An Article Submitted to

The B.E. Journal of Economic Analysis & Policy

Manuscript 3246

Discretion and Manipulation by Experts: Evidence from a Vehicle Emissions Policy Change

Lamar Pierce* Jason A. Snyder†

*Washington University in St. Louis, pierce@wustl.edu
†University of California, Los Angeles, jason.snyder@anderson.ucla.edu

Copyright ©2012 De Gruyter. All rights reserved.

Discretion and Manipulation by Experts: Evidence from a Vehicle Emissions Policy Change*

Lamar Pierce and Jason A. Snyder

Abstract

Environmental regulation seeks to limit pollution through strict emissions thresholds for existing cars, yet it remains unclear how frequently inspectors enforce these regulations and what impact test manipulation has on policy efficacy. We demonstrate (1) that there is a distinct discontinuous drop in the distribution of emissions results at the regulatory threshold (2) that when the state tightens emissions standards, over 50% of the vehicles newly at risk for failure now pass instantaneously after the regulation changes. These improvements cannot be explained by legitimate repairs but are consistent with facilities exploiting procedural discretion in order to help consumers evade the strengthened regulations.

KEYWORDS: Fraud, Corruption, Forensic Economics, Environmental Policy, Moral Hazard, Regulation, Ethics

*Thanks to seminar participants at Washington University St. Louis, University of California Los Angeles, Carnegie Mellon, and the Southern California Conference in Applied Microeconomics. Volodymyr Bilotkach and Matt Kahn provided us with thoughtful comments. John Elder provided excellent copyediting. All mistakes are ours alone. Lamar Pierce is an associate professor at the Olin Business School, Washington University in St. Louis. He is the corresponding author and can be reached at pierce@wustl.edu or at One Brookings Drive Box 1133, St. Louis, MO 63130. Jason Snyder is an assistant professor at the Anderson School of Management, UCLA.

I. Introduction

Vehicle emissions are one of the most important determinants of environmental quality worldwide, accounting for over half of carbon monoxide emissions, 29% of hydrocarbon emissions, and 10% of total suspended particulate emissions in the United States (Ernst et al. 2003; Currie and Walker 2011). The severity of these externalities makes government regulation necessary for optimally limiting mobile source emissions. While one effective regulation is to control emissions levels in new cars, governments must also inspect older cars to ensure that deterioration does not generate excessive pollution.[1] This regulatory testing of used cars typically employs strict emissions thresholds to assign a passing or failing status to a vehicle. Under strict application of the regulation, all vehicles emitting particulate levels above the threshold for that pollutant would be repaired or decommissioned. When the government desired a lower level of pollution, it could simply lower the thresholds. Marginal vehicles that had barely passed prior to the threshold decrease would then fail the emissions inspection and either be repaired or be replaced by cleaner vehicles. This marginal change in emissions thresholds would thereby improve air quality.

In a world with precise regulatory enforcement, the marginal benefit of such a policy change would be easily predictable. Yet agency concerns pervasive in the regulatory apparatus suggest otherwise. In the United States, most states delegate emissions inspections to private sector facilities.[2] The historical debate on privatization has largely revolved around trade-offs between the operational efficiency of the private sector and the reduced incentives for fraud in the public sector (Lazare 1980; Voas and Shelley 1995). Private inspection facilities are responsible for testing vehicles to ensure they are below the emissions thresholds and for reporting those vehicles that are not. Inspectors at these facilities have considerable discretion in what constitutes a fair test. In order to avoid reporting inaccurate results, they are allowed to stop or repeat any test they feel does not accurately represent the vehicle's true emissions. While this discretion may avoid some false positive results, it also provides considerable opportunity for moral hazard on the part of the expert agent. This moral hazard could produce behavior of questionable legality, as when an inspector repeatedly tests a marginal vehicle until it finally appears to

[1]Davis (2008) and Currie and Walker (2011) show that other policies, namely driving restrictions and E-ZPass, can also reduce total emissions and improve health.

[2]A survey of state websites found that 27 states outsource inspection to multiple licensed private firms, 11 use state-run facilities, one uses both state-run and private facilities, and one uses a licensed private monopoly.

be passing. It may also produce outright fraud, as when the inspector substitutes a cleaner vehicle in the test or diverts exhaust away from the testing equipment.[3]

There is ample anecdotal (Lambert 2000; States News Service 2010) and academic (Hubbard 1998, 2002; Pierce and Snyder 2008; Oliva 2012) evidence suggesting that private sector inspectors fraudulently pass vehicles that should fail in exchange for bribes or implicit promises of future business. While existing studies estimate *average levels* of emissions testing fraud, we take a *marginal approach* by examining only those vehicles which tightened emissions standards have put newly at risk of failing. More specifically, our paper examines how severely these agency concerns undermine the ability of a state government to decrease pollution by tightening emissions standards. Understanding this impact is critical for policy because it helps us to predict the marginal effect of inspector discretion on future policy changes in mobile emissions standards, rather than estimating the average effects, as has been done in previous work. Our work further contributes to research on the gaming and efficiency of "notched" policies (Blinder and Rosen 1985; Sallee and Slemrod 2010) and to the growing field of forensic economics (Zitzewitz 2012).

Using data from private emissions facilities in New York State, we examine the impact of the agency problem on the state's implementation of stricter emission standards in 2003. A key dilemma in uncovering the extent of this manipulation is that it is intentionally kept secret. Although it is difficult to determine if any individual test is fraudulent, the distribution of all test results can offer evidence of systematic manipulation. We focus on each vehicle's first test of each year, since the first test is when the inspector can first observe the failing condition of a vehicle prior to diagnosis and repair. Absent manipulation, one would expect the distribution of these initial emissions test scores to be continuous throughout the domain of possible scores. In fact, we find sharp discontinuities at the regulatory thresholds.[4] It looks as if many vehicles which would have had to be repaired in order to pass have simply disappeared.

We investigate what happens to these missing vehicles by studying how the distribution of initial test scores changes when the regulation becomes more stringent. We find that when the stringency of the regulation is tightened for two of the pollutants (hydrocarbons and nitrogen oxide), for approximately 50% of the vehicles that would have passed the old test but are now at risk of failing the new one, test results have somehow shifted to the passing range. The observed distributions are consistent with inspectors illegally helping these vehicles pass.

[3]We note that while the emissions testing facility is not defrauding the customer, it is nevertheless engaging in fraud in the same way that financial auditors do when validating false financial statements for their clients.

[4]For other uses of regression discontinuity in forensic economics, see Duggan and Levitt (2002), Wolfers (2006), Snyder (2010, 2012), and Forbes et al. (2011).

An alternative explanation is that, in response to the new regulation, vehicles newly at risk are preemptively repaired prior to their first test. Our results cast doubt on this hypothesis by demonstrating a dramatic shift of vehicles from above to below the new threshold immediately after the policy change. We observe this large shift in tests occurring within a day and even within several hours of the policy shift. If these vehicles were being preemptively repaired, we would expect a more gradual change in the test results, taking into account the time needed to schedule and carry out repairs. Furthermore, there is little incentive for the owner to request legitimate preemptive repairs, since a second test following the repair of a car that has failed is free. This evidence casts considerable doubt on the alternative hypothesis of preemptive repair and suggests, instead, that a substantial number of vehicles pass as a result of inspectors abusing their discretion in direct response to the increase in regulatory stringency.

Finally, we find no evidence that owners of automobiles newly at risk game the policy changes by having their cars inspected right before the change. We find that older vehicles—those more likely to fail the new threshold—were no more likely than newer cars to be inspected in the days leading up to the policy change.

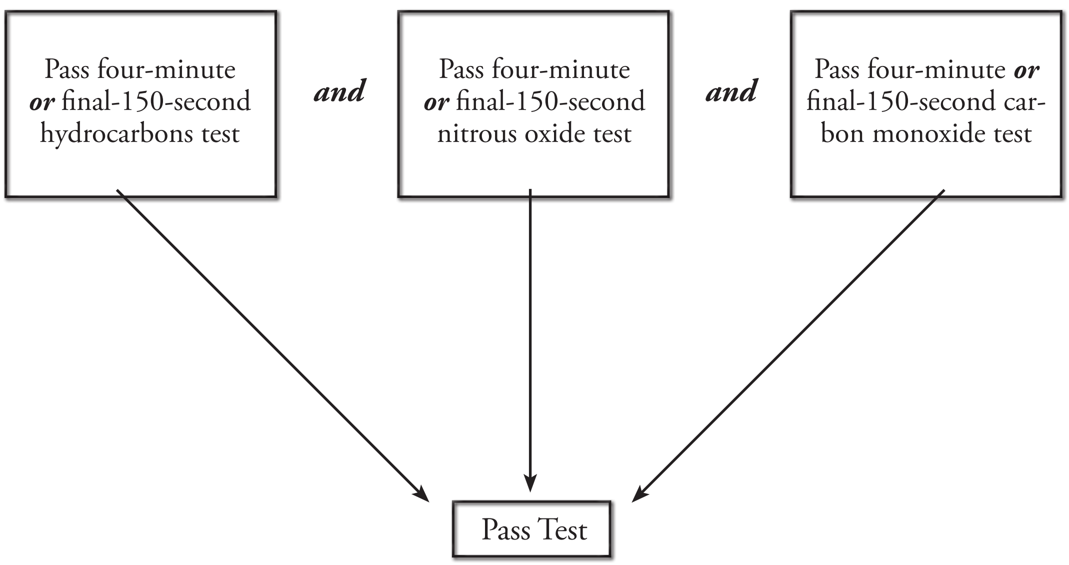
This paper proceeds as follows: Section II describes the market for emissions testing; Section III describes our data; Section IV provides our methodology and results; Section V concludes.

II. The market for emissions testing

The federal Environmental Protection Agency (EPA) mandates that those states with serious air pollution problems must institute vehicle emissions programs, yet leaves the implementation of these programs to those state governments. Some states directly test vehicles at state-owned facilities, but many outsource some or all testing to licensed privately owned firms. Emissions inspectors working at these private facilities are legally required to follow strict testing procedures, but there is still ample opportunity to cheat. First, the procedures themselves allow for some discretion in how the tests are conducted. Second, the dynamometer-based tailpipe testing common in many regions allows skilled mechanics to make temporary adjustments that allow almost any vehicle with excess emissions to pass an emissions test without addressing the underlying causes of the excess pollution. Third, even the most polluting cars can be certified clean when inspectors simply substitute other cars during the test. Evidence from Hubbard's (1998) study of California inspections suggests that such fraud is quite common, while a 2001 covert audit program in Salt Lake City, Utah found nearly 10% of the facilities overtly testing one car in place of another (Groark 2002). Massachusetts found that vehicles retested by the state had substantially higher levels of emissions than had been reported by the testing facilities.[5]

In our focal state, a driver registering a vehicle weighing less than 8,500 pounds and newer than 1981 must have it tested for hydrocarbons (HC), carbon monoxide (CO), and nitrogen oxide (NOx), choosing any licensed facility to conduct the test. In order to pass, the vehicle's emission score for each of these pollutants must fall below a given threshold. Two measurements are taken for each pollutant: one covering the full four-minute dynamometer test and one covering only the last 150 seconds of the test. As long as the vehicle passes one of these measures, it passes that part of the test. If a vehicle were to fail the four-minute hydrocarbons test, for example, but at least pass the 150-second test, it would be given a passing score on hydrocarbons. The only way to fail is to fail both of these tests.[6] We represent these pass criteria in Figure 1.

Figure 1: Passing an emissions inspection



While many areas in the United States now test post-1996 model cars using OBDII, an on-board diagnostic system, the less accurate and more easily manipulated tailpipe testing with a dynanometer is still used in much of the world, including much of the United States.[7] Since vehicles made prior to 1996 generate con-

[5]This report, published by the Massachusetts Office of the Inspector General, can be found at http://www.mass.gov/ig/publ/emissrpt.pdf.

[6]These two measures are highly correlated, although the four-minute tests are easier to pass and thus tend to be the binding constraints for a passing result.

[7]Thirty-four states still use tailpipe testing, although most tailpipe tests are now restricted to pre-1996 vehicles, many of which are the worst polluters in operation.

siderable mobile source pollution, dynamometer testing is still a critical means for providing environmental regulation even in states with OBDII testing.

A licensed inspector conducting a dynamometer test follows a state-prescribed sequence that measures emissions output at different engine loads (revolutions per minute). The inspector is allowed two tests to pass a vehicle, with the second chance designed to account for margin of error in test results and for the possibility that the engine was not properly warmed up prior to testing.[8] Inspectors strictly following protocol would conduct the first test, the results of which would be provided to the state via electronic transmission. If the car passes, it is issued an inspection sticker and is then legal for registration. If the car fails the first test, the inspector can choose to run a second test. Cars passing the second test are issued stickers and can legally register.

Cars that fail the second test must be repaired before registration, with repairs reported to the state prior to the vehicle being retested at the customer's choice of facility. Repairs may involve simple (but time-intensive) engine seal replacement or more expensive catalytic converter replacement. These repairs often result in a vehicle passing its retest, since they correct the fundamental mechanical problem causing excessive pollutants. In some cases, however, the car will continue to fail, because either the mechanical problems are too extensive to fix or the repairs are too expensive for the customer to afford. Consequently, if the vehicle has received over $450 in repairs, the owner is allowed a one-year exemption from the Department of Motor Vehicles (DMV). The facility must retain receipts for these repairs on-site in case of a DMV audit. The facility could potentially write fraudulent receipts, but the risk of more serious tax fraud penalties makes this behavior unlikely. If a vehicle has already received a one-year exemption and still cannot pass, the owner can either junk it or resell it to a region with less stringent emissions requirements.[9]

The loophole in this sequence is that the inspector is allowed to abort the first test twice, just in case he or she fails to follow the designated RPM trace or the testing machine malfunctions, both of which are unusual events. Aborted tests do not have to be reported to the state. So while this option may keep a few cars from failing unfairly, it also makes it easy for inspectors to avoid failing a vehicle. They are warned that a car is likely to fail—and how near it is to the threshold—before they have to report any results to the state.[10]

[8]Catalytic converters are ineffective at low temperatures.

[9]This option has interesting implications for the market for heavy-emitting vehicles. Davis and Kahn (2008), for instance, show that NAFTA caused many high-emissions vehicles to move to Mexico, where they replaced cars with even worse emissions.

[10]In some cases, the inspector might not even need pretesting to predict failure, in which case he or she could illegally alter the car before starting the test.

An inspector willing to fraudulently pass vehicles can stop the test and make temporary adjustments to change the results. The inspector may use fuel additives, adjust the tailpipe probe, or divert exhaust before it reaches the tailpipe. Even more nefariously, he or she may use a technique called "clean-piping," conducting the official test on another car known to be clean.[11] In theory, a mechanically inclined customer could make some of these adjustments before coming in for the inspection, but since inspectors are required to check for such tampering, it is difficult for customers to cheat without the complicity of the mechanic.

An inspector intending to help a gross polluter pass could just let the vehicle fail the first test, then run a second test after making illegal adjustments. Yet this is risky, as the state might well take note of such different test results without any reported repairs to account for them. It is much easier and less risky for the inspector to simply abort the first test, make the illegal adjustments, and report the fraudulent passing results as the first test.

Inspectors have strong incentives to do things like this. Hubbard (2002) found that customers are more likely to return to inspection stations that have passed them. Firms in the emissions testing market tend to profit from passing older cars, in particular, as this ensures that cars most likely to need future repairs will remain on the road. Customers whose cars fail emissions tests are likely either to retest the vehicle at another facility[12] or to buy new or newer cars that need little if any repair work. In testimony to the power of these incentives, the California Bureau of Automotive Repair noted that "it appears, based on BAR enforcement cases, that some stations improperly pass vehicles to garner more consumer loyalty for delivering to consumers what they want: a passing Smog Check result" (California Bureau of Automotive Repair 2011, p. 22).

In contrast, the incentives to fail a vehicle are weak, even for facilities that might retain the customer for immediate repair work. Emissions repair bills are limited to the $450 necessary to receive a one-year waiver, which is much less than the annual service and repair bill the facility could charge in subsequent years. Edmunds.com estimates the annual service and repair costs of a five-year-old Chevrolet TrailBlazer, for example, at $2,089, with older vehicles having even higher annual revenue potential.

There is a clear cost to society of allowing polluting cars to pass emissions tests: significantly greater air pollution in urban areas. The three tested pollutants—CO, HC, and NOx—all have proven health consequences. Carbon monoxide, an odor-

[11]The use of clean-piping is further discussed in Hubbard (1998) and Oliva (2012) and has been confirmed anecdotally.

[12]This is similar to the "audit shopping" studied in the accounting literature (e.g., Davidson et al. 2006). Bennett et al. (2013) find that a higher number of proximate emissions facilities increases both a facility's pass rate and car owners' switching behaviors.

less poisonous gas, inhibits the transport of oxygen from blood into tissues and can cause general difficulties in the cardiovascular and neural systems (Utell et al. 1994). HC and NOx, when combined in the presence of sunlight, form ground-level ozone that can aggravate respiratory problems, especially in children, and may cause permanent lung damage (Utell et al. 1994). A 10-year study of children, conducted by the University of Southern California, linked air pollution to higher school absenteeism due to respiratory problems, reduced lung function growth, and asthma (Gauderman et. al 2002). Health costs from vehicle emissions were estimated to be between $29 billion and $530 billion in 2001 in the United States (U.S. Environmental Protection Agency 2001).

III. Data

Our dataset comes from the New York Department of Environmental Conservation. In New York, emissions testing is conducted by licensed private firms. We use vehicle inspections conducted on gasoline-powered vehicles in downstate New York[13] under 8,500 pounds just before and after the state implemented a policy change on April 1, 2003 that lowered the threshold for two of the measured pollutants. We segment our sample by vehicle model years because the policy decreased the threshold most severely for vehicles with model years between 1983 and 1990, less so for vehicles manufactured between 1991 and 1995, and not at all for vehicles manufactured in 1996 or afterward. Furthermore, we limit our sample to vehicles tested three months before and after the policy shift.[14]

Our data include vehicles owned by individuals, corporations, fleets, and government agencies, although we are unable to directly observe vehicle ownership. The data collected during inspections include inspection date, inspection time, vehicle identification number (VIN), facility identifier, inspector identifier, and inspection results. These data allowed us to uniquely identify vehicles, including characteristics such as make, model, year, and odometer reading. The detailed information on the time and location of an inspection and on the vehicle's characteristics allows us to control for most predictors of vehicle deterioration and likely emissions.

[13]There are different regulations for vehicles in upstate New York.

[14]Our results do not change if we alter the window size. We also present results for a two-week window.

IV. Empirical methods and results

To establish evidence of discretionary passing by repair stations, we first examine the entire distribution of emissions results for vehicles tested by the state in 2003. The inspection generates six measures: the amount of HC, NOx, and CO pollutants for both four-minute and 150-second components of the test. A bright-line rule determines whether the vehicle has passed or failed each component of the test. If a measure is below a threshold designated by the state, it passes that component of the test; otherwise, it fails.

We exploit a policy change in April 2003 that lowered the threshold such that many cars that had passed in the previous year would not meet the new criteria. This policy change was an attempt by the state to reduce overall vehicle emissions and airborne particulates by requiring vehicles that had barely passed emissions tests in the previous years to be repaired or replaced. This policy change defines cars into three categories: never pass, pre-pass, and always pass. Never pass vehicles failed under both the old and new policies, pre-pass cars passed only under the old policy, and always pass vehicles passed under both policy regimes. Table 1 provides summary statistics on all tests within three months before and after the policy shift, separating the tests into our three model-year samples (1983-1990, 1991-1995, and 1996-2002). Based on observable vehicle characteristics, there appears to be little difference in the vehicles brought in before and after the change in threshold, a point we will focus on later in our identification.

Table 1A: Summary statistics for model-years 1983-1990

	Time period: Three months before policy change		Time period: Three months after policy change	
Variable	Observations	Mean	Observations	Mean
Result	94,945	.853	117,857	.791
Always pass	94,945	.708	117,857	.791
Pre-pass	94,945	.145	117,857	.059
Never pass	94,945	.147	117,857	.150
Odometer	94,945	114,393	117,857	111,267
Model-year	94,945	1988.128	117,857	1988.097

Table 1B: Summary statistics for model-years 1991-1995

	Time period: Three months before policy change		Time period: Three months after policy change	
Variable	Observations	Mean	Observations	Mean
Result	161,115	.919	200,608	.911
Always pass	161,115	.899	200,608	.911
Pre-pass	161,115	.02	200,608	.004
Never pass	161,115	.081	200,608	.085
Odometer	161,115	109,557	200,608	105,443
Model-year	161,115	1993.17	200,608	1993.17

Table 1C: Summary statistics for model-years 1996-2002

	Time period: Three months before policy change		Time period: Three months after policy change	
Variable	Observations	Mean	Observations	Mean
Result	238,564	.979	117,857	.979
Always pass	238,564	.979	117,857	.979
Pre-pass	238,564	0	117,857	0
Never pass	238,564	.021	117,857	.021
Odometer	238,564	58,164	117,857	55,587
Model-year	238,564	1998.701	117,857	1998.713

In Figures 2a and 2b, we show what happens to the distribution of the four-minute and 150-second hydrocarbons results when the threshold is lowered for model-year 1983-1990 vehicles. The dots represent the distribution of test results within three months before the policy shift (between January 1 and March 31 of 2003) and the X's represents the distribution of test results within three months after the policy shift (between April 1 and June 30 of 2003).[15] Vehicles with a hydrocarbon emissions reading just below the earlier (higher) threshold would have passed the hydrocarbon test prior to the policy shift but, without repairs, would have failed that test after the policy shift. Similarly, Figures 3a - 4b present the distributions of test results for nitrogen oxide and carbon dixoide before and after the policy change.

[15] For each density function we calculate the percentage of observations that fall within a given bin. Extreme observations to the right of the distribution have been trimmed. In specification (1), Y is the percentage of observations for a given emissions level. For example, in Figure 2 there are 658 vehicles with a four-minute hydrocarbon reading of 0 prior to the policy change, which represents 0.6% of the entire sample. Therefore, Y=0.006 when X=0.

Figure 2a: Impact of policy change on distribution of results of four-minute hydrocarbons tests for model-years 1982-1990

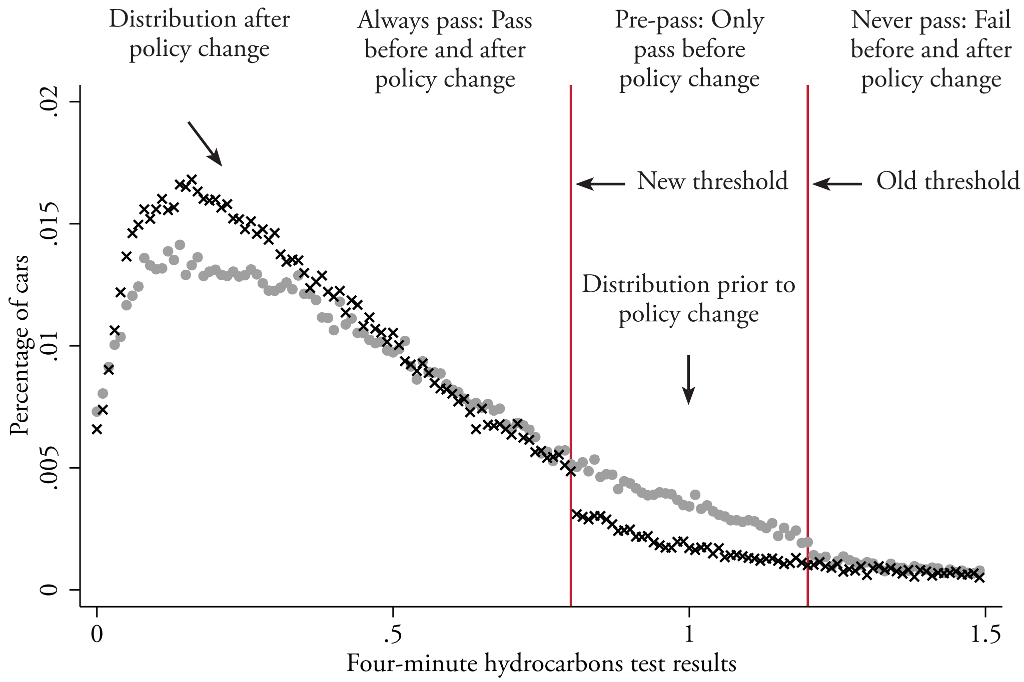


Figure 2b: Impact of the policy change on distribution of results of 150-second hydrocarbons tests for model-years 1983-1990

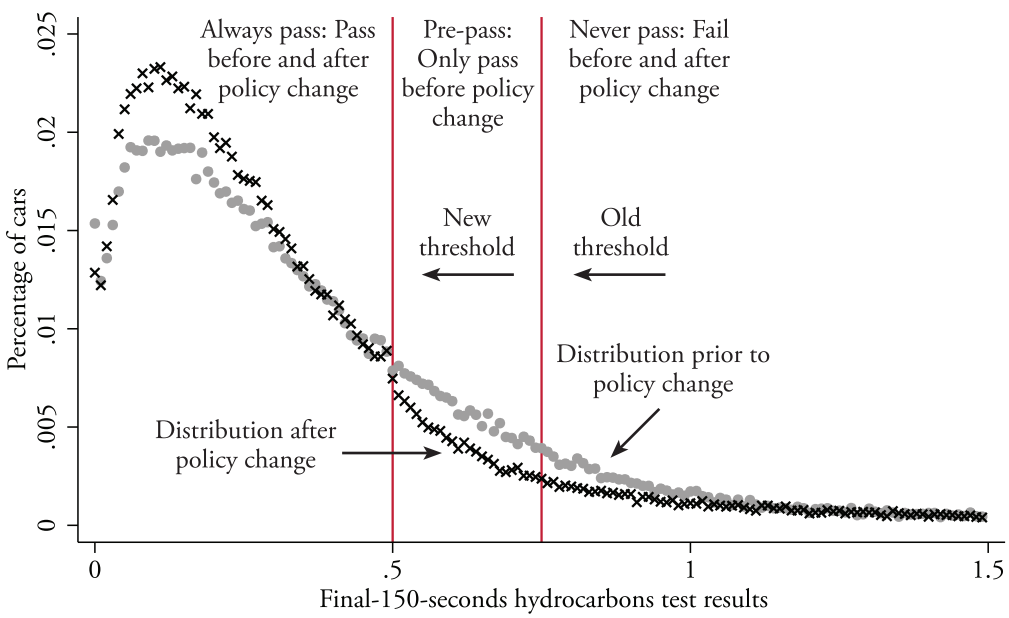


Figure 3a: Impact of policy change on distribution of results of four-minute nitrogen oxide tests for model-years 1983-1990

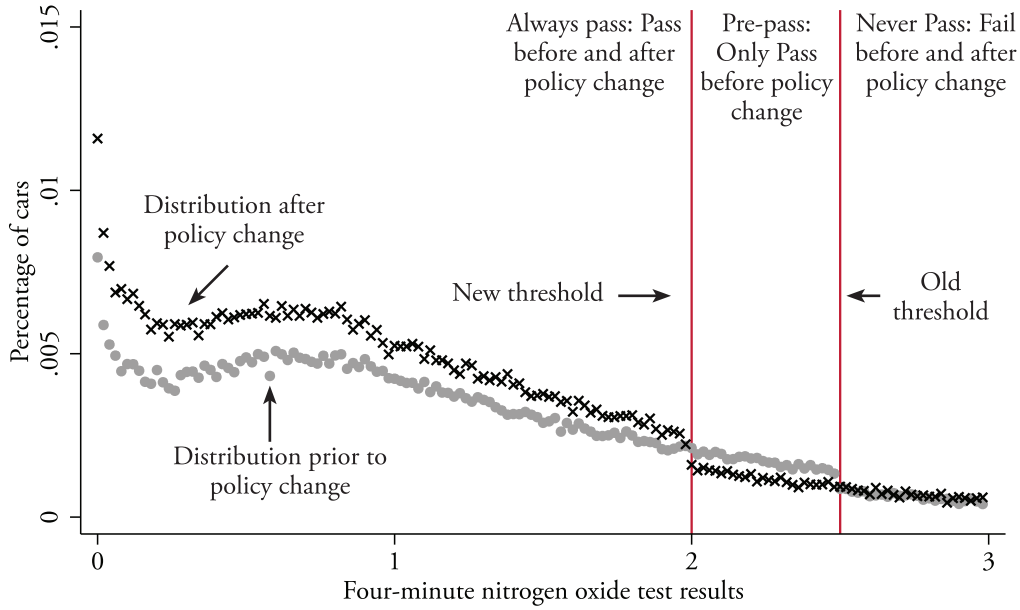


Figure 3b: Impact of policy change on distribution of result of 150-second nitrogen oxide tests for model-years 1983-1990

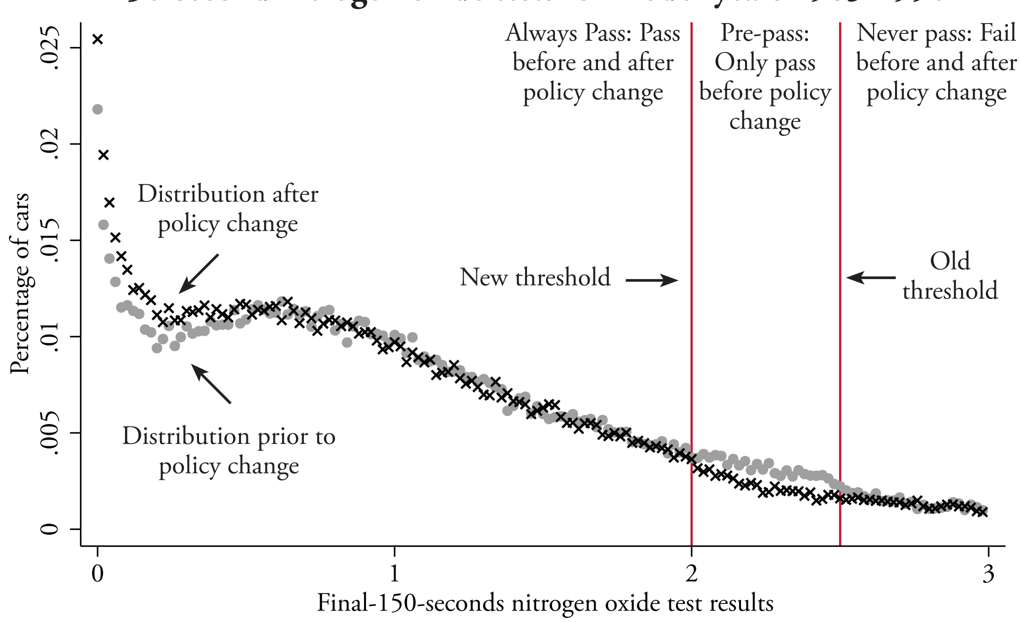


Figure 4a: Impact of policy change on distribution of results of four-minute carbon monoxide tests for model-years 1983-1990

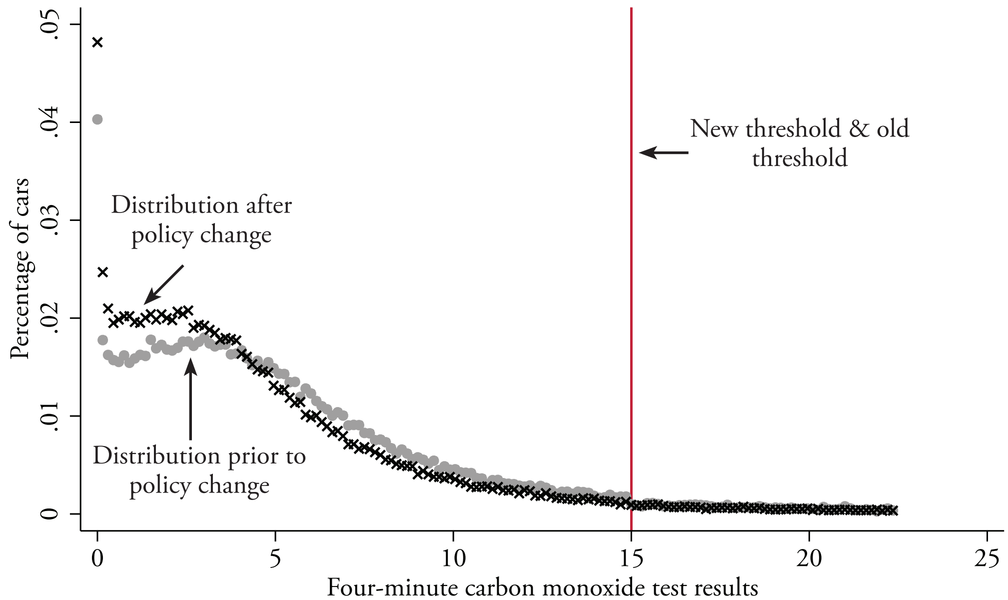
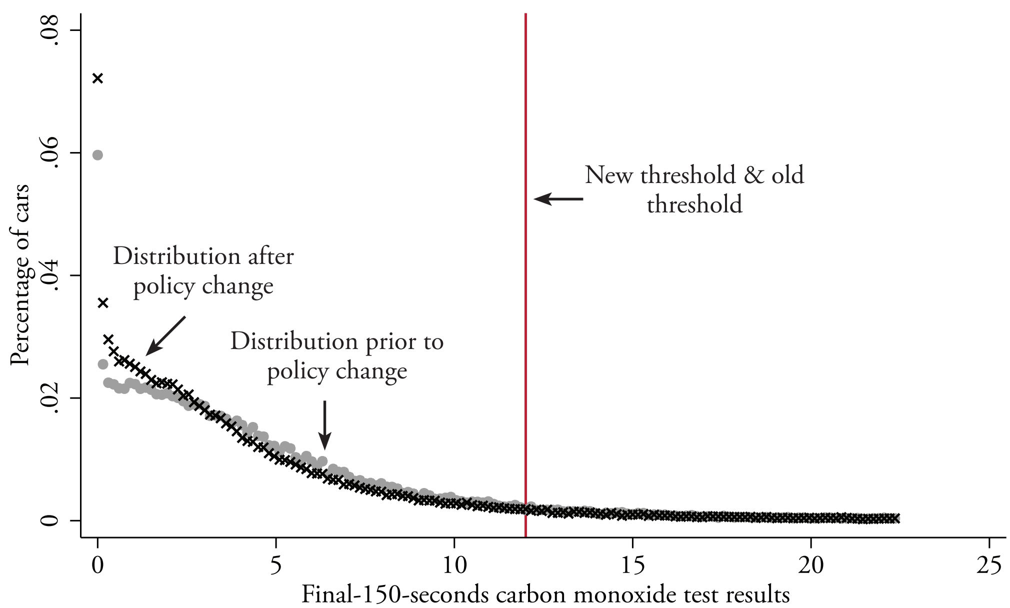


Figure 4b: Impact of policy change on distribution of results of 150-second carbon monoxide tests for model-years 1983-1990



For the tests prior to the change in regulations, the figures show small discontinuities at the old thresholds.[16] But for hydrocarbon and nitrogen oxide tests after the change in regulations, the figures clearly show significant discontinuities at the new regulatory thresholds. Since our data represent the first tests conducted to completion by the inspectors, these shifts in the discontinuity cannot have resulted from repairs conducted in response to failed tests, but instead likely represent inspector manipulation—possibly in response to aborted and therefore unreported tests. One can see that the magnitude of this manipulation has been significant, since the increased mass in the distribution is at very low emissions levels. This suggests that inspectors are unlikely to be simply warming up vehicles or adjusting the tailpipe probe in order to nudge failing cars to the passing side of the threshold; it is much more likely that they are using more serious manipulations such as clean-piping (Oliva 2012). We also note that the discontinuities are more defined for the four-minute tests, which reflects the fact that these are the easier of each pair of tests to pass and thus form the binding constraint for a passing test result.

As presented in Table 2, we estimate the significance of these discontinuities for all six emissions measurements, both before and after the threshold changes, using the following specification:

$$(1)\ Y = \alpha + \tau \cdot T + \sum_{i=1}^{3} [\beta_{i,1}(X-c)^n + \beta_{i,2} \cdot T \cdot (X-c)^n] + \varepsilon$$

In equation (1), T is an indicator equal to one if the observations of the density function are to the right of the thresholds in Figures 2a-4b. The constant c is used to center the polynomials at the cutpoint in the distribution. We estimate the density function using cubic polynomials estimated separately on each side of the threshold. The magnitude of the discontinuity in each distribution is given by the parameter τ. Our results show that the discontinuity is significant for all the pre-April 1 tests, which use the old thresholds. The discontinuity is also significant for many of the tests after the policy change, which use the new thresholds. The tests are consi stent with the visual evidence in Figures 2a-4b.

[16] These discontinuities in the pretest period are likely due to a combination of leniency and the legitimate attrition of failing cars from operation in prior years.

Table 2a: Estimated discontinuities for model-years 1983-1990 prior to the policy changed

Measure	Passing threshold	Bin size	Time period	Estimated jump at threshold	Number of observations
Four-minutes hydrocarbons	1.2 g/mi	.01 g/mi	January 1 - March 30	-.0017 (.0003)**	150
Final-150-second hydrocarbons	.75 g/mi	.01 g/mi	January 1 - March 30	-.0015 (.0004)**	150
Four-minute nitrous oxide	2.5 g/mi	.02 g/mi	January 1 - March 30	-.0007 (.0001)**	147
Final-150-second nitrous oxide	2.5 g/mi	.02 g/mi	January 1 - March 30	-.0007 (.0002)**	147
Four-minute carbon monoxide	15 g/mi	.15 g/mi	January 1 - March 30	-.0024 (.0003)**	148
Final-150-second carbon monoxide	12 g/mi	.15 g/mi	January 1 - March 30	-.0013 (.0003)**	146

Note: * and ** indicate significance at the 5% and 1% confidence levels, respectively. Parentheses contain standard errors. All discontinuities estimated using equation (1). Slight differences in observations are a function of different bin sizes, which were chosen to make the results approximately comparable. The abbreviation g/mi stands for grams per mile.

Table 2b: Estimated discontinuities for model-years 1983-1990 after the policy changed

Measure	Passing threshold	Bin size	Time period	Estimated jump at threshold	Number of observations
Four-minutes hydrocarbons	.8 g/mi	.01 g/mi	April 1 - June 30	-.0030 (.0004)**	150
Final-150-second hydrocarbons	.5 g/mi	.01 g/mi	April 1 - June 30	-.0030 (.0006)**	150
Four-minute nitrous oxide	2.0 g/mi	.02 g/mi	April 1 - June 30	-.0008 (.0001)**	147
Final-150-second nitrous oxide	2.0 g/mi	.02 g/mi	April 1 - June 30	.0001 (.0002)	147
Four-minute carbon monoxide	15 g/mi	.15 g/mi	April 1 - June 30	-.0020 (.0004)**	148
Final-150-second carbon monoxide	12 g/mi	.15 g/mi	April 1 - June 30	-.0009 (.0002)**	146

Note: * and ** indicate significance at the 5% and 1% confidence levels, respectively. Parentheses contain standard errors. All discontinuities estimated using a seventh-degree confidence interval. Slight differences in observations are a function of different bin sizes, which were chosen to make the results approximately comparable. The abbreviation g/mi stands for grams per mile.

In order to identify manipulation by emissions inspectors, we implement a research design that exploits the sharp policy shift in the regulatory thresholds. Using only those emissions tests within three months of the policy change, we define three categories of vehicle. The first group, *never pass*, is those cars whose emissions results for each of the three pollutants always (both in the focal year and the previous year) exceed both the old and new thresholds on either the four-minute or 150-second measurement.[17] These vehicles would fail the test both before and after the environmental standards were tightened. The second group, *pre-pass*, is those cars for which the emissions scores were between the old and new thresholds. They

[17] As explained earlier, a vehicle passes an emissions test so long as it passes either the four-minute or the 150-second test for each of the three emissions dimensions (HC, NOx, and CO).

would pass if tested before April 1, but would fail afterward. The third group, *always pass*, is those cars whose emissions results were low enough to pass in either regulatory period. Figures 2 and 3 show these three categories for the hydrocarbons and nitrogen oxide tests. Figure 4, which presents the unchanged carbon monoxide threshold, has no *pre-pass* group.

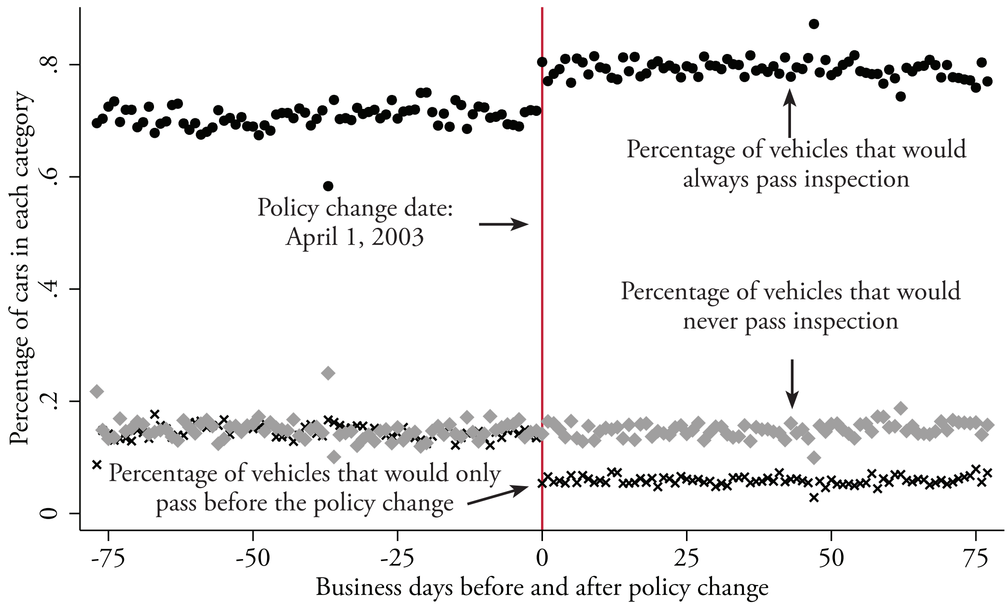
The key to our analysis is to identify the probability that a vehicle falls into each of these three ranges of emissions. If inspectors are not manipulating emissions tests, we would expect the probability for each category to remain approximately constant across short periods of time for the first emissions test observed. The number of cars in the *always pass* category, for example, should be about the same on any two consecutive days. While overall emissions levels may gradually decrease as vehicles are repaired and replaced, this change should be relatively continuous.

If, however, customers or inspectors are manipulating tests in response to the policy change, we can expect a discontinuity between the day before and the day of the policy switch. Inspectors who observe a vehicle falling into the *pre-pass* range on March 31 will allow the test to run to completion and legitimately pass the car. On or after April 1, however, some inspectors, observing a likely *pre-pass* result, would stop the test, manipulate the vehicle into the *always pass* range, and rerun the test, with the car appearing clean in the reported results. We would then observe that the number of *pre-pass* vehicles discontinuously drops on April 1, while the number of *always pass* vehicles discontinuously increases.

Such a discontinuous change is important for identification due to the possibility of preemptive repair. Since legitimately and permanently repairing vehicles takes time to schedule and complete, particularly when parts need to be ordered, it is an unlikely alternative explanation for such a discrete shock. We would likely observe no change to cars in the *never pass* range, as inspectors would manipulate the tests for these vehicles at similar rates both before and after the policy change.

Figure 5 presents the daily frequencies of model-year 1983-1990 vehicles falling into each of the three categories, measured between January 1 and March 31, 2003. The discontinuity in the *always pass* category is immediately evident. On April 1, 2003, the frequency of vehicles below the new threshold discretely increases, the frequency of vehicles in the *pre-pass* range drops dramatically, and the frequency for the *never pass* range appears unchanged.

Figure 5: Daily distribution of test result groups for model-years 1983-1990



In Figures 6a and 6b, we repeat our frequency plot for the *always pass* and *pre-pass* categories using percentages of cars tested before and after the policy change for each hour during which at least 100 cars were tested.[18] These figures address the possible explanation of preemptive repair, since vehicles tested in the morning hours of April 1 are unlikely to have been identified as failing and then legitimately repaired and retested all in the early morning. These vehicles could, however, easily have been identified as failing and then manipulated and retested all in the early morning, using common techniques such as clean-piping. Since these cars would have passed on March 31, they are also unlikely to have been tested a few days earlier and preemptively repaired.[19] The discontinuities for *always pass* and *pre-pass* are still obvious at the hourly level. The visual results for *always pass* and *pre-pass* vehicles in Figures 6a and 6b are consistent with inspectors beginning to manipulate tests for *pre-pass* vehicles exactly when these cars fall below the regulatory threshold: the first hour of business on April 1, 2003.

[18]Hours with fewer tests were highly imprecise. Larger gaps represent weekends.

[19]Vehicle owners have some flexibility on when to test their vehicle, so long as the test occurs prior to the expiration of the previous year's certification.

Figure 6a: Hourly percentage for always-pass for model-years 1983-1990

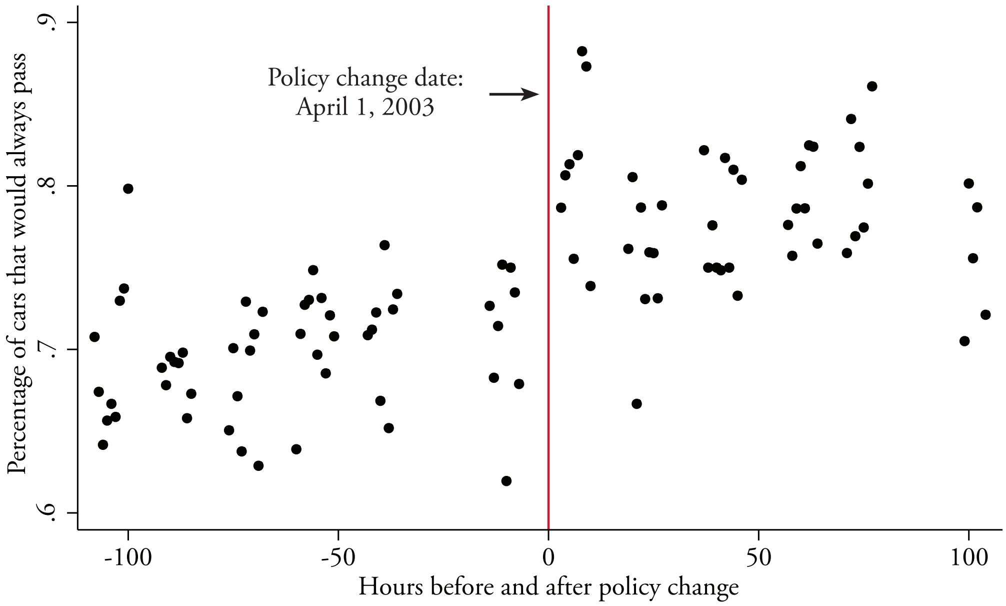
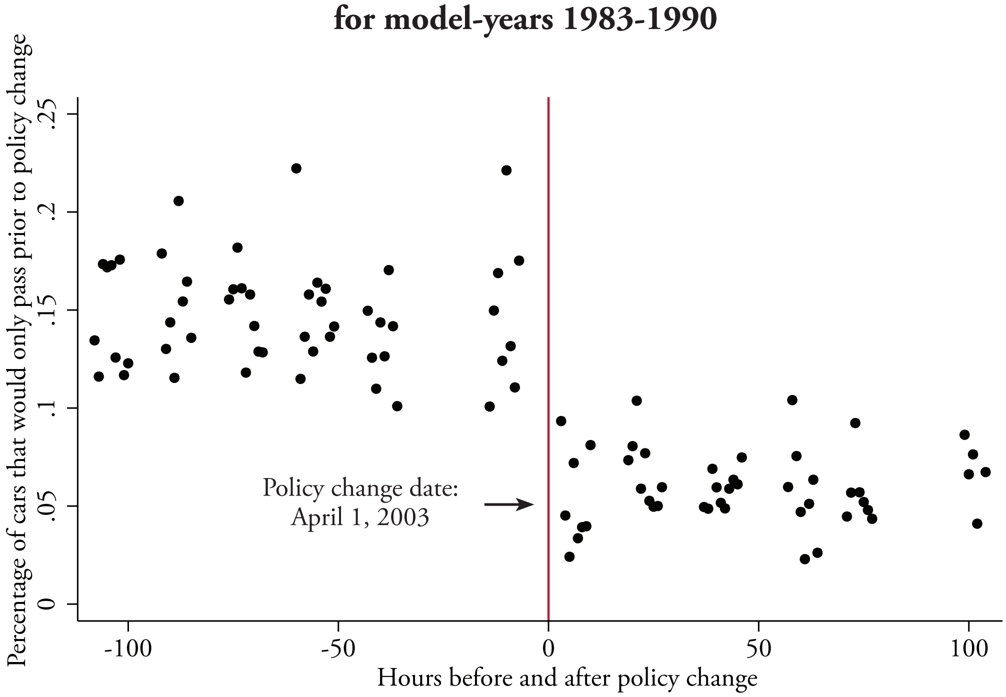


Figure 6b: Hourly percentage for pre-pass for model-years 1983-1990



To statistically test for the presence of a jump on April 1, we run linear probability models measuring the likelihood of vehicles appearing in each of the three categories. We use a standard approach similar to specification (1), regressing a dummy variable indicating whether the vehicle fell into one of the possible results categories on (a) a dummy variable indicating if the test was done after the policy change and (b) relevant control variables. Our specification includes fixed effects for eight-digit vehicle identification number (VIN) as well as a vehicle odometer cubic. The first eight digits of the VIN identify the manufacturer, year, model, and body and engine specifications and thus control for most vehicle-specific characteristics. In our most extensive specifications we use separate cubic time functions on either side of the threshold for polynomial smoothing.[20] We also include controls for time of day (morning, afternoon, or evening) and day of week.[21] The identification strategy involves demonstrating that the treatment variable causes a discrete increase in the probability of falling into the always pass category and a discrete decrease in the probability of falling into the pre-pass category, while controlling for time trends and other potential predictors. For this model, we use the sample of model-year 1983-1990 vehicles within three months of the policy change, since these cars faced the largest increase in test stringency.

We present our results in Table 3. Columns 1 and 2 represent linear probability models predicting the likelihood of falling into the *always pass* category, with Column 2 including cubic polynomial smoothing and control variables. Standard errors are clustered at the facility level. Our results suggest a 9% increase in the likelihood of falling into the *always pass* category starting April 1 and are robust to the inclusion of control variables. This suggests that vehicles are immediately shifted below the new regulatory threshold on the day of the policy change. Columns 3 and 4 represent similar linear probability models predicting the likelihood of a vehicle falling into the *pre-pass* category between the two regulatory thresholds. In each model we observe a 9% drop in the likelihood on the day of the regulatory change. Columns 5 and 6 repeat the linear probability model for the *never pass* category. These results show no change in the likelihood of falling into this category after the policy change. The results from these regression discontinuity models strongly support the visual evidence in Figures 5 and 6 and are consistent with inspectors manipulating emissions tests on *pre-pass* vehicles the moment the new threshold takes effect.

[20] The specification is similar to (1) except that the assignment variable is time and the threshold is April 1, 2003.

[21] Morning is defined as before 10 a.m., afternoon as between 10 a.m. and 3 p.m., and evening as after 3 p.m. Results are robust to these choices.

Table 3: Test results three months before and after the policy change for model-years 1983-1990

Independent variables	Dependent variable: Always pass		Dependent variable: Pre-pass		Dependent variable: Never pass	
	(1)	(2)	(3)	(4)	(5)	(6)
Post policy change	.084 (.003)**	.092 (.008)**	-.086 (.002)**	-.085 (.006)**	.003 (.002)	-.007 (.006)
Cubic polynomial of days	No	Yes	No	Yes	No	Yes
Cubic polynomial of days * Post policy change	No	Yes	No	Yes	No	Yes
Cubic odometer controls	No	Yes	No	Yes	No	Yes
VIN group effects	No	Yes	No	Yes	No	Yes
Day of the week effects	No	Yes	No	Yes	No	Yes
Time of day effects	No	Yes	No	Yes	No	Yes
Manufacture-year effects	No	Yes	No	Yes	No	Yes
Observations	212,802	212,802	212,802	212,802	212,802	212,802
Clusters	3,711	3,711	3,711	3,711	3,711	3,711

Note: * and ** indicate significance at the 5% and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the facility level. Saturday and Sunday are jointly coded as one business day. Time of day is an indicator for morning (before 10 a.m.), mid-day, or evening (after 3 p.m.). The results are robust to this choice.

Table 4 supplements this analysis by estimating the model using only tests from one week before and after the policy change, with similar results.[22] The one difference in this model is an estimated decrease in *never pass* vehicles. This discontinuity could be interpreted to suggest that inspectors increase their leniency even toward the most polluting cars.

[22] We exclude controls for day of the week in this model since we have only two of each in this narrow time period. The number of facilities—and thus the number of clusters—is reduced, since some facilities did not perform inspections during this more narrow time period.

Table 4: Test results one week before and after the policy change for model-years 1983-1990

Independent variables	Dependent variable: Always pass		Dependent variable: Pre-pass		Dependent variable: Never pass	
	(1)	(2)	(3)	(4)	(5)	(6)
Post policy change	.084 (.007)**	.127 (.029)**	-.084 (.005)**	-.079 (.021)**	.000 (.005)	-.048 (.024)*
Cubic polynomial of days	No	Yes	No	Yes	No	Yes
Cubic polynomial of days * Post policy change	No	Yes	No	Yes	No	Yes
Cubic odometer controls	No	Yes	No	Yes	No	Yes
VIN group effects	No	Yes	No	Yes	No	Yes
Day of the week effects	No	Yes	No	Yes	No	Yes
Time of day effects	No	Yes	No	Yes	No	Yes
Manufacture-year effects	No	Yes	No	Yes	No	Yes
Observations	17,069	17,069	17,069	17,069	17,069	17,069
Clusters	3,122	3,122	3,122	3,122	3,122	3,122

Note: * and ** indicate significance at the 5% and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the facility level. Saturday and Sunday are jointly coded as one business day. Time of day is an indicator for morning (before 10 a.m.), mid-day, or evening (after 3 p.m.). The results are robust to this choice.

One could worry that self-selection might be biasing the magnitude of our results. If owners of cars at risk of failing attempt to take the test before the policy change, our parameter might be incorrectly estimated. Intuitively, a driver would need to be unusually sophisticated and well-informed to engage in this strategy. Nevertheless, we present evidence that suggests that people did not rush to get their vehicles tested before the policy changed. If this had been the case, we would expect a significant bump in the number of 1983-1990 models being taken in for inspection just prior to April 1 to avoid the new test which they might not pass.

In Figure 7, we present the daily volume of inspections three months before and after the April 1 policy change; we see no considerable difference. The 1983-1990 models—those most affected by the policy change—show no difference relative to newer vehicles. In fact, the pattern of inspections for those vehicles and for the 1991-1995 vehicles, which had smaller threshold changes, is similar to the pattern of inspections for post-1995 vehicles, for which there was no change in regulation at all.

Figure 7: Daily percentage of vehicles from each model-year group

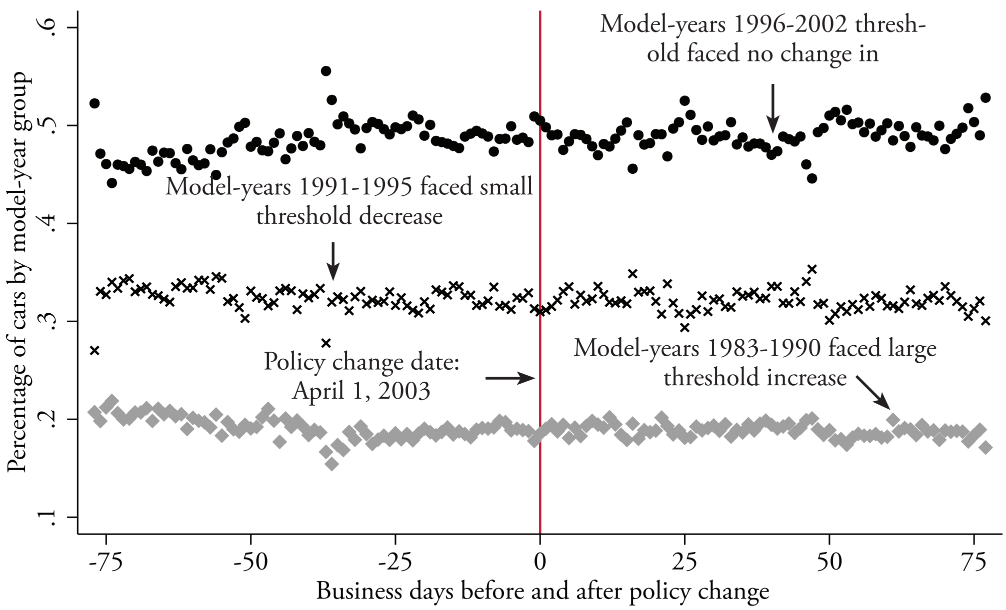


Table 5 follows suggestions by McCrary (2008) to test endogenous sorting more formally; it presents regression discontinuity models similar our prior models but with dummy variables for the three model-year categories as dependent variables.[23] Models with polynomial smoothing in Table 5 show no identifiable discontinuity following the policy change, consistent with our argument that customers are not preemptively testing at-risk cars prior to the policy change. Finally, in unreported results, we find no economically significant differences in the odometer readings, testing weight, or age of the vehicles being inspected before and after the policy change, consistent with the implication from Table 1.

[23]These models exclude vehicle characteristics, which would be endogenous regressors.

Table 5: Predicting the percentage of cars tested each day based on level of exposure to a more stringent policy

Independent variables	Dependent variable: Model-year: 1983-1990		Dependent variable: Model-year: 1991-1995		Dependent variable: Model-year: 1996-2002	
	(1)	(2)	(3)	(4)	(5)	(6)
Post policy change	-.004 (.001)**	-.004 (.010)	-.005 (.001)**	-.001 (.004)	.009 (.001)	.004 (.004)
Cubic polynomial of days	No	Yes	No	Yes	No	Yes
Cubic polynomial of days * post policy change	No	Yes	No	Yes	No	Yes
Observations	1,120,688	1,120,688	1,120,688	1,120,688	1,120,688	1,120,688
Clusters	3,837	3,837	3,837	3,837	3,837	3,837

Note: * and ** indicate significance at the 5% and 1% confidence levels, respectively. Parentheses contain standard errors clustered at the facility level.

V. Conclusions

In this paper we use a regression discontinuity design to identify the strategic manipulation of emissions tests by licensed private inspectors. As state-licensed experts charged with enforcing environmental regulations, inspectors are given discretion to gather and use information in the process of testing for the purpose of ensuring the fairness and accuracy of their results. Yet they can also use this discretion and information to manipulate the test by aborting it in process and then making temporary adjustments or substitutions that allow the vehicle to fraudulently pass. This manipulation is motivated by the moral hazard of attracting and retaining repeat repair-and-service business (Hubbard 2002). The many older cars in our sample provide consistent business for mechanics, who are likely to suffer financially if customers take their business to more lenient facilities or choose to replace their

cars with newer, more reliable ones.
We exploit a policy change that, on April 1, 2003, lowered the pass threshold based on the vehicle model-year, thereby immediately putting hundreds of thousands of previously compliant vehicles at risk for failure.

Given the considerable public welfare implications of increased mobile emissions (Ashenfelter and Greenstone 2004; Currie and Neidell 2005; Currie et al. 2009a, 2009b; Agarwal et al. 2010; Currie and Walker 2011; Fowlie et al. 2012), the argument for granting inspectors the discretion to stop tests in process seems tenuous. Although eliminating this discretion seems an obvious remedy, inspectors motivated by financial gain will likely find other ways to manipulate the system, albeit less precisely as they try to predict which vehicles require pretest manipulation. And given the low number of facilities detected and penalized by the state (States News Service, 2010), current levels of enforcement seem insufficient to deter manipulation. Our results therefore seem to justify increased investment in the monitoring of private inspection facilities, particularly when the state is implementing decreased thresholds that immediately put new cars at risk for failure.

Privatizing government regulatory enforcement may yield efficiency gains from competition and customer choice, but these gains come with considerable social cost through moral hazard for leniency. The state must give experts discretion in order to make full use of their knowledge and expertise, but must also acknowledge that this discretion can be abused for profit. Inspectors' discretion has strong implications for the state's ability to achieve regulatory goals by strengthening environmental standards; the private-market experts charged with inspection may have little incentive to enforce tighter standards and significant opportunity to circumvent them. Privatization in enforcement may therefore yield an additional cost not normally considered in outsourcing decisions: losing the option value of ratcheting up regulatory standards in the future.

It is important to note that the test manipulation we observe could actually serve the public welfare if the threshold were too low, such that the welfare benefits of repairing the marginal *pre-pass* vehicles were less than the costs. We find this to be unlikely in our setting, given the welfare calculation from Currie et al. (2009). Their estimates that increased standards in New Jersey saved 449 infant lives worth $2.2 billion annually suggest that the benefits to New York far outweigh the foregone costs of legitimate repairs. Given that our model parameters suggest that 9% of all fleet vehicles were illegitimately passed after the policy change, this welfare benefit could fund a lot of repairs for these polluting vehicles. For example, it could fund $1,100 in repairs for two million cars, which would be 9% of a total fleet of 25 million vehicles, far more than there actually are in New Jersey or downstate New York. Given such numbers, it is difficult to believe that manipulating the test results for these marginal vehicles could bring a net public welfare gain.

We note that, while our identification strategy involves estimating the fraudulent passing of only those vehicles newly affected by the policy change, the implications of our results go far beyond these "marginal" polluters. If emissions inspectors were only passing marginal polluters because they were close to the threshold, then the social welfare impact of this fraud might be small. It would be somewhat analogous to police officers giving warnings to drivers only slightly exceeding the speed limit while arresting more serious offenders.[24]

While we cannot directly observe leniency toward more extreme "gross" polluters, it is hard to believe that the same strong incentives that motivate inspectors to fraudulently pass marginal polluters do not also motivate leniency toward gross polluters. Furthermore, extensive remote testing procedures by analytical chemists at the University of Denver show that the worst five percent of all vehicles produce half of total emissions, with the worst one percent producing over twenty percent (Stedman 2002; Stedman et al. 2009).[25] These distributions are clearly inconsistent with the emissions distributions in Figures 2-4, suggesting that many gross polluters are indeed being fraudulently passed and that the social welfare impact of fraudulent testing may be much larger than we can identify in our analysis.

While this paper's implications for environmental policy and fraud are discouraging, there might seem to be a positive implication of our results; namely, that experts are willing to work hard to serve their clients, which seems to contrast with some evidence of moral hazard by experts, particularly car mechanics (Schneider 2012). But in reality, our focal expert—the inspector—serves two principals: the government that authorizes and licenses him and the customer who selects and pays him. What our results may therefore suggest is that, much like the supervisor in Tirole's (1986) three-tiered agency model, any expert willing to break the law will serve whichever player offers the best reward. So even if our results suggest that licensed experts are indeed willing to break laws to serve their clients, the welfare implications of this are not necessarily positive. Equivalent client-serving behaviors—such as doctors or pharmacists illicitly providing narcotics, attorneys suborning perjury, or private schools facilitating test fraud—all carry similarly troublesome welfare implications.

Finally, our results raise questions about whether emissions testing programs are

[24]There are notable differences, however, between police leniency and inspector leniency. First, police do not have the strong incentives to provide leniency that inspectors have. Second, drivers are unable to choose which police officer pulls them over, nor do police officers need to compete with one another for this opportunity. Finally, police leniency for marginal offenses is arguably efficient because marginal speeding tickets are likely to be thrown out in court if challenged unless the officer testifies—time the officer could better spend on duty stopping more egregious violations.

[25]These statistics are taken from multiple states with active testing programs, including California, which has one of the most strict and active programs. To the best of our knowledge, such tests have never been conducted in our focal state.

worth their considerable cost, given the widespread fraud suggested here. We note that the considerable welfare gains from reduced emissions (Currie et al. 2009b) and the relative efficiency of reducing pollution through mobile versus other sources (Fowlie et al. 2012) suggest that, even with such reduced efficiency, emissions testing is likely to be welfare-enhancing. The larger implication, however, is for how to adjust these programs to make them both more effective and more fair to the drivers who are unwilling to solicit fraud yet who suffer its environmental and health costs. Perhaps the most promising direction is technology that takes discretion away from highly incentivized private inspectors and also reduces the time burden to vehicle owners. Remote sensing detectors, which can be installed on freeway ramps, seem like promising solutions to these problems if they can be applied beyond the limited number of states that currently use them (Burgard et al. 2006).

References

Agarwal, Nikhil, Chanont Banternghansa, and Linda Bui. 2010. "Toxic Exposure in America: Estimating Fetal and Infant Health Outcomes from 14 Years of TRI Reporting." *Journal of Health Economics* 29: 557-574.

Ashenfelter, Orley, and Michael Greenstone. 2004. "Using Mandated Speed Limits to Measure the Value of a Statistical Life." *Journal of Political Economy* 112: 226-267.

Bennett, Victor, Lamar Pierce, Jason Snyder, and Michael Toffel. 2013. "Customer-Driven Misconduct: How Competition Corrupts Business Practices." *Management Science*. In press.

Blinder, Alan S., and Harvey S. Rosen. 1985. "Notches." *American Economic Review* 75: 736-747.

Burgard, Daniel A., Gary A. Bishop, Ryan S. Stadtmuller, Thomas R. Dalton, and Donald H. Stedman. 2006. "Spectroscopy Applied to On-Road Mobile Source Emissions." *Applied Spectroscopy* 60: 135A-148A.

California Bureau of Automotive Repair. 2011. *Final statement of reasons for Article 5.5 of Chapter 1 of Division 33 of Title 16 of the California Code of Regulations*. June 10.

Chay, Ken Y., and Michael Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession." *Quarterly Journal of Economics* 118: 1121-1167.

Collins, Glenn. 2003. "A Week into April, and Wondering Where that Snow Shovel is Stored." *New York Times*, April 8: D1.

Currie, Janet, Eric A. Hanushek, Megan E. Kahn, Matthew Neidell, and Steve G. Rivkin. 2009a. "Does Pollution Increase School Absences?" *Review of Economics and Statistics* 91: 682–694.

Currie, Janet, and Matthew Neidell. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics* 120: 1003-1030.

Currie, Janet, Matthew Neidell, and Johannes F. Schmieder. 2009b. "Air Pollution and Infant Health: Lessons from New Jersey." *Journal of Health Economics* 28: 688-703.

Currie, Janet, and Reed Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics* 3: 65-90.

Davidson, Wallace, III, Pornsit Jiraporn, and Peter DaDalt. 2006. "Causes and Consequences of Audit Shopping: An Analysis of Auditor Opinion, Earnings Management, and Auditor Changes." *Quarterly Journal of Business and Economics* 45: 69-87.

Davis, Lucas W. 2008. "The Effect of Driving Restrictions on Air Quality in Mexico City." *Journal of Political Economy* 116: 38-81.

Davis, Lucas W., and Matthew E. Kahn. 2010. "International Trade in Used Vehicles: The Environmental Consequences of NAFTA." *American Economic Journal: Economic Policy* 2: 58-82.

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

Ernst, Michelle, James Corless, and Ryan Greene-Roesel. 2003. "Clearing the Air: Public Health Threats from Cars and Heavy Duty Vehicles—Why We Need to Protect Federal Clean Air Laws." Washington, DC: Surface Transportation Policy Partnership. http://www.transact.org/report.asp?id=227.

Forbes, Silke, Mara Lederman, and Trevor Tombe. 2011. "Quality Disclosure Programs with Thresholds: Misreporting, Gaming, and Employee Incentives." Unpublished working paper, University of California, San Diego. http://weber.ucsd.edu/~sjanusze/www/FLT_Gaming.pdf

Fowlie, Meredith, Christopher Knittel, and Catherine Wolfram. 2012. "Sacred Cars? Cost-Effective Regulation of Stationary and Non-stationary Pollution Sources." *American Economic Journal: Economic Policy* 4: 98-126.

Gauderman W., F. Gilliland, H. Vora, E. Avol, D. Stram, R. McConnell, D. Thomas, F. Lurmann, H. Margolis, E. Rappaport, K. Berhane, and J. Peters. 2002. "Association Between Air Pollution and Lung Function Growth in Southern California Children: Results from a Second Cohort." *American Journal of Respiratory and Critical Care Medicine* 165: 76-84.

Groark, Virginia. 2002. "An Overhaul for Emissions Testing." *New York Times,* June 9: 4.

Hubbard, Thomas N. 1998. "An Empirical Examination of Moral Hazard in the Vehicle Inspection Market." *RAND Journal of Economics* 29: 406-426.

Hubbard, Thomas N. 2002. "How Do Consumers Motivate Experts? Reputational Incentives in an Auto Repair Market." *Journal of Law and Economics* 45: 437-468.

Lambert, C. 2000. "Payoffs Alleged as Emissions Tests End." *Palm Beach Post*, June 30: 1A

Lazare, D. 1980. "Jersey Weighs Its Shift to Private Inspections." *New York Times,* October 26: S21.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142: 698-714.

Oliva, Paulina. 2012. "Environmental Regulations and Corruption: Automobile Emissions in Mexico City." http://www.econ.ucsb.edu/~oliva/Docs/Smog_Checks_Jan2012.pdf.

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

Sallee, James M., and Joel Slemrod. 2010. "Car Notches: Strategic Automaker Responses to Fuel Economy Policy." NBER Working Paper No. 16604.

Schneider, Henry. 2012. "Agency Problems and Reputation in Expert Services: Evidence from Auto Repair." *Journal of Industrial Economics*. In press.

Snyder, Jason A. 2010. "Gaming the Liver Transplant Market." *Journal of Law, Economics, & Organization* 26: 546-568.

Snyder, Jason A. 2012. "Discontinuous Electoral Distributions" Working paper, University of California, Los Angeles.

States News Service. 2010. "40 Facilities Cited for Falsifying 20,000+ Vehicle Emissions Inspections." February 18. http://www.dec.ny.gov/press/62821.html

Stedman, Donald. 2002. "Remote Sensing: A New Tool for Automobile Inspection and Maintenance." Independence Institute Issue Paper 1-2002. http://www.feat.biochem.du.edu/reports.html.

Stedman, Donald, Gary Bishop, and Allison Peddle. 2009. "On-Road Motor Vehicle Emissions Including NH3, SO2, and NO2." *Final Report to the California Air Resources Board*. http://www.feat.biochem.du.edu/reports.html.

Tirole, Jean. 1986. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations." *Journal of Law, Economics, & Organization* 2: 181-214.

U.S. Environmental Protection Agency. 2001. *Our Built and Natural Environments: A Technical Review of the Interactions between Land Use, Transportation and Environmental Quality, Development Community and Environmental Division*. EPA 231-R-01-002, Washington, DC, January, www.epa.gov/piedpage/pdf/built.pdf.

Utell, Mark J., Jane Warren, and Robert F. Sawyer. 1994. "Public Health Risks from Motor Vehicle Emissions." *Annual Review of Public Health* 15: 157-178.

Voas, S., and P. Shelly. 1995. "Ridge Kills New Emissions Tests." *Pittsburgh Post-Gazette*, October 19: 1A

Wolfers, Justin. 2006. "Point Shaving: Corruption in NCAA Basketball." *American Economic Review* 96: 279-283.

Zitzewitz, Eric. 2012. "Forensic Economics." *Journal of Economic Literature* 50: 731-769