How the Clean Air Act “Hits Home” Through the I&M Program

Douglas S. Noonan
Georgia Institute of Technology

AEI Working Paper # 2011-09
December 2011

This project is a product of the National Research Initiative, an AEI program that is designed to support, publish, and disseminate research by university-based scholars and other independent researchers who are engaged in the exploration of important public policy issues.
How the Clean Air Act “Hits Home” through the I&M Program

Douglas S. Noonan, Ph.D.
Anupit Supnithadnaporn, Ph.D.

Air quality policy in the United States centers on the Clean Air Act (and its amendments) and its enforcement by the EPA and state agencies. The CAA includes a variety of regulatory tools and targets, such as regulating emitting industries, emissions control technologies for new cars, and forcing metropolitan areas to implement a host of measures to reduce local emissions under threat of “nonattainment status” (and the loss of federal transportation dollars) when local ambient air quality exceeds federal standards. Research has shown that the CAA has had a significant effect on ambient concentrations (Greenstone 2004, Auffhammer et al. 2009), on regulator behavior (Auffhammer et al. 2009), on the location decisions of industrial emitters (List et al. 2003, Greenstone 2002), on infant mortality (Chay and Greenstone 2003), and even on the amount of pollution emitted into waterways (Greenstone 2003). The CAA amendments of 1990, which established a large-scale “cap and trade” system for reducing sulfur dioxide emissions, has been shown to affect R&D into abatement technologies as well (Popp 2003).

While CAA has been shown to have some far-reaching implications for society, the sweeping nature of the CAA is often thought to have only minimal direct impact on typical U.S. households. From the household’s perspective, the CAA can be attributed with bringing catalytic converters to all of their cars, altering the quality of the air they breathe, and altering the industry (and available job) mix in their area. Yet these changes are all indirect impacts on the household as a result of the regulation affecting others’ (e.g., auto manufacturers, industrial emitters) behavior.

This report emphasizes one way that regulations following from the CAA directly influences individuals’ behavior: through the Inspection and Maintenance (I&M) programs. Metropolitan areas found to be in nonattainment typically must institute them, making these programs commonplace. Most residents own vehicles subject to these regulations. Thus, the I&M program component of CAA implementation has particularly widespread impact on American households. This report describes the ways in which the I&M program affects vehicle owners in Atlanta, a large automobile-dominated metropolitan area with serious air quality concerns. By focusing attention on how people respond to the testing program, we addresses concerns about the equity of the burden of I&M programs and mechanisms by which the fleet is “greened.”

As detailed below, this report shows several major impacts:

1. The I&M program is effectively screening dirtier vehicles out of the fleet. It must be noted, however, that failure rates are sufficiently low for the bulk of the tested fleet that the vast majority of tests accomplish little and the reliance on diagnostic computers (rather than actual tailpipe emissions) is problematic as vehicles age – precisely when better screening is most important.
2. Households tend to “shop” for inspection stations to minimize their costs – thus picking stations nearby their homes and those that are more lenient.
3. The I&M program imposes a burden on those who own dirty cars, and those cars are clearly owned by poorer households.
These findings hold several lessons for policymakers. First, air quality regulations can have large impacts directly on households, even if they are ostensibly aiming to regulate technologies. These regulations are shown to affect the composition of the vehicle fleet, which means affecting who owns what kinds of cars and how fast households replace their vehicles. Policymakers would do well to keep in mind these kinds of far-reaching impacts that might result from similar future regulations, such as technology mandates associated with carbon emissions. Second, these regulations’ impacts on households behavior is significant, predictable, and fraught with unintended consequences. Two significant implications are the burden of compliance falling most heavily on the poor and the persistent effort of households to avoid compliance costs. Third, implementing rules that seek to regulate millions of households’ vehicles poses massive administrative costs and challenges. These should not be underestimated.

The findings have implications for those seeking to tweak or more radically reform vehicle emissions policy. We show that the validity of I&M tests declines for older, dirtier cars. Casting some doubt on the testing mechanism itself points to alternative testing technologies. Showing how households “shop” for inspection stations attests to the impact of I&M on households and can guide regulators in permitting and auditing the hundreds of stations in the metro area. Moreover, we show that the I&M program in Atlanta is working to “green” the vehicle fleet, but its costs fall predominantly on poor households and it is not motivating greater maintenance of cars’ emission control systems as we might expect. This raises the possibility of reshaping the incentives and softening the blow on poor households. Further discussion of these policy reforms appears in the Conclusion.

The rest of the report is organized as follows. First, some background on the I&M program is presented. Then, three different aspects of the I&M program’s impacts on households are reviewed: (1) how I&M “greens” the fleet and who passes the test; (2) how households “shop” for testing stations; and (3) who bears the costs of repair and keeping cars clean.

1. I&M Background

The I&M program in Atlanta is a prominent measure to control air pollutant emissions in an attempt to achieve compliance with the Clean Air Act. Nearly all vehicle owners – and thus the vast majority of Atlanta residents – are subject to the regulation under the I&M. The idea behind I&M programs is to test the fleet periodically (annually for Atlanta) and require vehicles with tailpipe emissions exceeding set standards to receive repairs before returning to the fleet. All eligible vehicle owners must pass an I&M test annually as a precondition of renewing their vehicle registration. Certain vehicles are exempted from the I&M program. For Atlanta’s enhanced program, the fleet of vehicles to be tested includes cars and light trucks (weighing up to 8,500 pounds) of model years between 4 – 24 years ago. The three most recent model years and “antique” vehicles over 25 years old are exempted from testing. Additional exemptions are available for vehicles powered by diesel or exclusively alternative fuels and for vehicles owned by senior citizens (aged 65 and older).

Vehicles that fail the test will not be permitted to renew their vehicle registration (their license plate tags will be expired, etc.) until they undertake successful repairs and pass the test. The concern about individuals and firms responding strategically to circumvent the test or its costs implicitly suggests that the costs are nontrivial. There is evidence of moving high-emitting vehicles out of the I&M jurisdiction but still in the area (Stedman et al. 1998) possibly via “wash sales” – using out-of-area “pseudo addresses” to enable owners to keep driving in the area (Zia et al. 2006). Owners who can demonstrate that they have spent over a threshold (currently $787) on emissions control system repairs
following a failed test can receive an exemption from the test for that year, however this exemption is one-time-only and nonrenewable. Inspectors can recommend repairs even prior to an inspection if they notice an illuminated MIL, saving the owner some expense. Re-tests of vehicles that failed initially are also free of charge. Inspection stations are closely regulated, including a price ceiling, and audited extensively.

I&M program rules have evolved over time (Harrington et al. 1996, 2000). I&M programs began with the CAA of 1970 and became mandatory for certain areas under the CAAA of 1977. The CAAA of 1990 brought two tiers of I&M programs (for moderate and for serious nonattainment areas), differentiated by the type of testing employed. The enhanced category, which Atlanta fell into by 1997, required testing of all eligible vehicles in the region using one of three tests. Prior to 2002, the test was either the Acceleration Simulation Mode (ASM)\(^1\) or, if the vehicle could not be ASM-tested, the Two Speed Idle (TSI)\(^2\). (A fuel cap check and visual inspection also must follow either emissions test.) After 2002, all vehicles of model year 1996 or later tested using the (on-board diagnostic) OBDII test. Earlier model years used the TSI; the ASM is no longer in use.

Of special interest here is the OBDII test. Federal rules require OBDII and sensor systems installed on all vehicles up to 14,000 pounds manufactured since 1996. More than just a technology mandate for automakers, the OBDII now dictates pass or fail to most car owners. The OBDII monitors those sensors and, when emission control components of the engine system malfunction and might generate emissions exceeding 150% of the tailpipe emissions standards, turns on a malfunction indicator light (MIL).\(^3\) The earlier detection is aimed at reducing detection and repair costs as well as enabling repairs sooner. It might also be harder to cheat or fool the OBDII computer than it is to spoof an ASM or TSI test (e.g., by using a different vehicle). I&M station inspectors simply connect a computer to the vehicle's OBDII system to read the diagnostics. Passing the annual emissions test using OBDII requires three conditions be met: (1) the MIL must be off, (2) no fault codes stored in the computer, and (3) the readiness monitor is on. The OBDII test yields only a binary “pass / fail” outcome or anything actually measuring tailpipe emissions. (The validity of OBDII tests as a measure of actual vehicle emissions is considered below.) As with pre-1996 vehicles, a fuel cap pressure test is still required and failure may indicate a “dirty” vehicle even if its tailpipe emissions are negligible.

2. **Who Passes the Test in Atlanta?**

Which vehicles pass the test, which part of Atlanta's vehicle fleet are dirty, and how well do I&M test results match up with on-road performance are all critical questions in assessing the performance of the program. Moreover, answers to these questions all directly bear on Atlanta households: those whose car fails the test must go without their car or pay for costly repairs. Previous work (Zia et al. 2006) compiled voluminous data on Atlanta's fleet and which vehicle attributes are associated with higher on-road emissions. Their data – which are used and expanded below – combine two primary sources of data: the Continuous Atlanta Fleet Evaluation (CAFE) data collected on roadside monitors that measure tailpipe emissions at key locations around the city, and the state of Georgia's vehicle registration database that includes I&M test results. The CAFE and the registration databases are linked by the Vehicle Identification Number (VIN), which is essential to the registration and is also obtained for the CAFE remote sensing data (RSD) via photographs of the vehicle's license plate. In this way, attributes

---

1 The ASM measures tailpipe emissions from vehicles placed on a treadmill-like dynamometer while simulating realistic driving conditions.
2 The more primitive TSI measures tailpipe emissions from vehicles with engines idling at two different speeds.
3 This is the “Check Engine Light” for many of us.
of the car (e.g., year, make, model) can be compared to measured on-road emissions and to I&M testing status.

Which vehicles are dirty?
The Zia et al. (2006) results, summarized below, estimate how EMISSIONS depends on STATUS, VEHICLE, YEAR, and ENVIRONMENT. The variable EMISSIONS refers to carbon monoxide (CO), hydrocarbons (HC), or nitrogen oxide (NO), which the RSD measures. STATUS refers to the vehicle observed in the RSD as whether it is exempt from I&M testing (e.g., a waived class of vehicles such as very old ones, out-of-state vehicles, or in-state vehicles from elsewhere in Georgia), whether it is new to the region and not yet been tested, or whether it is in the testing fleet. It also refers to whether the vehicle has previously failed and attempted a re-test in that year. Altogether, there are many combinations of vehicle status in that study. The VEHICLE variable above refers to attributes of the vehicle (e.g., age, make, model, emissions control technology). The ENVIRONMENT variable refers to the physical conditions in which the RSD collected the data. YEAR indicates the year of RSD measurement, which ranged from 1997-2001. (Note that this data predate the OBDII testing.) The above relationship is estimated in a nonlinear fashion, and the reader is referred to Zia et al. (2006) for more detailed discussion. Factors associated with cleaner or dirtier vehicles, relative to an “average” car, are summarized in Table 1. (Some factors are omitted for space considerations; see Zia et al. (2006) for full details.)

Table 1: Select factors associated with more or less emissions (for different pollutants)

<table>
<thead>
<tr>
<th>Cleaner</th>
<th>CO</th>
<th>HC</th>
<th>NO</th>
</tr>
</thead>
<tbody>
<tr>
<td>country of manufacture</td>
<td>Japan, Germany, Sweden</td>
<td>Japan, Germany, Sweden</td>
<td>GM, Honda, Toyota,</td>
</tr>
<tr>
<td>engine technology*</td>
<td>II</td>
<td>II</td>
<td>Nissan, Mercedes,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Volvo, VW</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Japan, Germany</td>
</tr>
<tr>
<td>test status</td>
<td>retested &amp; passed or</td>
<td>retested &amp; passed or</td>
<td>retested &amp; passed</td>
</tr>
<tr>
<td></td>
<td>failed, ineligible</td>
<td>failed, ineligible</td>
<td>older</td>
</tr>
<tr>
<td></td>
<td>older</td>
<td>older</td>
<td>older</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>light trucks &amp; SUVs</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Mitsubishi</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dirtier</td>
<td>Honda, Toyota, Volvo</td>
<td>Nissan, Mazda,</td>
<td>Korea</td>
</tr>
<tr>
<td></td>
<td>, Nissan, Mazda,</td>
<td>Mitsubishi, VW</td>
<td>Korea, Mexico</td>
</tr>
<tr>
<td></td>
<td>Mitsubishi, Isuzu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>country of manufacture</td>
<td>Mexico</td>
<td>Korea</td>
<td>Korea, Mexico</td>
</tr>
<tr>
<td>engine technology</td>
<td>I, III, IV</td>
<td>I, IV</td>
<td></td>
</tr>
</tbody>
</table>

* I = Air injection reactor system, II = Exhaust gas circulation, III = Oxidation (two-way) catalyst, IV = Thermostatic air cleaner.

All comparisons are relative to a Ford passenger vehicle, made in the USA with 0 years of age, that has passed the I&M test, and was measured in 1997. See Zia et al. (2006), Table 4, for additional details.

The factors predicting dirtier vehicles differ across the various pollutants. The results nonetheless
permit a few general observations. Vehicles in the Atlanta fleet are getting cleaner over time. In any
given year, older vehicles are dirtier, although the decay rate is nonlinear and differs across pollutants.
Light trucks look cleaner than cars for HC and CO, but not for NO. There is considerable variation in
emissions depending on the make and country-of-final-assembly. Some manufacturers’ vehicles are
cleaner than others, even after controlling for age, engine technology, and the vehicle’s status in the
I&M program. Some European – Mercedes and VW (except for HC) – and Japanese – Honda and
Toyota (except for CO) – makes fare well. Country of manufacture also seems to play a significant
role, with Japanese and European builds appearing much cleaner than similar vehicles assembled
elsewhere. Vehicles manufactured in Mexico, Canada, and Korea tended to be dirtier than others.

Most interesting in these results are significant differences in emissions for vehicles where none might
be expected. Some major differences across I&M program status are evident. Relative to vehicles that
have passed their initial inspection, and controlling for the other factors listed in Table 1, vehicles
registered elsewhere in the state of Georgia and outside the I&M program jurisdiction tend to emit
more CO and HC (but not more NO). Owners in I&M areas apparently keep cleaner cars than those
outside. This could be due to better maintenance or exporting dirtier cars outside the area. Vehicles
being retested are much dirtier. These results indicate an effective I&M program. More troublingly,
emissions from retested vehicles that pass their retest do not appear any cleaner than those that fail their
retest. Owners retesting their vehicles look to be doing the bare minimum to obtain a pass.

Which vehicles pass the OBDII test?
The use of on-board diagnostic computers to monitor emissions marks a major shift in I&M program
operations. The Zia et al. (2006) analysis relies on data from the pre-OBDII testing era, when
inspections were exclusively based on actual tailpipe emissions. More recent RSD and I&M program
data can be used to ascertain the effectiveness of the OBDII system. Data from 2002 – 2005 are
employed to carry out this analysis. Similar to the earlier model, the on-road pollution level
\( \text{EMISSIONS INDEX} \) of each vehicle is predicted using OBDII test status, \( \text{YEAR} \) of measurement,
\( \text{VEHICLE} \), and \( \text{ENVIRONMENT} \). This analysis collapses the three different pollutants into a single
measure of emissions intensity that is constructed to be consistent with the I&M test principles.
Specifically, the \( \text{EMISSIONS INDEX} \) is the maximum value of the (normalized) HC, CO, and NO
measurements from the RSD, where each measure is scaled by the emissions standard that applies to
that vehicle type for that pollutant. Thus, the \( \text{EMISSIONS INDEX} \) exceeds a value of 1.0 if at least one
of the pollutants exceeds the threshold, which should indicate a “dirty” vehicle. Our statistical model
estimates how the emissions intensity is explained by OBDII test results, vehicle characteristics, and
RSD measurement conditions.\(^4\)

Table 2: Select factors associated with more or less emissions intensity

<table>
<thead>
<tr>
<th>Factor affecting EMISSIONS INDEX</th>
</tr>
</thead>
</table>

\(^4\) A few unavoidable challenges occur in constructing this test with existing data. Unfortunately, vehicles do not have an
RSD measure just prior to their visit to an inspection station. Repairs and modifications to the vehicle – in response to
or in anticipation of the inspection – are unobserved in the data and could explain inconsistencies between RSD
measures and OBDII results. To minimize this concern, only RSD measures that occur before the inspection are used –
thus ruling out post-test repairs impacting RSD observations. Moreover, for cars that appear multiple times in CAFE
database, only the most recent RSD measurement is used (on average 3.5-4 months before the test) in order to reduce the
opportunity for confounding factors to arise during the elapsed time between RSD and OBDII measurements.
The results presented in Table 2 show variation in emissions intensity can be explained by variation a vehicle characteristics. In particular, vehicles that failed their OBDII test tended to have significantly higher emissions on their previous on-road measurement – even after controlling for a variety of other factors. The emissions intensity is 0.132 higher on average for failing cars than passing cars. Put another way, a car whose worst pollutant is right at the threshold (i.e., EMISSIONS INDEX = 1.0) and passes the OBDII can be expected to have 13.2% higher emissions if it were to fail that OBDII test. In that regard, the results strongly suggest that the OBDII test is capturing high-emissions vehicles.\(^5\)

Other factors of interest show how vehicle age is positively associated with emissions, even after controlling for (future) OBDII test results and miles driven per year. Unsurprisingly, vehicles with more miles-per-year tend to emit more as emissions control systems wear down. In 2002, a model-year 2000 car would emit 6.7% more than an otherwise identical 1999 car, partly due to age and partly due to a new model. There is a “greening” of the fleet indeed (Kahn and Schwartz 2008). SUVs, light trucks, and vans tend to emit significantly more than cars also.\(^6\) Also similar to the results in Table 1, the location of manufacture, make, and engine technology factor prominently. Emissions are declining over time as well, even after controlling for these other factors. The results also verify that household choices, in particular how much they drive their car, greatly affect relative emissions – above and beyond what the OBDII test indicates.

**Does the OBDII validity decay?**

Although the OBDII test does detect the vehicles with higher emissions, patterns in the data raise serious concerns about the long-term validity and reliability of OBDII systems installed in cars. Older vehicles tend to pollute much more than newer cars, and the possibility that the diagnostic computer could also degrade over time – just as the vehicle’s emissions control system decays over time – poses a major concern for the I&M program effectiveness. Degraded OBII tests may allow dirty cars to pass or fail clean cars.

To assess the possibility of decaying validity in OBDII tests, a test of whether the OBDII results are consistent with the RSD measures is employed. An indicator variable, AGREE, is defined such that it

---

\(^5\) If vehicle owners with high on-road emissions before the test knew about it, presumably they would reduce rather than increase their emissions prior to the OBDII test date (in anything). This response would reduce the effect observed in Table 2, biasing the results downward. The results in Table 2 can be interpreted as lower-bounds on the OBDII effect.

\(^6\) Recall that the EMISSIONS INDEX already factors in the higher emissions allowances afforded heavier non-car vehicles.
takes a value of unity for vehicles whose OBDII test results “agree” with the RSD measures. If (on-road) tailpipe emissions exceed the standard (i.e., $EMISSIONS INDEX > 1.0$), an agreement implies that the OBDII test is failed. If $EMISSIONS INDEX < 1$ and the OBDII test indicates a pass, an agreement also occurs. Otherwise, $AGREE$ takes a value of 0. For this sample, 28% of the cars had RSD and OBDII measures that disagreed. This is predominantly due to those cars that “fail” their on-road emissions (perhaps due to very “noisy” RSD measures) and pass their OBDII test (24% of all observations).

A logistic model to predict the likelihood of $AGREE = 1$ is estimated, with explanatory factors including those used in Table 2 and an $ELAPSE$ variable. Figure 1 highlights the relationships of interest from this analysis: the effect of age and wear-and-tear (horizontal axis) on the likelihood of agreement between the on-road tailpipe emissions and the OBDII test results (i.e., validity). The graph shows an extremely high likelihood of agreement for young cars (largely because they all tend to be clean). This likelihood of agreement declines with age, especially quickly between ages 4 and 10. Also, vehicles that are driven more have a lower likelihood of agreement. The likelihood of agreement for a typical car actually falls below 50% before the vehicle reaches 10 years old.

![Graph](https://via.placeholder.com/150)

Figure 1: Vehicle age and predicted probability of the RSD-OBD agreement varying by mile driven per year

Clearly the ability of the OBDII to match with on-road emissions falls significantly with age and use. Whether this is a result of degraded performance of the OBDII system itself or due to the RSD measures themselves being more highly variable (and thus increasing the share of vehicles that have a “bad day” from the perspective of the RSD) is unknown at this point. This is clearly an important area for future research, given that older vehicles (and hence older OBDII systems) tend to have much higher emissions. But, regardless of the source of the disagreement, it does point to potential pitfalls in using on-road RSD or in-car OBDII technologies for monitoring emissions. A combination of the two may be warranted.

---

7 Other predictors include an $ELAPSE$ variable (the number of days between the car’s RSD observation and its OBDII test) and interactions between $ELAPSE$ and the vehicle age and the vehicle miles-per-year. As expected, larger $ELAPSE$ times are associated with greater likelihood of disagreement, although the effect is not remotely statistically significant (suggesting minimal bias from using RSD measurements that are not concurrent with OBDII tests).
Another observation is worth emphasis. The *ELAPSE* variable is basically uncorrelated with the disparity between RSD measures and OBDII test results. This is partly by design; the analysis was constructed to minimize the potential influence of confounding factors that might have occurred during that elapsed time. Nonetheless, the possibility that vehicle owners might alter their vehicle prior to OBDII testing cannot be ruled out. The evidence suggests that any modifications they make are not affecting their OBDII test results, on average. From this analysis, there is no evidence of owners effecting pretest repairs to achieve compliance.

3. Matching Households to Inspection Stations

Beyond the impact of I&M stations on repair decisions and fleet composition, the I&M program affects other household behaviors. In particular, the program constructs a new market for inspections. The resulting “shopping” behavior by vehicle owners, and the possibility of strategic behavior on the parts of owners and inspectors warrants some investigation. Households might be expected to economize over their travel costs to I&M stations. Deviations from the 'most convenient' station, however, might also be expected if there are some lenient inspection stations that are worth traveling the extra mile to visit. Price, congestion, and other factors may also play a role. On the other side of the market, inspection stations may sort themselves to target or serve certain parts of the market, implying a two-way 'matching' game between owners and stations. Evidence of sorting indirectly reveals the importance of the new market (created by the I&M regulation) to households. Shopping in this market is consistent with significant transaction costs, cost-minimizing behavior by motorists, and competition among inspection stations.

We start with the premise that vehicle owners simply want to pass the I&M requirements at a minimum cost. The primary costs associated with taking the test are the time and money spent on the inspection. The time cost includes travel time to the station and waiting times at the station. Moreover, if the vehicle fails the inspection, repair costs can be considerable. This suggests that vehicle owners may seek stations that are more lenient than others, perhaps even incurring additional expense to avoid failing. The leniency of stations may not be directly or easily observable to vehicle owners. Nonetheless, other forces may attract owners to lenient stations indirectly. For example, if car owners tended to return to (or recommend to others) stations that passed their car in the previous year and tended to switch stations following a “bad experience” of failing, then stations that tend to pass cars more often will tend to draw more customers. One of the primary empirical questions raised here is the extent to which customers are able to identify, even implicitly, and act upon station leniency.

The other side of the market – the inspection stations – also plays an active role in making these matches between motorist and inspector. The inspection market is a big business in Georgia, with more than 1,000,000 transactions at roughly 800 inspection stations in 2001 alone. While they compete with each other for motorists' business, they are constrained in how they compete. Inspection stations frequently have other (primary) businesses as gas stations or car dealerships. This analysis classifies inspections according to whether they are an independent store (53\% of all stations) or part of a chain and according to whether they are test-only (53\% of all stations) or test-and-repair stations.

---

8 The major auditing efforts of the state would be largely in vain if motorists could detect lenient stations while the auditor could not.

9 Georgia's I&M program includes a price ceiling of $25 for an inspection. Inspectors must also receive certification from the state so as to regulate a uniformity in the testing procedure. Many Atlanta stations offer discounted inspections and distribute coupons, bringing the fee down as low as $10 in some instances. Non-price competition includes offering free ancillary services (e.g., car wash) for inspected vehicles, prize drawings, or longer hours or more inspectors (thus reducing wait time).
Such concerns partly motivate our statistical analysis that explains vehicle owners' choice of which inspection station to use in Atlanta. We test whether stations' observable characteristics (e.g., travel cost, fee, inspection station type) and unobservable characteristics (e.g., failure rate, false-pass frequency) factor significantly in owners' choices. A conditional logit model is estimated to explain which station is chosen for each of the almost 40,000 vehicles for which home street addresses are known.10

Because station proximity to the home may not be as important as proximity to workplace, additional information about a sample of 8,069 Atlantans is obtained from the Atlanta Household Travel Survey (AHTS). The data from that survey, which includes home address and place of work, are matched with the home addresses from the CAFE and I&M program data for 2001. Knowing the work location allows tests for whether proximity-to-home or proximity-to-workplace better explains the choice of testing station. The results clearly show that distance-to-home is a much better predictor than distance-to-workplace. This exploration comes at a price: the sample size drops dramatically from roughly 40,000 to 465 when constrained to use only observations that can be matched using the AHTS. Because this small sample may not be representative, two sets of results are reported below for the two different samples.11 The explanatory factors used in this model measure characteristics of the station, characteristics of the vehicle owner, and characteristics of the vehicle.12 The full set of results is available upon request.

Table 3 summarizes the results for some key hypothesis tests, namely the impact of costs, competition, and station leniency on station selection by households. Travel costs, inspection fees, and waiting times are all expected to be inversely related with the likelihood that a station is chosen — and distance and fees are significantly negatively related to selection in both samples. Waiting time is also negatively related for the full sample. More nearby substitute stations, as expected, reduce the likelihood that any particular station is chosen. Leniency appears to be best measured by the failure ratio — the share of vehicle tests that are ‘fails’ — and does appear to significantly discourage selection by vehicle owners. Interestingly, once controlling for the failure ratio, the false pass odds (likelihood a ‘dirty car passes) is negatively related station selection. The failure ratio, arguably more observable to

---

10 Computational limitations require restricting the choice set for each household in the statistical analysis. This is commonly done in recreational and transportation demand analyses. A simple random subset of 9 stations (in addition to the one actually chosen) is used as the choice set for each vehicle owner when examining all 40,000 observations.

11 When using the smaller sample, however, computational limits that restrict the choice set to 10 stations do not apply and the full choice set across all stations can be used.

12 Of course, because the statistical model essentially explains the choice, by the motorist, among the many possible stations that they could select, each of the explanatory variables is ultimately an attribute of the station. Specifically, station characteristics include: inspection results (i.e., false pass odds, failure ratio), proxies for competition (i.e., density of stations in two-mile radius, station density interacted with inspection results, log of the total number of tests by the station, average daily tests per inspector, indicators for whether it is a chain store or a test-only station, total stores in chain, an indicator of station newness, the lowest testing fee charged by that station), mean age of vehicles, share of vehicles that are cars, share that are US-made, share that used the ASM test, and indicators for whether the station is open on Saturdays or Sundays), motorist characteristics (i.e., distance from the home and distance interacted with the log of the vehicle’s mileage, with the false pass odds, and with the failure ratio), and vehicle characteristics (i.e., interactions of mileage, whether the vehicle is a car, whether the vehicle is US-made, and whether the vehicle used the ASM test each with the false pass odds and with the failure ratio). The false pass odds is defined as the ratio of false passes (the number of vehicles inspected at that station that passed despite exhibiting emissions exceeding the ASM25 standards according to RSD measurements in 2001) to the “true” passes. This variable measures a station’s propensity to give favorable test results and, if known to motorists, might attract them. The failure ratio, on the other hand, is simply the proportion of total inspections at the station that fail (regardless of whether the failure was warranted). The failure ratio is another measure of station leniency.
motorists, appears to work much better as the proxy for station leniency. The evidence here suggests that it does have some influence in motorists choosing their I&M stations. Again, this is likely more due to repeat business (i.e., motorists who pass tend to return to the same station, failing motorists may be more likely to switch next year) or word-of-mouth than to any formal investigation by motorists.

Table 3: Summary of results for factors predicting motorists’ choice of a station

<table>
<thead>
<tr>
<th>Factor</th>
<th>expectation: more(+) or less(–) likely</th>
<th>Proxy variable</th>
<th>Partial sample (N=465)</th>
<th>Full sample (N=39,489)</th>
</tr>
</thead>
<tbody>
<tr>
<td>travel cost &amp; time</td>
<td>–</td>
<td>distance from house</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>inspection cost</td>
<td>–</td>
<td>min. test fee</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>waiting time</td>
<td>–</td>
<td>avg. daily tests per inspector</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>reputation</td>
<td>+</td>
<td>ln(total tests)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>competition</td>
<td>–</td>
<td>station density</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>complements</td>
<td>?</td>
<td>no repair</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>leniency</td>
<td>–</td>
<td>failure ratio</td>
<td>–**</td>
<td>–</td>
</tr>
<tr>
<td>leniency</td>
<td>+</td>
<td>false pass odds</td>
<td>–</td>
<td>–**</td>
</tr>
</tbody>
</table>

**. *** indicates significance at the 5% and 1% levels, respectively

This raises the possibility that vehicle owners might incur additional costs to use more lenient stations, possibly undermining the I&M program itself. The impact of lower failure ratio in attracting motorists is muted for stations in areas of high (station) density. Apparently the shopping for lenient stations involves selecting relatively isolated stations. Overall, the evidence does suggest (1) that vehicle owners are attracted to more lenient stations but not (2) that those stations tend to attract vehicles from especially far away.

4. Incidence of I&M Compliance Costs

The impact of the I&M program receives a good deal of attention largely because it impacts such a large population base. The “incidence” of I&M program costs has been discussed elsewhere (e.g., Harrington et al. 2000, Harrington and McConnell 1999). The costs of the program to vehicle owners can arise in three major categories: (1) repair, (2) maintenance, and (3) capital costs. Most of the research measuring its incidence has focused on repair costs. These costs include the time and travel costs associated with the test, as discussed in the previous section, and the costs associated with failing the test (e.g., Harrington et al. 2000, Kahn 1996a). Failing results in at least additional inconvenience, and typically some repair expenses or retirement and replacement of the vehicle in some cases. Repair costs are the easiest to observe, especially with administrative data on the program. There may be substantial other costs associated with program compliance, however, if households bear additional routine or preventative maintenance costs in order to keep their vehicle in compliance. These pre-test repair costs go undetected if only post-test (or failure-related) repair costs are assessed. The I&M program might increase the total cost of ownership by raising ongoing maintenance expenses and up-front or capital expenses if the program makes compliant vehicles more expensive. This third sort of program cost may be evident if cleaner cars sell for a premium (or if dirtier cars’ resale value is discounted).

While all vehicle owners bear some I&M program costs, individuals’ costs vary widely. The equity of these costs and their proportionality to income is difficult to assess due to data limitations. The repair costs – typically proxied by the incidence of inspection failure – are the most easily observed. Linking those repair costs to household attributes like income or race, however, is not as easy because demographics are not typically collected as part of vehicle registration or inspection data. Previous
studies (e.g., Zia et al. 2006, Kahn and Schwartz 2008) exploring the incidence of failures have typically used aggregate data from the Census to assess the equity implications of the policy. The results of this research have typically shown that failure rates negatively correlate with income, even after controlling for many vehicle characteristics. One interpretation of these results is that the program administration is somehow discriminatory or biased against the poor. Given that the inspection is a test of the vehicle and not the individual, it is difficult to see how such a program bias could emerge. Instead, the results may be explained by income being correlated with some unobserved aspect of the car (e.g., pre-test repairs, pre-test modifications) that enhances its prospects of passing.

These interpretations require making some very strong assumptions in having group-level variables (e.g., Census tract income) proxy for individual level variables (e.g., the owner/resident’s income). This well-known problem of “ecological inference” plagues many applications in social sciences (Robinson 1950) where aggregate data is used to make inferences about individual-level relationships. The negative correlation between income and failing that is observed when income is measured in the aggregate (e.g., tract, block group) may not hold if the individual’s income was known and used in the analysis.

We examine this with two different approaches and a rich dataset for the Atlanta area. First, we assess the sensitivity of the observed correlation between income and failing to weaker assumptions about how well group-level income proxies for individual-level income. This is accomplished via a Monte Carlo analysis that effectively simulates individual-level demographic data (calibrated to the match with the Census aggregates). Such an approach explicitly recognizes that the median income of the neighborhood in which the vehicle’s owner resides is proxy or a very noisy measurement of the owner’s actual income. Kahn (1996a) and Kahn and Schwartz (2008) conduct their statistical analyses as though group income is an exact measurement for individual income, implying that more information about income exists than is actually known. Our simulation approach explicitly accounts for the fact that all we know is that the individual comes from a neighborhood that has a particular income distribution. Armed with the income distributions in each census block group (the smallest Census unit that reports income), we take each individual’s income as a random draw from that distribution. This incorporates the noisiness of the proxy directly into the statistical analysis.

Second, we use the AHTS dataset of individual households in Atlanta, matched to their I&M testing results, to identify the relationship between income and failure rates at the individual level. Although this approach is straightforward in directly estimating the relationship between the two variables of interest, a large penalty is paid in terms of sample size when the survey data must be used (instead of the entire universe of I&M tests). The sample of households from the AHTS is only 465, compared to 685,714 from the administrative data. Thus, the second approach has precision and possible bias whereas the first approach lacks precision but should be unbiased.

The statistical model to assess the incidence of the I&M program on households focuses on the likelihood of failing an inspection. Of primary interest is whether higher income predicts lower failure rates. Two versions of the model are estimated to identify the effect of income: (1) conditional on vehicle characteristics, and (2) unconditional. The former effect is expected to be zero, as the inspection is of vehicles rather than owners. The latter effect could be nonzero if wealthier owners are more or less likely to fail due to owning different types of vehicles. These models can be formalized in the following logistic framework explaining the log-odds of failure for vehicle $i$:

$$
(1) \ln\left(\frac{p_i}{1 - p_i}\right) = \alpha_0 + \alpha_1 M_i + \alpha_2 V_i + \varepsilon_i
$$

$$
(2) \ln\left(\frac{p_i}{1 - p_i}\right) = \beta_0 + \beta_1 M_i + \theta_i
$$

where $p$ is the probability of failing, $M$ is the individual income, $V$ is a set of vehicle characteristics,
and $\varepsilon$ and $\theta$ are error terms. Naturally, vehicle characteristics should be strong predictors of failure. If $V$ is correlated with $M$, and dirtier vehicles are not distributed in the population independently of income, then $M$ should proxy for the $V$ omitted in the unconditional model (2) and $\beta_1$ should be nonzero. In other words, wealthier owners are less likely to fail because their cars are cleaner. An estimate of $\beta_1 < 0$ would indicate that failure, and hence post-test repair costs, tends to happen to the poor at least partly because their cars are dirtier. When those vehicle attributes are controlled for, as in the conditional model (1), the effect of income is less clear. Some (e.g., Kahn and Schwartz 2008, Harrington and McConnell 1999) have hypothesized that wealthier owners take better care of their car and thus $M$ proxies for unmeasured car quality in model (1), suggesting that $\alpha_1$ will be negative. This would be an important result, too, because it would suggest that another type of I&M program costs (maintenance costs) tend to fall on the wealthy. The role of aggregate or neighborhood income is unclear once individual income is controlled for. If there is a significant relationship between group income and individual failure rates, this raises interesting questions about peer effects, the spatial sorting of lenient inspection stations, and possible discrimination at the inspection stations. These will be discussed later.

Table 4 presents the results for models (1) and (2). $M$ is obtained using the two different approaches described above (simulated income and individually matched income). The simulated income results use the vast dataset of individual vehicles’ first attempt to pass the I&M test. The matched income results use the AHTS survey data linked to those I&M test results. The conditional model results are on the top half of Table 4; the unconditional results on the bottom half. The shaded columns represent the results under the more common approach of using simple aggregate income (the vehicle owner’s Census block group) as though it were the individual’s income.

The estimates here paint a stark picture of unconditional inequity but relative conditional equity. Using the simulated individual-level income, the conditional model shows $\alpha_1 = -0.03$. Although statistically significant, this effect is economically insignificant. For a given car and other neighborhood characteristics, increasing the income from the 1st percentile to the 90th percentile of the income distribution would predict an increase of only 0.45 percentage points (roughly from 7% to 7.5%) in the probability of failing the first inspection. The model using matched (actual) income measures echoes this result, where the $\alpha_1$ coefficient is estimated so imprecisely that it does not appear significantly different than zero.

The results are dramatically different when not controlling for vehicle characteristics. The bottom half of Table 4 shows much stronger effects of income on the probability of failing than the top half. For the simulated income data, income is negatively and significantly associated with the probability of failure. These results are consistent with Kahn (1996b). Relative to the conditional results in model (1), the effect is much larger but, again, not a very substantial effect. (The small effect may be primarily due to the nature of the simulated data, which essentially introduces measurement error into the income measure and thus attenuates the income effect in the model. This downward bias is consistent with the results for the matched income.) Using actual income measures shows a larger, negative, and marginally statistically significant (at the 10% level) effect of income.
Table 4: OLS Regression Results Summary

<table>
<thead>
<tr>
<th>source of income:</th>
<th>simulated</th>
<th>aggregate</th>
<th>matched</th>
<th>aggregate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>coef.</td>
<td>z-stat</td>
<td>coef.</td>
<td>z-stat</td>
</tr>
<tr>
<td>model:</td>
<td>(1) conditional</td>
<td></td>
<td>(2) unconditional</td>
<td></td>
</tr>
<tr>
<td>income effect</td>
<td>-0.0308</td>
<td>-5.28</td>
<td>-0.261</td>
<td>-16.83</td>
</tr>
<tr>
<td>vehicle characteristics</td>
<td>included</td>
<td>685,714</td>
<td>included</td>
<td>685,714</td>
</tr>
<tr>
<td>number of obs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>income effect</td>
<td>-0.1179</td>
<td>-22.91</td>
<td>-0.572</td>
<td>-49.57</td>
</tr>
<tr>
<td>vehicle characteristics</td>
<td>not included</td>
<td>685,714</td>
<td>not included</td>
<td>685,714</td>
</tr>
<tr>
<td>number of obs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Income measured as ln(income). Vehicle characteristics include vehicle’s age, ln(mileage), displacement, and controls for type (car, van), where it was produced (Europe, Asia, other), fuel induction type (EFI, FI, or MFI), and exhaust gas recirculation and thermostatic air cleaner technologies. The conditional models also include block-group percent black and percent other minority (plus percent latino, percent male, and median age for the simulated model).

The shaded columns in Table 4 indicate the results that would have been obtained if the analysis used only aggregate income data as though it were individual income, similar to previous studies. In all cases, using the median income from the block group of the vehicle’s owner yields much stronger (i.e., more negative) income effects – consistent with previous research. This points to the dangers of ecological fallacy if analysts were to draw inferences about individual-level relationships (e.g., wealthier individuals are less likely to fail their inspection) based on group-level income. The weakening of the relationship between income and emissions when using richer income data is a remarkable result in contrast with Kahn (1996a) and Kahn and Schwartz (2008).13 This might be due to a genuine effect of neighborhood wealth, regardless of the car owner’s income, on the chance of failure.

The apparent inequity in the unconditional model is consistent with vehicle quality, and cleanliness in particular, being distributed inequitably across the population of Atlanta. Poorer people tend to have cars more likely to fail. But, once you control for the type of car, poorer people appear no more likely to fail their inspection than wealthier people. This mixed message is good news in that the program implementation at the stage of the inspection station seems blind to income, but it is bad news in that the distribution of cars across social class places a heavier burden on the poor in terms of post-test repairs. How much of this inequitable distribution existed prior to the I&M program (e.g., wealthier people have long tended to owner newer and hence cleaner cars) and how much is a result of the I&M program (e.g., dirtier cars filter down to poorer owners as the wealthy pay a premium for cars that will pass) is beyond the scope of this assessment.

The inequity in the distribution of (observably) dirty cars raises the possibility that other burdens of the I&M program, aside from repair costs, are distributed inequitably. Recall the three types of household compliance costs for the I&M program: post-test repair, pre-test repair and the inspection itself, and capital (initial purchase) expense. Table 4 shows that poorer owners are somewhat more likely to fail, suggesting that they are more likely to bear the burden of dealing with out-of-compliance vehicles. This portion of the program’s incidence appears quite regressive. While inspection costs thus appear

---

13 Those authors use zip-code-level aggregate income, interpret it as though it proxied for household income, and argue that the use aggregate data will bias the effect of income downward (due to measurement error attenuating the effect). Replicating their approach with Atlanta data, however, show exactly the opposite to be the case: using richer data actually weakens the income-failure relationship.
somewhat regressive, the pre-test costs of keeping the vehicle better maintained and effecting repairs beforehand are a different story. Table 4 offers some indirect evidence on this matter. If additional (pre-test) maintenance expense was correlated with an individual’s income, then we would expect a significant role for income in model (1) where failure rates are conditional on observable vehicle characteristics. What Table 4 shows, instead, is no evidence that pre-test repair expenses are distributed inequitably with respect to income. Wealthier owners seem to take no better care of their cars than poorer owners. They just have cleaner cars (based on observable characteristics, like age, mileage, technologies) to start with. If they pay a premium for those cleaner cars, then the incidence of the capital expense component of the I&M program may fall disproportionately on the wealthy. Future research would do well to consider and measure those costs.

5. Conclusion

The I&M regulation is often thought to induce households to practice better car maintenance to keep their cars clean. It also aims to push households to export their dirty cars away from the polluted urban areas and replace them with newer, cleaner autos. The indirect evidence here suggests a few things:

- Households’ maintenance activity are not influenced by the inspections.
- There does not appear to be much pre-test repair, judging by the lack of correlation between passing the inspection and elapsed time (from RSD measurement to inspection) and between passing the inspection and household income. More time and resources does not seem to help people respond to the regulation with pre-test maintenance.
- They do react, however, by conserving compliance costs in getting the inspection.
- The post-test repair expenses can be substantial and they tend to fall on lower-income households.
- The inequity in the distribution of these compliance costs hints at another cost – the cost of owning cleaner vehicles – that may also be inequitably distributed. Wealthier households tend to own much cleaner cars.

I&M programs like Atlanta’s do appear to have substantial impact on households. These costs of modern I&M programs have been observed elsewhere as well (e.g., Ando et al. 2007). The command-and-control nature of the regulation has some advantages and disadvantages. The advantages come in their demonstrated effectiveness in reducing emissions and in overcoming problems associated with a decentralized system where local regulators may be too lax (Harrington et al. 1996). The I&M approach also is relatively simple and administratively easy to implement. On the other hand, mandating technologies (e.g., catalytic converters, OBDII) has drawbacks like inefficiency arising from one-size-fits-all prescriptions that are common to command-and-control policies. I&M programs designed to enforce an upper-limit to the emissions flow or concentration of each vehicle neglects that vehicle’s total emissions. Many of the advantages of decentralized, fee-based policies akin to a “polluter pays” policy may not be realized for vehicle emissions because neither the owners nor the regulators have good information about each vehicle’s emissions or effectiveness of repair efforts (Harrington and McConnell 1999, Ando et al. 2007). The tradeoffs between uncertainty and transaction costs in I&M programs are evident here as well. The worst offenders, older vehicles, are also the most difficult to measure by lower cost measurement approaches – on-road RSD or by in-car OBDII technologies. The uncertainty to households about the effectiveness of repair efforts can explain the absence of evidence of pre-test repair activities. These kinds of fundamental uncertainties can undermine the theoretical superiority of market-based approaches to vehicle emissions control.

14 Inspection costs are fairly uniform in that (nearly) everyone must get inspected and the fee is subject to a price ceiling, but time costs are likely closely linked to income (Calfee and Winston 1998).
Vehicle emissions control policy demands careful analysis of costs and benefits to households. A companion report explores evidence of the benefits of air quality improvements. Connecting the program implementation to actual air quality improvements has been done elsewhere (e.g., Harrington et al. 2000, NRC 2001, Stedman et al. 1997). The value of those air quality improvement may be substantial, but so are the I&M program costs. I&M programs have typically had smaller environmental benefits and greater implementation costs than projected (Harrington et al. 2000, NRC 2001). That dirtier cars in Atlanta tend to be registered out-of-area (Table 1) and that households choose closer, cheaper, and faster inspection stations (Table 3) all point to sizeable I&M compliance costs to households.

I&M programs have drawn considerable political pressure. This is due in no small part to its attempt to regulate an environmental problem caused by a difficult-to-detect small minority\textsuperscript{15} of millions of individual motorists. Applying the EPA’s conventional command-and-control regulations (i.e., mandating technologies and technological performance standards) at the household level has proven politically difficult and economically costly, although its benefits in terms of emissions reduction and alteration of fleet composition are substantial. The I&M program experience highlights the difficulty in regulating at the household level, instead of regulating a handful of producers.\textsuperscript{16} The rarity of high-emitters compounds this by making the compliance costs highly inequitable and, when high-emitting vehicles tend to be owned by the poor, regressive.

It also highlights the practical realities and importance of transaction costs in environmental regulation. The difficulty in measuring emissions (as a concentration coming out of a tailpipe and also in arguably more relevant “per year” terms) should not be understated. The political, economic, and technical pressures against RSD must be considered relative to the same challenges facing inspection stations. As technologies change, new opportunities arise. The tradeoffs between uncertainty and cost are likely to remain paramount, however. Choosing among different testing mechanisms has important implications for both uncertainty and cost. Switching to OBDII over ASM, for instance, brings lower costs but also represents a shift from measuring actual emissions to monitoring the performance of the emissions control technology. If the problem arises from the degradation of (federally mandated) emissions control technologies, the solution of another (federally mandated) OBDII technology may prove far from perfectly effective.

This opens the door to search for more flexible policy approaches. Combining RSD, OBDII, and inspection stations may offer efficiency gains. Instituting an emissions fee instead of emissions standards may also bring efficiency gains, but it does not address the fundamental measurement problems associated with vehicle emissions. If that policy problem is a market failure arising from the negative externality of vehicular emissions, and monitoring proves especially difficult, a mixed system of regulation and market-based incentives may present still further advantages. Harrington and McConnell (1999) usefully discuss alternative liability and property rights regimes in a Coasian context. Assigning liabilities for emissions control to other parties (e.g., vehicle manufacturers via warranties) and allowing for transfer of liability can leverage market forces while also reducing transaction costs associated with pollution control by targeting policy at a more concentrated group (e.g., manufacturers).

\textsuperscript{15} NRC (2001) notes that “typically, less than 10% of the fleet contributes more than 50% of the emissions for any given pollutant.”

\textsuperscript{16} Notice how one of the biggest advances in I&M program cost effectiveness came from the OBDII test, which is available because of regulating the producers.
Effectively pricing the pollution for numerous mobile-source emitters when their emissions are difficult to detect involves both costly monitoring and a recognition that the externality arises from the quantity, timing, and location of the emissions – not the concentration of emissions at a particular moment. Even an approximate Pigouvian tax would weight the emissions concentration by the amount of driving in the area. The technological innovations of primary interest to policymakers should be those that help put a price on emissions: monitoring emissions and levying fees. Once the market failure is corrected, the motorists will search for affordable behavioral changes or technological improvements to avoid the fees. The regulators need not “pick winners” for pollution control technologies. But they do need to explore ways to better detect emissions and price them. If the entitlement to pollute others’ airshed anonymously is removed and privacy concerns ameliorated, significant improvements can be made to reducing the uncertainties and to improving the cost-effectiveness of the air quality improvements. A revenue-generating market-based approach like Pigouvian taxes also has the potential to address some equity concerns via other targeted subsidy programs (see, e.g., Walls and Hanson 1999).
Bibliography


California Air Resources Board (CARB). 2000. *Final evaluation of California’s enhanced vehicle inspection and maintenance program (Smog Check II)*. Sacramento, CA.


