Environmental Regulations and Corruption:
Automobile Emissions in Mexico City

Online Appendices, Tables and Figures

Paulina Oliva

May, 2014
Online Appendix 1: Serial Correlation Coefficient under The Null Hypothesis

The statistical test from the regression methodology has null hypothesis:

\[ H_0 : r_{it} = \tilde{r}_{it} = x_{it} \beta + u_t, \]

and, according to (A1), alternative hypothesis:

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

Under assumptions (A1) - (A3), the expected value of the OLS coefficient estimator from regression (3) in the main paper, \( \hat{\gamma}_c \), under the null hypothesis is zero.

Proof: Under the null hypothesis:

\[ E^* (r_{it} | x_{it}, r_{i-1t-1}) = x_{it} \beta + E^* (u_{it} | x_{it}, r_{i-1t-1}) \]

Define \( w_i = [x_{it}, r_{i-1t-1}] \) and \( W = [w_{2i}, w_{3i}, ..., w_{N_1}, w_{N_1+2}, w_{N_1+3}, ..., w_{N_1+N_2}, ...] \), 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

\[ E^* (u_{it} | x_{it}, r_{i-1t-1}) = w_i E (w_i^T w_i)^{-1} E (w_i^T u_{it}) \quad (1) \]

Equation 1 will equal zero under the null hypothesis if all elements of \( E (w_i^T u_{it}) \), i.e. \( E(r_{i-1t-1} u_{it}) \) and \( E(x_{it}^T u_{it}) \), are zero. Under the null, \( E(r_{i-1t-1} u_{it}) = E(\tilde{r}_{i-1t-1} u_{it}) \). The first term equals zero because

\[ E(\tilde{r}_{i-1t-1} u_{it}) = E((x_{i-1t-1} \beta + u_{i-1t-1}) u_{it}) = E(x_{i-1t-1} \beta u_{it}) + E(u_{i-1t-1} u_{it}) = 0, \]
by assumptions (A2) and (A3). The second term, \( E(x_{it}^T u_{it}) \), equals zero by definition, since since \( u_{it} = \tilde{r}_{it} - E^*(\tilde{r}_{it} | x_i) = \tilde{r}_{it} - x_i \beta \).

In order to establish consistency of the test (asymptotic power 1), we need to add the following second moment assumptions:

(A4) \[ E \left( u_{it}^2 x_{it}^T x_{it} \right) = 0 \]

(A5) \[ E \left( u_{it}^2 x_{i-1t-1}^T x_{i-1t-1} \right) = 0 \]

(A6) \[ E \left( u_{it}^2 u_{i-1t-1}^2 \right) = 0 \]

Assumption (A4) is an homoskedasticity assumption, while assumption’s (A5) and (A6) require that the
variance of the error term does not vary with observed and unobserved determinants of neighboring vehicles’ emissions. Assumptions (A4)-(A6) guarantee the asymptotic normality of $\hat{\gamma}_c$, and deliver asymptotic power 1, or consistency of the test (Casella and Berger, 2002).

**Online Appendix 2: Robustness Checks on Cheating Test**

**Permutations Test for Cheating**

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

In order to measure the difference between consecutive tests, I consider all readings for each car, i.e. each pollutant in each of the different driving conditions. This amounts to 8 readings per test: $HC$ at 24kmh, $HC$ at 40kmh, $CO$ at 24kmh, $CO$ at 40kmh, $O_2$ at 24kmh $O_2$ at 40kmh, $NO$ at 24kmh and $NO$ at 40kmh. The measure of multivariate distance I will use is

$$d_t = (D_t V^{-1} D_t')^{1/2},$$

where $D_t = (r_{s,1,t}, r_{s,2,t}, ..., r_{s,8,t}) - (r_{s,1,t-1}, r_{s,2,t-1}, ..., r_{s,8,t-1})$, $r_{s,j,t}$ is a transformed measure of pollutant $j$ in test $t$, and $V$ is the estimated variance covariance matrix across standardized pollutants. This multivariate distance measure is commonly known as Mahalanobis distance and is similar to the Euclidean distance except that differences of each pollutant are weighted not only by the variance of each pollutant but also by the covariance across pollutants. Hence, if pollutants $HC$ and $CO$ are positively correlated, a large difference

---

1 The transformation used is given by $r_{s,j,t} = \sqrt{r_{j,t}}$. Other standardizations, such as $ln(r_{j,t})$, yield similar results.
in HC will have a smaller weight whenever the difference in CO is also large. Notice that t indexes tests in the order in which they occurred. The solid lines in Appendix Figure 1 show the distribution of $d_t$ for two different smog-check centers in Panels A and B.

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

A formal statistical test can be derived from this methodology. First, I compute a test statistic that describes the relevant part of the distribution. Since I’m looking for an excessive amount of difference close to zero, or an excessive amount of small $d_t$ draws, I chose the fifth percentile, $\hat{q}_{0.05}^d$, of the distribution of $d_t$. I will compare this statistic with the distribution of the fifth percentile of $d_{t,n}$ for $n = 1, ..., 1000$. I.e., I will generate an empirical distribution of $\hat{q}_{0.05}^d$ by performing 1000 different random permutations of the smog test ordering. The comparison yields p-values for the null hypothesis of no cheating. The test is performed at the center level for each of the 80 smog-check centers in the 2003-2nd-half round of D.F.’s smog-check center data. The p-values for all centers are less than 0.001, which rejects the null hypothesis of no cheating for all centers with 99.9 percent 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).

---

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

3 Other test statistics, such as the proportion of vehicles under $d_t = 1$ and the tenth percentile of $d_t$, yield similar results.
Online Appendix 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 pollutants standardized by the variance covariance matrix (Mahalanobis distance).

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.
**Time between tests and serial correlation**

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

---

**Online Appendix Figure 2: Serial Correlation by Time Window within Consecutive Tests**

Notes: This figure shows the coefficients (and confidence intervals) of the interactions between time window dummies and lagged emissions in a seemingly unrelated equations model.
Cheating and the relationship between emissions and car characteristics

Finally, notice that when donor cars are more frequently used, vehicle emissions are more likely to be matched to the “wrong” car characteristics. Therefore, the final set of robustness tests evaluates whether the relationship between car characteristics and emissions varies across centers by the level of corruption. Specifically, I classify centers in 10 groups according to the $p$-value of the joint test for whether emissions are serially correlated. Centers with the lowest $p$-values are more likely to engage in cheating using donor cars, while centers with higher $p$-values are less likely to engage in cheating. I then run a simple linear regression of emissions on model-year and a constant for each group. The estimated coefficient on model-year should be more negative for the group of centers that, based on the statistical test, are unlikely to have cheating (group 1) and less negative for centers that are more likely to have cheating (group 10).

Web Appendix Figure 3 depicts the strength of the estimated relationship between the year the car was made and the measured emissions by how likely a center is to be engaged in cheating. It confirms that serial correlation is related to cheating and not other factors, and suggests that centers that do not use donor cars are less likely to cheat in other ways. To see this, note that centers were grouped in Web Appendix Figure 3 according to a measure of serial correlation, which reflects the prevalence of cheating with donor cars. However, other types of cheating could also produce a discordance between car characteristics and emissions. Therefore, an additional implication of Web Appendix Figure 3 is that centers that are classified as having low cheating of the form of donor cars also have low prevalence of other forms of cheating.

The validity of this falsification test can be further verified by checking whether the negative slope in the graphs disappears when an environmental predictor of emissions, such as temperature, is used instead of a car attribute. The environmental determinant of emissions is always matched to the right emission readings despite the presence of corruption, hence the coefficient should not fall as corruption increases. A figure analogous to Web Appendix Figure 3 where I use temperature instead of model year (Web Appendix Figure 4) confirms this prediction.

Another validity check for this robustness test consists on holding age distribution constant. Because of potential self selection into cheating and non-cheating centers, vehicle characteristics may differ across centers with different levels of cheating (see Section 7 and Web Appendix Figure 5). If the effect of age on emissions is non-linear, we would expect the marginal effect of age to differ across centers as well. Web Appendix Figure 5 replicates Web Appendix Figure 3 holding the age distribution of vehicles constant across centers. The positive relationship between the coefficients and the level of cheating rules out that non-linear effects of age are driving the results of the robustness test.
Notes:

1. This figure tests 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 panel, 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 center. The horizontal axis describes each group of centers by the size of the p-value for the joint significance test of serial correlation. Each number corresponds to a different percentile in the p-value, with group 1 being the one with the lowest p-value and lowest evidence of cheating.
Notes:

1. This figure tests whether temperature at the time of the emission reading shows the same serial correlation patterns depicted in Online Appendix Figure 3. A positive or negative slope in the above graphs would suggest that test determinants other than cheating may constitute an alternative explanation to Online Appendix Figure 3.

3. For a more detailed explanation of how this figure was constructed see notes for Web Appendix Figure 3.
Online Appendix Figure 5: Robustness Test using Constant Age Distribution Across Cheating Groups

Notes:

1. This figure tests for whether the relationship between cheating and predictive power of age on emissions found in Online Appendix Figure 3 is attributable to differences in the age distribution of vehicles across centers. If the effects of age on emissions are non-linear and different centers have different age distributions of vehicles, Online Appendix Figure 3 could reflect a relationship between serial correlation and age distribution of vehicles. To test for this alternative explanation of Online Appendix Figure 3, I reweight the samples of vehicles from each cheating group (1-15) to keep the age distribution of vehicles constant. Using these reweighted samples, I replicate Online Appendix Figure 3; finding that the predictive power of car characteristics is still negatively correlated with the strength of the serial correlation. Hence, nonlinear effects in age are not responsible for Online Appendix Figure 3 (for a more detailed explanation of how this figure was constructed see notes for Online Appendix Figure 3).
Online Appendix 3: Existence and Uniqueness of Fixed Point for Expected Values

In order to prove the existence of a fixed point in the expected value equations

\[ \mathbb{E}V_{t}^{od} = \tau \left( k + \log \left( \exp \left( \frac{-\omega - c - \beta}{\tau} \right) + \exp \left( \frac{-\omega - (1 - P)\mathbb{E}V_{t+1}^{ev}}{\tau} \right) + \exp \left( \frac{-\delta f}{\tau} \right) \right) \right) \]  \hspace{1cm} (2)

\[ \mathbb{E}V_{t}^{ev} = \tau \left( k + \log \left( \exp \left( \frac{-\omega - \beta}{\tau} \right) + \exp \left( \frac{-\omega + (1 - P)\mathbb{E}V_{t+2}^{od}}{\tau} \right) + \exp \left( \frac{-\delta f}{\tau} \right) \right) \right) \]  \hspace{1cm} (3)

It is sufficient to prove that the above system of equations is a contraction. In order to prove uniqueness of the fixed point, we also need to show that there is a unique intersection point of the expected value equations. The following lines prove that the conditions for existence and uniqueness are satisfied. In addition, I will prove that in equilibrium, \( \mathbb{E}V_{t}^{od} < \mathbb{E}V_{t+1}^{ev} \).

**Proposition 1: Existence**

The vector of dynamic functions \( \mathbb{E}V_{t}^{od} = f_1(\mathbb{E}V_{t+1}^{ev}) \) and \( \mathbb{E}V_{t+1}^{ev} = f_2(\mathbb{E}V_{t+2}^{od}) \) has a fixed point.\(^4\)

**Proof:**

Existence can be proved by showing that the system of equations is a contraction in \( \mathbb{R}^2 \), i.e. that \( 0 < \frac{\partial \mathbb{E}V_{t}^{od}}{\partial \mathbb{E}V_{t+1}^{ev}} < 1 \) and \( 0 < \frac{\partial \mathbb{E}V_{t+1}^{ev}}{\partial \mathbb{E}V_{t+2}^{od}} < 1 \).

Note that

\[ \frac{\partial \mathbb{E}V_{t}^{od}}{\partial \mathbb{E}V_{t+1}^{ev}} = (1 - P) \frac{\exp \left( \frac{-\omega - c + (1 - P)\mathbb{E}V_{t+1}^{ev}}{\tau} \right)}{\exp \left( \frac{-\omega - c - \beta}{\tau} \right) + \exp \left( \frac{-\omega - c + (1 - P)\mathbb{E}V_{t+1}^{ev}}{\tau} \right) + \exp \left( \frac{-\delta f}{\tau} \right)} \]

and

\[ 0 < \frac{\partial \mathbb{E}V_{t}^{od}}{\partial \mathbb{E}V_{t+1}^{ev}} < (1 - P) < 1 \]

Similarly,

\[ \frac{\partial \mathbb{E}V_{t+1}^{ev}}{\partial \mathbb{E}V_{t+2}^{od}} = (1 - P) \frac{\exp \left( \frac{-\omega + (1 - P)\mathbb{E}V_{t+2}^{od}}{\tau} \right)}{\exp \left( \frac{-\omega - \beta}{\tau} \right) + \exp \left( \frac{-\omega + (1 - P)\mathbb{E}V_{t+2}^{od}}{\tau} \right) + \exp \left( \frac{-\delta f}{\tau} \right)} \]

and

\[ 0 < \frac{\partial \mathbb{E}V_{t+1}^{ev}}{\partial \mathbb{E}V_{t+2}^{od}} < (1 - P) < 1 \]

\(^4\)Note that \( f_1(.) \) and \( f_2(.) \) denote functions, while \( f \) with no subscripts denotes a known constant, the fine for non compliance.
Using the unitary metric (the sum of the absolute values of the differences between $E_{V_t}^{ev}$ and $E_{V_{t+1}^{ev}}$ and between $E_{V_t}^{od}$ and $E_{V_{t+1}^{od}}$) to measure the distance between two points in $\mathbb{R}^2$, these conditions are enough for existence of a fixed point, $(E_{V_t}^{ev}, E_{V_t}^{od})$.

□

**Proposition 2: Uniqueness**

The fixed point of the mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$, $E_{V_t}^{od} = f_1(E_{V_t}^{ev})$ and $E_{V_{t+1}^{ev}} = f_2(E_{V_{t+1}^{od}})$, is unique.

Proof:

We can solve for $E_{V_{t+1}^{ev}}$ in equation (2) to get

$$E_{V_{t+1}^{ev}} = f_1^{-1}(E_{V_t}^{od}) = \frac{\tau \log \left( \exp \left( \frac{E_{V_{t}}^{od} - k}{\tau} \right) - \exp \left( \frac{-\omega - c - \beta}{\tau} \right) - \exp \left( \frac{-\delta f}{\tau} \right) \right) + \omega + c}{1 - P}$$

(4)

The fixed point, or the point at which $E_{V_t}^{ev}$ and $E_{V_t}^{od}$ are constant, occurs when $E_{V_t}^{od} = E_{V_t}^{od}$, and $E_{V_{t+1}^{ev}} = E_{V_{t+1}^{ev}}$ for every $t$. Substituting $E_{V_{t+2}^{od}}$ and $E_{V_{t+1}^{od}}$ for $E_{V_t}^{od}$, and $E_{V_{t+1}^{ev}}$ for $E_{V_{t+1}}^{ev}$, in equations (4) and (3) yields a system of two equations and two unknowns:

$$E_{V_t}^{ev} = f_1^{-1}(E_{V_t}^{od}) = \frac{\tau \log \left( \exp \left( \frac{E_{V_{t+1}}^{od} - k}{\tau} \right) - \exp \left( \frac{-\omega - c - \beta}{\tau} \right) - \exp \left( \frac{-\delta f}{\tau} \right) \right) + \omega + c}{1 - P}$$

(5)

$$E_{V_t}^{ev} = f_2(E_{V_t}^{od}) = \tau \left( k + \log \left( \exp \left( \frac{-\omega - \beta}{\tau} \right) + \exp \left( \frac{-\omega + (1 - P) E_{V_t}^{od}}{\tau} \right) + \exp \left( \frac{-\delta f}{\tau} \right) \right) \right)$$

(6)

that has a solution at the intersection of these equations. There will be a unique fixed point if equations (5) and (6) are both contractions and if they intersect only once.

Uniqueness can be proved by showing that

$$\frac{\partial f_1^{-1}(E_{V_t}^{od})}{\partial E_{V_t}^{od}} > \frac{\partial f_2(E_{V_t}^{od})}{\partial E_{V_t}^{od}}$$

(7)

for all values of $E_{V_t}^{od}$.

First, note that

$$\frac{\partial f_1^{-1}(E_{V_t}^{od})}{\partial E_{V_t}^{od}} = \frac{1}{1 - P} \frac{\exp \left( \frac{E_{V_t}^{od} - k}{\tau} \right) - \exp \left( \frac{-\omega - c - \beta}{\tau} \right) - \exp \left( \frac{-\delta f}{\tau} \right)}{\exp \left( \frac{E_{V_t}^{od} - k}{\tau} \right) - \exp \left( \frac{-\omega - c - \beta}{\tau} \right) - \exp \left( \frac{-\delta f}{\tau} \right)}$$

(8)

and
\[
\frac{\partial f_2(E_{\text{Vod}})}{\partial E_{\text{Vod}}} = (1 - P) \frac{\exp \left( -\omega + (1 - P)E_{\text{Vod}}/\tau \right)}{\exp \left( -\omega + (1 - P)E_{\text{Vod}}/\tau \right) + \exp \left( -\omega - \beta \right) + \exp \left( -\delta f \right)}
\] (9)

Hence, condition (7) translates into

\[
\frac{1}{1 - P} \frac{\exp \left( \frac{E_{\text{Vod}}}{\tau} - k \right)}{\exp \left( \frac{E_{\text{Vod}}}{\tau} - k \right) - \exp \left( -\omega - c - \beta \right) - \exp \left( -\delta f \right)} > (1 - P) \frac{\exp \left( -\omega + (1 - P)E_{\text{Vod}}/\tau \right)}{\exp \left( -\omega + (1 - P)E_{\text{Vod}}/\tau \right) + \exp \left( -\omega - \beta \right) + \exp \left( -\delta f \right)},
\]

which can be rewritten as

\[
\frac{\exp \left( \frac{E_{\text{Vod}}}{\tau} - k \right)}{\exp \left( \frac{1 - P)E_{\text{Vod}}}{\tau} - \frac{\omega}{\tau} \right)} > (1 - P)^2 \frac{\exp \left( \frac{E_{\text{Vod}}}{\tau} - k \right) - \exp \left( -\omega - c - \beta \right) - \exp \left( -\delta f \right)}{\exp \left( \frac{1 - P)E_{\text{Vod}}}{\tau} - \frac{\omega}{\tau} \right) + \exp \left( -\omega - \beta \right) + \exp \left( -\delta f \right)}
\] (10)

Inspection of (10) demonstrates that the inequality, necessary for uniqueness, holds: the numerator of the LHS is larger than the numerator of the RHS, while the denominator of the LHS is smaller than the denominator on the RHS. In addition, the RHS is multiplied by the square of a fraction.

\[\square\]

**Proposition 3:** \(E_{\text{Vod}} < E_{\text{Vev}}\)

**Proof:**

We proceed by contradiction. Assume \(E_{\text{Vod}} - E_{\text{Vev}} = \epsilon \geq 0\), then

\[
\frac{\exp \left( E_{\text{Vod}}/\tau \right)}{\exp \left( E_{\text{Vev}}/\tau \right)} \geq 1
\]

which can also be expressed as

\[
\frac{\exp(k) \left( \exp \left( -\omega - c - \beta \right) + \exp \left( -\omega - c + (1 - P)E_{\text{Vev}}/\tau \right) + \exp \left( -\delta f \right) \right)}{\exp(k) \left( \exp \left( -\omega - \beta \right) + \exp \left( -\omega + (1 - P)E_{\text{Vod}} + (1 - P)\epsilon /\tau \right) + \exp \left( -\delta f \right) \right)} \geq 1
\] (11)

but (11) is false for \(c > 0\). Then, it must be that \(E_{\text{Vod}} < E_{\text{Vev}}\).

\[\square\]
Online Appendix 4: Identification of Econometric Model

Magnac and Thesmar (2002) show that the generic class of discrete choice dynamic models to which this model belongs is non-parametrically unidentified. However, they also provide useful insights on what parametric and non-parametric restrictions can deliver partial identification. The identification proof below uses their framework to establish which structural objects are semiparametrically identified and to prove parametric identification. In addition, Heckman and Navarro (2007) show conditions under which dynamic structural models are semiparametrically identified in the case of finite horizon models. My infinite horizon assumption prevents me from non-parametrically identifying the distribution of the random utility shocks in each period, as I cannot observe payoffs in the final period. However, the variation in the probability of passing the test provided by different car characteristics is akin to the variation in school costs that allow Heckman and Navarro to identify unobservable period by period payoffs without confounding selection.

Basic Model Assumptions

When faced with the smog-check requirement, the car owner chooses between three alternatives, \( i = \{X, B, A\} \), where \( X \) represents missing the test, \( B \) represents bribing, and \( A \) represents lawfully carrying out the emissions test. There are two possible states characterized by whether the testing round is odd or even. I will denote the state variables in odd states as \( s \) and in even states as \( s' \).

Following Magnac and Thesmar (2002), I start by assuming additive separability between the deterministic and random components of the instantaneous utility functions (AS), perfect expectations on the law of motion from odd to even rounds and vice versa (i.e. agents have full knowledge on their probability of passing the test) (PS), and conditional independence of the random utility components across testing rounds.

Assumption (AS) allows me to write the instantaneous utility function for choice \( i \) in odd and even testing rounds as

\[
    u_i(s, \tilde{\varepsilon}) = u_i^*(s) + \varepsilon_i, \quad \text{and} \quad u_i(s', \tilde{\varepsilon}') = u_i^*(s') + \varepsilon_i',
\]

respectively, where \( \tilde{\varepsilon} = (\varepsilon_X, \varepsilon_B, \varepsilon_A) \) and \( \tilde{\varepsilon}' = (\varepsilon_X', \varepsilon_B', \varepsilon_A') \). I assume that the random utility components of the current state can be observed by the individual in the current state but not by the econometrician. Random utility components in future states cannot be observed by either.

Given these assumptions, the Bellman Equation for choice \( i \) and in an odd testing round can thus be written as

\[
    v_i(s) = u_i^*(s) + \varepsilon_i + \mathbb{E} \left[ \mathbb{E}_{s'} \left( \max_j (v_j(s')) \mid s' \right) \right], \quad (12)
\]

where the inner expectation in (12) is taken with respect to the joint distribution of random error terms in an even period, while the outer expectation is taken with respect to the transition probabilities from odd to
even periods, i.e. the probability of passing the test. Equivalently, the Bellman equation for choice $i$ in an even testing round can be written as

$$v_i(s') = u_i^*(s') + \varepsilon_i + \mathbb{E}\left(\max_j (v_j(s)) \mid s\right)$$

**Structural Model Restrictions**

Assume I have two state variables such that $s = (s_1, s_2)$ and $s' = (s'_1, s'_2)$, where $s_1$ represents money and $s_2$ represents time. Money and time requirements vary across choices $i = \{X, B, A\}$. Hence, I can modify the instant utility function notation such that $u_i^*(s) = \delta^{1(i=X)}u^*(s_{i,1}, s_{i,2})$, where $\delta$ is the product of the six-month discount rate and the probability of getting caught. The smog-check requirement structure offers some information about money and time requirements for each decision:

(i) Missing the test is not time consuming and, by definition, will result in no additional testing rounds. Hence,

$$\mathbb{E}_{\tilde{\mathbb{E}}} (v_X(s)) = \mathbb{E}_{\tilde{\mathbb{E}'}'} (v_X(s')) = u_X^*(s) = \delta u^*(-f, 0)$$

where $f$ is the (known) fine for not complying with the smog-check, which will be paid in the next smog-check period with a probability less than or equal to one.

(ii) Bribes are competitive and bribing will guarantee passing the test, therefore leading to no additional testing rounds. Hence,

$$\mathbb{E}_{\tilde{\mathbb{E}}} (v_B(s)) = u_B^*(s) = u^*(-\beta - c, -\omega), \text{ and}$$

$$\mathbb{E}_{\tilde{\mathbb{E}'}'} (v_B(s')) = u_B^*(s') = u^*(-\beta, -\omega),$$

where $c$ is the cost of odd tests and $\omega$ is the time it takes to perform a smog-check.

Restrictions (i) and (ii) are akin to absorbing state restrictions, and restriction (i) doubles as an exclusion restriction since the utility level associated with $X$ is constant across rounds. The first part of assumption (i) is straightforward and implies that missing the test carries no additional costs except the expected cost of being caught. The second part follows from the fee structure imposed by the local environmental authority.\(^5\)

Assumption (ii), perfect competition, implies that there will be an equilibrium bribe. Competition across centers is likely strong. Even the most remote area (the southeast area of the city) has about 8 centers

---

\(^5\)It is likely the case that centers operate according to the rules issued by the Ministry of Environment with respect to the fee structure. If they were to charge for retests, for instance, motorists would have a strong incentive to report misconduct. Evidence from newspapers confirm this: a newspaper article mentions the suspension of a smog-check center after a motorist reported that they were charging for the retest (SINEMBARGO, May 24, 2012 http://www.sinembargo.mx/24-05-2012/243837).
within a 9 km radius. An individual who lives in the middle of the city could choose among 30 to 40 centers within this range. In addition, the assumption of constant bribes across rounds implies that centers cannot charge higher bribes after motorists fail once.\(^6\)

(iii) The probability of passing the test is a function of car-specific characteristics. Assumption (iii) also implies that this probability is the same across testing rounds, i.e. \(P(s) = P(s') = P\).

Taking the test lawfully may lead to failing the test with probability \((1 - P)\). Failing the test on an odd testing round will lead to an even decision round and vice versa. Therefore,

\[
\mathbb{E}_{\tilde{\epsilon}}(v_A(s)) = u^*(-c, -\omega) + (1 - P)\mathbb{E}_{\tilde{\epsilon}'} \left( \max_j(v_j(s')) \mid s' \right)
\]

\[
\mathbb{E}_{\tilde{\epsilon}'}(v_A(s')) = u^*(0, -\omega) + (1 - P)\mathbb{E}_{\tilde{\epsilon}} \left( \max_j(v_j(s)) \mid s \right)
\]

Assumptions (i)-(iii) results in stationarity of the dynamic model: individuals switch back and forth between odd and even states, but all odd testing rounds are indistinguishable from each other and all even rounds are indistinguishable from each other.

Assumptions (iii) and (iv) allow us to use predicted probabilities of passing as transition probabilities, as well as in computing the continuation value in individual utilities. These assumptions rule out learning (i.e. individuals do not update their probability of passing the test upon failing) and the possibility that individuals may have more information than the econometrician regarding their probability of passing.\(^7\)

Restrictions (i), (ii) and (iii) narrow down the list of objects to be identified to:

\[
b = \{u^*(-c, -\omega), u^*(0, -\omega), u^*(-\beta - c, -\omega), u^*(-\beta, -\omega), \delta u^*(-f, 0),
F_{\tilde{\epsilon},\tilde{\epsilon}'}, P, \mathbb{E}_{\tilde{\epsilon}}(\max_j(v_j(s)) \mid s), \mathbb{E}_{\tilde{\epsilon}'}(\max_j(v_j(s')) \mid s')\}
\]

Data

Define \(I = \{X, B, A\}\) and \(i \in I\) (\(i' \in I\)) a choice made in an odd (even) testing round. The standard discrete dynamic model assumes that all components in the following function are observed empirically (Magnac

---

\(^6\)A center manager would be tempted to threaten with demanding a higher bribe in the second test in order to induce motorists to pay the bribe in the first attempt and save the center’s resources in the retest. However, serial correlation is just as strong in retests as it is in first tests, suggesting no deterrence effect of this sort.

\(^7\)The fact that the within-vehicle variance in emissions is large implies that the amount of learning that occurs between the first test and the retest is rather small. Given that the model is estimated using information on the first test and retest exclusively, learning is unlikely to play an important role in vehicle owner behavior. To confirm this, I use simulations to explore the impact that ignoring learning would have on the model estimates. Findings suggest that some model parameters, such as the bribe, would be slightly downward biased and that model predictions on bribing rates would be largely unaffected. A similar procedure can be used to explore the effect of private information on model estimates. In this case, findings suggest that even large amounts of private information will have little impact on model estimates and predictions.
and Thesmar, 2002)

\[ \forall (i, i') \in I \times I, \forall (s, s') \in S \times S', \]

\[ \Pr(i', s', i|s) = \Pr(i'|s') \cdot \Pr(s'|s, i) \cdot \Pr(i|s) \]

Under model restrictions (i)-(iii),

\[ \Pr(s'|s, i) = \begin{cases} 
0 & \text{if } i = X \\
0 & \text{if } i = B \\
1 - P & \text{if } i = A 
\end{cases} \]

Also, note that bribing decisions are unobserved. Therefore only the left side of the following probabilities is empirically observable

\[ g_3(P) = \Pr(X|s) \]
\[ g_1(P) = \Pr(A|s)(1 - P) \]
\[ g_4(P) = \frac{\Pr(X|s')(1 - P) \Pr(A|s)}{\Pr(A|s)(1 - P)} = \Pr(X|s') \]
\[ g_2(P) = \frac{\Pr(A|s')(1 - P)^2 \Pr(A|s)}{\Pr(A|s)(1 - P)} = \Pr(A|s')(1 - P) \]

These represent the probability of missing the test requirement before the first test round, the probability of failing the first test round, the probability of missing the test requirement after failing the first test round, and the probability of failing the second test round. These moments are a function of \( P \), the probability of passing the test, which is treated as known and varies across vehicles. In practice, I use \( \hat{P} \), which is predicted from a first stage regression of the probability of passing on car characteristics. Hence, there are as many values of \( P \) as a car type.

Equations (13) can therefore be estimated for each car type with conditional sample moments, \( \tilde{g}_1(P) \), \( \tilde{g}_2(P) \), \( \tilde{g}_3(P) \), and \( \tilde{g}_4(P) \), and constitute the basic elements of identification.

**Additional Restrictions**

Magnac and Thesmar (2002)’s non-identification result would apply to this model even if the probability of bribing and the transition probability, \( P \), were observed from the data. However, the same paper offers some insights into which further restrictions can provide identification of the instantaneous utility functions given a parametric assumption of the distribution of the random utility components and observable transition probabilities.
(iv) Because the transition probability, $P$, is unobservable from the choice and outcome data, these results cannot be applied without an exogenous source of identification. Hence, this model relies on the identification of $P$ as a function of vehicle characteristics from a reduced form model using data from the non-cheating centers.

(v) Furthermore, we assume that the random utility components have an extreme value distribution with mean zero and scale parameter $\tau$, i.e. $F_{\tilde{\varepsilon}, \tilde{\varepsilon}'}$ is known up to a single parameter. This restriction replaces the CI assumption.

Restrictions (i)-(v) allow us to invert (13) as a function of the remaining structural objects plus $\tau$:

$$b' = \{ u^*(c, -\omega), u^*(0, -\omega), u^*(-\beta, -\omega), u^*(-\beta - c, -\omega), \delta u^*(-f, 0),$$

$$\tau, \mathbb{E}_{\tilde{\varepsilon}} (\max_j (v_j(s)) | s), \mathbb{E}_{\tilde{\varepsilon}'} (\max_j (v_j(s')) | s') \}$$

Identification

The identification result boils down to the inversion of (13) given assumptions (AS) and (i)-(v). Note that assumption (v) implies that $F_{\tilde{\varepsilon}} = F_{\tilde{\varepsilon}'}$. Forth, I denote the expectation operators $\mathbb{E}_{\tilde{\varepsilon}}$ and $\mathbb{E}_{\tilde{\varepsilon}'}$ as $\mathbb{E}$. These probabilities can now be rewritten as a function of the structural objects to be identified:

- The probability that an individual fails in the first test attempt, $Pr(A|s)(1 - P)$, or

$$g_1(P) = \frac{\exp \left( \frac{\mathbb{E}_{v_A}(P,s)}{\tau} \right)}{\exp \left( \frac{\mathbb{E}_{v_A}(P,s)}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_B}(P,s)}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_X}(P,s)}{\tau} \right)}(1 - P)$$

- The probability that an individual fails in the second testing round conditional on failing the first time around, $Pr(A|s')(1 - P)$, or

$$g_2(P) = \frac{\exp \left( \frac{\mathbb{E}_{v_A}(P,s')}{\tau} \right)}{\exp \left( \frac{\mathbb{E}_{v_A}(P,s')}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_B}(P,s')}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_X}(P,s')}{\tau} \right)}(1 - P)$$

- The probability that an individual misses the test in the first testing round, $Pr(X|s)$, or

$$g_3(P) = \frac{\exp \left( \frac{\mathbb{E}_{v_X}(P,s)}{\tau} \right)}{\exp \left( \frac{\mathbb{E}_{v_A}(P,s)}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_B}(P,s)}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_X}(P,s)}{\tau} \right)}$$

- The probability that an individual misses the test in the second testing round conditional on failing the first time around, $Pr(X|s')$, or

$$g_4(P) = \frac{\exp \left( \frac{\mathbb{E}_{v_X}(P,s')}{\tau} \right)}{\exp \left( \frac{\mathbb{E}_{v_A}(P,s')}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_B}(P,s')}{\tau} \right) + \exp \left( \frac{\mathbb{E}_{v_X}(P,s')}{\tau} \right)}$$
Results 1 through 3 prove identification of structural objects $\frac{u^*(-c, -\omega)}{\tau}$, $\frac{u^*(0, -\omega)}{\tau}$, $\frac{u^*(-\beta - c, -\omega)}{\tau}$, $\frac{u^*(-\beta, -\omega)}{\tau}$, and $\frac{\delta u^*(-f, 0)}{\tau}$; while Corollaries 1.1 and 1.3, stated below, prove identification of parameters $\tau, \omega, \beta$, and $\delta$ when I assume linearity of the utility function.

**Result 1** The scaled difference between two instantaneous utilities in consecutive test rounds can be written as a function of $g_1, g_2, g_3$, and $g_4$.

\[
\frac{u^*(0, -\omega) - u^*(-c, -\omega)}{\tau} = \log \left( \frac{g_3(P)}{g_4(P)} \right) (2 - P) + \log \left( \frac{g_2(P)}{g_1(P)} \right)
\]

**Proof**

**Part A** Given the assumption (v) of the distribution of the unobservable utility components, then we can rewrite the expected value functions as

\[
\mathbb{E}(\max(v_j(P, s))) = \tau \left( k + \log \left( \frac{\mathbb{E}v_A(P, s)}{\tau} \right) + \frac{\mathbb{E}v_B(P, s)}{\tau} + \frac{\mathbb{E}v_X(P, s)}{\tau} \right)
\]

(18)

\[
\mathbb{E}(\max(v_j(P, s'))) = \tau \left( k + \log \left( \frac{\mathbb{E}v_A(P, s')}{\tau} \right) + \frac{\mathbb{E}v_B(P, s')}{\tau} + \frac{\mathbb{E}v_X(P, s')}{\tau} \right)
\]

(19)

where $k$ is Euler’s constant.

Using (18) and (19), the denominator of $g_1$ and $g_3$ can be rewritten as $\exp \left( \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s)) - k) \right)$ and the denominator of $g_2$ and $g_4$ can be rewritten as $\exp \left( \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s')) - k) \right)$, where $k$ is Euler’s constant.

Hence, we can write $g_3$ and $g_4$ as

\[
g_3(P) = \frac{\exp \left( \frac{\mathbb{E}v_X(P, s)}{\tau} \right)}{\exp \left( \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s)) - k) \right)}
\]

(20)

\[
g_4(P) = \frac{\exp \left( \frac{\mathbb{E}v_X(P, s')}{\tau} \right)}{\exp \left( \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s')) - k) \right)}
\]

(21)

Note that assumption (i) implies that the decision to miss the test yields the same level of utility regardless of the testing round. More precisely, missing the test is equally costly regardless of the state variable, so $\mathbb{E}v_X(P, s') = \mathbb{E}v_X(P, s) = \mathbb{E}v_X(P)$. Also, because individuals who decide to miss the test in a given testing period do not experience further testing rounds, $\mathbb{E}v_X(P) = \delta u^*(-f, 0)$, i.e. it doesn’t depend on the transition probability, $P$.

Substituting these restrictions in (20) and (21), I obtain
\[ \log(g_3(P)) = \left( \frac{\delta u^*(f, 0)}{\tau} \right) - \frac{\mathbb{E}(\max(v_j(P, s))}{\tau} + k \]  

(22)

\[ \log(g_4(P)) = \left( \frac{\delta u(-f, 0)}{\tau} \right) - \frac{\mathbb{E}(\max(v_j(P, s'))}{\tau} + k \]  

(23)

Subtracting (22) from (23) yields

\[ \log \left( \frac{g_3(P)}{g_4(P)} \right) = \frac{\mathbb{E}(\max(v_j(P, s'))) - \mathbb{E}(\max(v_j(P, s))}{\tau} \]  

(24)

**Part B** Using (18) and (19) I can rewrite (14) and (15) as

\[ \log \left( \frac{g_1(P, s)}{1 - P} \right) = \frac{\mathbb{E}v_A(P, s)}{\tau} - \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s))) + k \]  

\[ \log \left( \frac{g_2(P, s')}{1 - P} \right) = \frac{\mathbb{E}v_A(P, s')}{\tau} - \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s'))) + k \]  

(25)

Solving for \( \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s))) \) and \( \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s'))) \) and taking subtracting yields

\[ \frac{\mathbb{E}(\max(v_j(P, s'))) - \mathbb{E}(\max(v_j(P, s))}{\tau} = \frac{\mathbb{E}v_A(P, s') - \mathbb{E}v_A(P, s)}{\tau} - \log \left( \frac{g_2(P)}{g_1(P)} \right) \]  

(26)

Substituting (24) in (26) I get

\[ \log \left( \frac{g_3(P)}{g_4(P)} \right) + \log \left( \frac{g_2(P)}{g_1(P)} \right) = \frac{\mathbb{E}v_A(P, s') - \mathbb{E}v_A(P, s)}{\tau} \]  

(27)

Using stationarity of the Bellman equations and assumption (iii), the difference between the \( v_A \) functions in two periods is given by

\[ \mathbb{E}v_A(P, s') - \mathbb{E}v_A(P, s) = u^*(0, -\omega) - u^*(-c, -\omega) - (1 - P) (\mathbb{E}(\max(v_j(P, s'))) - \mathbb{E}(\max(v_j(P, s)))) \]  

(28)

Substituting (24) into (28) and (28) into (27) we get

\[ \log \left( \frac{g_3(P)}{g_4(P)} \right) + \log \left( \frac{g_2(P)}{g_1(P)} \right) = -(1 - P) \log \left( \frac{g_3(P)}{g_4(P)} \right) + \frac{u^*(0, -\omega) - u^*(-c, -\omega)}{\tau} \]
Solving for the difference between the utility functions in two periods, we get

\[
\log \left( \frac{g_3(P)}{g_4(P)} \right) (2 - P) + \log \left( \frac{g_2(P)}{g_1(P)} \right) = \frac{u^*(0, -\omega) - u^*(-c, -\omega)}{\tau}
\]

(29)

□

**Corollary 1.1** If we assume that utility functions are linear in money and time, i.e. \( u^*(s_1, s_2) = s_1 + s_2 \), Result 1 implies that \( \tau \) is identified.

**Proof**

Adding linearity to the utility function lets me fully exploit the fact that the test is costly in odd periods and costless in even periods. Substituting the linear instantaneous utility functions, \(-\omega\) and \(-c - \omega\), in (29) yields

\[
\log \left( \frac{g_3(P)}{g_4(P)} \right) (2 - P) + \log \left( \frac{g_2(P)}{g_1(P)} \right) = \frac{c}{\tau}
\]

Since \( c \) is known, this equation allows us to recover \( \tau \).

□

**Result 2** The instantaneous utility from testing lawfully in the second test round and the utility of missing the test are identified from a system of linear equations with two unknowns involving a function of \( g_2(P), g_3(P) \) and \( g_4(P) \) and \( P \), all of which vary across car types and are empirically observable.

**Proof**

Substituting the Bellman equation in (25) and adding restriction (iii) I get

\[
\log \left( \frac{g_2(P)}{1 - P} \right) = u^*(0, -\omega) + (1 - P) \mathbb{E}(\max(v_j(P, s))) - \frac{1}{\tau} \mathbb{E}(\max(v_j(P, s')) + k)
\]

(30)

Substituting (22) and (23) in (30) I get

\[
\log \left( \frac{g_2(P)}{1 - P} \right) + (1 - P) \log(g_3(P)) - \log(g_4(P)) = \left( \frac{u^*(0, -\omega)}{\tau} + k \right) - P \left( \frac{\delta u^*(-f, 0)}{\tau} + k \right)
\]

(31)

Note that equation (31) describes a system of linear equations with two unknowns, \( u^*(0, -\omega) \) and \( \delta u^*(-f, 0) \).
There are as many equations as different car types, since \( P \) is known and varies with car characteristics. Hence, \( u^*(0, -\omega) \) and \( \delta u^*(-f, 0) \) are overidentified.

\[ \square \]

**Result 3**

(a) Results 1 and 2 identify the utility of testing lawfully in the first testing round, \( \frac{u^*(-c, -\omega)}{\tau} \).

(b) \( \frac{E(\max(v_j(P,s))}{\tau} \) and \( \frac{E(\max(v_j(P,s'))}{\tau} \) are identified.

(c) \( \frac{u^*(-\beta, -\omega)}{\tau} \) and \( \frac{u^*(-\beta-c, -\omega)}{\tau} \) are also identified.

**Proof**

(a) Subtract \( \frac{u^*(0, -\omega)}{\tau} \) from the left and right sides of (29) and multiply by \( -1 \).

(b) Subtract \( \frac{\delta u^*(-f, 0)}{\tau} \) from the left and right sides of (22) and (23) and solve for \( \frac{E(\max(v_j(P,s))}{\tau} \) and \( \frac{E(\max(v_j(P,s'))}{\tau} \).

(c) We can rewrite (19) as

\[
\exp\left(\frac{E(\max(v_j(P,s'))}{\tau} - k\right) = \exp\left(\frac{E v_A(P,s')}{\tau}\right) + \exp\left(\frac{E v_B(P,s')}{\tau}\right) + \exp\left(\frac{E v_X(P,s')}{\tau}\right),
\]

(32)

Assumptions (i)-(iii) allow us to write the above expression as

\[
\exp\left(\frac{E(\max(v_j(P,s'))}{\tau} - k\right) = \exp\left(\frac{u^*(0, -\omega)+(1-P)E(\max(v_j(P,s))}{\tau}\right) + \exp\left(\frac{u^*(-\beta, -\omega)}{\tau}\right) + \exp\left(\frac{\delta u^*(-f, 0)}{\tau}\right).
\]

(33)

Note that after substituting previous identification results, the only remaining unknown in this equation is \( \frac{u^*(-\beta, -\omega)}{\tau} \). Using an analogous expression to (32) for (18) allows us to identify \( \frac{u^*(-\beta-c, -\omega)}{\tau} \).

\[ \square \]

**Corollary 3.1** If we assume that utility functions are linear in money and time, i.e. \( u^*(s_1, s_2) = s_1 + s_2 \), Corollary 1.1 and Results 1 and 2 imply that \( \omega, \beta, \) and \( \delta \) are identified.

**Proof**

Linearity implies that \( \frac{u^*(-c, -\omega)}{\tau} = \frac{-c-\omega}{\tau} \), which identifies \( \omega \), since \( c \) is known and \( \tau \) is identified from Corollary 1.1. Linearity also implies that \( \frac{u^*(-\beta, -\omega)}{\tau} = \frac{-\beta-\omega}{\tau} \), which identifies \( \beta \), since \( \omega \) and \( \tau \) are identified.

Finally, linearity implies \( \frac{\delta u^*(-f, 0)}{\tau} = \frac{-\delta f}{\tau} \), which identifies \( \delta \), since \( \tau \) is identified and \( f \) is known.

\[ \square \]
Online Appendix 5: Policy Outcome Calculations

The benefits from each policy in Table 5 are calculated by simulating owners’ decisions under each policy scenario. The share of individuals who perform maintenance corresponds to the share of individuals whose expected cost of the program is lower if they perform maintenance than if they do not. Appendix Table 3 details every step of the calculation using Policy 5 from Table 5 as an example. My assumptions for each of the calculations shown in Appendix Table 3 is explained below.

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

Second, I use the fuel-based methodology for emission inventories proposed by Singer and Harley (1996), and applied by Schifter et al. (2005) for Mexico City, to convert emission concentrations at vehicles’ exhaust, which is recorded by the smog check readings in parts per million (ppm), to total emission tons released into the environment. This methodology uses fuel economy, miles driven and pollutant densities. The results of this calculation are given in Column 1 of Table 6.

Third, I follow Small and Kazimi (1995) in attributing ambient concentrations of air pollutants to vehicle emissions in tons. Health costs from pollution have been estimated for ambient concentrations of air pollutants such as particulate matter ($PM_{10}$) and ozone ($O_3$), but have not been estimated for $NO_x$ and $HC$, which are the pollutants measured in the smog-check readings. However, $PM_{10}$ and $O_3$ are partially formed from $NO_x$ and $HC$, so following Small and Kazimi (1995), I use the elasticities of conversion from the literature (e.g. Charlson and Wigley 1994) to estimate the concentration of air pollutants due to emissions in tons of primary pollutants. Also, I follow Song et al. (2010) to link $O_3$ to $HC$ emissions. They estimate that a one percent increase in volatile organic compounds ($VOC$) emissions ($HC$ are a sub-class of $VOC$) results in about a half percent increase in average $O_3$ ambient concentration. These calculations are in Column 2 of Table 6.\footnote{In their work within the MILAGRO project Song et al. (2010), they find that, contrary to what was previously believed (e.g. Molina and Molina, 2002), ozone formation in Mexico City is VOC-sensitive.}

Fourth, I calculate the health costs of emissions by summing up mortality costs from the different pollutants. I use Schwartz (1994) estimates of mortality from $PM_{10}$ concentration and Bell et al. (2004) estimates of $O_3$-related mortality. In the case of $CO$, I use the direct health impacts on infant mortality from Neidell and Currie (2005). To calculate the mortality costs in dollars, I use the value of statistical life used by Molina.
and Molina (2002) for Mexico City. The lives-per-ton calculations implied by these numbers are in Column 3 of Table 6.

Column 4 shows the deaths per ton of vehicle emissions and is the product of columns 2 and 3. Column 5 shows the cost in dollars of a ton of pollutant and is the product of column 4 and the statistical value of life, 650 thousand USD (Molina and Molina 2002). Column 6 is the product of columns 1 and 4. Finally, column 7 is the product of column 6 and the statistical value of life. The last two columns present the benefits of the set of policies considered in number of lives per year and in 2003 U.S. dollars respectively.
<table>
<thead>
<tr>
<th>Car</th>
<th>Coefficient</th>
<th>Interaction</th>
<th>p-value</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chrysler</td>
<td>0.49</td>
<td>log(KM)</td>
<td>-0.035</td>
<td>[1.41]</td>
</tr>
<tr>
<td>Ford</td>
<td>2.125</td>
<td>Chrysler*log(KM)</td>
<td>0.007</td>
<td>[5.79]**</td>
</tr>
<tr>
<td>GM</td>
<td>0.461</td>
<td>Ford*log(KM)</td>
<td>-0.014</td>
<td>[1.28]</td>
</tr>
<tr>
<td>Nissan</td>
<td>-0.504</td>
<td>GM*log(KM)</td>
<td>0.004</td>
<td>[1.56]</td>
</tr>
<tr>
<td>VW</td>
<td>-0.047</td>
<td>Nissan*log(KM)</td>
<td>0.05</td>
<td>[0.16]</td>
</tr>
<tr>
<td>Chrysler*Age</td>
<td>-0.056</td>
<td>VW*log(KM)</td>
<td>0.027</td>
<td>[2.72]**</td>
</tr>
<tr>
<td>Ford*Age</td>
<td>-0.147</td>
<td>Monday</td>
<td>-0.011</td>
<td>[6.87]**</td>
</tr>
<tr>
<td>GM*Age</td>
<td>-0.052</td>
<td>Tuesday</td>
<td>-0.01</td>
<td>[2.73]**</td>
</tr>
<tr>
<td>Nissan*Age</td>
<td>0.012</td>
<td>Wednesday</td>
<td>-0.017</td>
<td>[0.62]</td>
</tr>
<tr>
<td>VW*Age</td>
<td>-0.014</td>
<td>Thursday</td>
<td>0.005</td>
<td>[0.77]</td>
</tr>
<tr>
<td>Chrysler*Age2</td>
<td>0.001</td>
<td>Friday</td>
<td>0.006</td>
<td>[2.35]*</td>
</tr>
<tr>
<td>Ford*Age2</td>
<td>0.003</td>
<td>Saturday</td>
<td>0.013</td>
<td>[5.59]**</td>
</tr>
<tr>
<td>GM*Age2</td>
<td>0.001</td>
<td>Aug</td>
<td>0.066</td>
<td>[2.80]**</td>
</tr>
<tr>
<td>Nissan*Age2</td>
<td>-0.001</td>
<td>Sep</td>
<td>-0.022</td>
<td>[1.65]</td>
</tr>
<tr>
<td>VW*Age2</td>
<td>0.000</td>
<td>Oct</td>
<td>-0.184</td>
<td>[0.36]</td>
</tr>
<tr>
<td>Public</td>
<td>1.186</td>
<td>Nov</td>
<td>-0.128</td>
<td>[6.06]**</td>
</tr>
</tbody>
</table>

Notes: ***p < 0.01, **p < 0.05, *p < 0.1.
<table>
<thead>
<tr>
<th>Feature</th>
<th>Coefficient</th>
<th>SE</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility</td>
<td>0.124</td>
<td>0.087</td>
<td>[1.00]</td>
</tr>
<tr>
<td>Age</td>
<td>0.077</td>
<td>0.279</td>
<td>[4.56]**</td>
</tr>
<tr>
<td>Age2</td>
<td>-0.002</td>
<td>0.122</td>
<td>[4.25]**</td>
</tr>
<tr>
<td>Age*Transport</td>
<td>-0.121</td>
<td>0.27</td>
<td>[16.04]**</td>
</tr>
<tr>
<td>Age*Corporate</td>
<td>-0.017</td>
<td>0.135</td>
<td>[2.13]*</td>
</tr>
<tr>
<td>Temperature</td>
<td>0.008</td>
<td>0.119</td>
<td>[4.69]**</td>
</tr>
<tr>
<td>Model year &gt; 1991</td>
<td>0.119</td>
<td>0.005</td>
<td>[5.08]**</td>
</tr>
</tbody>
</table>

Notes:
2. Sample used for the estimation included 47,940 observations, which correspond to all tests to non-exemptible vehicles in the low cheating sample (see Section 6.1).
<table>
<thead>
<tr>
<th></th>
<th>Non-Cheaters</th>
<th>Cheaters</th>
<th>Difference</th>
<th>Non-Cheaters</th>
<th>Cheaters</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1st Decile)</td>
<td>(2nd to 10th Deciles)</td>
<td></td>
<td>(1st Decile)</td>
<td>(2nd to 10th Deciles)</td>
<td></td>
</tr>
<tr>
<td>Model-year</td>
<td>1997.35</td>
<td>1994.82</td>
<td>-2.533</td>
<td>84390</td>
<td>110143</td>
<td>25754</td>
</tr>
<tr>
<td>Kilometers</td>
<td>[6.558]</td>
<td>[7.839]</td>
<td>[126.91]**</td>
<td>[117413.6]</td>
<td>[152334.6]</td>
<td>[66.80]**</td>
</tr>
<tr>
<td>Size</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VW Sedan and Chevy</td>
<td>0.394</td>
<td>0.348</td>
<td>-0.046</td>
<td>0.310</td>
<td>0.341</td>
<td>-0.031</td>
</tr>
<tr>
<td></td>
<td>[0.489]</td>
<td>[0.476]</td>
<td>[37.32]**</td>
<td>[0.281]</td>
<td>[0.265]</td>
<td>[14.99]**</td>
</tr>
<tr>
<td>Mini-Compact</td>
<td>0.086</td>
<td>0.076</td>
<td>-0.010</td>
<td>0.030</td>
<td>0.070</td>
<td>0.040</td>
</tr>
<tr>
<td></td>
<td>[0.281]</td>
<td>[0.265]</td>
<td>[14.99]**</td>
<td>[0.179]</td>
<td>[0.163]</td>
<td>[13.68]**</td>
</tr>
<tr>
<td>Compart-Medium</td>
<td>0.033</td>
<td>0.027</td>
<td>-0.006</td>
<td>0.078</td>
<td>0.065</td>
<td>-0.013</td>
</tr>
<tr>
<td></td>
<td>[0.179]</td>
<td>[0.163]</td>
<td>[13.68]**</td>
<td>[0.268]</td>
<td>[0.246]</td>
<td>[19.75]**</td>
</tr>
<tr>
<td>Medium-Large</td>
<td>0.041</td>
<td>0.028</td>
<td>-0.013</td>
<td>46.673</td>
<td>59.517</td>
<td>12.844</td>
</tr>
<tr>
<td></td>
<td>[0.198]</td>
<td>[0.165]</td>
<td>[29.55]**</td>
<td>[161.256]</td>
<td>[173.943]</td>
<td>[28.76]**</td>
</tr>
<tr>
<td>Sport</td>
<td>0.037</td>
<td>0.027</td>
<td>-0.010</td>
<td>451.116</td>
<td>448.046</td>
<td>-3.070</td>
</tr>
<tr>
<td></td>
<td>[0.189]</td>
<td>[0.162]</td>
<td>[23.85]**</td>
<td>[600.098]</td>
<td>[599.76]</td>
<td>[0.38]</td>
</tr>
<tr>
<td>Minivan</td>
<td>0.006</td>
<td>0.006</td>
<td>0.000</td>
<td>0.381</td>
<td>0.469</td>
<td>0.088</td>
</tr>
<tr>
<td></td>
<td>[0.078]</td>
<td>[0.077]</td>
<td>[0.47]</td>
<td>[0.886]</td>
<td>[0.934]</td>
<td>[0.47]</td>
</tr>
<tr>
<td>Van</td>
<td>0.027</td>
<td>0.037</td>
<td>0.010</td>
<td>0.446</td>
<td>0.666</td>
<td>0.220</td>
</tr>
<tr>
<td></td>
<td>[0.163]</td>
<td>[0.189]</td>
<td>[20.90]**</td>
<td>[0.891]</td>
<td>[1.012]</td>
<td>[85.11]**</td>
</tr>
<tr>
<td></td>
<td>Mean (SD) 1</td>
<td>Mean (SD) 2</td>
<td>Difference</td>
<td>Pr(pass) All</td>
<td>Pr(pass) Non-Exempt</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>-------------</td>
<td>-------------</td>
<td>------------</td>
<td>-------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>Pick Up</td>
<td>0.022 (0.016)</td>
<td>0.026 (0.018)</td>
<td>0.004</td>
<td>0.925 (0.024)</td>
<td>0.895 (0.027)</td>
<td></td>
</tr>
<tr>
<td>Cylinders</td>
<td>3.462 (1.100)</td>
<td>3.542 (1.183)</td>
<td>0.079</td>
<td>0.759 (0.427)</td>
<td>0.787 (0.409)</td>
<td></td>
</tr>
</tbody>
</table>

Notes:

1. The first two columns of each panel report the mean value for the sample and the corresponding standard deviation in brackets. 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 low evidence of corruption is 162,907. The remaining tests, 1,343,762 belong to centers with high 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. The 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.
Online Appendix Table 3: Two Period Model Parameters and Demand for Bribes

Panel A: Parameter Estimates

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>Estimate</th>
<th>Bootstrapped</th>
<th>Implied estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>[1]</td>
<td>[2]</td>
<td>[3]</td>
</tr>
<tr>
<td>( w )-intercept</td>
<td>-34.42</td>
<td>80.52</td>
<td>Mean time cost 45.07</td>
</tr>
<tr>
<td>( w )-slope</td>
<td>7.57</td>
<td>12.73</td>
<td>Minimum time cost 26.52</td>
</tr>
<tr>
<td>( b )</td>
<td>230.67</td>
<td>81.73</td>
<td>Maximum time cost 69.21</td>
</tr>
<tr>
<td>( \delta )</td>
<td>0.624</td>
<td>0.089</td>
<td>SE of random shock 74.36</td>
</tr>
</tbody>
</table>

| \( t \)          | 57.98  | 47.14 |

Panel B: Fitted and predicted probabilities for each history

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>O1: Postpone</td>
<td>0.0832</td>
<td>0.0169</td>
</tr>
<tr>
<td>O2: Bribe/No bribe - Pass</td>
<td>0.755</td>
<td>0.7422</td>
</tr>
<tr>
<td>O3: No bribe-Fail-Postpone</td>
<td>0.0037</td>
<td>0.0001</td>
</tr>
<tr>
<td>O4: No bribe-Fail-Bribe/No bribe-Fail-No bribe-Pass</td>
<td>0.1142</td>
<td>0.1792</td>
</tr>
<tr>
<td>O5: No bribe-Fail-No bribe-Fail-bribe</td>
<td>0.0439</td>
<td>0.0516</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>All Histories</th>
<th>Predicted</th>
<th>Bootstrapped Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Postpone</td>
<td>0.0169</td>
<td>0.0080</td>
</tr>
<tr>
<td>Bribe</td>
<td>0.0476</td>
<td>0.0589</td>
</tr>
<tr>
<td>Decision History</td>
<td>Actual Probability</td>
<td>Fitted Probability</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>No bribe-Pass</td>
<td>0.6946</td>
<td>0.0903</td>
</tr>
<tr>
<td>No bribe-Fail-Postpone</td>
<td>0.0001</td>
<td>0.0002</td>
</tr>
<tr>
<td>No bribe-Fail-Bribe</td>
<td>0.0310</td>
<td>0.0302</td>
</tr>
<tr>
<td>No bribe-Fail-No bribe-Pass</td>
<td>0.1482</td>
<td>0.0275</td>
</tr>
<tr>
<td>No bribe-Fail-No bribe-Fail-Bribe</td>
<td>0.0516</td>
<td>0.0085</td>
</tr>
<tr>
<td>Total bribing</td>
<td>0.0786</td>
<td>0.1106</td>
</tr>
</tbody>
</table>

Notes:

1. Table A3 shows the results of the two-period maximum likelihood estimation of the bribing behavior model developed in Section 6 using a 5% random sample (17,365 vehicles). Panel A shows the model’s parameter estimates and Panel B shows the actual and fitted probabilities as well as the simulated probabilities for all decision histories, including the unobserved ones.

2. For notation, see notes on Table 5. The last row of Panel B offers the total amount of bribing, i.e. the sum of 0.0476 and 0.0310.

3. The standard errors for parameters and predictions correspond to bootstrapped standard errors. These errors were obtained from 100 replications of the test for cheating, estimation of the mapping between car characteristics and probability of passing, and MLE of car owner decision model.
Online Appendix Table 4: Conditional Actual and Fitted Probabilities for each Outcome

Panel A: Conditioning on Predicted Probability of Passing

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Predicted Probability of Passing</th>
<th>Actual</th>
<th>Fitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1 conditional on</td>
<td>&lt; 0.7</td>
<td>0.0018</td>
<td>0.0007</td>
</tr>
<tr>
<td></td>
<td>(0.7, 0.8)</td>
<td>0.0071</td>
<td>0.0039</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.8</td>
<td>0.0010</td>
<td>0.0004</td>
</tr>
<tr>
<td>O2 conditional on</td>
<td>&lt; 0.7</td>
<td>0.0363</td>
<td>0.0349</td>
</tr>
<tr>
<td></td>
<td>(0.7, 0.8)</td>
<td>0.2502</td>
<td>0.2465</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.8</td>
<td>0.0333</td>
<td>0.0344</td>
</tr>
<tr>
<td>O3 conditional on</td>
<td>&lt; 0.7</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>(0.7, 0.8)</td>
<td>0.0071</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.8</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>O4 conditional on</td>
<td>&lt; 0.7</td>
<td>0.0086</td>
<td>0.0108</td>
</tr>
<tr>
<td></td>
<td>(0.7, 0.8)</td>
<td>0.0508</td>
<td>0.0602</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.8</td>
<td>0.0046</td>
<td>0.0052</td>
</tr>
<tr>
<td>O5 conditional on</td>
<td>&lt; 0.7</td>
<td>0.0039</td>
<td>0.0042</td>
</tr>
<tr>
<td></td>
<td>(0.7, 0.8)</td>
<td>0.0204</td>
<td>0.0179</td>
</tr>
<tr>
<td></td>
<td>&gt; 0.8</td>
<td>0.0018</td>
<td>0.0007</td>
</tr>
</tbody>
</table>

Panel B: Conditioning on Log of Car Value

<table>
<thead>
<tr>
<th>Car Value</th>
<th>Actual</th>
<th>Fitted</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1 conditional on</td>
<td>&lt; mean</td>
<td>0.0068</td>
</tr>
<tr>
<td></td>
<td>&gt; mean</td>
<td>0.0058</td>
</tr>
<tr>
<td>O2 conditional on</td>
<td>&lt; mean</td>
<td>0.1884</td>
</tr>
<tr>
<td></td>
<td>&gt; mean</td>
<td>0.1930</td>
</tr>
<tr>
<td>O3 conditional on</td>
<td>&lt; mean</td>
<td>0.0000</td>
</tr>
<tr>
<td></td>
<td>&gt; mean</td>
<td>0.0000</td>
</tr>
<tr>
<td>O4 conditional on</td>
<td>&lt; mean</td>
<td>0.0428</td>
</tr>
<tr>
<td></td>
<td>&gt; mean</td>
<td>0.0322</td>
</tr>
<tr>
<td>O5 conditional on</td>
<td>&lt; mean</td>
<td>0.0181</td>
</tr>
<tr>
<td></td>
<td>&gt; mean</td>
<td>0.0130</td>
</tr>
</tbody>
</table>
Notes: See notes under Table 5.
Online Appendix Table 5: Predicted Probability of Passing by Cheating Group

<table>
<thead>
<tr>
<th>Deciles of cheating test</th>
<th>Observed passing rate</th>
<th>Mean predicted probability of passing</th>
<th>Number of tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.759</td>
<td>0.760</td>
<td>39,799</td>
</tr>
<tr>
<td>2</td>
<td>0.748</td>
<td>0.762</td>
<td>42,078</td>
</tr>
<tr>
<td>3</td>
<td>0.776</td>
<td>0.748</td>
<td>59,140</td>
</tr>
<tr>
<td>4</td>
<td>0.760</td>
<td>0.741</td>
<td>65,094</td>
</tr>
<tr>
<td>5</td>
<td>0.788</td>
<td>0.757</td>
<td>60,100</td>
</tr>
<tr>
<td>6</td>
<td>0.817</td>
<td>0.740</td>
<td>68,220</td>
</tr>
<tr>
<td>7</td>
<td>0.818</td>
<td>0.750</td>
<td>69,284</td>
</tr>
<tr>
<td>8</td>
<td>0.793</td>
<td>0.743</td>
<td>93,017</td>
</tr>
<tr>
<td>9</td>
<td>0.802</td>
<td>0.740</td>
<td>62,139</td>
</tr>
<tr>
<td>10</td>
<td>0.756</td>
<td>0.750</td>
<td>50,921</td>
</tr>
</tbody>
</table>
Online Appendix Table 6: Example of Benefit Calculation Corresponding to Policy 5 in Table 5

<table>
<thead>
<tr>
<th></th>
<th>Sum of emission differences</th>
<th>Contribution to average concentration (in mg/m³ or ppm) per ton of emissions</th>
<th>Deaths per year per unit of concentration (^{d,e})</th>
<th>Deaths per ton of emissions</th>
<th>Cost of a ton of pollutant in U.S. dollars (^{f})</th>
<th>Total number of lives saved per year</th>
<th>Total benefits from reduced emissions in $1000 (^{f})</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOx (PM10)</td>
<td>48</td>
<td>2.18E-05</td>
<td>104.400</td>
<td>2.28E-03</td>
<td>1,480.8</td>
<td>1.112</td>
<td>723</td>
</tr>
<tr>
<td>HC (PM10)</td>
<td>164</td>
<td>9.14E-06</td>
<td>104.400</td>
<td>9.55E-04</td>
<td>620.5</td>
<td>0.157</td>
<td>102</td>
</tr>
<tr>
<td>HC (O3)</td>
<td>164</td>
<td>7.34E-08</td>
<td>10.142</td>
<td>7.44E-07</td>
<td>0.5</td>
<td>0.000</td>
<td>0</td>
</tr>
<tr>
<td>CO</td>
<td>3,055</td>
<td>1.19E-06</td>
<td>634.463</td>
<td>7.53E-04</td>
<td>53.3</td>
<td>2.302</td>
<td>1,496</td>
</tr>
<tr>
<td>Total</td>
<td>3,708</td>
<td>3.571</td>
<td>2,321</td>
<td></td>
<td></td>
<td>3.571</td>
<td>2,321</td>
</tr>
</tbody>
</table>
Notes:

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

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

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

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

e. For NOx (PM10) and HC (PM10), these numbers are calculated using Schwartz (1994) estimate of 0.5222 deaths per 100,000 people per unit of TSP. For CO, I use infant mortality estimates from Arceo, Hanna and Oliva (2013) (166.4 per 100,000 births).

f. The Value of Statistical Life used is 650 thousand USD and is taken from Molina and Molina (2002).