expected utility cost of the test as a function of the probability of passing the test. The slope of this curve can be interpreted as the willingness to pay for car maintenance and car renewal.

6.1 Model set-up

In the following paragraphs I will develop structural model for the demand of cheating that can be estimated using the predicted probabilities of passing the test from the previous section and the car owner decisions observed from smog-check center data. The model for the demand of cheating assumes that each vehicle has an inherent probability of passing the test and that this probability is known by the owner. According to the model, the owner decides whether to bribe or not during her visit to the smog-check center based on the probability of passing and the costs associated with each option.

Since retesting is not restricted by the current regulation, the model assumes that car owners can retest their vehicles indefinitely. At the beginning of each testing or retesting round car owners decide between bribing the technician (B), not bribing the technician (A) and postponing the check beyond the deadline (X). The car owner will choose among these three alternatives based on the expected cost, or utility loss, associated with each of them.

The official test's price scheme is set by the government. The price of the first attempt to pass the test is 17.5 US$. The price of every subsequent even attempt is 0, and the cost of every subsequent odd attempt is 17.5 US$. Hence, attempts 2, 4, 6, etc. are free as long as the car owner returns to the smog check center where the previous attempt failed. The fine for postponing the smog check beyond the deadline is 80 US$, and it has to be paid to the bank before any subsequent visit to the smog check center. This fine is hard to avoid since the car owner cannot smog-check her car again unless she provides the proof of payment of the fine from an authorized bank. The proof consists on a form with a unique numeric code. During the next visit to a smog-check center, the authenticity of the code is verified. If the code is valid, the belated check can proceed. The fines can be accumulated if the vehicle owner skips more than one smog check cycle.

Figure 3 represents the car owner's alternatives and payoffs in the form of a decision tree.

Each circle represents the beginning of a decision round. Each branch represents a choice in each decision round. Payoffs appear at the end of each branch in the tree: w is the time cost associated with each visit, b is the amount of the bribe, p is the price of odd attempts (17.5 US$), z is the probability of passing the test, δ is a discount factor for "future" costs and f is the fine that applies to belated smog checks (85 US$).

This model will assume that the bribe, b, is constant across individuals. This assumption is consistent with a competitive model for bribing. Competition is a plausible assumption to make given the large number of smog check centers in Mexico City (80 in Distrito Federal) and the fact that a vehicle can choose any center to perform the smog check. Also, it is important to note that centers cannot compete in test prices, since these are fixed by the government. Hence, they may resort to lower bribes in order to attract costumers. The assumption of perfect competition in the bribing market will be further discussed in Section 7, where I propose a simple model for the market of bribing. The time cost of going to the center, w, will be allowed to vary across individuals.[26]

Figure 3 also represents the dynamics of smog check-related decisions. Every smog check cycle starts with an odd round that can occur in any point in time within the permissible two-month window. At this point, the car owner decides between choices X, A and B. If the choice is to postpone the test, X, then the car owner will transfer the smog check's cost to the future; hence, the cost will be discounted by the factor δ. Notice that in Figure 1, the choice X is followed by an arrow that ends at the beginning of a new round. This arrow is accompanied by the greek letter δ, that denotes the discount factor that will apply to the cost of starting a new round the following period. The cost of alternative X will also include the discounted value of the fine $-\delta f$.

If the car owner chooses to bribe, B, the smog-check cycle ends and the car owner receives her smog-check certificate upon paying the bribe, b, the smog check cost, p, and spending some time at the center, which costs w. If the car owner chooses not to bribe, A, then she will pay the cost, c, spend the same amount of time at the center, which has a cost of w, and pass the test with probability z. With probability $1 - z$, she will enter a new decision round. The following decision round is similar, except options A and B do not carry a check cost, p. The car owner can potentially enter an infinite amount of decision rounds if she chooses A every time. However, the larger the probability of passing, z, the fewer subsequent rounds she will face in expectation.

The decision framework simplified in Figure 3 can be further formalized in a utility maximization model. For simplicity, I will assume utility from choice S (where $S = X, B, A$), in smog check cycle m and decision round t, is linear in money. I will also assume that the utility from each

[26] When estimating the model, I try different specifications for the time cost that depend on either the cost of the car (as a proxy of income) or the randomly assigned period for smog checking (either June-July, July-August, August-September, etc.). This last specification tries to model variation in individual time availabiility across months.

option, S, has a random component $\epsilon^S_{m,t}$ that is independent across choices and across decision rounds. I assume that the random utility shocks, $\epsilon^S_{m,t}$, can only be observed at the beginning of round t. These utility shocks can be understood as unobserved events that may change the relative cost of each choice, for example: unforeseen time constraints that make postponing the test more attractive, an inspector's visit to the center that makes bribing very expensive, etc. For computational purposes, I will further assume that the shocks are extreme value distributed:

$$\epsilon^S_{m,t} \sim EV(1) \text{ for } S = X, A, B$$

Since the odd attempts cost p while the even attempts are free, the utility from each alternative will also vary between odd and even smog check rounds. The expected utility from each alternative when t is odd will be given by:

$$u^{odd,B}_{m,t} = -w - p - b + \sigma\epsilon^B_{m,t}$$

$$u^{odd,A}_{m,t} = -w - p + (1 - z) \cdot \mathbb{E}_{m,t}(V^{even}_{m,t+1}) + \sigma\epsilon^A_{m,t} \tag{5}$$

$$u^{odd,X}_{m,t} = \delta(-f + \mathbb{E}_{m,t}(V^{odd}_{m+1,1})) + \sigma\epsilon^X_{m,t},$$

The expected utility for each alternative when t is even will be given by:

$$u^{even,X}_{m,t} = \delta(-f + \mathbb{E}_{m,t}(V^{even}_{m+1,1})) + \sigma\epsilon^X_{m,t}$$

$$u^{even,B}_{m,t} = -w - b + \sigma\epsilon^B_{m,t} \tag{6}$$

$$u^{even,A}_{m,t} = -w + (1 - z) \cdot \mathbb{E}_{m,t}(V^{odd}_{m,t+1}) + \sigma\epsilon^A_{m,t}$$

where σ denotes the standard deviation of the utility shock.

In the utility functions above the terms $\mathbb{E}_{m,t}(V^{even}_{m,t+1})$ and $\mathbb{E}_{m,t}(V^{odd}_{m+1,1})$ denote the expected utility from the following decision round. For example, the utility of not-bribing on an even node, $u^{odd,A}_{m,t}$, is the cost of the test, $-w - p$, plus the probability of failing times the expected value of an even decision round at the smog check. The value of decision round t in cycle m is given by:

$$V^{odd}_{m,t} = \max(u^{odd,X}_{m,t}, u^{odd,B}_{m,t}, u^{odd,A}_{m,t} | \Omega_{m,t})$$

$$V^{even}_{m,t} = \max(u^{even,X}_{m,t}, u^{even,B}_{m,t}, u^{even,A}_{m,t} | \Omega_{m,t}),$$

where $\Omega_{m,t} = \{\epsilon^X_{m,t}, \epsilon^B_{m,t}, \epsilon^A_{m,t}\}$. Notice that the expectation operators in (5) and (6) are indexed by m, t, which means that the expected value is "computed" at cycle m and decision round t. The random utility shocks for $t' > t$ and $m' > m$ are unknown at m, t. I assume, however, that the car owner knows the distribution of the utility random shocks when computing the expected value.

The extreme value distribution assumption for the random shocks in the model has a very important advantage for the model's solution. Using McFadden and Domencich's result (1996), the extreme value distribution assumption results in a closed form expression for the expected value of subsequent rounds:

$$\mathbb{E}_{m,t}(V^{odd}_{m,t}) = \tau\left(\kappa + \log\left(\exp\left(\frac{u^{odd,X}_{m,t}}{\tau}\right) + \exp\left(\frac{u^{odd,B}_{m,t}}{\tau}\right) + \exp\left(\frac{u^{odd,A}_{m,t}}{\tau}\right)\right)\right),$$

and

$$\mathbb{E}_{m,t}(V^{even}_{m,t}) = \tau\left(\kappa + \log\left(\exp\left(\frac{u^{even,X}_{m,t}}{\tau}\right) + \exp\left(\frac{u^{even,B}_{m,t}}{\tau}\right) + \exp\left(\frac{u^{even,A}_{m,t}}{\tau}\right)\right)\right);$$

where κ is the Euler constant and τ is the scale parameter of the random utility shock distribution, which is related to σ by $\sigma = \left(\frac{\pi^2}{6}\tau\right)^{\frac{1}{2}}$. An infinite horizon for decision rounds, $t = 1, 2, ..., \infty$, implies that all the expected costs from all odd rounds are identical and the expected costs from all even rounds are identical as well. This will allow me to solve for the stationary expected value of all even and all odd nodes. The estimation of the model uses a nested fixed point algorithm to recompute the expected value of subsequent decision nodes as a subroutine to the standard maximum likelihood problem (Rust, 1987) [27]

The model assumes that car owners make optimal decisions given the values of the parameters and their probability of passing the test. For example, an individual will choose to bribe in an odd round if $u^{odd,B}_{m,t} > u^{odd,A}_{m,t}$ and $u^{odd,B}_{m,t} > u^{odd,X}_{m,t}$. Given the extreme value distribution of the random shocks, the probability that this occurs is given by a multilogit-like expression:

[27] Define $\mathbb{E}V^{ev} = \mathbb{E}_{m,t}(V^{even})$ $\forall$ all m, t s.t. t is even; and $\mathbb{E}V^{od} = \mathbb{E}_{m,t}(V^{odd})$ $\forall$ all m, t where t is odd. Then,

$$\mathbb{E}V^{ev} = \tau\left(\kappa + \log\left(\exp\left(\frac{u^{even,X}_{m,t}(\mathbb{E}V^{ev}))}{\tau}\right) + \exp\left(\frac{u^{even,B}_{m,t}}{\tau}\right) + \exp\left(\frac{u^{even,A}_{m,t}(\mathbb{E}V^{od}))}{\tau}\right)\right)\right)$$

$$\mathbb{E}V^{od} = \tau\left(\kappa + \log\left(\exp\left(\frac{u^{odd,X}_{m,t}(\mathbb{E}V^{od})}{\tau}\right) + \exp\left(\frac{u^{odd,B}_{m,t}}{\tau}\right) + \exp\left(\frac{u^{odd,A}_{m,t}(\mathbb{E}V^{ev})}{\tau}\right)\right)\right)$$

The values $\mathbb{E}V^{ev}$ and $\mathbb{E}V^{od}$ will be solved numerically such that the above two equations hold. I obtained informal evidence of uniqueness for the solution of the system above by comparing the solution across random starting values.

$$\Pr(S^{od}=B)=\frac{\exp\left(\frac{-w-p-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}{1+\exp\left(\frac{-w-p-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)+\exp\left(\frac{-w-p+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)},$$

where X is the base category. Similar probability expressions can be derived for each of the other choices from each type of round (see Appendix 3):

6.2 Estimation

The parameters to estimate are w, the discount factor, δ, the scale parameter of the extreme value distributed errors, τ, and the equilibrium bribe b. The smog check price, p, and the fine, f, are known and constant. The probability of passing the test, z, will be estimated by its predicted value given the car's characteristics and the parameter estimates from a low-cheating subsample.

The parameter estimates will be obtained my maximizing the likelihood of the observed individual decisions and outcomes given the model's probability of each decision-outcome history. A history of decisions and test outcomes for a given individual on a given smog check cycle is given by the sequence of bribe/no-bribe/postpone decisions and pass/fail test outcomes. In other words, a history of decisions and test outcomes is a possible path on the tree described in Figure 3, until a final payoff is reached. An example of a history of decisions and test outcomes would be: not bribing in the first test, failing, then bribing in the second test.

The complete set of decision and test outcome histories is described below:

$H1$ $\quad S_{1,m}=X$

$H2$ $\quad S_{1,m}=B$

$H3$ $\quad S_{1,m}=A,\ \text{test}_{\mathit{1}}=\mathit{pass}$

$H4$ $\quad S_{1,m}=A,\ \text{test}_{\mathit{1}}=\mathit{fail},\ S_{2,m}=X$

$H5$ $\quad S_{1,m}=A,\ \text{test}_{\mathit{1}}=\mathit{fail},\ S_{2,m}=B$

$H6$ $\quad S_{1,m}=A,\ \text{test}_{\mathit{1}}=\mathit{fail},\ S_{2,m}=A,\ \text{test}_2=\mathit{pass}$

$H7, H8, ...$ $\quad S_{1,m}=A,\ \text{test}_{\mathit{1}}=\mathit{fail},\ S_{2,m}=A,\ \text{test}_2=\mathit{fail}...$

The history described in the previous paragraph corresponds to $H5$. The sequence of potential histories continues to infinity, since a vehicle can get retested as many times as the car owner chooses. However, I will not need to consider every potential history for the estimation; the first few will suffice. The reason for this will become clear below.

An expression for the probability of each of these histories can be derived from the probabilities for each decision and the probability for passing the test. For example, the probability of history $H4$ can be derived from the model to be:

$$\Pr(H4) = \Pr(S^{od} = A) \cdot (1 - z) \cdot \Pr(S^{ev} = X),$$

where $\Pr(S^{od} = A)$ is the theoretical probability of attempting without bribing in the first node, $1 - z$ is the probability of failing the test and $\Pr(S^{ev} = X)$ is the theoretical probability of postponing the test to the next cycle. The model probabilities for each remaining histories are listed in Appendix 3.

Not all decision and test outcome histories are separately observed in the data. More specifically, it is impossible to distinguish between histories $H2$ and $H3$. In both cases the car owner will be observed to have a single try at the smog test and obtain a pass. In a similar fashion, it is impossible to distinguish between histories $H5$ and $H6$. In both cases we observed that the car owner fails her first try and then passes her second try. However, the model parameters can be identified from fitting the sum of probabilities corresponding to the confounded observed history. For example, the theoretical probability for the confounded set of histories $(H2, H3)$ is

$$\Pr(H2 \text{ or } H3) = \Pr(H2) + \Pr(H3) = \Pr(S^{od} = B) + \Pr(S^{od} = A) \cdot z$$

The probabilities for the remaining four observable histories are listed in Appendix 3.

The log-likelihood function for the estimation will be given by the following expression:

$$\begin{aligned} lnL(w, b, \delta, \tau | f, p, z_1, ..., z_N) = \\ \sum_{i=1}^{N} D_i^{H1} \cdot \Pr_i(\mathit{H1}) + D_i^{\mathit{H2,H3}} \cdot \Pr_i(\mathit{H2} \text{ or } \mathit{H3}) + D_i^{\mathit{H4}} \cdot \Pr_i(\mathit{H4}) + \\ D_i^{\mathit{H5,H6}} \cdot \Pr_i(\mathit{H5} \text{ or } \mathit{H6}) + D_i^{\mathit{H7,...}} \cdot \Pr_i(\mathit{H7}, ...), \end{aligned}$$

where i indexes vehicles, N is the total number of vehicles and D_i^H is a dummy variable that takes the value of one if i's history is described by H. The likelihood will be maximized with respect to parameters w, b, δ and τ.

7 A model of the market for bribes

The market of bribes at the smog-check centers is unique among other types of corruption in that the emission check requirement can be provided by many different agents. Although entry to this market is restricted since the government does not grant new smog check center licenses, there is a large number of centers among which a car owner can choose. This setup generates incentives among centers to compete for customers. However, centers cannot compete based on prices, since government sets the price. They can, however, compete on bribes; which would imply that the bribe should fall to the point where the market for bribes is cleared.

In the model proposed, I assume that the market of bribes is competitive and that cheating is limited by the number of donor cars available. The demand of cheating will be given by all those vehicle owners for which the expected utility of bribing is higher than the expected utility of not bribing. This demand can be derived from the previous model by adding up the probabilities of bribing across individuals. The supply in this market will be determined by the amount of donor cars available for cheating, which I will assume is proportional to the number of cars that do not bribe and pass the test.

I restrict the analysis of the market of bribes to vehicles that do not qualify for the exemption for the driving restriction program or non exemptible vehicles, *i.e.* model 1993 and older plus all transportation vehicles (see Table 1). This will allow me to focus on high-polluting vehicles. Cheating among exemptible vehicles may also be prevalent. However, cheating among them occurs because of policy incentives to meet a much stricter emission standard in order to avoid the driving restriction. Cheating is substantially more harmful among older vehicles, since their emissions could reach very high levels if unmonitored (Wenzel *et al.* 2004).

This said, vehicles that may qualify for the exemption play an important role in the market for bribes for non-exemptible cars. This is because both type of vehicles attend the same centers. More importantly, most non-cheating vehicles that have access to an exemption are excellent candidates for donor cars for non-exemptible vehicles, since their emissions are most likely lower than the limit faced by the later. Hence, when considering the supply of donor cars, I will include the proportion of vehicles with access to an exemption that appear to be non cheaters.

A model for the market of bribes will allow me to simulate changes in policies, such as tightening of the emission standards, and their consequences on the amount of bribing. The estimates from a cheating demand model can also be used to estimate the willingness to pay for car maintenance. This estimate can be helpful when evaluating the costs of programs that subsidize car maintenance or car scrappage. Finally, the structural model proposed suggests a cheap way to raise the price of the bribe and increase the incentives for car maintenance: separating clean and dirty cars across centers. The policy suggestion is based on exploiting the cheating's dependency on clean donor cars.

The remaining of the section will be structured as follows. First, I develop an estimable model for the demand of cheating that will deliver values for the structural parameters that govern demand and for the equilibrium bribe. Second, I propose a stalized model for the supply of cheating that will allow me to analyze changes in the equilibrium size of the bribe and the number of cheaters when simulating alternative policies.

7.1 A simple illustration of the market equilibrium

This subsection proposes a basic theoretical model that will simplify the understanding of the market equilibrium. The demand of cheating will be defined by all vehicle owners for whom bribing will result in higher utility than not bribing. Assume utility is linear in all costs. Then the utility of bribing will be defined by

$$u^B = -b - w - p,$$

where b is the bribe, w is the time cost from the trip to the center, and p is the government-fixed price of the certificate.

The utility from not bribing further depends on the probability of passing the test, z. In this simple version of the model, I will assume that vehicle owners have only one opportunity for retesting. The estimable version of the demand assumes that car owners can keep retesting indefinitely. For this model I will further assume that the cost of not getting the certificate is too high to risk failing on the retest. Therefore, everyone cheats on the retest. However, the government mandates that the second attempt is free. Hence, the first-time failing vehicles will only pay $w + b$ for cheating on the retest. The utility for not-bribing is then given by

$$u^A = -w - p + (1 - z)(-w - b),$$

where z is the predicted probability of passing the test, $w + p$ is the cost of the first test and $(1 - z)(w + b)$ is the expected cost of failing and retesting.

A car owner bribes if

$$u^B > u^A, \tag{7}$$

which in terms of costs implies

$$z < \frac{w}{w + b} = q.$$

q can be interpreted as the minimum probability for attempting to pass the test without bribing. The value of q is inversely correlated with the bribe. The demand for cheating will be defined by all vehicles for which $z < q$. For expositional purposes, assume the cumulative distribution of z is given by $F(z)$. Assume further that $F(z)$ is continuous and differentiable and that $F'(z) = f(z)$. The aggregate demand for cheating will then be given by

$$Q^D = N\left(F(q) + (1 - \overline{z}_A)(1 - F(q))\right), \tag{8}$$

where N is the total number of vehicles that do not have access to an exemption, and $\bar{z}_A$ is the mean probability of passing for non-cheaters:

$$\bar{z}_A = \frac{1}{1-F(q)} \int_q^1 z f(z) dz.$$

The demand for cheating given by equation (8) can be divided in two parts. The first term inside the parentheses is the proportion of vehicles that bribe right away. The second term in parentheses is the proportion of non-bribers that fail the test. Since only one retest is permitted, all first-attempt failing vehicles add up to the demand of cheating.

Notice that the demand is downward sloping since

$$\frac{\partial Q^d}{\partial b} = \frac{-w}{(w+b)^2} \cdot f(q) \cdot q < 0,$$

The supply of cheating will be defined by the number of times a technician will take a bribe in exchange for a "pass" certificate. For simplicity, I assume that the marginal cost of taking a bribe is zero for the technician and for the center. For this simple version of the model, assume that there is a single emission limit that all vehicles need to meet.

In order to provide the cheating service, the technician will require a non-bribing vehicle that has been shown to pass the test. The supply of cheating will then be restricted by the number of non-cheating vehicles that pass the emission test. More specifically, the model will assume that a center can only take λ bribers per donor car available. The number λ can be positive, but less than one if only some potential donor cars can be used for bribing, *e.g.* those for which the owner is easily convinced, or those who happen to arrive at the same time than a cheating car. The number λ can be greater than one if a single donor car can be used more than once for cheating. I assume all non cheating vehicles that have access to the exemption are good donor cars for non-exemptible vehicles. Assume this number is equal to E. Under this assumption, the aggregate supply is given by:

$$Q^s = \lambda \left(N \bar{z}_A (1 - F(q)) + E \right)$$

The supply is concave and may be downward sloping for some range of q, since

$$\frac{\partial \bar{z}_A}{\partial b} = \frac{-w}{(w+b)^2} \cdot \frac{f(q) \cdot (\bar{z}_A - q)}{1 - F(q)} < 0$$

$$\frac{\partial (1 - F(q))}{\partial b} = \frac{w}{(w+b)^2} \cdot f(q) > 0.$$

The equilibrium bribe will be at the point where aggregate supply equals aggregate demand. In order to illustrate the equilibrium, assume probabilities of passing have a uniform distribution: $F(z) = z$. Then the equilibrium amount of cheating is given by:

$$Q^* = \frac{N}{2} + \left(\frac{\lambda}{1+\lambda}(1+2\eta) - \frac{1}{1+\lambda}\right)\frac{N}{2},$$

where $E = \eta N$. The equilibrium given by the uniform distribution assumption offers some simple intuition. First take the case where there are no exemptible vehicles ($\eta = 0$) and each donor car can produce a single cheater ($\lambda = 1$). In this basic case, $Q^* = \frac{1}{2}N$. Since the mean probability of passing is $\frac{1}{2}$, half of the car fleet will pass and provide clean emissions for the other half. Notice that $\lambda > 1$ results increases the number of cheaters above 50 percent. Also, $\eta > 0$ will lead to a higher number of cheaters for any given λ.

Figure 1 illustrates the market equilibrium assuming $F(z)$ is uniform. For this figure, w is assumed to be 30 pesos ($3), $\lambda = 1.25$, $E = 0$ and $N = 100$. The market equilibrium is located where the supply and demand curves intersect. Increases in λ and η will shift the supply curve up, leading to a lower bribe and a higher number of cheaters.

The framework illustrated in Figure 2 will be helpful for analyzing changes in policy. For instance, one can simulate an exogenous increase in the bribe as a consequence of increased center monitoring. I can also simulate exogenous increases in the distribution of the probability of passing the test that may come from subsidized maintenance or car renewal. Finally, I can simulate a decrease in supply that may be reached by separating vehicles with a high probability passing from vehicles with a low probability of passing.

The following subsection computes the demand and supply curves of bribing from the structural model estimated in Section 6. Compared to the simple model presented in this subsection, the demand curve from structural estimation relaxes some important assumptions in order to better resemble the actual regulatory framework. For instance, it assumes that vehicles can retest indefinitely and that they may postpone the check if they want to. Hence, the always-bribe assumption for the retest is also relaxed.

7.2 Demand and supply for cheating from model estimation

As mentioned at the beginning of this section demand for cheating can be derived from the structural model once the key parameters of the model are estimated. Given estimates for w, δ and τ, I can approximate the steady state demand for cheating in a given smog check cycle as

$$
\begin{aligned}
\hat{Q}^d(\hat{b}) = \; & \textstyle\sum_{i=1}^{N} \Pr(S_i^{od} = B) + \Pr(S_i^{od} = A) \cdot (1 - z_i) \cdot \Pr(S_i^{ev} = B) + \\
& \textstyle\sum_{i=1}^{N} \Pr(S_i^{od} = X) \cdot \big(\Pr(S_i^{od} = B) + \Pr(S_i^{od} = A) \cdot (1 - z_i) \cdot \Pr(S_i^{ev} = B)\big) + \\
& \textstyle\sum_{i=1}^{N} \Pr(\text{H7}) \cdot \big(\Pr(S_i^{ev} = B) + \Pr(S_i^{ev} = A) \cdot (1 - z_i) \cdot \Pr(S_i^{od} = B)\big) \,,
\end{aligned} \tag{9}
$$

where the first line in equation (9) corresponds to all vehicles that bribe in the first and second visits; the second line corresponds to all vehicles that have postponed the smog check in the previous period and have decided to cheat in either the first or the second visit during the current period; and the third line corresponds to all vehicles that failed the second attempt at the test and decided to bribe in either their third or their fourth visit. The demand curve approximation excludes those vehicles that decide to bribe after the fourth visit or those vehicles that decide to postpone the check and then bribe after the second visit. However, the number of tests in these categories amount to less than 1 percent of the total. $\hat{Q}^d$ can be simulated for various values for b, in order to obtain a demand curve.

The supply will be composed of all non-exemptible vehicles that pass the test and all non-cheating exemptible vehicles. The estimates suggest that at least 20.5 percent of all exemptible vehicles are cheaters[28]. On the other hand, vehicles that may access exemption make 61 percent of the total carfleet in 2003. Since 79.5 percent of them are non-cheaters, the potential supply of donors from vehicles that may qualify for the exemption is 1.03 times as many as non-exemptible vehicles. Hence, we can express the supply of donor cars from exemptible vehicles as $1.03N$, where N is the number of non-exemptible vehicles.

The total supply is then given by the following expression:

$$
\begin{aligned}
\hat{Q}^s(\hat{b}) = \; & \lambda \left(1.03 \cdot N + \textstyle\sum_{i=1}^{N} \Pr(S_i^{od} = A) \cdot z_i + \Pr(S_i^{od} = A) \cdot (1 - z_i) \cdot \Pr(S_i^{ev} = A) \cdot z_i\right) + \\
& \lambda \left(\textstyle\sum_{i=1}^{N} \Pr(S_i^{od} = X) \cdot \big(\Pr(S_i^{od} = A) \cdot z_i + \Pr(S_i^{od} = A) \cdot (1 - z_i) \cdot \Pr(S_i^{ev} = A) \cdot z_i\big)\right) + \\
& \lambda \left(\textstyle\sum_{i=1}^{N} \Pr(\text{H7}) \cdot \big(\Pr(S_i^{ev} = A) \cdot z_i + \Pr(S_i^{ev} = A) \cdot (1 - z_i) \cdot \Pr(S_i^{od} = A) \cdot z_i\big)\right) ,
\end{aligned} \tag{10}
$$

where the parameter λ represents the productivity of potential donor cars or the number of cheaters that a potential donor car produces. The first line in equation (10) can be divided in two components. The first component, $\lambda(1.03 \cdot N)$, is the supply of cheating that originates from exemptible vehicles. The second component represents all vehicles that passed the test without bribing in the first and second visits. The second line in equation (10) originates from vehicles that postponed the test and then passed it within the first two visits, while the third line originates from vehicles

[28] The lower bound for the prevalence rate of corruption among exemptible vehicles is 20.5 percent from the IV estimation on oxygen (not reported in tables).

that pass the test in fourth or fifth visits.[29]

Since in equilibrium $Q^s(b) = Q^d(b)$, an estimate for λ can be calculated from equating the estimates for the demand and supply evaluated at $\hat{b}$. When simulating the supply for different values of b, I will assume λ remains constant and equal to the estimated value. I will also assume that the market for cheating of exemptible cars is unaffected by changes in the market for cheating of non-exemptible vehicles. This is a reasonable assumption to make since donor cars for exemptible vehicles need to meet stricter emission standards.

8 Structural model estimation and simulations

8.1 Parameter estimates and predicted probability of bribing

The decision model proposed in Section 6 is estimated using a five percent random sample of all non-exemptible vehicles.[30] The estimated parameters are given in Table 7. Time cost, w, is modeled as a linear function of the log-price of the car: $w = w_{intercept} + w_{slope} \cdot ln(price)$. I included the standard errors from the numerical Hessian of the likelihood function. However, these are incorrect since they do not account for the fact that z_i is itself an estimate. Bootstrapped errors will be included in a future version of the paper.

The left hand side of Panel A presents the parameter estimates as estimated by MLE. Notice that time cost varies positively with the price of the car and has some variation across periods. The mean time cost for each period varies little: between 30 and 35 pesos (0.3 US$ and 3.5 US$ respectively) as shown in the right hand side of Panel A.

The equilibrium bribe is 97.11 pesos (9.7 US$), which is a plausible value according to newspaper articles on the matter (50 to 230 pesos). The estimate for the discount factor is very low, 0.37, if we consider that the fine is usually paid a few months after the deadline expires. Such a high discount rate might signal present bias or a generalized procrastination problem. And, indeed, it is surprising to see that many people wait to the last days of the month to smog-check their cars even when the lines are longer during these days. However, the interpretation of δ as a behavioral parameter is imprecise, since it is assumed to be constant across periods. Another possibility is that individuals think they can find a way to avoid paying the fine in the future. In reality this is not easy, since they have to show a proof of payment in order to get their vehicle smog-checked; and further avoiding the smog-check may result in multiple traffic tickets.

[29] Vehicles that pass the test in a fourth or fifth attempt might have a relatively low probability of passing, which makes them poor candidates for donor cars. However, excluding these vehicles from the supply approximation in equation (10) changes the total amount of cheating available by less than three percent.

[30] The size of the sample was chosen for computational reasons, since each iteration of the maximizing algorithm includes an internal numerical search for the stationary solution to the expected value of future nodes. However, the estimates are robust to different random samples and different sizes of the sample (results not shown).

Finally, the estimate for the standard deviation of the random shock, 105.71, is even higher than the bribe. This may be a result of the simplicity of the model, since it does not account for much of the heterogeneity across individuals. It could also reflect an overly optimistic option value for the future choices. To see this, notice that the higher the variance of the error term in (5) and (6), the higher the probability that a large and positive shock will appear in any of the future choices. The large variance of the random shock decreases the expected cost of future nodes, since one can always choose the option with the highest payoff. Estimates are robust to changes in the specification.[31]

Panel B of Table 8 shows the fit of the model as well as the predicted probabilities for each history from H1 to H7. The fit of the model is best for the H3,H4 combined histories and for H4. Despite the low discount rate, the model fails to predict the number of postponed tests by 3.5 percent. The total cheating prevalence rate estimates from this method can be calculated as the sum of predicted histories H2 and H5. Notice that this estimate is 27 percent, which is consistent with the lower bound estimate from the regerssion methodology: 20 percent.

8.2 Reducing donor car availability

Figure 4 presents a simulation for the market of cheating. Estimates for the current demand and supply functions are given by the solid lines. The parameter λ was calibrated by equating $\hat{Q}^d$ to $\hat{Q}^s$ (evaluated at the equilibrium bribe, 97.11), and equals 0.16. This value implies that about 1.6 out of ten potential donor cars produce a cheater. The simulations of demand and supply are helpful for evaluating alternative policies. For example, reducing the availability of donor cars.

Assuming that this parameter is constant, I can predict what would happen to the equilibrium bribe and the demand for cheating if I separated vehicles that can obtain an exemption and vehicles that cannot across centers. This would lead to approximately 68 percent fewer donor cars available for cheating.[32] In Figure 4 the corresponding simulated supply corresponds to the dotted line. Notice that the reduction in supply of donor cars would lead to an increase in the equilibrium bribe of 75 percent and an fall in the number of cheaters by more than 50 percent. The increase in the equilibrium bribe is particularly important since it will positively affect the incentives for car maintenance.

[31] An alternative specification allows the time cost to vary by smog-check period. Every vehicle has a two-month period in each year half period to be taken to the smog-check. Period 1 corresponds to January-February and July-August, period 2 corresponds to February-March and August-September, and so on. Periods are assigned randomly across vehicles. However, some months of the year may have more costly time costs than others. Table A4 shows the results for a specification where the time cost is allowed to vary by period. The average time cost in this alternative specification is of 43 pesos and the bribe is 107.31 pesos, still within the expected range of values.

[32] The supply of donor cars from exemptible vehicles is estimated to be $1.03N$, where N is the number of non-exemptible vehicles. Figure 4 shows that this contraction of the supply amounts to approximately 68 percent less donor vehicles.

So far, the analysis of smog check efficiency proposed in this paper does not address the trade off between car maintenance and bribing since I do not observe any decisions on the matter. However, by estimating the model for the demand of cheating, I can simulate the willingness to pay for car maintenance as the difference between the expected cost of the smog check for a car with a high probability of passing and the expected cost of the smog check for a vehicle with a low probability of passing. Figure 5 shows how the expected cost of smog checking a vehicle varies with the probability of passing according to the model estimates. The solid line represents the expected cost of a smog check according to the estimated parameters. Notice that the expected cost of a smog check lies in the 130-160 Mexican pesos interval. Given that the cost of the first test is 175 Mexican pesos, the range of values for the expected cost of the test seem too low. The reason behind such a low expected cost is the high variance of the random shocks. Since random shocks in the utility function can be either positive or negative, a high variance of the shocks will increase the option value at each node and hence will reduce the expected cost of the overall smog check.

The dotted line shows what would happen to the expected cost of a smog checks if the equilibrium bribe increased by 75 percent, which corresponds to the restriction in the supply of donor cars. Notice that simulated expected cost is three times more sensitive to changes in the probability of passing with the higher bribe. This effect has important implications for the incentives to pay for car maintenance. According to the model assumptions, the resulting increase in the bribe from restricting donor car supply doubles the willingness to pay for lower emissions.

However, car maintenance costs are usually between 500 and 2000 pesos. Even with a bribe that was 75 percent higher, the willingness to pay for infreasing the probability of passing from 20 to 80 percent would be 60 pesos. A bribe of 160 pesos (new equilibrium in Figure 4) would still be much more attractive than paying for car maintenance.

8.3 Stricter emission standards

An alternative policy strategy could be the tightening of emission requirements for vehicles with no access to the exemption. For instance, we could change the emission limits to match the emission requirements for the exemption. [33]In contrast with the previous policy proposed, which affects only the supply side of the market, this policy would shift both demand and supply. A more strict emission standard will increase the demand for bribes as well as reducing the supply of donor cars. Hence, the effect of this policy is unpredictable without a model for the demand and supply elasticities. The structural framework allows me to simulate such a policy change by changing the

[33]The change would only affect non public transportation vehicles, since taxis and other public transportation vehicles of all model-years are already subject to the exemption standard, which is similar to the EPA Tier 1 standard.

probability of passing the test to the one associated with the tighter emission standards. Figure 6 shows the simulated demand and supply for bribing under the stricter emission standard (black lines); and compares it with the current setup (gray lines). As expected, the tighter standard increases the demand for bribes. However, less vehicles meet the standard, which shifts the supply of donor cars inwards. Since the supply of donor cars is very inelastic, the resulting equilibrium has a higher bribe and lower cheating prevalence rate than the current regulation. As a result, the bribe becomes twice as expensive and the number of cheaters falls by 20 percent.

The simulation depicted in Figure 6 keeps the assumption that all vehicles that can apply to the exemption are good donor cars for the non-exemptible car fleet. This assumption becomes less realistic when the emission standards are equal across the two groups. Relaxing this assumption would result in a more pronounced reduction in corruption and larger increase in the bribe, since the supply of bribes would shift further inwards.

As mentioned above, the reduction in corruption in the new equilibrium is partially due to the low elasticity of the supply of bribes. However, it is important to recognize that small magnitude of the elasticity is mostly due to the assumption that the proportion of suitable donor cars that are actually used for cheating, λ, is constant (the calibrated value is 0.16). An increase in this factor would result in a more elastic supply. This in turn would yield a smaller reduction in the corruption prevalence and a smaller increase in the bribe as a result of tightening the emission standards.

8.4 Incidence considerations

It is important to point out that the two policies proposed above would result in even higher burdens for individuals with low-cost, high emitting vehicles. Corruption ameliorates the regressiveness of environmental regulations by reducing time and maintenance costs for high-emitting vehicle owners. The government could counteract this regressiveness by subsidizing maintenance and car renewal for owners of high polluting vehicles. Moreover, the structural model suggests that investments in car maintenance are unlikely since the wilingness to pay for increasing the probability of passing lies below the minimum repair costs. In this context, a policy that subsidizes car maintenance would be progressive and more effective in emission reduction than the actual policies.

9 Conclusions and future work

This paper provides evidence of widespread cheating in the form of donor cars throughout Mexico City smog check centers. The methodology used to detect cheating relies on the identification of unusual serial correlation patterns that emerge as a consequence of measuring emissions repeatedly for low-emission vehicles or donor car. Results from a two different versions of the test for cheating, a permutations test and a regression model, suggest that at least 86 percent of centers are engaged in corruption. In addition, estimates from the OLS regression model can be interpreted as estimates of cheating prevalence. The most reliable estimates suggest that cheating is present in at least 20 percent of all tests.

The reduced form section of the paper also provides estimates for the amount of emissions that are underrepresented by smog check center data. Prevalence rate estimations can be obtained for each smog check center. By identifying centers that are less subject to this type of cheating, I predict true emissions for the remaining centers. The prediction is based on a set of observed car characteristics like age of the car, brand, size, service, mileage and time of the day when emissions were measured. Predicted emissions are most likely an underestimate of true predictrions since selection across centers may occur. Centers with little or no cheating available may attract cleaner vehicles. Nevertheless, predicted emissions are useful as a lower bound for emission missmeasurement. The results suggest that hydrocarbons are highly underrepresented by smog check center readings: mean predicted emissions for this gas are 17 percent above the observed level. Nitrogen oxides and carbon monoxide are underrepresented by about 9 percent.

The results from reduced form estimation are revealing and important to better understand the effectiveness of emission regulations. However, they offer little in terms of policy alternatives that can ameliorate this type of cheating and improve the effectiveness of emission monitoring. The efforts to eradicate fraud of smog check centers on behalf of the government have been focused in other forms of cheating such as car tampering and the black market of certificates. Donor type of cheating is harder to detect and prevent. The paper offers a simple model for the market of bribes that emphasizes the relationship between clean and dirty cars and how it affects the equilibrium quantity of fraud and the equilibrium bribe. Under competition between centers, cheating is constrained by availability of clean cars among the incoming customers. Restricting the supply of clean vehicles in some centers could result in a reduced amount of bribing and an increase equilibrium price for the bribe. This, in turn, would raise the incentives for car maintenance and reduce emissions. The supply restriction could be achieved by separating vehilcles that qualify for a driving restriction exemption from those that do not qualify for the exemption across centers.

Upon estimating the demand for cheating, I simulate the change in policy described above and predict the reduction in the total amount of cheating and the equilibrium bribe. The results suggest that separating exemptible and non-exemptible vehicles across centers would double the price of the bribe and will lead to a fall in corruption of about 50 percent. The structural model can also be used to look at the implication of the increased bribing cost on the willingness to pay for car maintenance. The results suggest that, when the supply of donor cars is restricted by this policy, vehicles are willing to spend about twice as much to lower vehicle emissions compared to the current status.

The structural model for the demand for cheating also helps simulate the effect of tightening emission standards on the amount of corruption. When the emission standards are tightened to an EPA Tier 1-like level, the demand for cheating increases. However, the availability of donor cars contracts more than enough to compensate for the demand increase, resulting in a net decrease in corruption and a net increase in the equilibrium bribe.

In future revisions of this paper the identification of the structural model can be strenghtened by the inclusion of additional car owner's characteristics that determine time costs. Since the data includes zip codes for the car owners, census block level demographics can be matched to the smog check center data. The cost of bribing can also be better identified by taking into account distance from the owner's house to the center. Since some centers are more prone to cheating than others, the vehicle owner may be forced to travel long distances to be able to bribe. Long traveling distances may play an important role in the cost benefit analysis of the individual when evaluating the cheating option.

References

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

[2] Kenneth Chay and Michael Greenstone, 2005, "Does Air Quality Matter? Evidence from the Housing Market", Journal of Political Economy, 2005, vol. 113, no. 2

[3] , G. S. and T. Feyzioglu, 1997a, "Rationing Can Backfire: The 'Day without a Car' in Mexico City," World Bank Economic Review, 11(3): 383-408.

[4] Eskeland, G. S. and T. Feyzioglu, 1997b, "Is Demand for Polluting Goods Manageable? An Econometric Study of Car Ownership and Use in Mexico," Journal of Development Economics, 53(2): 423-445.

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

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

[7] Good, Phillip (2000), Permutation Tests: A Practical Guide to Resampling Methods fro Testing Hypothesis, Springer Series in Statistics.

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

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

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

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

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

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

[14] 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.

[15] Shifter 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.

[16] Wenzel, T., Singer, B.C., Slott, R.S., 2000a. "Some issues in the statistical analysis of vehicle emissions." Journal of Transportation and Statistics 3, 1–14.

[17] Wenzel, T., Singer, B.C., Gumerman, E.Z., Slott, R.S., Sawyer, R.F., 2000b. 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.

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

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

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

Table 1: Emission restrictions in 2003 -2004

	Certificate= exemption	Certificate = non-exemption				
	[1]	[2]	[3]	[4]	[5]	[6]
	Only private and corporate vehicles, model >1992	Private vehicles, model>1990	Private vehicles, model<1991	Corporate vehicles, model<1994	Corporate vehicles, model>1993	Public transportation vehicles
HC	100	200	300	200	350	100
NO	1200	2500	2500	2500	2500	1200
CO	1	2	3	2	3	1
O2	15	15	15	15	15	15

Notes:
1. Source: Secretearía de Medio Ambiente, Distrito Federal
2. Limits are given in parts per million for HC and NO and in percent of volume for CO and O2.

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

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

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

Table 3: Tests for cheating from OLS

	[1]	[2]	[3]	[4]
HC24	0.039	0.024	0.019	0.018
	[0.002]**	[0.002]**	[0.002]**	[0.002]**
CO24	0.078	0.042	0.029	0.027
	[0.004]**	[0.003]**	[0.003]**	[0.003]**
NO24	0.153	0.145	0.086	0.072
	[0.007]**	[0.008]**	[0.005]**	[0.004]**
O2_24	0.18	0.14	0.12	0.11
	[0.008]**	[0.008]**	[0.007]**	[0.006]**
Observations	1754206	1754206	1754206	1754206
Car characteristics		X	X	X
Center FE			X	X
Testing equipment FE				X

Notes:

1. Standard errors in brackets. * Significant at 5%; **Significant at 1%
2. Source: Smog Check Center data for 2003 (2), Distrito Federal
3. Sample for estimation corresponds to a 30 percent random sample by center.
4. Car characteristics include dummies for brand, size and service, interactions of size and service dummies with model-year, and polynomials on model-year, odometer and time of the day.

Table 4: Selecting sample with no evidence of cheating

Center	HC	CO	NO	O2	Number of tests
Panel A: Centers for which test fails to reject no-cheating by all four specifications (p-values in brackets)					
1069	-0.063	-0.028	-0.058	0.060	263
	[0.297]	[0.355]	[0.139]	[0.327]	
1071	-0.016	0.050	0.034	0.039	364
	[0.461]	[0.202]	[0.352]	[0.278]	
9016	0.010	0.089	0.040	0.104	11013
	[0.317]	[0.051]	[0.217]	[0.095]	
9031	0.002	0.070	-0.004	0.120	14130
	[0.600]	[0.093]	[0.132]	[0.120]	
9034	0.036	0.015	0.008	0.043	13947
	[0.138]	[0.316]	[0.449]	[0.134]	
9038	0.026	0.015	0.003	0.403	17583
	[0.213]	[0.132]	[0.706]	[0.054]	
9055	0.026	0.036	0.013	0.180	15166
	[0.126]	[0.156]	[0.522]	[0.121]	
9057	0.007	0.061	0.004	0.066	14493
	[0.225]	[0.051]	[0.619]	[0.077]	
9059	-0.001	0.020	-0.001	0.077	10397
	[0.785]	[0.054]	[0.890]	[0.127]	
9064	0.002	0.044	0.005	0.057	4574
	[0.306]	[0.106]	[0.293]	[0.091]	
9074	0.003	0.037	-0.007	0.091	15145
	[0.596]	[0.092]	[0.238]	[0.104]	
Panel B: Centers for which test rejects no-cheating in all four specifications (p-values in brackets)					
9009	0.007	0.100	0.013	0.150	25399
	[0.020]	[0.006]	[0.004]	[0.028]	
9013	0.047	0.105	0.087	0.074	21269
	[0.020]	[0.002]	[0.005]	[0.020]	
9046	0.057	0.159	0.086	0.121	17560
	[0.014]	[0.004]	[0.005]	[0.027]	
9047	0.019	0.112	0.076	0.138	24233
	[0.011]	[0.000]	[0.000]	[0.004]	
9054	0.023	0.129	0.105	0.292	13731
	[0.030]	[0.011]	[0.025]	[0.005]	
9073	0.033	0.053	0.089	0.107	28096
	[0.017]	[0.016]	[0.003]	[0.012]	

Notes:

1. Source: OLS estimations of serial correlation by center (2003 (2)). Numbers in main rows are coefficients on lagged emissions.

2. Controls include brand categories, 8 size categories, 3 service categories, size and service interactions with model-year, model-year polynomial, mileage polynomial and time of the day polynomial.

3. Numbers in brackets are p-values.

4. All standard errors are clustered at the lane (or testing machine) level. The number of lanes per center varies between 1 and 7, with a mean of 4.

5. The remaining centers (omitted in this table) are those for which no cheating is rejected for 1, 2 or 3 specifications are omitted in this table. These groups have 9, 20 and 34 centers, respectively.

Table 5: Test outcomes by evidence of cheating

Panel A: Vehicle characteristics	Some evidence	No evidence	Difference	Panel B: Test outcomes	Some evidence	No evidence	Difference
Model-year	1986.8	1986.8	-0.018	HC40	96.0	109.6	11.704
	[7.591]	[7.465]	[0.030]		[226.7]	[294.6]	[0.922]**
Size (0-8)	1.5983	1.5319	-0.066	HC24	107.1	128.1	19.067
	[2.258]	[2.148]	[0.011]**		[262.2]	[340.6]	[1.066]**
VW Sedan and Chevy	0.3884	0.3718	-0.017	CO40	0.842	1.046	0.202
	[0.487]	[0.483]	[0.007]		[1.223]	[1.427]	[0.005]**
Mini-Compact	0.3367	0.3626	0.027	CO24	0.860	1.132	0.269
	[0.472]	[0.481]	[0.002]**		[1.239]	[1.503]	[0.005]**
Compart-Medium	0.0876	0.0909	0.004	NO40	453.9	677.5	225.543
	[0.283]	[0.287]	[0.001]**		[539.5]	[602.1]	[2.162]**
Medium-Large	0.0471	0.0500	0.003	NO24	620.7	887.2	270.79
	[0.212]	[0.218]	[0.001]**		[724.5]	[760.3]	[2.890]**
Sport	0.0140	0.0139	0.0001	O2_24	0.862	0.986	0.123
	[0.118]	[0.117]	[0.001]		[1.070]	[1.118]	[0.004]**
Minivan	0.0056	0.0045	-0.001	O2_40	1.053	1.186	0.133
	[0.075]	[0.067]	[0.000]**		[1.216]	[1.252]	[0.005]**
Van	0.0085	0.0053	-0.003	Pr(Pass)	0.7366	0.7059	-0.029
	[0.092]	[0.073]	[0.000]**	(includes retests)	[0.441]	[0.456]	[0.002]**
Pick Up	0.0870	0.0805	-0.01				
	[0.282]	[0.272]	[0.001]**				
Standard Pick Up	0.0251	0.0205	-0.004				
	[0.156]	[0.142]	[0.001]**				
Displacement	2530	2505	-26.866				
	[1450.3]	[1374.6]	[5.746]**				
Kilometers	184932	192424	7,492.36				
	[211582.8]	[230117.6]	[846.191]**				

Notes:

1. The first two columns of each panel report the mean value for the sample and the corresponding standard deviation in parenthesis. The third column reports the difference in means between the non-cheating centers and the rest of the sample. Standard errors for the difference are reported in brackets, and ** denotes the difference is statistically significant at 1 percent level.

2. The number of tests in centers with no evidence of corruption is 117,075. The remaining tests, 1,357,898 belong to centers with some evidence of corruption.

3. The probability of passing the test includes all tests and retests. This probability is lower than the probability of passing a single test, since retesters have a lower probability of passing.

4. Last eight size categories correspond to EPA categories according to gasoline mileage and emissions. These categories were matched to the data using the information on brand and model.

5. I assigned a separate size category for "VW Sedan and Chevy" since these are popular economic vehicles that constitute over 30 percent of the carfleet in Mexico City. These vehicles have lower maintenance costs and higher gasoline mileage than other vehicles.

6. Displacement is a measure of the volume of the engine.

Table 6: Using low-cheating centers to predict true mean emissions and probability of passing the test

		HC24	*HC40*	*CO24*	*CO40*	*NO24*	*NO40*	*O224*	*O240*	*Pr(Pass)*
		(ppm)	(ppm)	(% of vol.)	(% of vol.)	(ppm)	(ppm)	(% of vol.)	(% of vol.)	Emission threshold for non-exemptable vehicles
Panel A: Observed Mean Emissions										
Exemptible										
No	Mean	109.81	97.56	0.891	0.872	693.61	517.24	1.072	0.869	0.762
	SD	273.24	231.67	1.228	1.209	726.03	560.50	1.202	1.014	0.426
Yes	Mean	27.95	23.92	0.197	0.187	325.76	231.47	0.373	0.333	–
	SD	73.25	64.68	0.528	0.507	474.99	357.15	0.739	0.658	
Total	Mean	57.49	50.49	0.448	0.434	458.49	334.58	0.625	0.526	–
	SD	178.65	152.61	0.913	0.894	604.66	462.30	0.992	0.845	
Panel B: Predicted Mean Emissions. OLS										
Exemptible										
No	Mean	124.90	105.03	0.964	0.925	873.94	647.69	1.147	0.912	0.727
	SE	15.22	14.19	0.217	0.328	168.66	61.72	0.242	0.376	0.112
Yes	Mean	34.08	26.52	0.222	0.199	345.28	241.78	0.258	0.212	–
	SE	6.25	17.28	0.178	0.274	152.86	63.61	0.240	0.324	
Total	Mean	66.85	54.85	0.489	0.461	536.03	388.25	0.579	0.465	–
	SE	8.83	16.49	0.192	0.294	158.58	63.15	0.240	0.343	

Notes:

1.Panel A reports the mean and standard deviation of pollutants for each of the two speeds at which they are measured, 24kmh and 40kmh, from smog check center data of 2003, 1st half (D.F.). The last column reports the average probability of passing the emission threshold for non-exemptable cars on the first test. Notice that this probability is larger than the one reported in Panel A of Table 5 since retests are not included in the average.

2.Panel B reports the predicted emissions from an OLS model of pollutants estimated on the non-cheating sample of centers. The car characteristics included in the prediction model are brand, displacement and service categories, the interaction of the last two with model-year, model year polynomial, time of day and mileage polynomials.

3. The numbers in brackets in Panel B are average prediction errors.

4. The last column of Panel B shows the predicted probability of passing the emission threshold for non-exemptible vehicles. The predicted value is computed using a probit model estimation on non-cheating centers. The probit model contains similar covariates as the OLS model for pollutants (see note 2 above).

Table 7: Model parameters and demand for bribes

Panel A: Parameter Estimates

Model Parameters	Estimate	SE (see note 3)	Structural Parameters	
w-intercept	43.6	59.1292	Mean time cost	33.69
w-slope	-0.99	5.9355	Minimum time cost	30.05
b	97.11	8.8538	Maximum time cost	35.61
δ	0.37	0.0193	Bribe	97.11
τ	82.42	5.5076	Discount factor, δ	0.37
			SE of random shock (σ)	105.71

Panel B: Fitted and predicted probabilities for each history

Observed History	Actual	Fitted
H1: Postpone	0.0956	0.0594
H3,H4: Bribe/No bribe - Pass	0.7284	0.7456
H4: No bribe-Fail-Postpone	0.0051	0.0014
H5, H6: No bribe-Fail-Bribe/No bribe-Fail-No bribe-Pass	0.1266	0.1545
H7: No bribe-Fail-No bribe-Fail	0.0444	0.0391

History	Predicted
H1: Postpone	0.0594
H2: Bribe	0.2044
H3: No bribe-Pass	0.5412
H4: No bribe-Fail-Pospone	0.0014
H5: No bribe-Fail-Bribe	0.0648
H6 No bribe-Fail-No bribe-Pass	0.0897
H7: No bribe-Fail-No bribe-Fail	0.0391

Notes:

1. Table 7 shows the results of the maximum likelihood estimation of the bribing behavior model developed in Section 6. 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 10.15. 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 that is related to parameter τ by the formula given on footnote 25.

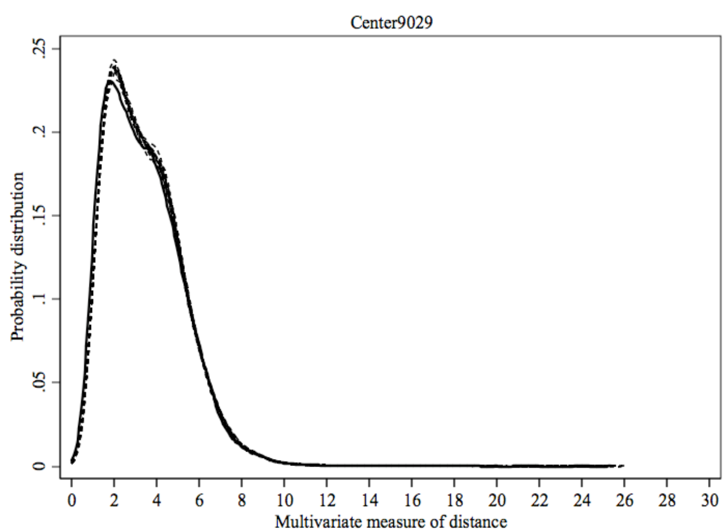
3. The standard errors in the second column of Panel A correspond to the squared root of the inverse numerical Hessian. These standard errors are inconsistent since they do not account for the fact that the probability of passing the test is a predicted value.

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

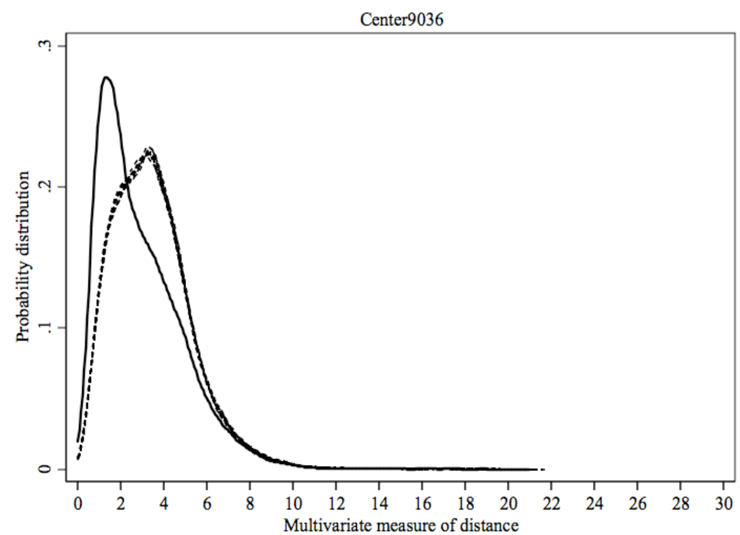
5. The bottom of Panel B shows mean Predicted probabilities for all histories up to H7. The percentage of tests with cheating can be cumputed by adding up the probability of histories H2 and H5: 27 percent. Further bribing could occur within history H7, however, this is likeley very small (less than 3.9 percent).

Figure 1: Nonparametric test for cheating

Panel A: Center with low cheating prevalence

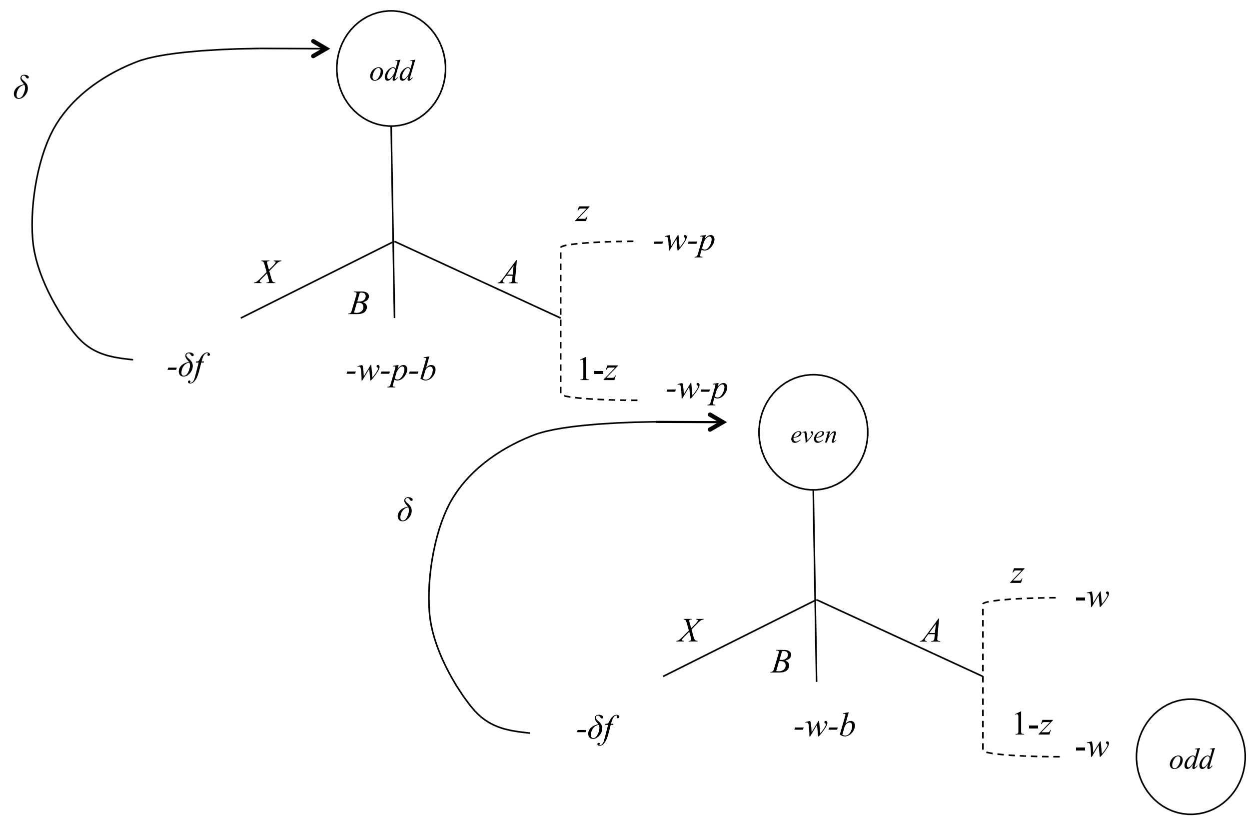


Panel B: Center with high cheating prevalence



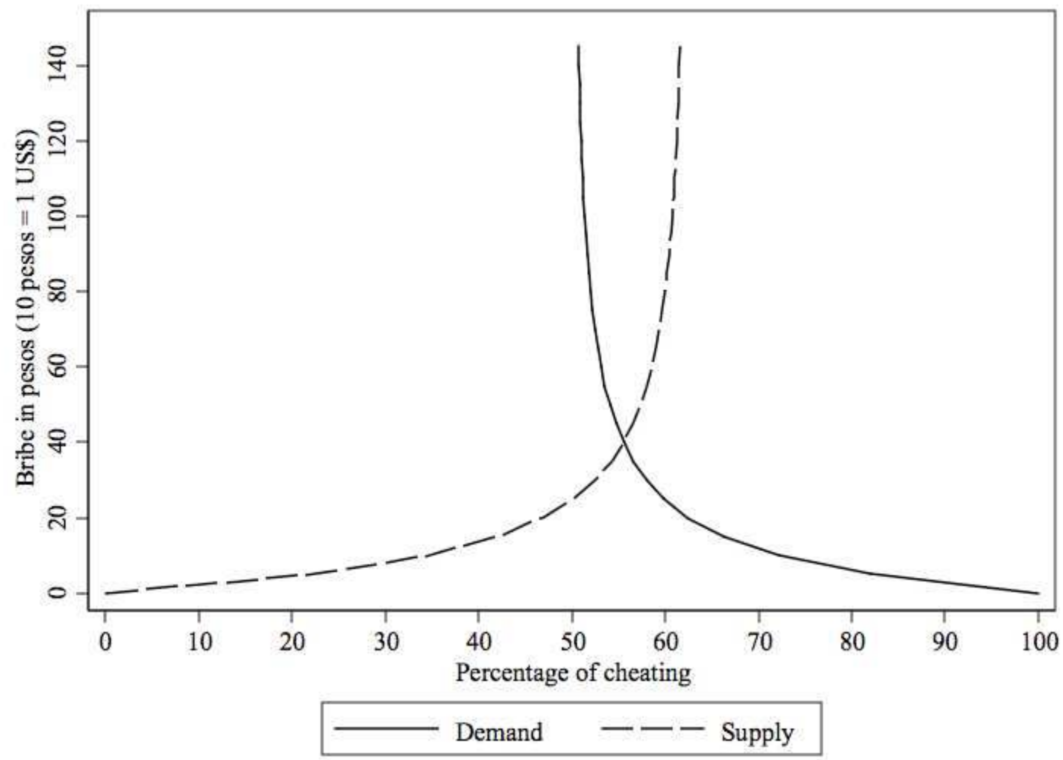
Notes:
1. Smog check center data from 2003 (2nd half), D.F.
2. The measure of distance corresponds to the square root of the sum of square differences between standardized pollutants.
3. Dotted lines correspond to each of 10 different random permutations of the original data. The solid line correspond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

Figure 2: Decision tree



Notes:
Each circle represents the start of a new attempt round. Since the choice sequence has an infinite horizon, all *odd* rounds are identical and all *even* rounds are identical. Every cycle starts with an *odd* round. The choices are specified along the tree branches. Dotted branches denote uncertain outcomes and the probability associated with each uncertain outcome is written alongside each dotted branch. The payoffs for each choice are specified at the end of each branch. In the case of the choice *X*, the payoff $-\delta f$ adds up to the discounted payoff from restarting the game at the end of the arrow.

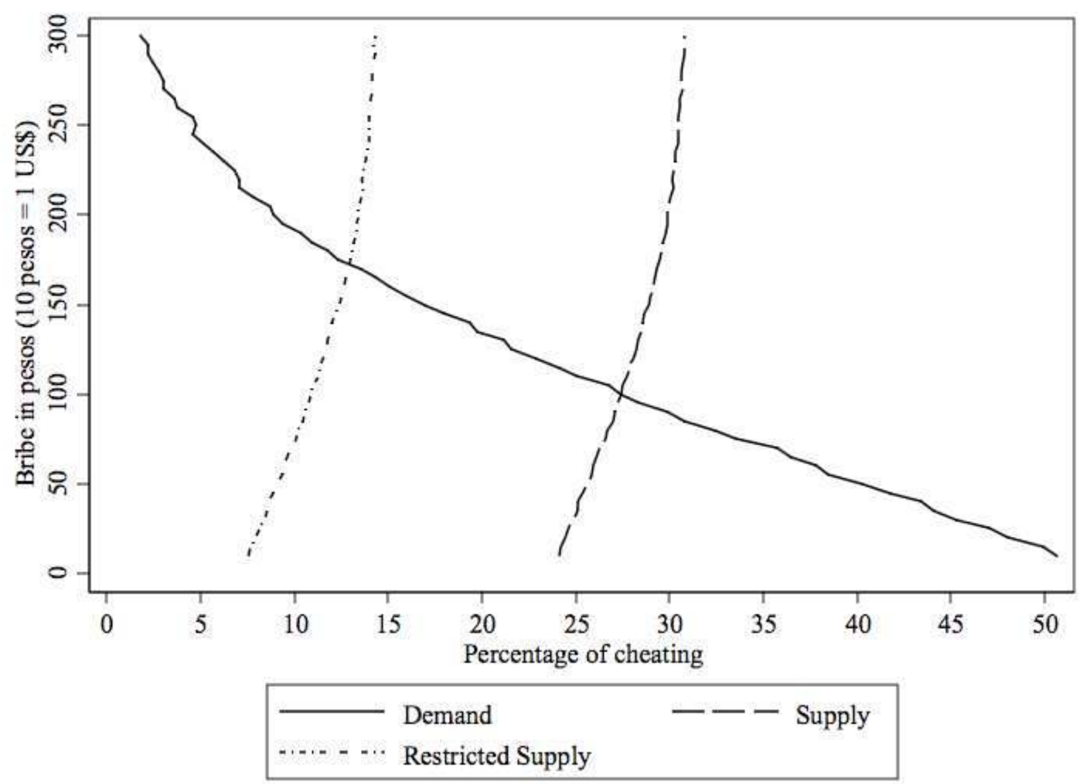
Figure 3: Example of an equilibrium in the market for bribes



Notes:

1. The figure assumes that the probability of passing is distributed uniformly between 0 and 1; that the constant productivity of each godmother car, λ, equals 1.25; and that the number of *godmother* cars coming for the exemptable car fleet, E , equals 0.

Figure 4: Simulated equilibrium. Reducing the supply of donor cars.

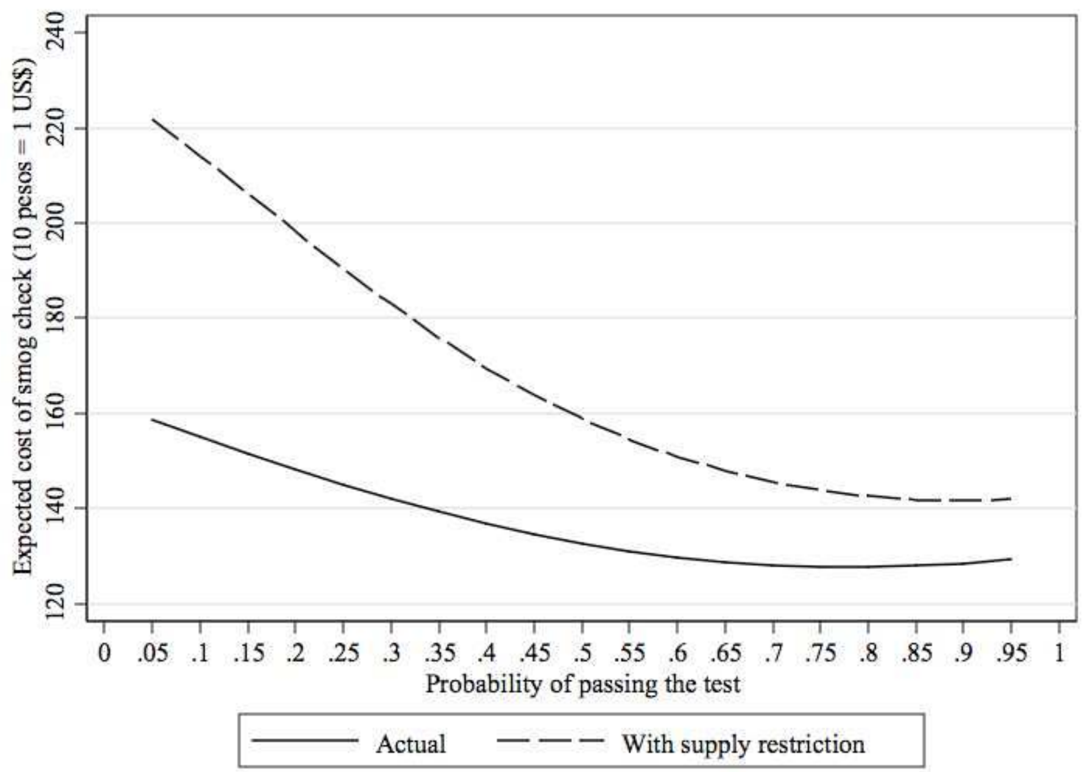


Notes:

1. The figure uses the imputed probability of passing and the estimated parameters, δ, τ and ω to simulate the demand and supply for bribing at each value of the bribe. The parameter λ is calibrated such that the demand intersects the supply at the estimated value for the bribe (91 pesos) which is assumed to be the equilibrium value.

2. The dashed line represents the actual supply of bribing, which includes all vehicles that may apply for the driving restriction exemption. The dotted line excludes exemptable vehicles from the supply of donor cars.

Figure 5: Willingess to pay for car maintenance

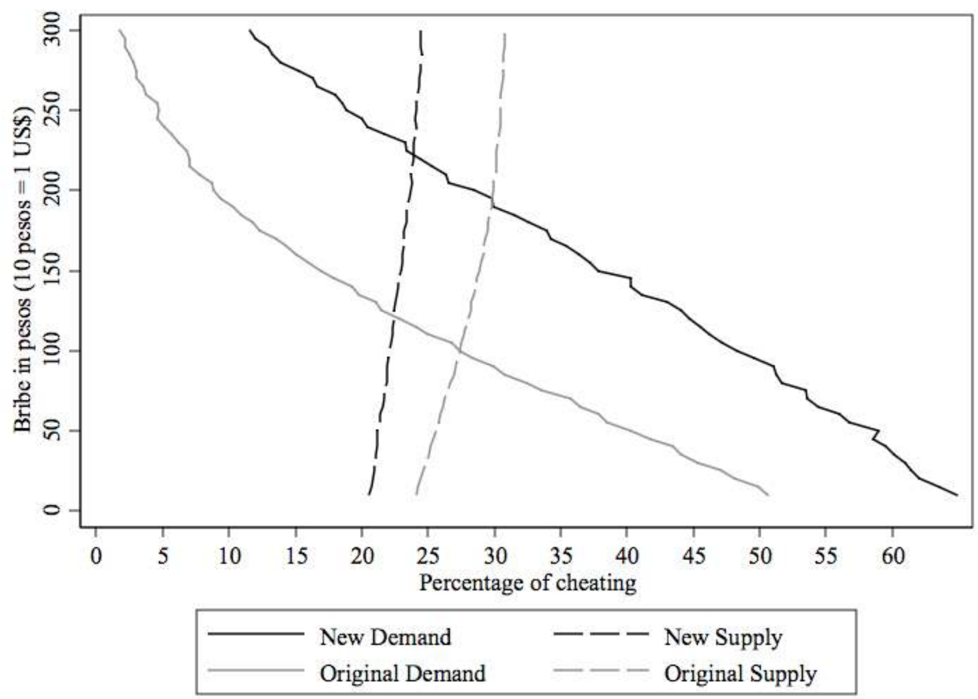


Notes:

1. This figure depicts the expected value of odd nodes as a function of the probability of passing. The solid line corresponds to a bribe of 91 pesos (9 US$), while the dashed line corresponds to a bribe of 190 pesos (19 US$).

2. The time cost parameter is assumed to be constant and equal to the mean time cost from the structural estimation: 42.40 pesos. The variance and intertemporal discount parameters, τ and δ, are assumed to be equal to their structural estimates: 94.14 and 0.4 respectively.

Figure 6: Simulated equilibrium. Tightening the emission standards.



Notes:
1. The figure uses the imputed probability of passing and the estimated parameters, δ, τ and ω to simulate the demand and supply for bribing at each value of the bribe. The parameter λ is calibrated such that the demand intersects the supply at the estimated value for the bribe (91 pesos) which is assumed to be the equilibrium value.
2. The gray lines correspond to demand and supply under the current emission standards. The black lines correspond to the demand and supply under tighter emission standards (similar to EPA Tier 1).

Appendix 1: Consistency of test for cheating

The statistical test from the regression methodology has null hypothesis:

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

and alternative hypothesis:

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

Under assumptions (A1) and (A2), the OLS coefficient $\hat{\gamma}_c$ from regression 3 is a consistent test statistic of null hypothesis.

Proof:

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

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

$$\mathbb{E}^*(u_{it}|\mathbf{x}_i, r_{i-1,t-1}) = \mathbf{w}_i\mathbb{E}(\mathbf{w}_i\mathbf{w}_i')^{-1}\mathbb{E}(\mathbf{w}_i u_{it}) \tag{11}$$

Equation 11 will equal zero under the null hypothesis if all elements of $\mathbb{E}(\mathbf{w}_i u_{it})$, $\mathbb{E}(r_{i-1,t-1}u_{it})$ and $\mathbb{E}(\mathbf{x}_i u_{it})$, are zero. The first term,

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

by assumptions (A1) and (A2). The second term,

$$\mathbb{E}(\mathbf{x}_i u_{it}) = \mathbf{0},$$

since $\underline{u_{it} = r_{it} - \mathbb{E}^*(r_{it}|\mathbf{x}_i)}$.

Appendix 2: Interpreting the serial correlation coefficient as the cheating prevalence rate

Given equation 2 for observed emissions in the presence of cheating, the serial correlation coefficient may have a prevalence rate interpretation. The consistent OLS estimation of p_0 in regression equation 4 requires mean independence between $r_{i-1,t-1}$ and $\nu_{i,t}$. Mean independence between $r_{i-1,t-1}$ and $\nu_{i,t}$ is unlikely in the current setup for the following reasons. First, notice that the substitution of $r_{i-1,t-1}$ for $r_{i-1,t}$ creates a measurement error problem since their components $e_{i-1,t}$ and $e_{i-1,t-1}$ are different. Measurement error in the right hand side variable of interest will generate downward bias in the estimate of corruption prevalence.

Second, there is a correlated random coefficient problem since c_i and $r_{i-1,t-1}$ are negatively correlated. This occurs because donor cars should have lower emissions, hence $\mathbb{E}(r_{i-1,t-1}|c_{i=1}) < \mathbb{E}(r_{i-1,t-1})$. Borrowing from Chay and Greenstone's (2005) notation for correlated random coefficients, define $c_i = p_0 + p_i$, where p_0 is the overall cheating prevalence rate or average c_i and p_i is a random deviation from the average. Notice that p_i is mean zero, *i.e.* $\mathbb{E}(p_i) = p_0(1-p_0) + (1-p_0)(-p_0) = 0$. Using this new notation, correlated random coefficient bias will arise from the correlation between the varying component of c_i, p_i, and $r_{i-1,t-1}$.

The following subsections will elaborate on the two sources of bias. I will start by analyzing the measurement error problem exclusively, by assuming away the correlation between c_i and $r_{i-1,t-1}$. Then, I will analyze the correlated random coefficient problem by assuming that $r_{i-1,t}$ is observed in order to simplify the analysis of correlated random coefficients bias.

A2.1 Measurement Error and Instrumental Variables

Measurement error arises from the fact that fake emissions, r_{i-1t}, are unobserved. However, we do observe emissions from donor cars in their own test, r_{i-1t-1}, which are related to r_{i-1t} by the following equation:

$$r_{i-1,t} = r_{i-1,t-1} + e_{i-1,t} - e_{i-1,t-1} \tag{12}$$

In this subsection I will analyze how the measurement error in the right hand side variable affects the consistency of $\hat{p_0}$. In order to do so, I will abstract form the correlated random coefficients problem, *i.e.* I will assume in this subsection that $c_i \perp r_{i-1,t-1}$; which is equivalent to assume that $p_i \perp r_{i-1,t-1}$. The joint effect of correlated random coefficients and measurement error is analyzed in Appendix 2.

Substituting 12 in 4 yields:

$$r_{i,t} = p_0 r_{i-1,t-1} + \mathbf{x}_i \gamma'_x + \nu'_{i,t}, \tag{13}$$

where $\gamma'_x = (1-p_0)\beta$, and $\nu'_{it} = c_i(e_{i-1,t} - e_{i-1,t-1}) + (1-c_i)(a_i + e_{i,t}) - p_i \tilde{r}_{i,t} - p_i r_{i-1t-1}$.

Under the assumptions of this subsection, consistent estimation of p_0 requires mean independence between $r_{i-1,t-1}|\ \mathbf{x}_i$ and $\nu'_{i,t}$: $\mathbb{E}(\nu'_{i,t}|r_{i-1,t-1}, \mathbf{x}_i) = \mathbb{E}(\nu'_{i,t}|\mathbf{x}_i)$. However, for presentation purposes, I will devote the following paragraphs to verify a stronger condition for consistency:

$$\mathbb{E}(\nu'_{i,t}|r_{i-1,t-1}) = \mathbb{E}(\nu'_{i,t}),$$

which implies the former one. This will allow me to omit $\mathbf{x}_i$ in the conditional expectations below.

The first component of $\nu'_{i,t}$ will violate this independence condition by introducing classic measurement error bias. The presence of this term in the regression's residual will bias the estimate of p_0 towards zero:

$$\mathbb{E}(c_i(e_{i-1,t} - e_{i-1,t-1})|r_{i-1,t-1}) = -p_0\mathbb{E}(e_{i-1,t-1}|r_{i-1,t-1})$$

The last three components of $\nu'_{i,t}$ will meet the independence condition under the simplifying assumption of this subsection: $p_i \perp r_{i-1,t-1}$.[34]

Notice that if $c_i = 0$ for all i, then $\mathbb{E}(\nu'_{i,t}|r_{i-1,t-1}) = 0$. This result implies that a significant coefficient on lagged emissions is evidence of corruption even in the presence of measurement error. Hence, equation 13can be used to test for corruption even if there is measurement error in $r_{i-1,t-1}$.

Next, I propose a simple instrumental variable framework can be used to eliminate the bias associated with measurement error. In the absence of other sources of bias, an ideal instrument should be correlated with $r_{i-1,t-1}$ but uncorrelated with $e_{i-1,t-1}$. As mentioned in the previous section, each test performed at the center offers two measures for each gas: the first one is taken at 25 rpm and 40 kph and the second one at 50 rpm and 24 kph. Within each pair of same-gas measures, each measure is a great predictor of the other one. In addition, since they are taken from different smoke samples, they do not share the time specific error component of emissions: e_{it}.

Define the lag of the reading at 25 kph as $r_{i-1,t-1}$ and the lag of the reading at 40 kph as $r'_{i-1,t-1}$. To simplify the analysis in this subsection, assume also that $c_i \perp r'_{i-1,t-1}$. The first stage of the IV estimation will be given by the following equation:

$$r_{i-1,t-1} = \theta r'_{i-1,t-1} + \delta\mathbf{x}_i + u_{i-1,t-1},$$

where $u_{i-1,t-1} = e_{i-1,t-1} - e'_{i-1t-1} + b_{i-1,t-1} - b'_{i-1,t-1}$. The term b_{i-1t-1} is a component of emissions that varies with speed and rpm, and the apostrophe indicates that the emission component belongs to the reading at 40 kph. The variable r'_{i-1t-1} will be a suitable instrument for r_{i-1t-1} if $\theta \neq 0$ and

[34] Notice that, since this subsection assumes p_i is independent of $r_{i-1,t-1}$, the last three components of error term $\nu'_{i,t}$ are not correlated with $r_{i-1,t-1}$and will not cause bias in the estimation of p_0. More specifically, $\mathbb{E}((1-c_i)(a_i + e_{i,t})|r_{i-1t-1}) = -\mathbb{E}(c_i(a_i+e_{i,t})) = \kappa_1$, $\mathbb{E}(p_i r_{i-1,t-1}|r_{i-1t-1}) = \mathbb{E}(p_i)r_{i-1,t-1} = 0$, and $\mathbb{E}(p_i\tilde{r}_{i,t}|r_{i-1t-1}) = \mathbb{E}(p_i\tilde{r}_{i,t}) = \kappa_2$, where κ_1 and κ_2 are two constant terms.

$$\mathbb{E}(\nu'_{i,t}|r'_{i-1,t-1}) = \mathbb{E}(\nu'_{i,t}). \tag{14}$$

Under the IV assumption

(A3) $\quad \mathbb{E}(e_{i-1,t-1}|r'_{i-1,t-1}) = 0,$

and the simplifying assumption that c_i and $r'_{i-1,t-1}$ are independent, the condition given by 14 is met.

Summing-up, under assumption (A3) and in the absence of correlated random coefficient bias, IV will yield a consistent estimate of p_0. Section A2.3 will discuss the results from IV specifications.

A2.2 Correlated Random Coefficients and donor car selection.

As mentioned at the beginning of this section, correlated random coefficient bias may arise since donor cars are presumably picked among those cars that have reliably low emissions. The selection of donor cars based on low emissions will generate a negative correlation between c_i and $r_{i-1,t}$: when $c_i = 1$, $r_{i-1,t}$ is likely to be below the emission limit for i. The negative correlation between the random coefficient, c_i, and the main regressor, $r_{i-1,t}$, will generate downward bias in the estimate of $\mathbb{E}(c_i)$ (Wooldridge, 2002).

In order to focus on this source of bias, assume for simplicity that emissions of car $i-1$, when used as a donor car, are directly observed from the data, or that $r_{i-1,t-1} = r_{i-1,t}$. Borrowing from Chay and Greenstone's (2005) notation for correlated random coefficients, define $c_i = p_0 + p_i$, where p_0 is the overall cheating prevalence rate or average c_i and p_i is a random deviation from the average. Notice that p_i is mean zero, *i.e.* $\mathbb{E}(p_i) = p_0(1-p_0) + (1-p_0)(-p_0) = 0$. With this notation and the assumption that $r_{i-1,t-1} = r_{i-1,t}$, the equation to estimate becomes:

$$r_{i,t} = p_0 r_{i-1,t} + \mathbf{x}_i \gamma''_x + \nu''_{i,t}, \tag{15}$$

where $\nu''_{i,,t} = p_i r_{i-1,t} + (1-p_0)(a_i + e_{i,t}) - p_i \tilde{r}_{i,t}$, $p_0 = \mathbb{E}(c_i) = \Pr(c_i = 1)$. The consistent estimation of p_0 under the current assumptions, would require $\mathbb{E}(\nu''_{i,t}|r_{i-1,t-1}, \mathbf{x}_i) = \mathbb{E}(\nu''_{i,t}|\mathbf{x}_i)$. However, for presentation purposes, I will verify instead that the more stringent condition $\mathbb{E}(\nu''_{i,t}|r_{i-1,t-1}) = \mathbb{E}(\nu''_{i,t})$ holds. This will allow me to omit $\mathbf{x}_i$ from the conditional expectations below.

Under the assumptions of this subsection, the conditional expectation $\mathbb{E}(\nu''_{i,t}|r_{i-1,t-1})$ can be decomposed as:

$$\mathbb{E}(\nu''_{i,t}|r_{i-1,t}) = \mathbb{E}(p_i|r_{i-1,t})r_{i-1,t} + (1-p_0)\mathbb{E}(a_i + e_{i,t}|r_{i-1,t}) - \mathbb{E}(p_i\tilde{r}_{i,t}|r_{i-1,t}). \tag{16}$$

Equation 16 shows that the test for corruption is reliable even in the presence of correlated random coefficient bias. To see this, notice that if $c_i = 0$ for all i, then $p_i = 0$ and $\mathbb{E}(\nu''_{i,t}|r_{i-1,t}) = 0$.

However, equation 16 shows that $\mathbb{E}(\nu''_{i,t}|r_{i-1,t}) \neq \mathbb{E}(\nu''_{i,t})$ whenever c_i is correlated with $r_{i-1,t}$. This constitutes a problem for the interpretation of $\hat{p}_0$ as the cheating prevalence rate. Notice that all three terms in equation 16 vary with $r_{i-1,t}$, and will induce bias in $\hat{p}_0^{OLS}$ if $\Pr(c_i = 1|r_{i-1,t}) \neq \Pr(c_i = 1)$. First, $\mathbb{E}(p_i|r_{i-1,t}) \neq 0$, since p_i is just $c_i - p_0$ and c_i varies with $r_{i-1,t}$. More specifically: $\mathbb{E}(p_i|r_{i-1,t}) = (1 - p_0)\Pr(c_i = 1|r_{i-1,t}) + (-p_0)\Pr(c_i = 0|r_{i-1,t})$. Second, $\mathbb{E}(a_i + e_{i,t}|r_{i-1,t}) \neq 0$, since at least the permanent component of emissions, a_i, may be a determinant of the decision to cheat, and c_i is correlated with $r_{i-1,t}$. Third, $\mathbb{E}(p_i\tilde{r}_{i,t}|r_{i-1,t}) \neq \text{Cov}(p_i, \tilde{r}_{i,t})$ since both, p_i and $\tilde{r}_{i,t}$ are related to the decision to cheat, c_i. More specifically:

$$\mathbb{E}(p_i\tilde{r}_{i,t}|r_{i-1,t}) = \Pr(c_i = 1|r_{i-1,t})(1 - p_0)\mathbb{E}(\tilde{r}_{i,t}|c_i = 1) \quad (17)$$

$$+\Pr(c_i = 0|r_{i-1,t})(-p_0)\mathbb{E}(\tilde{r}_{i,t}|c_i = 0)$$

Again, equation 17 differs from $\text{Cov}(p_i\tilde{r}_{i,t}) = p_0(1-p_0)\mathbb{E}(\tilde{r}_{i,t}|c_i = 1) - p_0(1-p_0)\mathbb{E}(\tilde{r}_{i,t}|c_i = 0)$, since $\Pr(c_i = 1|r_{i-1,t}) \neq p_0$.

The correlated random coefficients bias stems from the fact that clean cars make good donor cars and dirty cars don't. A more careful examination of the rule for choosing a donor car will help understand the magnitude of this bias. First, notice that donor cars should meet the standards in all 4 different gases: HC, CO, NO and O_2. The standards in some of these gases are easier to meet than in others. For example, less than 0.005 of vehicles fail to meet the O_2 standard.[35] Hence, we would expect O_2 to be fairly similar across donor and average cars. This would result in less correlated random coefficient bias when using O_2 as the left hand side variable in 15[36]In summary, we would expect correlated random coefficient bias to be less important for those gases whose limits are easier to attain. Section A2.3 will show results using all for all four gases as right hand side variables; and will verify the magnitude of the prevalence estimate from each specification is associated with the relative stringency of the test on each corresponding gas.

A2.3 Comparing OLS and IV estimates of cheating prevalence

Table A1 compares OLS and IV serial correlation estimates for all four pollutants. Section A2.1 argues that IV corrects for measurement error bias under assumption A3. Since measurement error causes downward bias in the serial correlation coefficient, we would expect IV estimates to be

[35]As mentioned in Section 2, the emission test includes a limit in O_2 in order to avoid tampering with the engine, which usually consists on changing the proportion of oxygen used during fuel combustion.

[36]Notice that there could still be substantial correlated random coefficient bias if O_2 was strongly correlated with other gases on which selection is more stringent, like HC. The correlation of O_2 and HC is the strongest between any pair of gases with a correlation coefficient of 0.22.

higher than OLS estimates. Table A1 shows that this is the case for every pollutant. Measurement error seems to be more important for oxygen (fourth row of each panel).

Section A2.2 argues that those pollutants that are unimportant for donor car selection will yield better estimates of cheating prevalence since they will be less subject correlated random coefficient bias. In the case of oxygen, correlated random coefficient bias is thought to be minimal since 99.7 of vehicles are under the threshold compared to 92.5, 93 and 95 for hydrocarbons, carbon monoxide and nitrogen oxides, respectively. The estimates in Table A1 are consistent with this observation, since oxygen yields the highest estimates of cheating prevalence. Notice also that the second highest estimates correspond to nitrogen oxides, for which maximum levels are the second easiest to attain.

In conclusion, when interpreting the serial correlation coefficient as cheating prevalence rate, the most reliable estimates correspond to the IV estimates in the oxygen specification (bottom row of Table A1), which range from 20 to 25 percent.[37]

[37]The specification with not car characteristic controls yields an estimate of 32 percent, however vehicle sorting across centers or time of the day could be producing some additional correlation. Hence, estimates that control for car and test characteristics are more reliable.

Appendix 3: Likelihood Function from Structural Estimation

The probabilities of each decision at an odd decision node are given by:

$$\Pr(S^{od}=B)=\frac{\exp\left(\frac{-w-p-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}{1+\exp\left(\frac{-w-p-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)+\exp\left(\frac{-w-p+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}$$

$$\Pr(S^{od}=X)=\frac{1}{1+\exp\left(\frac{-w-p-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)+\exp\left(\frac{-w-p+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}$$

$$\Pr(S^{od}=A)=\frac{\exp\left(\frac{-w-p+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}{1+\exp\left(\frac{-w-p-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)+\exp\left(\frac{-w-p+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}$$

Similarly, the probabilities of each decision at an even decision node are given by:

$$\Pr(S^{ev}=X)=\frac{1}{1+\exp\left(\frac{-w-p-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)+\exp\left(\frac{-w+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}$$

$$\Pr(S^{ev}=B)=\frac{\exp\left(\frac{-w-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}{1+\exp\left(\frac{-w-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)+\exp\left(\frac{-w+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}$$

$$\Pr(S^{ev}=A)=\frac{\exp\left(\frac{-w+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}{1+\exp\left(\frac{-w-b-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)+\exp\left(\frac{-w+(1-z)\mathbb{E}V^{ev}-\delta(-f+\mathbb{E}V^{od})}{\tau}\right)}$$

The probabilities for each history of decisions and test outcomes are given by:

$$\Pr(H1)=\Pr(S^{od}=X)$$

$$\Pr(H2)=\Pr(S^{od}=B)$$

$$\Pr(H3)=\Pr(S^{od}=A)\cdot z$$

$$\Pr(H4)=\Pr(S^{od}=A)\cdot(1-z)\cdot\Pr(S^{ev}=X)$$

$$\Pr(H5)=\Pr(S^{od}=A)\cdot(1-z)\cdot\Pr(S^{ev}=B)$$

$\Pr(H6) = \Pr(S^{od} = A) \cdot (1 - z) \cdot \Pr(S^{ev} = A) \cdot z$

$\Pr(H7, H8, ...) = \Pr(S^{od} = A) \cdot (1 - z) \cdot \Pr(S^{ev} = A) \cdot (1 - z)$

The probabilities in the list above sum up to one. Notice that the last probability in the list includes all histories after $H6$. Hence, I do not need to derive an expression for the probability of each subsequent history. The choice of where to stop considering subsequent histories is, to some extent, arbitrary. However, a good reason for stopping at $H7$ is that 95 percent of the observed histories are described by $H1$-$H6$.

Since not all histories are observed, the model's probabilities for each observed outcome are given by:

$\Pr(H1) = \Pr(S^{od} = X)$

$\Pr(H2, H3) = \Pr(S^{od} = B) + \Pr(S^{od} = A) \cdot z$

$\Pr(H4) = \Pr(S^{od} = A) \cdot (1 - z) \cdot \Pr(S^{ev} = X)$

$\Pr(H5, H6) = \Pr(S^{od} = A) \cdot (1 - z) \cdot \Pr(S^{ev} = B) + \Pr(S^{od} = A) \cdot (1 - z) \cdot \Pr(S^{ev} = A) \cdot z$

$\Pr(H7, H8, ...) = \Pr(S^{od} = A) \cdot (1 - z) \cdot \Pr(S^{ev} = A) \cdot (1 - z)$

Table A1: Prevalence rate estimations from OLS and IV

	[1]	[2]	[3]	[4]
Panel A: OLS				
HC24	0.039	0.024	0.019	0.018
	[0.002]**	[0.002]**	[0.002]**	[0.002]**
CO24	0.078	0.042	0.029	0.027
	[0.004]**	[0.003]**	[0.003]**	[0.003]**
NO24	0.153	0.145	0.086	0.072
	[0.007]**	[0.008]**	[0.005]**	[0.004]**
O2_24	0.18	0.14	0.12	0.11
	[0.008]**	[0.008]**	[0.007]**	[0.006]**
Panel B: IV				
HC24	0.058	0.032	0.029	0.027
	[0.003]**	[0.003]**	[0.002]**	[0.002]**
CO24	0.093	0.032	0.034	0.032
	[0.004]**	[0.014]*	[0.003]**	[0.003]**
NO24	0.166	0.153	0.087	0.076
	[0.008]**	[0.010]**	[0.005]**	[0.005]**
O2_24	0.322	0.248	0.218	0.199
	[0.016]**	[0.016]**	[0.014]**	[0.013]**
Observations	1754206	1754206	1754206	1754206
Car characteristics		X	X	X
Center FE			X	X
Testing equipment FE				X

Notes:

1. Standard errors in brackets. * Significant at 5%; **Significant at 1%
2. Source: Smog Check Center data for 2003 (2), Distrito Federal
3. Sample for estimation corresponds to a 30 percent random sample by center.
4. Car characteristics include dummies for brand, size and service, interactions of size and service dummies with model-year, and polynomials on model-year, odometer and time of the day.

Table A2: Robustness test 1. Is serial correlation higher when consequtive tests are closer in time?

	Hydrocarbons		Carbon Monoxide		Nitrogen Oxides		Oxygen	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Cheating prevalence rate by time between tests								
0-5 min	0.032	0.026	0.044	0.035	0.158	0.087	0.154	0.129
	[0.008]**	[0.007]**	[0.009]**	[0.008]**	[0.019]**	[0.010]**	[0.016]**	[0.015]**
5-10 min	0.017	0.012	0.036	0.028	0.139	0.086	0.17	0.142
	[0.003]**	[0.002]**	[0.005]**	[0.005]**	[0.013]**	[0.008]**	[0.014]**	[0.012]**
10-15 min	0.018	0.015	0.036	0.031	0.11	0.073	0.25	0.23
	[0.006]**	[0.005]**	[0.006]**	[0.006]**	[0.011]**	[0.008]**	[0.090]**	[0.092]*
15-20 min	0.02	0.017	0.029	0.023	0.096	0.062	0.149	0.128
	[0.006]**	[0.005]**	[0.008]**	[0.008]**	[0.011]**	[0.010]**	[0.014]**	[0.013]**
more than 20 min	0.015	0.01	0.028	0.022	0.076	0.047	0.109	0.094
	[0.004]**	[0.004]*	[0.006]**	[0.005]**	[0.008]**	[0.007]**	[0.024]**	[0.021]**
Direct effect of time between tests on emissions								
5-10 min relative to 0-5 min	5.268	4.365	0.007	0.02	49.669	6.581	-0.051	-0.009
	[2.530]*	[1.802]*	[0.016]	[0.011]	[27.537]	[9.169]	[0.021]*	[0.011]
10-15 min relative to 0-5 min	9.207	5.715	0.036	0.037	92.495	22.258	-0.147	-0.081
	[3.299]**	[2.029]**	[0.020]	[0.013]**	[34.556]**	[11.313]	[0.065]*	[0.065]
15-20 min relative to 0-5 min	10.313	6.316	0.052	0.05	106.494	27.656	-0.076	-0.011
	[3.547]**	[2.496]*	[0.022]*	[0.016]**	[35.216]**	[12.984]*	[0.026]**	[0.018]
more than 20 min relative to 0-5 min	9.818	5.841	0.059	0.055	120.839	29.464	-0.048	0.009
	[3.149]**	[2.224]*	[0.021]**	[0.015]**	[35.572]**	[13.453]*	[0.026]	[0.020]
Car characteristics	X	X	X	X	X	X	X	X
Center FE		X		X		X		X
Observations	466303	466303	466303	466303	466303	466303	466303	466303
R-squared	0.14	0.06	0.33	0.18	0.47	0.21	0.41	0.19

Notes:

1. Source: OLS regressions on smog-check center data.

2. Excluded category on direct time-effect dummies is 0-5 minutes.

Table A3: Robustness Test 2
Comparing cheating evidence between passed and failed tests

	[1]	[2]	[3]	[4]
	Just passed		Just failed	
Panel A: OLS				
HC24	0.432	0.453	-0.065	-0.033
	[0.089]**	[0.039]**	[0.046]	[0.062]
	5745	5745	3801	3801
CO24	0.448	0.454	0.118	0.114
	[0.038]**	[0.036]**	[0.028]**	[0.025]**
	4267	4267	3675	3675
NO24	0.312	0.312	0.06	0.061
	[0.059]**	[0.049]**	[0.020]**	[0.018]**
	4762	4762	5142	5142
Panel B: IV				
HC24	0.596	0.773	-0.092	0.021
	[0.103]**	[0.068]**	[0.071]	[0.053]
	5745	5745	3801	3801
CO24	0.962	0.234	0.149	-2.571
	[1.287]	[0.271]	[0.031]**	[68.755]
	4267	4267	3675	3675
NO24	0.41	0.412	0.035	0.058
	[0.057]**	[0.055]**	[0.024]	[0.217]
	4762	4762	5142	5142
Car characteristics	X	X	X	X
Centro FE		X		X

Notes:

1. Standard errors in brackets. * Significant at 5%; **Significant at 1%. Number of observations in regression below standard errors.
2. Source: Smog Check Center data for 2003 (2), Distrito Federal
3. Sample for panels A and D corresponds to all observations for which lagged hydrocarbons are within 0.05 standard deviations from the threshold faced by the vehicle. The sample is also constrained to tests that have all other lagged pollutants below the limits faced by the vehicle.
4. Sample for panels B and E corresponds to all observations for which lagged carbon monoxides are within 0.3 standard deviations from the threshold faced by the vehicle. The sample is also constrained to tests that have all other lagged pollutants below the limits faced by the vehicle.
5. Sample for panels C and F corresponds to all observations for which lagged carbon monoxides are within 0.1 standard deviations from the threshold faced by the vehicle. The sample is also constrained to tests that have all other lagged pollutants below the limits faced by the vehicle.
6. Sample in columns [1] and [2] is further constrained to tests for which lagged emissions of the corresponding pollutant are below the threshold (*i.e.* the previous vehicle would have passed the test under the current vehicle standards)
7. Sample in columns [3] and [4] is further constrained to tests for which lagged emissions of the corresponding pollutant are above the threshold (*i.e.* the previous vehicle would not have passed the test under the current vehicle standards)
8. Car characteristics include dummies for brand, size and service, interactions of size and service dummies with model-year, and polynomials on model-year, odometer and time of the day.

Table A4: Model parameters and demand for bribes

Panel A: Parameter Estimates

Model Parameters	Estimate	SE	Structural Parameters	
w-period1	54.11	15.9572	Time cost	
w-period2	38.33	15.7234	period 1	54.11
w-period3	41.44	14.4959	period 2	38.33
w-period4	39.91	15.6888	period 3	41.44
w-period5	42.96	15.9474	period 4	39.91
b	105.71	9.6788	period 5	42.96
δ	0.38	0.0206	Bribe	105.71
τ	83.67	5.6474	Discount factor, δ	0.38
			SE of random shock	107.31

Panel B: Fitted and predicted probabilities for each history

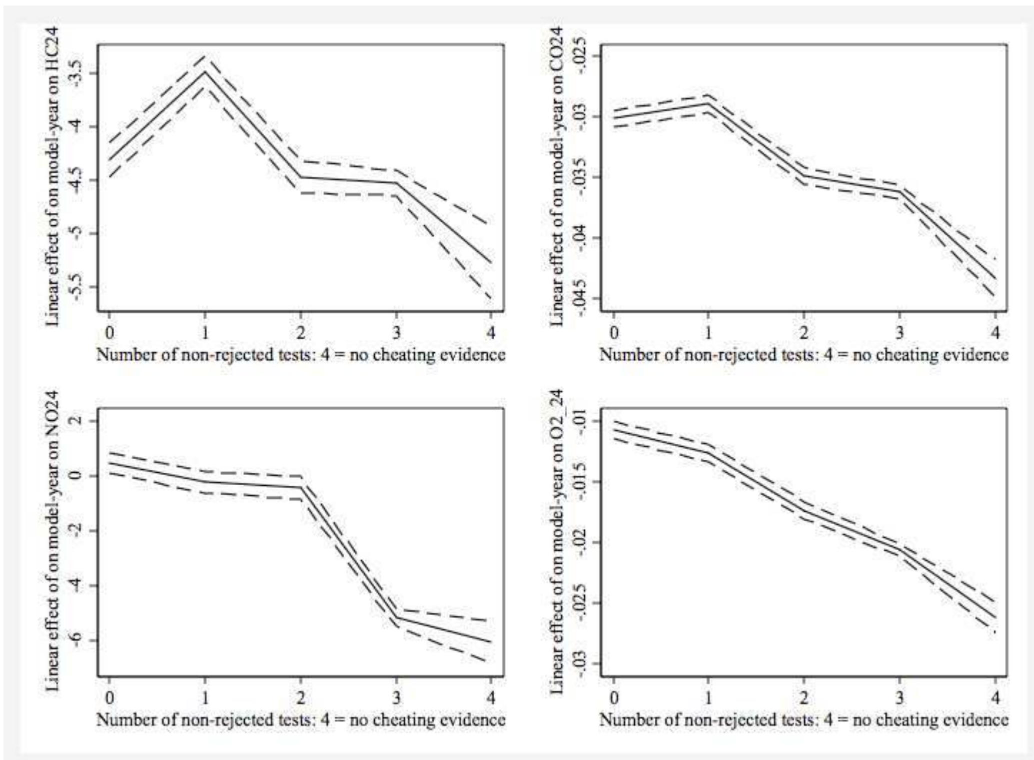
Observed History	Actual	Fitted
H1: Postpone	0.0956	0.0603
H3,H4: Bribe/No bribe - Pass	0.7284	0.7432
H4: No bribe-Fail-Postpone	0.0051	0.0016
H5, H6: No bribe-Fail-Bribe/No bribe-Fail-No bribe-Pass	0.1266	0.155
H7: No bribe-Fail-No bribe-Fail	0.0444	0.0399

History	Predicted
H1: Postpone	0.0603
H2: Bribe	0.1967
H3: No bribe-Pass	0.5465
H4: No bribe-Fail-Pospone	0.0016
H5: No bribe-Fail-Bribe	0.0628
H6 No bribe-Fail-No bribe-Pass	0.0922
H7: No bribe-Fail-No bribe-Fail	0.0399

Notes:
1.The first five model parameters correspond to the time costs associated with being assigned to each of the five smog-check two-month windows available. Period 1 corresponds to June-July, period 2 to July-August, etc. Vehicles are randomly assigned to periods.

2. See notes of Table 7

Figure A1: Falsification test for non-cheating center sample



Notes:

1. This figure tests for whether vehicle characteristics are more highly correlated with emissions as the evidence of cheating weakens. The figure restricts the analysis to the most important predictor of emissions: model-year. Each panel shows how the linear relationship between emissions and model-year becomes more negative as the evidence of cheating falls.

2. In each of the panels, the solid links the slope coefficient estimates from 5 OLS regressions of emissions on model-year and a constant. The dashed lines link the confidence intervals of the slope coefficients.

3. Each number in the horizontal axis corresponds to a different sample of centers, where centers where classified according to the strength of the cheating evidence in each centers. The horizontal axis describes each group of centers by the number of cheating tests (one for each pollutant) for which the no-cheating hypothesis was not rejected. Zero means evidence of cheating was found in all 4 tests, while 4 means no evidence was present in any of the four tests. Hence the group described by number 4 corresponds to the non-cheating sample of centers that is used in the paper for predicting emissions and the probability of passing the test.

Table A5: Probit Model for the Probability of Passing

Manufacturer	Probit coefficient	Standard Error
Chrysler / Dodge	423.249	[523.75]
Ford	423.540	[523.75]
GM	423.375	[523.75]
Nissan	423.338	[523.75]
Renault	423.162	[523.76]
VAM	423.352	[523.76]
VW	423.314	[523.75]
Jeep	423.301	[523.76]
Mercedes	423.438	[523.75]
Dina	423.046	[523.76]
BMW	422.708	[523.75]
Honda	422.221	[523.75]
Porsche	423.204	[523.76]
Volvo	422.608	[523.74]
Toyota	423.474	[523.75]
Fiat	422.240	[523.74]
Service and interactions		
Taxi	-95.065	[13.45]
Coorporate	-81.627	[17.42]
Taxi*Modelo	0.048	[0.007]
Coorp*Modelo	0.041	[0.008]
Engine displacement and interactions		
2000 Disp	-74.154	[16.02]
3000 Disp	-19.878	[7.86]
4000 Disp	-137.969	[18.18]
5000 Disp	5.798	[9.67]
6000 Disp	-41.403	[11.73]
2000 Disp*Modelo	0.037	[0.008]
3000 Disp*Modelo	0.010	[0.004]
4000 Disp*Modelo	0.069	[0.009]
5000 Disp*Modelo	-0.003	[0.005]
6000 Disp*Modelo	0.021	[0.006]
Modelo	-0.430	[0.530]
$Modelo^2$	$1.08x10^{-4}$	[$1.339x10^{-4}$]
km	$-3.50x10^{-6}$	[$3.75x10^{-7}$]
km^2	$9.05x10^{-12}$	[$2.01x10^{-7}$]
km^3	$-1.21x10^{-17}$	[$3.63x10^{-18}$]
km^4	$5.78x10^{-24}$	[$2.03x10^{-24}$]
time (seconds since 12:00am)	$2.10x10^{-4}$	[$1.95x10^{-5}$]
$time^2$	$-4.21x10^{-9}$	[$4.04x10^{-10}$]
$time^3$	$2.72x10^{-14}$	[$2.75x10^{-15}$]

Notes:
1. Standard errors in brackets.
2. The model does not include a constant. No category is exluded from manufacturers
3. Private vehicle category exluded from services
4. Displacement of 1000 exluded from engine displacement