Journal of Environmental Economics and Management 67 (2014) 371–395

Contents lists available at ScienceDirect

Journal of Environmental Economics and Management

journal homepage: www.elsevier.com/locate/jeem

# Cars on crutches: How much abatement do smog check repairs actually provide?

Pierre Mérel [a,*], Aaron Smith [a], Jeffrey Williams [a], Emily Wimberger [b]

[a] University of California, Davis, 1 Shields Ave, Davis, CA 95616, USA and Giannini Foundation of Agricultural Economics
[b] Bren School of Environmental Science & Management, UC Santa Barbara, 2400 Bren Hall, Santa Barbara, CA 93106, USA

ARTICLE INFO

*Article history:*
Received 12 March 2013
Available online 18 January 2014

*Keywords:*
Pollution abatement
Abatement persistence
Repair durability
Inspection and maintenance
Smog check

ABSTRACT

Not as much abatement as has been presumed. Smog check programs aim to curb tailpipe emissions from in-use vehicles by requiring repairs whenever emissions, measured at regular time intervals, exceed a certain threshold. Using data from California, we estimate that on average 41% of the initial emissions abatement from repairs is lost by the time of the subsequent inspection, normally two years later. Our estimates imply that the cost per pound of pollution avoided is an order of magnitude greater for smog check repairs than alternative policies such as new-vehicle standards or emissions trading among industrial point sources.



*What's broken can always be fixed…What's fixed will always be broken.*

—Jens Lekman

## Introduction

Local air pollution arising from economic activity causes significant human health and environmental damage (Chay and Greenstone, 2005; Muller et al., 2011).[1] Motor vehicles represent the largest single source of nitrogen oxides (NOx), volatile organic compounds (including hydrocarbons, HC) and carbon monoxide (CO) in the United States (Environmental Protection Agency, 2008). NOx and volatile organic compounds react with sunlight to form ground-level ozone, leading to respiratory problems and damage to crops and sensitive vegetation and ecosystems. Human exposure to high concentrations of CO inhibits the blood's ability to transport oxygen and can result in nausea, angina, and even death (Environmental Protection Agency, 2011b).

Policy intervention can reduce the damages from local air pollution. A common view among economists, generally supported by empirical evidence, is that market-based mechanisms such as emissions trading or emissions taxes are preferable to command and control (CAC) approaches such as performance standards (Freeman, 2002; Fowlie, 2010; Fowlie et al., 2012a).

While taxes or caps represent a workable option for point sources, to this date actual tailpipe emissions from motor vehicles cannot reliably and economically be measured, as would be required to implement a market-based approach (Fullerton and West, 2002, 2010; Fowlie et al., 2012b). Gasoline taxes have been suggested as a second-best price instrument to control tailpipe emissions, but because emissions crucially depend on vehicle characteristics, including the condition of

* Corresponding author. Fax: +1 530 752 5614.
*E-mail addresses:* merel@primal.ucdavis.edu (P. Mérel), asmith@primal.ucdavis.edu (A. Smith), williams@primal.ucdavis.edu (J. Williams), emilywimberger@gmail.com (E. Wimberger).

[1] For example, Muller et al. (2011) estimate that several industries, including solid waste combustion and oil and coal-fired power plants, entail air pollution costs larger than their value added.

0095-0696/$ - see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jeem.2013.12.006

emissions control components, these taxes entail significant efficiency losses compared to a first-best emissions tax (Fullerton and West, 2010).[2]

Not surprisingly, then, current regulations aimed at curbing vehicles′ tailpipe emissions typically fall into the CAC category. Gasoline content regulations intended to reduce emissions of ozone precursors from mobile sources can, in principle, make gasoline burn more cleanly through changes in chemical composition (Aufhammer and Kellogg, 2011). However, such standards cannot address engine failures or defective emissions control components, hence the need for regulations pertaining to motor vehicles themselves. For new vehicles, regulations dictate maximum emissions rates and specific emissions control technologies to manufacturers.

For in-use vehicles, the focus of this paper, standards typically provide an emissions rate ceiling by vehicle class. In the United States, the 1977 Amendments to the Clean Air Act require all areas of the country that fail to meet National Ambient Air Quality Standards (NAAQS) to implement inspection and maintenance (I/M) programs to reduce tailpipe emissions from in-use cars and light duty trucks.[3] Vehicles must undergo emission-related repairs whenever their tailpipe emissions, measured at regular yet distant time intervals—typically two years—exceed established thresholds. I/M programs are ubiquitous across the U.S. and most developed countries, where they represent an essential lever to control emissions from mobile sources (Harrington et al., 2000).

The present paper seeks to document the emissions reductions attributable to I/M programs and to provide reasonable estimates of their cost effectiveness. Despite their prevalence and the empirical significance of the pollution they address, these programs have received little attention in the economics literature. Notably, it is not known whether they are cost effective relative to other policies targeting local air pollution.

The existence of I/M programs is predicated upon the expectation that repairs conducted on non-compliant vehicles durably reduce tailpipe emissions. An assumption commonly made in calculating benefits from I/M programs is that the initial repair-induced abatement is *persistent* throughout the I/M cycle (Harrington et al., 2000; Ando et al., 2000; California Air Resources Board, 2009; Environmental Protection Agency, 2012). Under the *persistence* assumption, the cumulative reduction in emissions is computed by multiplying the drop in a vehicle′s emissions rate per mile attributable to repairs by the expected number of miles driven during the cycle. However, if the benefits from emission-related repairs in terms of avoided emissions decrease over vehicle use, that is, if the abatement lacks persistence, benefits calculated based on the persistence assumption will clearly be overstated (Harrington and McConnell, 1994; Hubbard, 1997; Singer and Wenzel, 2003). Empirical evidence regarding the actual degree of abatement persistence is lacking (Environmental Protection Agency, 2002; Singer and Wenzel, 2003). As a result, attempts to relax the persistence assumption in the evaluation of I/M benefits have been rare and have relied on *ad hoc* assumptions.[4]

This study is the first to provide econometric evidence of the lack of persistence of repair-induced emissions abatement throughout an I/M cycle. Our inference is based on a large cross section of vehicles from California for which we have detailed and reliable information on tailpipe emissions and emission-related repairs performed at high-quality stations over the course of the past decade. California is a particularly well-suited setting to investigate the effect of repairs on the emissions deterioration of in-use vehicles due to its relatively old fleet and the extensiveness of its I/M program, known as the California Smog Check Program. We find that on average, 41% of the abatement observed at the time of repair is lost by the subsequent inspection. The magnitude of this loss is shown to be robust to various model specifications and various ways to aggregate targeted pollutants.

Two sources of lack of abatement persistence should be distinguished. First, because the emissions control components of vehicles deteriorate with use, their tailpipe emissions rates tend to increase secularly. We refer to this increase in a vehicle′s emissions rate over use as its *emissions trajectory*. To the extent that this trajectory is concave—we provide empirical evidence that it is—successful repairs will bring vehicles back to a portion of their emissions trajectory where their emissions accrual rate with use is higher (their emissions trajectory is steeper). Consequently, the gap in emissions intensity between a repaired vehicle and a counterfactual vehicle that did not receive repairs will decrease over vehicle use, i.e., the repair-induced abatement will not be persistent. Unless vehicles can be repaired to perform better than they historically did, this type of abatement deterioration with use is technically unavoidable.

The second source of lack of abatement persistence is poor repair durability. Only durable repairs can bring back vehicles to an earlier point in their emissions trajectory, in the sense that the vehicle will follow its historical path of emissions accumulation once again. Non-durable repairs, by contrast, will decrease the emissions rate of a vehicle while increasing its emissions accrual rate relative to its historical trajectory, further eroding the persistence of repair-induced abatement. Arguably, I/M programs can be improved to increase the durability of repairs, for instance by improving technician training, increasing the performance standards of stations, or mandating centralized testing and repairs. As such, it is important from a policy perspective to distinguish between the two aforementioned sources of abatement deterioration.[5]

[2] Several authors have investigated the merits of more complicated tax schemes to replicate or approximate the first-best tax on emissions. See for instance Innes (1996) and Fullerton and West (2002, 2010).

[3] Six pollutants are covered by NAAQS: CO, lead, NOx, ozone (a secondary pollutant formed by the interaction of NOx and volatile organic compounds in the presence of sunlight), particle pollution, and sulfur dioxide (Environmental Protection Agency, 2010b).

[4] The only instances we are aware of are the newer iteration of California′s mobile emissions model EMFAC (California Air Resources Board, 2001b) and the study of Singer and Wenzel (2003).

[5] California′s EMFAC model currently accounts for the first source of abatement loss—but not for the second one—when deriving the emissions trajectories of vehicle groups subject to I/M over time. Even so, abatement persistence may still be assumed by the California Air Resource Board when performing cost–benefit analysis (California Air Resources Board, 2004, 2009).

Our empirical model allows us to decompose the loss in repair-induced abatement between inspections into these two elementary components. We find that about two-thirds of the 41% abatement loss is directly attributable to the nonlinearity in a vehicle′s normal emissions trajectory. We investigate the possibility that this nonlinearity effect might be under-estimated due to omitted variables bias and conclude that any bias is likely small. The remaining abatement loss reflects a pure lack of repair durability. That is, our estimates suggest that repairs made on non-compliant vehicles, especially older vehicles, fall short of being "as good as new." The fact that our abatement loss is derived for a sample of repairs conducted at high-quality stations probably means that average Smog Check repairs, or repairs performed in anticipation of a Smog Check inspection without formal emissions testing, would show even lower durability. This finding illustrates the broader proposition that fixes made on defective durable goods cannot completely erase the effects of time and use.

Having detailed information on repair costs for our sample of vehicles, we use our empirical estimates to calculate, under several scenarios, the cost of repairs per pound of pollution avoided over the course of the I/M cycle. We make some simplifying assumptions to generate these calculations, but even our most conservative estimates suggest that repairs made in response to the California Smog Check Program are an expensive way to reduce pollution.

In our most conservative scenario, we estimate that each pound of CO, NOx, and HC avoided requires $2.17 in repairs for the average vehicle and $1.27 for gross polluters, a class of high-emitting vehicles. Of these three pollutants, CO is the least toxic (Small and Kazimi, 1995). If we only credit NOx and HC abatement, our estimated cost rises by a factor of ten to $17.91 on average and $12.49 for gross polluters. These estimates understate the total cost of the program because they use only the direct cost of emission-related repairs and exclude factors such as administrative costs and the opportunity cost of drivers′ time.

The results of Fowlie et al. (2012b) provide some context for our estimates. These authors compute engineering-based estimates of the cost of reducing NOx emissions from power plants and newly manufactured motor vehicles that are much less than our cost estimates for reducing in-use vehicle emissions through I/M. Their preferred estimate for the marginal abatement cost of NOx emissions from new vehicles through federal emissions standards is $0.45/lb of NOx, while that from power generation is $0.95/lb, implying that sizable efficiency gains could be realized by shifting more of the abatement burden towards emissions standards for new cars. Our results suggest that their conclusions do not extend to I/M standards for in-use vehicles, as the associated abatement costs far exceed those of new cars and electricity generation on the margin.

California Assembly Bill 2289, which took effect in 2013, aims to improve the overall cost effectiveness of the Smog Check Program by alleviating testing procedures for newer model year vehicles and directing high-emitting vehicles towards stations with enhanced performance standards (California Air Resources Board, 2012b). Our results confirm that this move will likely improve the cost effectiveness of emission-related repairs at the fleet level, because cost-effectiveness ratios are much lower for high-emitting vehicles than for the broader set of vehicles failing an emissions test. However, our findings also indicate that cost-effectiveness ratios look worse once the lack of abatement persistence is taken into consideration. As a result, refocusing the program towards high-emitting vehicles, while certainly desirable, would likely not by itself bring the pollution abatement cost from mobile sources down to that from point sources.

The paper is organized as follows. In the Background section, we provide an overview of the California Smog Check Program and the California Consumer Assistance Program (CAP). This information is critical to understanding our choice of data and empirical strategy. The Elementary decomposition of the total effect section formalizes the decomposition of the total effect of repair-induced abatement on subsequent emissions accrual into the two elementary effects described above. The Empirical strategy section explains our empirical strategy to identify these two effects. The Estimation results section presents results from our main specification and the Robustness checks section offers a series of robustness checks to address potential bias due to omitted variables, measurement error, and pollutant aggregation. In the Cost effectiveness analysis section we use our empirical estimates to derive indicators of cost-effectiveness for I/M repairs and compare them to available estimates pertaining to emissions reductions from other sources. The last section concludes.

## Background

Within the United States, 33 states and the District of Columbia have some form of I/M program (Environmental Protection Agency, 2003). Under the direction of EPA, non-attainment regions of the country must implement I/M programs to reduce vehicle tailpipe emissions. States are given jurisdiction over the design and administration of their programs. Inspection frequency, type of testing facility, inspection method, and regulated pollutants vary not only by the state but also by the intra-state region. The penalty for non-compliance depends on the specific I/M program, with most states using loss of vehicle registration or monetary fines to encourage participation.

### *The California Smog Check Program*

California has the most stringent vehicle tailpipe emissions requirements and the largest I/M program in the nation. The Bureau of Automotive Repair (BAR) administers California′s I/M program. Most of California′s vehicle fleet is subject to the Enhanced Smog Check Program, which requires inspections biennially and upon change of vehicle ownership. The enhanced program consists of a three-part inspection encompassing a visual, functional, and emissions test. The visual test ensures that all emissions control components are present, while the functional test certifies the functionality of the emissions control components, including the malfunction indicator light, gas cap, and the On-Board Diagnostic II system (for all 1996 and newer vehicles). The emissions test, meanwhile, measures the concentration of HC, CO, and NOx emitted from the vehicle tailpipe at two different speeds using the Acceleration Simulation Mode (ASM) procedure.

Vehicles must have emissions below make-, model year- and model-specific cut-points in order to pass the emissions component of the inspection.[6] Vehicles that fail any of the three components must be repaired by certified facilities until they pass the inspection.[7]

Through 2004, all vehicles older than four model years of age were subject to biennial testing. As of January 1, 2005 a modification to the program increased the age of exemption to six model years. As of April 1, 2005, a rolling 30-year exemption of older vehicles was replaced by a mandate for all 1976 and newer model years to be subject to the Smog Check Program enforced in their geographic jurisdiction.

Unlike many states, I/M inspections in California are conducted both by Test-Only facilities and Test & Repair stations. Vehicles most likely to fail the Smog Check inspection are directed to receive inspection at high performing stations in an attempt to better target high-emitting vehicles.[8] Beginning in 2002, directed vehicles were required to receive inspections at Test-Only stations. In 2003, in response to the increased time cost of testing and repairing directed vehicles at different stations, the Gold Shield station category was introduced. Gold Shield stations were high-performing stations that were able to both test *and* repair directed vehicles.

California Assembly Bill 2289, which was enacted in 2010, marked the first major improvements to the Smog Check Program since the mid-1990s. As of January 1, 2013, the Gold Shield classification has been replaced by STAR certification, which is granted to stations that meet enhanced performance standards and gives them the right to inspect directed vehicles (California Air Resources Board, 2012a). Starting in 2013, all 1999 and older model year vehicles will be directed to STAR stations. In contrast, in 2014, newer vehicles will only receive visual and functional tests utilizing On-Board Diagnostic systems to monitor emissions, thereby reducing inspection costs for consumers and equipment costs for stations (California Air Resources Board, 2012b). The new policy will thus redirect the inspection effort towards vehicles most likely to be high emitters.

*The Consumer Assistance Program*

In 1997, Assembly Bill 57 required BAR to provide financial assistance to qualified low-income motorists needing emission-related repairs, broadly defined as repairs required to pass a Smog Check inspection. Under CAP, which began operation in 1998, qualified consumers can receive repair assistance of up to $500 or choose to retire their vehicle for a stipend. Along with low-income motorists, defined as households whose income does not exceed 185% of the federal poverty level, the repair assistance program was open to all vehicles directed to Test-Only stations that failed their biennial inspection. This support for owners of directed vehicles regardless of income ended in 2010. Low-income motorists must contribute a $20 copay for repair assistance, while owners of directed vehicles were required to contribute $100 to receive up to $500 in emission-related repair assistance, including diagnostic expenses. If eligible under both options, a rational motorist would choose to apply under the low-income option.

Prior to each biennial inspection, Smog Check stations are required to inform motorists about the Consumer Assistance Program and eligibility requirements. CAP repairs must be conducted at authorized Gold Shield or STAR stations. Consumers are required to apply for the program within 180 days after the expiration of vehicle registration and must have their vehicles repaired within 90 days of receiving authorization.[9] A technician at a CAP-sanctioned repair shop proposes to BAR a specific repair and its anticipated costs. Approval is not automatic; sometimes several repairs are authorized over several days. CAP does not cover repairs made upon changes of ownership, which require a Smog Check unless one has been performed within 90 days. From 2000 to 2010, about 40,000 vehicles were repaired annually under CAP; another 10,000 applications were denied annually.[10]

**Elementary decomposition of the total effect**

Even with reasonable maintenance, vehicle emissions increase with mileage and age due to deterioration of the engine and emissions control components. The literature has acknowledged the importance of accounting for this secular deterioration of a vehicle's emissions status over usage (White, 1982; Singer and Wenzel, 2003), and emissions simulation models used at the federal and state levels typically account for such deterioration (California Air Resources Board, 2001a,b; Environmental Protection Agency, 2001, 2011a).

Given that a vehicle's tailpipe emissions worsen with usage, the calculated benefits from repair-induced abatement depend on the difference in the vehicle's emissions deterioration rates with and without repairs. One controversial but still commonly used assumption is that a vehicle's emissions deterioration rate is essentially unaffected by repairs, so that the

[6] Historically, cut-points have not always been model-specific. Until March 30, 2010, cut-points depended on vehicle size, model year, and weight.

[7] Vehicles may be issued a one-time repair cost waiver by a referee if they do not pass inspection after spending $450 or more on emission-related repairs. Waivers provide a two-year extension during which the vehicle must undergo additional repairs in order to pass its next biennial Smog Check.

[8] Vehicles are directed using the High Emitter Profile (HEP), which determines the probability that a vehicle will fail the inspection based on vehicle characteristics and prior inspection history.

[9] This last rule is not always enforced.

[10] Since 2010, CAP repairs are only available to low-income motorists and the number of repaired vehicles per year has declined.

initial abatement persists through usage—an assumption we call the *persistence hypothesis*. This assumption has been used—sometimes implicitly—in the literature to derive the benefits of various I/M programs (McConnell, 1990; Bishop et al., 1993; Lawson, 1995; Harrington et al., 2000; Ando et al., 2000), as well as in early iterations of California's mobile emissions model EMFAC (California Air Resources Board, 2001b) and in EPA's MOBILE model (Harrington and McConnell, 1994; National Research Council, 2001; Environmental Protection Agency, 2002). It is still present in EPA's latest emissions model, MOVES.[11]

Little empirical evidence supports the persistence hypothesis. Tailpipe emissions have been argued to increase, sometimes rapidly, after a fail-pass series of inspections, suggesting that repairs may be short-lived and that the emissions deterioration rate after repairs may exceed that in the absence of repairs (Lawson, 1993; National Research Council, 2000; Wenzel, 2001; Sierra Research, 2009). Wenzel (2001, p. 388) writes that "more research is necessary to estimate what emissions would have been in the absence of an I/M program." Singer and Wenzel (2003) follow this logic in their assessment of California's Smog Check Program. But short of empirically estimating emissions trajectories for repaired and non-repaired vehicles, they consider a series of scenarios regarding the emissions accumulation process over time, providing a range of possible benefits from I/M. Not surprisingly, the authors conclude that their estimates "are highly sensitive to assumptions about vehicle deterioration in the absence of Smog Check" (p. 2588).

The present study fills this research gap, by showing that emissions deteriorate faster after a repair than they would have without a repair, thereby eroding the cumulative environmental benefits from repair-induced abatement.

Fig. 1 depicts the possible benefits over one I/M cycle from an emission-related repair. The figure shows three abatement-persistence scenarios, assuming a locally concave counterfactual emissions path, i.e., the growth rate in emissions would have slowed without the repair. (This concavity is consistent with our empirical findings.) In each panel of Fig. 1, cumulative emissions savings through the I/M cycle are represented as the shaded area. Panel (a) depicts an abatement that is *persistent*; at each point on the emissions path after the repair, emissions are below the counterfactual by an amount equal to the repair abatement. Panel (b) shows repairs that are *as good as new*. In this scenario, a repair effectively shifts a car back to an earlier point in its emissions history. Because the emissions trajectory is concave, the resulting abatement fails to be persistent and the cumulative emissions reductions are lower than in panel (a). Panel (c) shows repairs that fall short of being as good as new, i.e., a repaired vehicle accumulates emissions faster than a non-repaired vehicle that begins at the same emissions level. Overall, Fig. 1 illustrates the proposition that assumptions regarding the counterfactual emissions path and the effect of repairs on emissions accumulation play an essential role in assessing expected environmental benefit. Even though these graphs were drawn with vehicle tailpipe emissions in mind, they illustrate the broader issue of properly accounting for the cumulative benefits of an intervention whose instantaneous effects may not persist over time due to the effects of nonlinearity and durability.

The scenarios depicted in Fig. 1 suggest decomposing the total effect of repairs on a vehicle's emissions deterioration into two elementary effects. First, repairs can affect the rate of emissions deterioration if the vehicle's emissions trajectory is not linear. If the emissions path is concave in usage, as in Fig. 1, and the repairs are as good as new, the rate of emissions deterioration will become higher as a vehicle is brought back to an earlier point in its emissions trajectory. Following this logic, newer iterations of the California Air Resources Board's EMFAC model have indeed assumed that the rate of emissions deterioration after an I/M repair is equal to the historical rate of emissions deterioration at the post-repair emissions level (California Air Resources Board, 2001b).

In addition to this first effect, repairs may *rotate* the emissions trajectory of a vehicle as soon as they are not as good as new. That is, at the same emissions level, the emissions deterioration rate of a repaired vehicle could be higher than that of an otherwise similar non-repaired vehicle.[12]

More formally, denote by $\Delta E_j$ the increment in emissions for vehicle $j$ between two consecutive Smog Checks, that is, the difference between the emissions level at the first test of the second cycle ($E_j^2$) and that recorded at the successful post-repair test of the first cycle ($E_j^1$). We write the increment $\Delta E_j$ as a function of the emissions level $E_j^0$ at the beginning of the first cycle, the amount of emissions abatement following repairs $A_j = E_j^0 - E_j^1$, vehicle characteristics $\mathbf{X}_j$, and usage intensity between cycles $\mathbf{Z}_j$:

$$\Delta E_j = \phi\left(E_j^0, A_j, \mathbf{X}_j, \mathbf{Z}_j\right). \tag{1}$$

Note that, because repairs are discrete and depend on discretionary actions taken by repair technicians, the emissions level after repairs $E_j^1$ need not be equal to that dictated by the regulatory cut-points, and in fact our data clearly suggest that there is remaining variation in $E_j^1$ after controlling for the vehicle characteristics that determine cut-points. Since these characteristics are included in $\mathbf{X}_j$, it is this residual variation that allows $A_j$ to affect $\Delta E_j$ independent of $E_j^0$ and $\mathbf{X}_j$.

The total effect of a unit of repair-induced abatement on the emissions increment between inspections is the partial derivative $\partial\phi/\partial A_j$. This effect captures the difference in emissions accumulation between two vehicles that are identical with

[11] EPA's MOBILE and MOVES models are used by jurisdictions other than California to compute mobile emissions inventories pursuant to the Clean Air Act. In MOBILE6, the persistence assumption is used when predicting the emission rates of vehicles subject to a biennial I/M program in the year they are not inspected (Environmental Protection Agency, 2002). MOVES computes the effects of I/M on vehicle emission rates based on I/M adjustment factors directly derived from MOBILE6 runs (Environmental Protection Agency, 2010a). With MOBILE/MOVES, emissions reductions attributable to I/M in a calendar year are computed by multiplying differences in emissions factors between I/M and non-I/M scenarios by vehicle activity levels, again assuming abatement persistence.

[12] Repairs could have a preventative effect and result in a lower-than-before rate of emissions deterioration, for the same emissions level. This case would correspond to a "better than new" scenario, but we fail to find empirical evidence of such an effect.

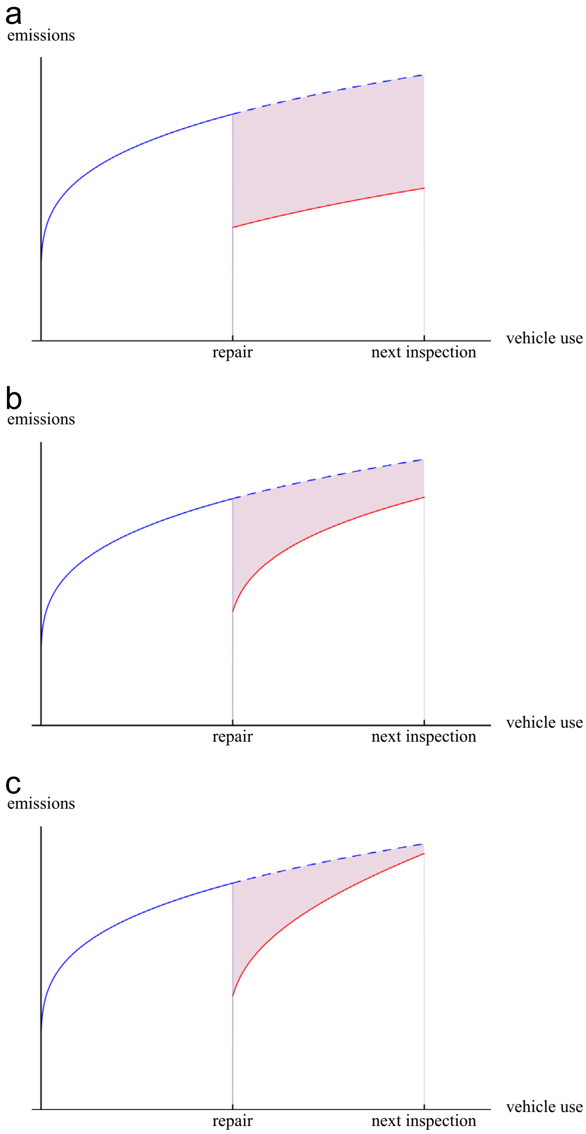


**Fig. 1.** Effect of repairs on I/M cycle emissions reductions. (a) Persistent abatement, (b) "As-good-as-new" repairs and (c) Not "as-good-as-new" repairs.

respect to their use, characteristics and *pre-repair* emissions levels, but one of which underwent an extra unit of abatement relative to the other.

Under the hypothesis that repairs are as good as new, a unit increase in abatement should have the same effect as a unit decrease in pre-repair emissions, that is, $\phi(E_j^0, A_j, \mathbf{X}_j, \mathbf{Z}_j) = \phi(E_j^0 - A_j, 0, \mathbf{X}_j, \mathbf{Z}_j)$, an identity that implies $\partial\phi/\partial A_j = -\partial\phi/\partial E_j^0$. If the emissions path of the vehicle is locally concave, as depicted in Fig. 1, $\partial\phi/\partial E_j^0 < 0$, and thus $\partial\phi/\partial A_j > 0$ for as-good-as-new repairs. In the more likely case that repairs are not as good as new, $\partial\phi/\partial A_j > -\partial\phi/\partial E_j^0$.

This last inequality implies that the total effect $\partial\phi/\partial A_j$ can be written as the sum of (i) a "trajectory effect" equal to $-\partial\phi/\partial E_j^0$, capturing the change in the emissions increment that would result from an as-good-as-new repair that brought emissions down by one unit, and (ii) a residual effect that accounts for the remainder of the change in the emissions increment. This residual effect essentially captures differences in emissions accumulation between two vehicles that are identical with respect to their use, characteristics and *post-repair* emissions level, but one of which underwent an extra unit of repair-induced abatement relative to the other. It is reasonable to call this residual the "repair durability effect."

## Empirical strategy

The ideal experiment to identify the effect of repair-induced abatement on subsequent emissions accrual would compare the emissions accrual of two identical vehicles having undergone different levels of abatement. Since a given vehicle can only be repaired once, the best we can hope for is to use naturally occurring variation in abatement intensity while controlling for a large number of potentially confounding factors, that is, vehicle attributes that are potentially correlated with pre-repair emissions and/or repair-induced abatement and partially explain the emissions accrual process.

To this end, we develop a stylized cross-sectional econometric model of emissions accumulation between two I/M cycles to estimate empirically (i) the effect of repair-induced abatement on the emissions increment between cycles and (ii) the extent to which repairs performed on non-compliant vehicles are less than "as good as new" in the sense of the Elementary decomposition of the total effect section. We estimate this model on a subset of vehicles for which we have reliable data on repairs and emissions, and for which we believe that the potential for omitted variable bias is minimal. Being aware of the inherent limitations of using cross-sectional variation, we provide multiple robustness checks that lend credence to our estimates.

### *Econometric model*

Our econometric specification is the linearized version of (1)

$$\Delta E_j = \alpha + \beta_0 E_j^0 + \beta_1 A_j + \mathbf{X}_j \boldsymbol{\gamma} + \mathbf{Z}_j \boldsymbol{\delta} + \varepsilon_j \tag{2}$$

where $j$ denotes the index for one observation and the coefficients of interest are $\beta_0$ and $\beta_1$. The parameter $\beta_1$ captures the effect of repair-induced abatement on the emissions increment, while the extent to which repairs fail to be as good as new equals $\beta_0+\beta_1$. The fact that Eq. (2) is linear does not imply that the emissions trajectory itself is linear; the presence of $E_j^0$ on the right-hand side implies that the emissions increment (the slope) is allowed to change along the trajectory. Nonetheless, this specification assumes that the trajectory is either everywhere concave ($\beta_0 \le 0$) or everywhere convex ($\beta_0 \ge 0$); in robustness checks, we allow for more flexible specifications.

The set of controls $\mathbf{Z}_j$ includes two variables reflecting the intensity of usage between the passing test of the first cycle and the first test of the second cycle: distance traveled, inferred from the odometer readings recorded at the two inspections, and the time elapsed between tests. The set of controls $\mathbf{X}_j$ contains, in our most flexible specification, Vehicle Identification Number (VIN) Prefix fixed effects, which are essentially make $\times$ model $\times$ model-year $\times$ drivetrain $\times$ engine-size fixed effects. These fixed effects are fine enough to control for cut-points, therefore identification of $\beta_1$ relies on actual variation in $E_j^1$ conditional on cut-points. In less flexible specifications, we use model-year, class-category, make, and engine-size fixed effects. Vehicle class includes 6 categories from sedan to van and make includes 20 groupings by manufacturer. All these variables represent time-invariant characteristics that can be expected to influence emissions deterioration. The set $\mathbf{X}_j$ also includes one variable capturing a vehicle's usage history, the odometer reading at the time of the repair.

We estimate specification (2) using ordinary least squares (OLS). Parameters $\beta_0$ and $\beta_1$ are distinguished by means of cross-sectional variation in the level of emissions recorded at the beginning and the end of the first repair cycle and that recorded at the subsequent inspection, approximately two years later.

Notably, the coefficient $\beta_1$ benefits from two sources of identification. First, vehicles with similar pre-repair emissions and subject to the same emissions cut-points may nonetheless receive different repairs and different abatement, because the cause of test failure may differ, the repair process is essentially discrete, it is impossible to bring emissions precisely to the cut-points, and different technicians may make different repair decisions. In the Estimation results section we provide empirical evidence that there is indeed significant variation in post-repair emissions conditional on vehicle characteristics (and thus cut-points) and pre-repair emissions. This source of identification reflects intensive margin variation in the amount of abatement received, conditional on pre-repair emissions and observable controls.

Second, our sample includes vehicles that did not fail the emissions test but instead failed the visual or functional test. These vehicles, which represent 19% of our sample, received almost no abatement on average. They thus constitute a useful pseudo control group and a source of extensive margin variation in abatement. We show in the Robustness checks section that reducing the sample to only those vehicles having failed the emissions test does not affect our estimates by much.

Our identification assumptions are that $E_j^0$ and $A_j$ are uncorrelated with $\varepsilon_j$. Although $\mathbf{X}_j$ and $\mathbf{Z}_j$ contain numerous controls for vehicle characteristics and vehicle usage, it is possible that unobserved factors, such as driver behavior, may cause omitted variable bias in the OLS estimates. We address this possibility by conducting a series of robustness checks in the Robustness checks section.

### *Data set*

To estimate model (2), we use a sample of Smog Check repairs funded by CAP between July 2000 and August 2010. This time period is the largest for which we have access to CAP repair information. Our data set uses information from four distinct sources: BAR's CAP database, BAR's Repair database, BAR's Vehicle Information Database (VID) and the DMV's Vehicle Registration Database. We observe data on all CAP repairs made in California during these years. The majority of vehicles included in the sample appear only once, and we treat the sample as a cross section. Details about the construction of the data set are given in Appendix A.

Our choice to use CAP-funded repairs is largely dictated by the need to limit, to the extent possible, the influence vehicle owners may have on the quantity and quality of repairs (Lawson, 1993; Harrington et al., 1998). Because we want to estimate the effect of repairs on emissions, we need repairs to be exogenous to the error term. If vehicle owners were to bargain with repair technicians for limited fixes (low-cost repairs that would bring emissions below but close to the cut-points), then our "treatment" variable $A_j$ could be endogenous. We could reasonably expect "bad" cars to systematically receive less abatement than other vehicles, which would bias $\hat{\beta}_1$ towards zero, assuming that bad cars accrue tailpipe emissions at a faster rate. A related concern is that the repairs themselves may vary in quality. Again, bad cars would tend to receive limited *and poor* repairs, exacerbating the previous bias.

We believe that vehicle repairs funded through California's CAP are unlikely to suffer much from these issues. First, owners benefiting from the subsidy must send their vehicles to a Gold Shield station. Gold Shield stations are arguably less subject to quality variation than other stations. They cannot have had disciplinary actions against either the station or any technician, are required to meet quarterly testing and repair standards, and must submit to periodic performance reviews. Second, BAR itself must approve the repair suggested by the station. Third, the fact that vehicle owners benefit from the $500 subsidy implies that they have less incentive to influence the extent and quality of repairs. Overall, CAP-funded repairs are likely to represent the "state of the art" in terms of emission-related repairs. The fact that these repairs are arguably of higher quality than average means that our estimate of abatement persistence—or lack thereof—will likely overstate the extent of abatement persistence for the broader set of vehicles undergoing repairs pursuant to the California Smog Check Program.

Focusing on vehicles repaired through CAP further addresses the potential issue of reversion to the mean that would have arisen had we chosen to use a broader sample of vehicles. Due to inherent variability in emissions measurements, a non-negligible number of recorded fail-pass sequences can be categorized as mere "re-tests," consecutive tests done in the hope that the first recorded fail was an abnormally high measurement and that the second test will better reflect the purported low emissions status of the vehicle (Wenzel et al., 2004). Because applications sent to CAP correspond to "true fails" which necessitate costly repairs, we do not worry about re-tests. That is, we can be confident that actual repairs were performed between the first failing test and the final passing test, which would not be the case if we were to use the broader set of Smog Check records available in the VID.

Another important reason for using CAP-funded repairs is that participating vehicle owners are much less likely to have paid for pre-inspection repairs, that is, unrecorded repairs performed in anticipation of an I/M test. Eligible motorists are expecting the majority of their expenses to be covered by the subsidy, and only those expenditures incurred after a failing test are covered by the program. The prevalence of pre-inspection repairs in the general fleet could be high. Using a fleet of vehicles from California, Singer and Wenzel (2003) estimate the emissions reductions attributable to pre-inspection repairs to be commensurate with those attributable to repairs conducted after a vehicle has failed an inspection.[13] Being confident that no pre-inspection repairs were performed prior to the recorded measurements is important as our identification strategy relies on comparing vehicles having undergone different levels of (recorded) abatement.

For each pollutant (HC, CO and NOx), we convert emissions from ASM readings to grams per mile using the Federal Test Procedure (FTP) equations outlined in the Carl Moyer Program guidelines and developed by Eastern Research Group for the Bureau of Automotive Repair (California Air Resources Board, 2008).[14] Pollutants are then aggregated by FTP weight to yield a single measure of a vehicle's tailpipe emissions.

As all three criteria pollutants are regulated under California's Enhanced Smog Check Program, aggregating them to a single measure of air pollution is defensible, if only for the fact that vehicle owners incur costs for abating all of them. Pollutant-specific versions of model (2) could technically be estimated, but the question then arises of whether to include information on the other two pollutants (pre-repair emissions and abatement) in each pollutant-specific regression. Omitting such information is problematic as the emissions status of the vehicle with respect to other pollutants may influence the accrual process of any given pollutant, resulting in omitted variable bias. Adding such information also causes problems, because the interpretation of a partial effect of abatement of one pollutant, holding constant the other two pollutants, does not make sense from a physical standpoint. The reason is that it is not possible to abate one pollutant without influencing the emissions of the other two pollutants. We therefore believe that the best way to capture the inter-relatedness of pollutant emissions in a parsimonious specification is to aggregate them to a single measure. We discuss the implications of choosing a particular aggregation rule in the Robustness checks section.[15]

Table 1 summarizes the main characteristics for our sample of 209,603 CAP-funded repairs. 39,073 vehicles passed the emissions component but failed the Smog Check for visual or functional reasons. We report summary statistics separately for them. We also report summary statistics for vehicles having failed the emissions test at the so-called "gross polluter" level. Gross polluters are vehicles with HC, CO, or NOx emissions in excess of emissions standards established by BAR which typically correspond to more than twice the cut-points for regular fails.[16] (These vehicles are included in the broader set of vehicles having failed the emissions test.)

[13] Our results are, in a way, providing information regarding the benefits to be expected from pre-inspection repairs, as there is no reason to believe that such repairs would last longer than those performed under CAP.

[14] These conversion equations are used by Knittel and Sandler (2011) to evaluate the co-benefits of carbon pricing in transportation.

[15] To be sure, we did run gas-specific regressions, with and without controlling for information on other pollutants. The coefficient estimates were in line with those reported in the Robustness checks section for the alternative aggregation rules.

[16] Vehicles identified as gross polluters cannot be certified by Test & Repair stations. They must be certified by a Test-Only station or by a Gold Shield station. Under Assembly Bill 2289, they need to be certified by a STAR station.

**Table 1**
Summary statistics for CAP repairs.

| Observations | Failed emissions 170,530 | | Passed emissions 39,073 | | Gross polluters 70,504 | |
|---|---|---|---|---|---|---|
| Numerical variables | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| Vehicle characteristics | | | | | | |
| Model year (MY) (year) | 1990.36 | (4.85) | 1995.77 | (4.56) | 1989.00 | (4.79) |
| Odometer (miles) | 156,960 | (66,572) | 135,653 | (53,849) | 155,850 | (69,459) |
| Usage characteristics | | | | | | |
| Distance traveled (miles) | 12,988 | (12,445) | 15,332 | (12,424) | 12,290 | (12,615) |
| Days elapsed (days) | 641 | (139) | 653 | (133) | 639 | (139) |
| Aggregate emissions (g/mile) | | | | | | |
| $E_j^0$ | 21.188 | (23.483) | 5.390 | (6.220) | 31.862 | (30.608) |
| $A_j$ | 11.383 | (19.791) | 0.456 | (2.413) | 20.520 | (26.632) |
| $\Delta E_j$ | 4.827 | (13.383) | 0.678 | (4.794) | 7.104 | (17.250) |
| HC emissions (g/mile) | | | | | | |
| $E_j^0$ | 1.433 | (1.876) | 0.380 | (0.466) | 2.094 | (2.603) |
| $A_j$ | 0.677 | (1.675) | 0.034 | (0.198) | 1.200 | (2.425) |
| $\Delta E_j$ | 0.324 | (1.186) | 0.059 | (0.613) | 0.442 | (1.403) |
| CO emissions (g/mile) | | | | | | |
| $E_j^0$ | 18.318 | (22.165) | 4.507 | (5.445) | 28.121 | (29.276) |
| $A_j$ | 10.081 | (19.000) | 0.364 | (2.177) | 18.563 | (25.843) |
| $\Delta E_j$ | 4.235 | (12.674) | 0.564 | (4.427) | 6.359 | (16.344) |
| NOx emissions (g/mile) | | | | | | |
| $E_j^0$ | 1.437 | (0.896) | 0.503 | (0.427) | 1.647 | (1.074) |
| $A_j$ | 0.625 | (0.781) | 0.058 | (0.196) | 0.757 | (1.024) |
| $\Delta E_j$ | 0.267 | (0.563) | 0.055 | (0.244) | 0.304 | (0.653) |
| Repair cost ($) | 461 | (232) | 457 | (258) | 472 | (218) |
| Categorical variables | % | | % | | % | |
| Class categories | | | | | | |
| Sedan | 61.44 | | 56.64 | | 59.99 | |
| Station wagon | 3.51 | | 2.11 | | 3.85 | |
| Pick-up | 16.77 | | 18.07 | | 18.14 | |
| SUV | 9.53 | | 12.59 | | 9.45 | |
| Minivan | 6.91 | | 9.45 | | 6.46 | |
| Van or motorhome | 1.84 | | 1.14 | | 2.10 | |
| Cylinders | | | | | | |
| 4 or less | 52.54 | | 44.10 | | 53.53 | |
| 6 | 31.39 | | 41.91 | | 30.62 | |
| 8 or more | 15.49 | | 13.15 | | 15.38 | |
| Rotary | 0.57 | | 0.84 | | 0.47 | |
| Manufacturer | | | | | | |
| US | 44.23 | | 51.88 | | 43.85 | |
| Asian | 46.78 | | 41.70 | | 46.74 | |
| European | 8.99 | | 6.42 | | 9.42 | |

Table 1 shows that for each pollutant, the average reduction in emissions from repairs represents about half of the average pre-repair emissions for vehicles that failed the emissions test. Therefore, the average abatement recorded at the time of repair is substantial. For vehicles that passed the emissions test but failed one of the other two components of the inspection, the average abatement is positive but an order of magnitude smaller.

*Generalizability to the California fleet*

Our empirical estimates are based on the subset of the California fleet of vehicles that received funds from CAP between July 2000 and August 2010 and survived through to the next Smog Check inspection. Here we compare our sample to the broader fleet of vehicles that failed a Smog Check inspection and presumably underwent emission-related repairs during the same period.

Using data from BAR's VID, we show in Table 2 the average characteristics of all California vehicles that failed an ASM Smog Check inspection during the same period and that survived to the next I/M cycle. The table also shows statistics for the subset of vehicles directed to a Test-Only station. There were 6,103,866 failing vehicles in the California fleet. Among those, 2,839,913 vehicles had been directed to a Test-Only station.

**Table 2**
Comparison of California vehicles that failed an ASM Smog Check between July 2000 and August 2010 and survived to the next inspection.

| Observations | (1) CAP 209,603 | (2) Directed to Test-Only 2,839,913 | (3) All CA 6,103,866 |
|---|---|---|---|
| Model year (year) | 1991.37 | 1990.78 | 1992.22 |
| Weight (lb) | 3397 | 3441 | 3542 |
| Odometer (miles) | 152,988 | 147,570 | 136,759 |
| Emissions ($E_j^0$) | | | |
| HC (g/mile) | 1.24 | 1.24 | 1.09 |
| CO (g/mile) | 15.74 | 15.31 | 13.56 |
| NOx (g/mile) | 1.26 | 1.24 | 1.12 |

Vehicles repaired under CAP appear quite similar to the vehicles directed to Test-Only stations and are only slightly older and more polluting than the full fleet of vehicles that failed a Smog Check inspection and survived to the next Smog Check inspection. The average vehicle in our sample has model year 1991.37, which is about half a year newer than the average Test-Only vehicle (1990.78) and less than a year older than the average vehicle in the CA fleet that failed a Smog Check (1992.22). CAP vehicles have similar average weight and odometer readings to the Test-Only fleet. They are slightly lighter, and have higher mileage than the average vehicle in the California fleet that failed a Smog Check. They exhibit 13–16% greater emissions than the California fleet that failed a Smog Check. However, their emissions levels match closely those of the set of failing cars directed to Test-Only stations. As such, we expect our results to translate well to the fleet of vehicles directed to a Test-Only station. Directed vehicles represent a sizable portion of the California fleet, 36% of the fleet subject to Smog Check as of 2002 (California Air Resources Board, 2003). During the sample period, directed vehicles represented about half of the vehicles that failed a Smog Check inspection. Since vehicles are directed to Test-Only stations based on BAR's High Emitter Profile, these vehicles are arguably responsible for a very large share of fleet-level emissions. Projected changes to the California Smog Check Program due to Assembly Bill 2289 indicate that the inspection and repair effort will be refocused towards older model-year and high-polluting vehicles. Therefore, our fleet of vehicles seems to be representative of the set of vehicles targeted by the new program.

## Estimation results

Before estimating model (2), it is useful to gauge the extent of variation in post-repair emissions conditional on other covariates. As explained above, this variation is crucial to distinguishing the effect of abatement on subsequent emissions accrual, that is, the coefficient $\beta_1$. A regression of post-repair emissions $E_j^1$ on $E_j^0$ and VIN Prefix fixed effects shows that about 28% of the variation in $E_j^1$ remains unexplained. Adding a set of dummies for the type of repair performed does not decrease this figure by much (25%). Reassuringly, variation in $E_j^1$ conditional on $E_j^0$ and VIN Prefix fixed effects appears to be much smaller on the subsample of vehicles that did not fail the emissions test (7%).[17]

Acknowledging this conditional variation in post-repair emissions, we proceed to estimating model (2) on the full sample of CAP repairs. We then re-estimate (2) on sub-samples based on model-year groupings, in order to better account for secular changes in the emissions performance of vehicles and attendant heterogeneity in the emissions accrual process. The potential for omitted variable bias is addressed in the Robustness checks section.

### *Main sample*

In Table 3 we present results from a series of regressions that differ only in the controls that are included. In all specifications, we include as an explanatory variable the odometer at the time of the first inspection, in order to capture the vehicle's usage history. The canonical model of emissions accumulation discussed in the Elementary decomposition of the total effect section, whereby a vehicle's emissions increase secularly with usage according to a well-defined emissions trajectory, essentially implies that adding information reflecting past vehicle use should be redundant when controlling for emissions at the time of the repair. Said differently, given the emissions trajectory, all the information relevant to decipher how a vehicle's emissions will deteriorate with use (the slope of the trajectory) should be fully captured by its current emissions level. However, vehicles have a history of emission-related repairs. Consequently, their pre-repair emissions are not only reflective of their past usage, but also of previous repairs. As such, there could be explanatory power left in the usage history of a vehicle beyond that provided by the pre-repair emissions level. To control for such effects, we include odometer at the time of repair in all regressions, so that we essentially compare vehicles that have the same pre-repair emissions and the same usage history.

[17] Regressions based on each individual pollutant lead to very similar conclusions.

**Table 3**
Results for the full sample of 209,603 repairs.

| Dep. var. $\Delta E_j$ Obs. 209,603 | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| $E_j^0$ | 0.081*** | −0.135*** | −0.229*** | −0.244*** | −0.265*** | −0.278*** | −0.279*** | −0.282*** | −0.252*** |
| | (0.007) | (0.011) | (0.012) | (0.012) | (0.011) | (0.012) | (0.012) | (0.013) | (0.014) |
| $(E_j^0)^2$ | – | – | – | – | – | – | – | – | −3.01e−04** |
| | | | | | | | | | (8.76e−05) |
| $A_j = E_j^0 - E_j^1$ | 0.143*** | 0.315*** | 0.392*** | 0.394*** | 0.404*** | 0.408*** | 0.408*** | 0.428*** | 0.408*** |
| | (0.009) | (0.010) | (0.012) | (0.011) | (0.011) | (0.012) | (0.012) | (0.014) | (0.012) |
| Odometer | 3.50e−06*** | 4.54e−06*** | 4.95e−06*** | 6.14e−06*** | 6.81e−06*** | 7.03e−06*** | 6.00e−06*** | 6.14e−06*** | 5.81e−06*** |
| | (4.64e−07) | (4.65e−07) | (4.63e−07) | (4.98e−07) | (5.04e−07) | (5.28e−07) | (5.43e−07) | (5.45e−07) | (5.45e−07) |
| MY f. e. | No | Yes | – | – | – | – | – | – | – |
| MY/Class f. e. | No | No | Yes | – | – | – | – | – | – |
| MY/Class/Make f. e. | No | No | No | Yes | – | – | – | – | – |
| MY/Cl./Mk./Eng. f. e. | No | No | No | No | Yes | – | – | – | – |
| VIN Prefix f. e. | No | No | No | No | No | Yes | Yes | Yes | Yes |
| Distance traveled | | | | | | | 3.17e−05*** | 3.13e−05*** | 3.12e−05*** |
| | | | | | | | (2.54e−06) | (2.55e−06) | (2.54e−06) |
| Time elapsed | | | | | | | 0.002*** | 0.002*** | 0.002*** |
| | | | | | | | (1.82e−04) | (1.82e−04) | (1.82e−04) |
| $A_j \times$ Oxygen Sensor | | | | | | | | −0.029*** | |
| | | | | | | | | (0.007) | |
| $A_j \times$ Catalytic Conv. | | | | | | | | −0.018* | |
| | | | | | | | | (0.008) | |
| $A_j \times$ Fuel Evap. Syst. | | | | | | | | −0.044** | |
| | | | | | | | | (0.017) | |
| $A_j \times$ Ignition Syst. | | | | | | | | 0.020* | |
| | | | | | | | | (0.009) | |
| $A_j \times$ Other Compon. | | | | | | | | −0.024* | |
| | | | | | | | | (0.010) | |
| $R$-squared | 0.125 | 0.148 | 0.157 | 0.176 | 0.209 | 0.289 | 0.291 | 0.292 | 0.291 |

Reported standard errors are heteroskedasticity-robust. In column (9), $E_j^0$ is demeaned, so that $\hat{\beta}_0$ represents the marginal effect of pre-repair emissions evaluated at the sample mean of $E_j^0$.
* Statistical significance at the 5% level.
** Statistical significance at the 1% level.
*** Statistical significance at the 0.1% level.

Column (1) shows the estimates of $\beta_0$ and $\beta_1$ with only the control for usage history. The coefficient on $E_j^0$ is slightly positive, suggesting a convex emissions trajectory, while that on $A_j$ is also positive, indicating that repairs are not as good as new. Of course, this first specification is naive, as it ignores the fact that different vehicle models are likely to have different emissions trajectories. Given that the emissions performance of new vehicles has evolved substantially over time due to more stringent emission standards, we expect vehicle model year to influence the emissions increment. A 2001 volume published by the National Academy of Sciences notes the following:

> Along with tightening the new-vehicle-emissions certification standard, vehicle manufacturers are also required to extend the time provided for emissions component warranties up to 80,000 miles. This mandate, as well as the availability of more robust technology such as advanced catalytic converters, has helped to reduce significantly the in-use deterioration rate of emissions control components (National Research Council, 2001, p. 195).

Other characteristics, such as the manufacturer or the class of the vehicle, appear as logical controls. The next series of controls we add seeks to capture time-invariant vehicle characteristics that could be relevant in explaining the emissions accrual process. To the extent that these controls are relevant and correlated with $E_j^0$ and $E_j^1$, adding them as regressors should affect the coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$.

The progression in Table 3 shows that a few time-invariant controls appear sufficient to stabilize the coefficients of interest, despite the relatively large number of available controls. Including model-year fixed effects affects the coefficients on $E_j^0$ and $A_j$ markedly. In particular, it suffices to reverse the sign of the coefficient $\hat{\beta}_0$, now indicating a concave trajectory which will be robust to the addition of further controls. Refining these fixed effects by class category further impacts $\hat{\beta}_0$ and $\hat{\beta}_1$, as shown in column (3). Further refining the fixed effects by make and subsequently by engine size affects $\hat{\beta}_0$ by a small amount and $\hat{\beta}_1$ by an even smaller amount (columns (4) and (5)). Column (6) corresponds to the most flexible specification of vehicle characteristics, allowing for 19,378 VIN Prefix fixed effects. The move from model-year × class × make × engine-size fixed effects (4784 categories) to the much finer VIN Prefix fixed effects specification results in a substantial increase in the model R-squared, yet it only affects the coefficients of interest slightly. This fact alone is reassuring regarding the potential for omitted variable bias, as we would expect time-invariant vehicle characteristics to be an important determinant of the emissions accrual process.

The addition of the two available controls capturing vehicle usage between inspections, distance traveled and time elapsed, leaves the two estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ essentially unchanged (column (7)). Yet, one could have expected distance traveled, a proxy for usage intensity, to be positively related to both the emissions increment and pre-repair emissions, resulting in upward bias on $\hat{\beta}_0$ when omitted from the specification. As such, the fact that the coefficients of interest do not move when we add these two controls is again reassuring regarding the size of the potential bias from omitting information regarding, say, whether the vehicle was driven in an urban or rural environment.

In column (8), we investigate the extent to which the effect of repair-induced abatement on emissions accumulation depends on the type of repair performed. Within the CAP database, repairs are classified into 18 categories. We add interaction terms between $A_j$ and repair type dummies to specification (7). Because a significant number of repairs are reported to involve only diagnosis, the baseline effect of $A_j$ on the emissions increment is that of a repair that consists solely of diagnosis, while the coefficients on interacted terms indicate adjustments to the baseline effect due to a particular repair type.[18] For brevity, we only report coefficients on interaction terms that are statistically significant at the 5% level or better. The results indicate that the average loss in abatement between inspections is only slightly dependent on repair type.

Finally, in column (9) we report estimates for a more flexible specification of the emissions trajectory, where $E_j^0$ is allowed to influence the future emissions increment both linearly and quadratically. This specification thus allows for trajectories with changing convexity. Yet, the estimated coefficients on $E_j^0$ and its square both have a negative sign, supporting the view that within the range of emissions observed in our sample, the trajectory is everywhere concave.[19]

In our preferred specification, reported in column (7) of Table 3, the coefficient on abatement implies that a one unit increase in abatement through repairs will be matched, on average, by a 0.41 unit recuperation in emissions by the subsequent inspection. This means that 59% of the unit decrease in emissions will have persisted. Notably, the counterfactual here is not the pre-repair emissions level, but rather the emissions level that the vehicle would have achieved by the next inspection in the absence of the unit decrease in emissions. The coefficient on $E_j^0$ implies that about 28% of the abatement is lost due to the concavity in a vehicle′s emissions trajectory. Even if repairs were as good as new, this part of the initial abatement would be lost. The remaining 13% loss in abatement is due to the fact that repairs are not, in fact, as good as new.

Given the large number of VIN Prefix categories, many include only a few observations. Table 4 reports estimates based on subsamples for which VIN Prefix categories with only a small number of observations have been excluded. The table shows that the coefficients of interest are very stable across these subsamples. In addition, some vehicles appear more than once in the sample, meaning that for a (comparatively small) subset of observations the variable $E_j^0$ can be considered to be a lagged dependent variable. Fixed-effect estimation of dynamic panels produces inconsistent estimates for short panels

[18] The categories are Oxygen Sensor, Catalytic Converter, Carburetor, Fuel Evaporation System, Ignition System, EGR System, Other Related Emissions Components, AIS System, Thermostatic Air Cleaner, Wiring to Sensors, Vacuum Lines, Fuel Injection, Fuel Cap, Cooling System, Sensors, PCV, Vacuum Switches, and Spark Control System.

[19] We also estimated a cubic specification. The coefficient on the cube of $E_j^0$ was not statistically significant.

**Table 4**
Samples with minimum size of fixed-effects categories and with no repeated vehicles.

| Dep. var. $\Delta E_j$ | (1) $N_1 \geq 3$ | (2) $N_1 \geq 5$ | (3) $N_1 \geq 10$ | (4) $N_1 \geq 50$ | (5) w/o repeated veh. |
|---|---|---|---|---|---|
| $E_j^0$ | $-0.282^{***}$ | $-0.277^{***}$ | $-0.291^{***}$ | $-0.302^{***}$ | $-0.291^{***}$ |
| | (0.012) | (0.012) | (0.013) | (0.020) | (0.013) |
| $A_j$ | $0.412^{***}$ | $0.406^{***}$ | $0.419^{***}$ | $0.439^{***}$ | $0.422^{***}$ |
| | (0.012) | (0.012) | (0.012) | (0.019) | (0.012) |
| Usage history | Yes | Yes | Yes | Yes | Yes |
| VIN prefix f. e. | Yes | Yes | Yes | Yes | Yes |
| Usage characteristics | Yes | Yes | Yes | Yes | Yes |
| # of f. e. categories | 11,348 | 8249 | 4914 | 848 | 19,372 |
| Obs. | 198,889 | 188,277 | 166,136 | 81,148 | 193,736 |
| $R$-squared | 0.242 | 0.218 | 0.205 | 0.196 | 0.301 |

Reported standard errors are heteroskedasticity-robust. $N_1$ denotes the minimum number of observations in a fixed-effect category.
*** Statistical significance at the 0.1% level.

**Table 5**
Results for the model-year groupings.

| Dep. var. $\Delta E_j$ | (1) 1975–1980 | (2) 1981–1983 | (3) 1984–1986 | (4) 1987–1992 | (5) 1993–1995 | (6) 1996–2000 | (7) 2001–2003 | (8) 2004 |
|---|---|---|---|---|---|---|---|---|
| $E_j^0$ | $-0.187^{***}$ | $-0.229^{***}$ | $-0.294^{***}$ | $-0.391^{***}$ | $-0.407^{***}$ | $-0.359^{***}$ | $-0.320^{***}$ | $-0.362$ |
| | (0.031) | (0.037) | (0.021) | (0.013) | (0.017) | (0.019) | (0.063) | (0.206) |
| $A_j$ | $0.379^{***}$ | $0.387^{***}$ | $0.410^{***}$ | $0.491^{***}$ | $0.487^{***}$ | $0.415^{***}$ | $0.442^{***}$ | $0.462^{*}$ |
| | (0.034) | (0.037) | (0.020) | (0.011) | (0.011) | (0.016) | (0.049) | (0.200) |
| Usage history | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| VIN prefix f. e. | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Usage characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Failed emissions (%) | 92.1 | 96.2 | 95.6 | 92.2 | 87.6 | 51.4 | 21.1 | 19.2 |
| Obs. | 4844 | 9056 | 28,000 | 71,979 | 46,097 | 44,204 | 5053 | 370 |
| $R$-squared | 0.263 | 0.299 | 0.204 | 0.206 | 0.169 | 0.224 | 0.396 | 0.512 |

Reported standard errors are heteroskedasticity-robust.
* Statistical significance at the 5% level.
*** Statistical significance at the 0.1% level.

(Nickell, 1981). Although we include VIN Prefix rather than vehicle-level fixed effects, our estimates could potentially be contaminated by this bias. Column (5) shows that our results are robust if we only include the first observation of vehicles that appear more than once in the sample.

Our finding that the emissions path is locally concave is consistent with the notion that there is a physical limit to the emissions intensity of a vehicle, and this limit dictates a path of emissions over usage that is increasing at a decreasing rate. Interestingly enough, this concavity in the emissions trajectory of a vehicle was posited early on by Downing (1973) based on expert engineering advice. It is reflected, to a certain extent, in currently available emissions models (California Air Resources Board, 2001b; Environmental Protection Agency, 2011a,c). In particular, EPA indicates that its MOVES model specifies base emission rates by vehicle age group, and that the most rapid change in emissions occurs when vehicles are between 4 and 9 years of age. As such, emissions deterioration would be slower after 9 years, resulting in a locally concave emissions path. Given the 4 or 6 year exemption from Smog Check for newer vehicles, this concave portion of the emissions path would be the relevant one.

### *Model-year groupings*

To investigate the extent to which the parameters $\beta_0$ and $\beta_1$ may differ across model years, we re-estimate (2) on subsamples of model-year groups. Results are presented in Table 5. The groupings are based on the structure of the California Smog Check ASM cut-points for the period and reflect advancements in the emissions control technology (e.g., introduction of the three-way catalytic converter in 1981, implementation of the On-Board Diagnostic II system in 1996) as well as the secular tightening of federal and state emissions standards (e.g., Tier 0 and Tier 1 EPA federal vehicle emissions standards).[20]

[20] Model-year groupings are slightly different for passenger cars and trucks. The headings of Table 5 reflect the groupings for passenger cars, but the groupings themselves take account of the class of each vehicle.

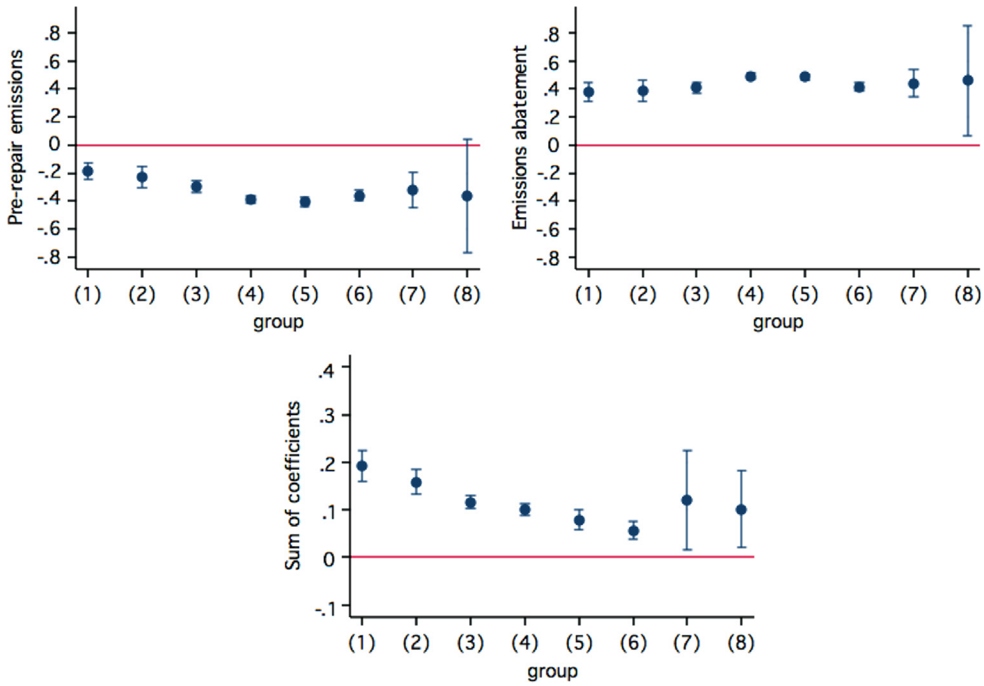


**Fig. 2.** Estimates for the model-year groupings.

Table 5 shows that the signs and magnitudes of the coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ are comparable across model-year groupings, particularly for $\hat{\beta}_1$. Fig. 2 provides a visual summary of these results. The point estimate of $\beta_1$ changes from one group to the other, without a clear pattern across model years, varying between 0.38 and 0.49. The coefficient $\hat{\beta}_0$ varies non-monotonically between $-0.19$ and $-0.41$. The sum $\hat{\beta}_0+\hat{\beta}_1$ decreases for more recent model-year groups, indicating that repairs made on newer model-year vehicles are more effective, in the sense that they are closer to bringing vehicles to their historical emissions deterioration rate. (There is an increase in the sum $\hat{\beta}_0+\hat{\beta}_1$ for the two most recent model-year groups but the estimates are not precise for these two groups due to the lower sample size.) That is, a larger share of the total effect can be attributed to less-than-perfect repairs for older model-year vehicles.[21]

## Robustness checks

The various robustness checks we report in this section aim at addressing three potential concerns regarding the estimation of model (2). The first concern is that unobserved vehicle and driver characteristics that matter in explaining the emissions increment $\Delta E_j$ may be correlated with $E_j^0$ or $E_j^1$, resulting in omitted variable bias. The second concern relates to possible measurement error in $E_j^0$ and $E_j^1$. The last concern relates to our choice to aggregate all three pollutants by FTP weight.

### *Unobserved heterogeneity*

Heterogeneity in vehicle quality or driver behavior that is not captured by included controls could result in omitted variable bias. For brevity, we summarize this unobserved heterogeneity in a proxy we call "emissions lemon." An emissions lemon is a vehicle with higher than average emissions for reasons that we do not observe; such vehicles have greater $E_j^0$ and $\Delta E_j$ than non-lemons. To understand the signs of any bias in $\hat{\beta}_0$ and $\hat{\beta}_1$ due to the unobserved emissions-lemon characteristic, it is convenient to consider the modified regression

$$\Delta E_j = \alpha + b_0 E_j^0 + b_1 E_j^1 + \mathbf{X}_j \boldsymbol{\gamma} + \mathbf{Z}_j \boldsymbol{\delta} + \varepsilon_j. \tag{3}$$

The presence of emissions lemons causes $E_j^0$ to be positively correlated with $\varepsilon_j$, which implies plim $(N^{-1}\mathbf{E}^{0'}\boldsymbol{\epsilon}) > 0$ and a positive bias in the OLS estimate $\hat{b}_0$. In regression (2), the coefficient on abatement $\hat{\beta}_1$ is only identified from the post-repair emissions level $E_j^1$, and as such $\hat{\beta}_1 = -\hat{b}_1$, while $\hat{\beta}_0 = \hat{b}_0 + \hat{b}_1$. Thus, positive bias in $\hat{b}_0$ induces positive bias in $\hat{\beta}_0$, while bias in $\hat{b}_1$ would generate bias in both $\hat{\beta}_0$ and $\hat{\beta}_1$.

[21] This finding is robust to alternative aggregation rules, as indicated in Table 10 in Appendix B.

**Table 6**
Results for the reduced samples.

| Dep. var. $\Delta E_j$ | (1) Failed emissions test | (2) Low income applicant | (3) Directed applicant | (4) 1 std. dev. | (5) 1/2 std. dev. | (6) 1/4 std. dev. | (7) Cost $<400$ | (8) Cost $<300$ |
|---|---|---|---|---|---|---|---|---|
| $E_j^0$ | $-0.278^{***}$ | $-0.257^{***}$ | $-0.311^{***}$ | $-0.254^{***}$ | $-0.263^{***}$ | $-0.207^{**}$ | $-0.293^{***}$ | $-0.270^{***}$ |
| | (0.013) | (0.020) | (0.016) | (0.014) | (0.023) | (0.055) | (0.020) | (0.029) |
| $A_j$ | $0.402^{***}$ | $0.375^{***}$ | $0.442^{***}$ | $0.447^{***}$ | $0.447^{***}$ | $0.430^{***}$ | $0.449^{***}$ | $0.437^{***}$ |
| | (0.013) | (0.019) | (0.016) | (0.012) | (0.016) | (0.024) | (0.019) | (0.028) |
| Usage history | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| VIN prefix f.e. | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Usage charact. | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Failed emiss. (%) | 100 | 77.2 | 85.1 | 79.9 | 82.5 | 84.5 | 79.4 | 78.0 |
| Obs. | 170,530 | 98,622 | 110,981 | 178,293 | 97,320 | 47,457 | 84,020 | 46,573 |
| *R*-squared | 0.284 | 0.362 | 0.331 | 0.301 | 0.371 | 0.437 | 0.354 | 0.425 |

Reported standard errors are heteroskedasticity-robust.
** Statistical significance at the 1% level.
*** Statistical significance at the 0.1% level.

The existence of bias in $\hat{b}_1$ depends on whether and how the emissions-lemon characteristic affects post-repair emissions $E_j^1$. We expect a weak effect of emissions-lemon status on $E_j^1$ because the extent of abatement depends largely on the Smog Check cut-points, which are based on time-invariant vehicle characteristics captured in the VIN Prefix fixed effects, and on repair decisions made by technicians. Moreover, repairs are discrete; a vehicle owner cannot choose a repair that moves her car just below the cut-point and typically has little discretion over the extent of repairs. In our sample, the potential for owners to affect $E_j^1$ is even less than in the general population for reasons discussed in the Data set section. Nonetheless, emissions lemons that fail emissions tests with a higher emissions level could also have higher $E_j^1$. For example, two cars may get an identical repair that abates 12 g/mile of emissions, but if one of those cars is an emissions lemon that started with higher emissions, then it will also have higher post-repair emissions. Even in this case, $\hat{b}_1$ remains unbiased as long as emissions lemons affect post-repair emissions only through $E_j^0$, i.e., if controlling for $E_j^0$ absorbs the effect of the emissions-lemon characteristic.[22] For these reasons, we think the bias in $\hat{b}_1$ is likely to be small.[23]

We conduct three series of robustness checks to address the potential for omitted variable bias due to emissions lemons. In the first check, we report estimates for subsamples that we expect *a priori* to have different distributions of the emissions-lemon characteristic. In the second check, we gradually reduce variation in pre-repair emissions levels towards zero and thereby reduce the emissions-lemon induced variation in our sample. Finally, we reduce the potential for endogenous repairs by using samples for which the CAP subsidy fully covered the cost of repairs.

First, we estimate (2) on various subsamples for which we believe that the proportion of poorly functioning or poorly maintained vehicles would significantly differ from that in the full sample. The idea behind this check is as follows: to the extent that the unobserved emissions-lemon variable causes $\hat{\beta}_0$ and $\hat{\beta}_1$ to be biased, the magnitudes of the biases should change when the model is estimated on subsamples that contain very different shares of emissions lemons. We would therefore be inclined to believe that if the coefficients of interest are, in fact, comparable across such subsamples and the entire sample, the size of the bias should be small. We perform this first series of checks on three different subsamples and report the results in columns (1)–(3) of Table 6.

In column (1) of Table 6, we estimate Eq. (2) on the subsample of vehicles that failed the emissions test, that is, we remove the pseudo control group and rely only on intensive margin variation among vehicles having directly received emission-related repairs. Because 19% of vehicles passed the emissions test, the proportion of lemons should be substantially higher in the subsample of observations that failed the emissions test than in the full sample. However, the coefficients of interest are almost identical to those in column (7) of Table 3.

Next, we estimate model (2) on the subsample of CAP repairs that were funded through the low-income application option. The prevalence of emissions lemons is likely to be higher among low-income motorists than in the full sample, due to assumed lower vehicle maintenance. Again, the coefficients of interest are similar for this subpopulation, which suggests that the emissions-lemons bias, if it exists, is likely small (column (2)). Column (3) shows the results of estimating the main model on the subpopulation of vehicles for which repairs were eligible for CAP because owners were directed to a Test-Only station. Again, the results seem comparable to those obtained for the full sample.

[22] See Appendix C for a formal proof of this claim.

[23] Another potential source of bias in our estimates may be technician skill, irrespective of the emissions-lemon characteristic of vehicles. That is, if skilled technicians perform both larger and longer-lasting abatements relative to unskilled technicians, then vehicles in our sample having received more abatement will systematically have lower emissions increments, biasing the coefficient estimate $\hat{\beta}_1$ toward zero. We are not too worried about this potential source of bias. First of all, since we find a rather large effect of abatement on subsequent emissions accrual, we can interpret our result as a lower bound to the underlying, uncontaminated effect. Second, raising overall technician skill may be one way to achieve emissions reductions. As such, the effect we capture by not controlling for technician quality still has policy relevance.

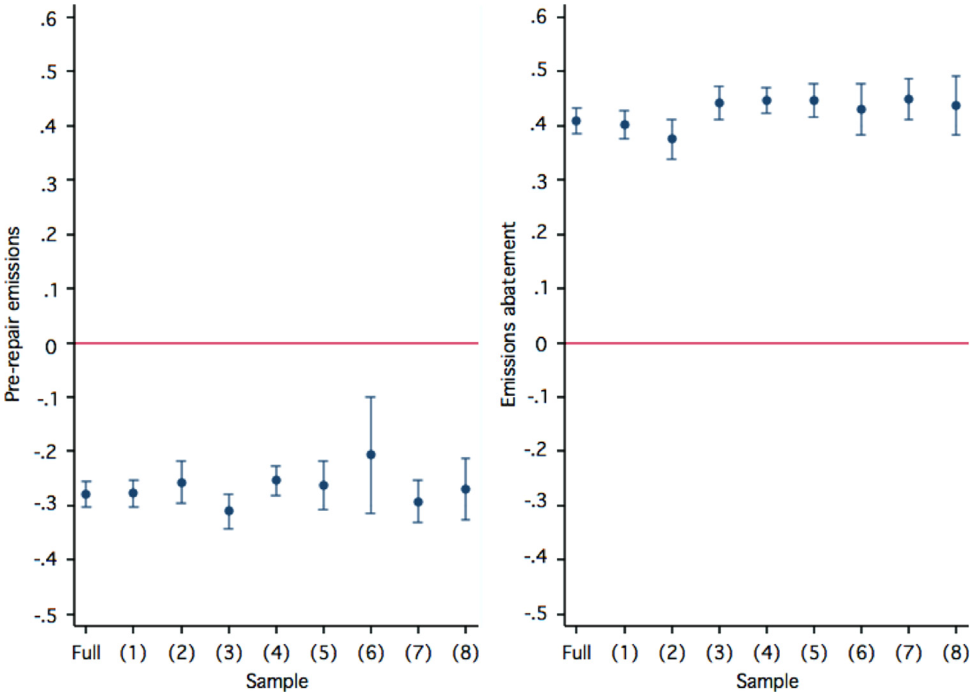


**Fig. 3.** Estimates for the full and reduced samples. *Note*: Sample (1) corresponds to vehicles that failed the emissions test only. Sample (2) corresponds to repairs funded through the low-income option. Sample (3) corresponds to repairs funded through the directed to Test-Only option. Samples (4)–(6) correspond to subsets of vehicles with reduced variation in $E_j^0$. Samples (7) and (8) correspond to subsets of repairs with total cost smaller than \$400 and \$300, respectively.

The second check consists of estimating (2) on a series of subsamples that sequentially reduce variability in pre-repair emissions by removing observations that have emissions far from the mean. Because pre-repair emissions levels vary considerably by model year, we perform this sample reduction on a model-year basis; that is, for each model year, we remove observations for which pre-repair emissions lie one, one-half, and one-fourth standard deviation away from the model-year specific mean. For the most stringent criterion, this exercise reduces the sample to 23% of its initial size.

The idea behind this series of robustness checks is similar to that underlying a regression discontinuity design, in which variation in the variable responsible for assignment to treatment is reduced to such a low level that unobserved relevant variables can be safely assumed to be distributed randomly between the treated and untreated groups. In our case, we posit that if, conditional on model year, there is little variation in pre-repair emissions levels across sample observations, then unobserved vehicle and driver characteristics that both affect $E_j^0$ and $\Delta E_j$ will be randomly assigned across vehicles that failed the emissions test and those that did not (extensive margin), as well as across vehicles that received high abatement and those that received low abatement (intensive margin). Of course, when interpreting such regressions, we should be cautious about the estimated coefficient on $E_j^0$, which becomes less precise as we remove variation in this variable. This series of robustness checks should thus shed more light on the potential bias on $\hat{\beta}_1$ than that on $\hat{\beta}_0$.[24]

Columns (4)–(6) of Table 6 show that our estimate $\hat{\beta}_1$ is only slightly affected by the sample reduction. The estimate $\hat{\beta}_0$ is slightly lower for the reduced samples than for the full sample, but does not vary much across the reduced samples. We also ran regressions on the same set of subsamples adding the square of $E_j^0$ as a regressor, in order to capture possible nonlinearities in the relationship between the emissions increment and the pre-repair emissions level. In each case, the coefficient on abatement was unaffected by this change.[25] We take this evidence as supportive of our OLS estimates for the full sample.

Although we expect CAP repairs to be state of the art, with only limited influence from vehicle owners on the extent and quality of repairs, we investigate the potential for endogenous repairs by estimating model (2) on subsamples of vehicles for which the total cost of repairs would have entirely been covered by the CAP subsidy. If vehicle owners did influence the extent and/or quality of repairs, this influence should be much weaker or nonexistent for the subset of repairs that were funded, at the margin, by the subsidy. Results are presented in the last two columns of Table 6 for repairs less than 400

[24] The main reason why we do not implement a regression discontinuity design based on the Smog Check passing emissions threshold is that there are a total of six cut-points, three for each pollutant under two acceleration simulation modes. As a result, a vehicle's emissions can be close to one cut-point but far from the others. Instead, we prefer to compare vehicles for which *aggregate* emissions are arbitrarily close to the model-year-specific mean.

[25] The coefficient and its standard error were the same up to the third digit.

**Table 7**
Assessment of the measurement error bias.

| Dep. var. $\Delta E_j$ Obs. 93,893 | (1) First measurement | (2) Average measurement |
|---|---|---|
| $E_j^0$ | $-0.294$*** | $-0.292$*** |
| | (0.018) | (0.018) |
| $A_j$ | 0.416*** | 0.420*** |
| | (0.018) | (0.018) |
| Usage history | Yes | Yes |
| VIN prefix f. e. | Yes | Yes |
| Usage characteristics | Yes | Yes |
| $R$-squared | 0.350 | 0.350 |

Reported standard errors are heteroskedasticity-robust.
*** Statistical significance at the 0.1% level.

dollars and 300 dollars, respectively.[26] The means and standard deviations of $E_j^0$ for these sub-samples are quite similar to those of the full sample, while those of $A_j$ are lower, as expected. The estimates of $\beta_0$ and $\beta_1$ are comparable to those obtained for the full sample.

To summarize, Fig. 3 depicts the estimates of $\beta_0$ and $\beta_1$ obtained with the full sample and the reduced samples of Table 6. The figure shows that our results are robust across specifications and that the emissions-lemon bias is likely to be small.

*Measurement error*

If emissions are measured with error, then our estimates may be biased. Measurement error in emissions test results comes mostly from intermittently functioning emissions control components rather than errors in the test itself (Wenzel et al., 2000). This intermittency means that emissions tests provide a snapshot that may over- or under-estimate the average emissions level of the vehicle. This measurement error is likely to affect $E_j^0$ much more than $E_j^1$ because a vehicle is less likely to have poorly functioning emissions control components after it has been repaired. In addition, Bishop et al. (1996) show that measurement variability is much greater for high-emissions vehicles, which also implies smaller measurement error in $E_j^1$.

The smooth emissions curves in Fig. 1 are best thought of as showing average emissions of the vehicle during some time window, and it is these curves that our regression models aim to estimate. If the measurement error in $E_j^0$ is uncorrelated with true average emissions and with the other covariates, then we have classical, or textbook, measurement error and OLS estimates of the parameter $b_0$ in regression (3) are biased towards zero. The larger is the measurement error as a proportion of the variance in measured emissions across vehicles, the larger is this bias. However, if the measurement error is correlated with average emissions or with other covariates, then the bias could be of either sign. For example, if the tests were conducted in such a way that intermittently functioning components were more likely to malfunction during the test than on the road, then $E_j^0$ would tend to overestimate the average emissions of high-polluting vehicles, causing OLS to overestimate the concavity of the emissions trajectory.

We address the potential for bias arising from measurement error in $E_j^0$ by estimating our preferred specification on a subsample of vehicles for which we have two measures of pre-repair emissions conducted at two different stations, with no reported repair in between. This subsample is possible to assemble because for part of the sample period the first failing test had to be conducted at a Test-Only station. CAP applicants would then drive to a CAP-sanctioned repair station where a diagnostic test (our second measure of pre-repair emissions) would be conducted. We have 93,893 such observations in our sample. With average emissions of 17.44 g/mile and standard deviation of 21.02 g/mile, these observations have similar mean and standard deviation to our main sample. The second measurement has a slightly smaller mean than the first (17.39 vs 17.49). Although statistically significant ($t=3.06$), this difference is small in magnitude, so only a negligible amount of repairs may have occurred between tests.

Using a standard variance decomposition, we estimate that 10% of the variance in emissions comes from within-car variation, i.e., measurement error. Under classical measurement error, this would imply a 10% downward bias in our estimate of $b_0$.[27] If the measurement error is correlated with average emissions or with the other covariates, then the bias in $b_0$ could be greater or smaller than 10%. To estimate the potential effect of measurement error, we report in Table 7

[26] The mean of the repair cost in the total sample is $460. Given the size of the subsidy and the size of the co-pays, we believe that repairs conducted for 400 dollars or less are repairs that were genuinely cheaper, not repairs that were kept artificially cheap to avoid additional costs on the part of applicants. We also ran these regressions excluding repairs less than 50 dollars, and the estimates were similar.

[27] To make similar calculations for the post-repair emissions variable $E_j^1$, we used vehicles which failed the smog check but not the emissions test, and which received less than one hundred dollars in repairs. For these vehicles, we considered pre-repair emissions to be an alternative measure of $E_j^1$. We found that 7% of the variance in emissions comes from within-car variation, so we would expect the bias in $b_1$ to be of a similar order of magnitude to that in $b_0$.

**Table 8**
Results for alternative pollutant aggregation rules.

| Dep. var. $\Delta E_j$ Obs. 209,603 | (1) $E=0.1E_{CO}+0.495E_{NOx}+0.405E_{HC}$ | (2) $E=E_{NOx}+E_{HC}$ | (3) $E=3.33E_{NOx}+E_{HC}$ |
|---|---|---|---|
| $E_j^0$ | $-0.313$*** | $-0.433$*** | $-0.395$*** |
| | (0.010) | (0.010) | (0.007) |
| $A_j$ | 0.437*** | 0.507*** | 0.514*** |
| | (0.009) | (0.007) | (0.005) |
| Usage history | Yes | Yes | Yes |
| VIN prefix f. e. | Yes | Yes | Yes |
| Usage characteristics | Yes | Yes | Yes |
| $R$-squared | 0.291 | 0.241 | 0.291 |

Reported standard errors are heteroskedasticity-robust.
*** Statistical significance at the 0.1% level.

regression results for this subsample using two alternative pre-repair emissions estimates. Column (1) reports the estimates when using the first measurement done at the Test-Only station, which corresponds to our measure of $E_j^0$ for the full sample. The results are almost identical to those for the full sample reported in column (7) of Table 3. The second column uses the average of the pre-repair emissions to construct the variables $E_j^0$ and $A_j$, thereby reducing the variance of the measurement error by half. Under classical measurement error, the bias in the OLS estimate of $b_0$ in column (2) would be half that in column (1). Thus, classical measurement error implies that the OLS estimate of $b_0$ should be about 5% greater in column (2) than column (1).

From Eq. (3), the OLS estimate of $b_0$ equals the sum of the coefficients on $E_j^0$ and $A_j$. This estimate is $-0.294+0.416=0.122$ in the model that uses the first measurement and $-0.292+0.420=0.128$ in the model that uses the average measurement. This increase of 5% in the coefficient estimate corresponds exactly to that predicted by classical measurement error. Thus, our estimate of $b_0$ may be biased by 10%, but it is clear from Table 7 that this bias does not change our main results. Our estimates of the two parameters of interest remain the same to two decimal places.

*Aggregation of pollutants*

In the above, we aggregated the three regulated pollutants by simply adding their FTP weight, expressed in grams per mile. Here we test the sensitivity of our results to the choice of aggregation rule.

Not all three pollutants have the same effects on health and the environment. For instance, 1 g of CO emitted into the air has less severe health consequences than 1 g of NOx or HC (Small and Kazimi, 1995; Delucchi et al., 2002).[28] This fact could suggest a different aggregation rule for pollutants, in which the relative weight of CO would be lowered to account for its lower health toxicity. Fullerton and West (2010) use the following aggregation rule: $E=0.1E_{CO}+0.495E_{NOx}+0.405E_{HC}$. Harrington et al. (2000) propose an alternative rule based on pollutant damage estimates from Small and Kazimi (1995): $E=3.33E_{NOx}+E_{HC}$. Recently, Mérel and Wimberger (2012) suggest dropping CO while keeping equal weights for NOx and HC. We implement all three options.

Results reported in Table 8 show that whatever the aggregation rule, the decomposition of the total effect into the trajectory and repair durability effects still holds. The estimated total effect increases slightly in magnitude once the relative weight of CO is lowered. In all specifications, the estimate of $\beta_0$ remains negative, indicating that the emissions trajectory is locally concave. The size of the implied repair durability effect is 7–12%, slightly lower than the 13% estimate obtained with equal emissions weights.

## Cost effectiveness analysis

This section derives measures of repair cost effectiveness based on our empirical estimates and information from the CAP sample.

*One-cycle benefits and costs*

Depending on model-year group, 38–49% of repair-induced abatement does not persist through the next Smog Check inspection. Unfortunately, the exact timing of this abatement loss cannot be known. Although we would expect the emissions recuperation after repairs to occur as the vehicle is being driven, adding an interaction term between abatement and distance traveled, say, $A_j \times DISTANCE_j$ to the main specification barely changes the estimates of $\beta_0$ and $\beta_1$, and the

[28] Although CO is regulated by EPA, no region of California currently exceeds the NAAQS for CO.

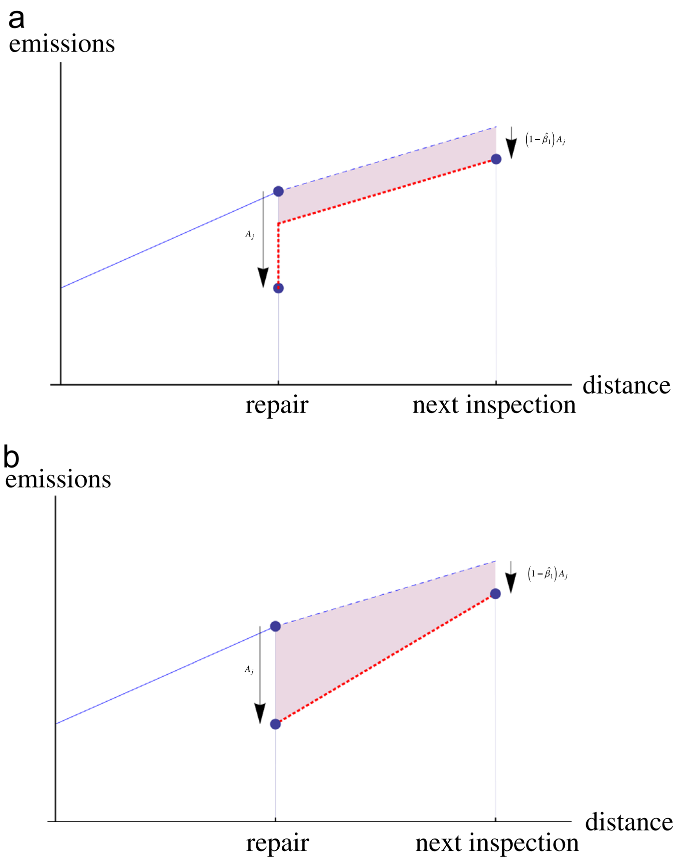


**Fig. 4.** Cumulative emissions savings under different durability assumptions. (a) Short-lived repairs and (b) Smooth recuperation.

estimate on the interaction term turns out to be very small and not statistically significant.[29] This finding could be interpreted as meaning that most of the observed abatement loss occurs shortly after repairs. Alternatively, if the abatement loss is due to a one-time repair failure that happens at a random point in time between inspections, there need not exist a clear relationship between distance traveled and the amount of abatement loss, which could explain why the coefficient on $A_j \times DISTANCE_j$ is insignificant.

Clearly, these two scenarios imply different cumulative emissions benefits throughout the I/M cycle. To keep the discussion simple, assume that the interval between inspections is short enough for the counterfactual emissions path to be locally linear, that is, absent repairs emissions would accrue at a constant rate. The overall concavity in the counterfactual emissions trajectory is then the result of different rates of emissions accrual before and after the repair. This approximation to a smooth emissions path by a piecewise linear function is depicted in Fig. 4.

In the case where the entire emissions recuperation relative to the counterfactual occurs shortly after repairs, cumulative emissions reductions should be calculated as $(1-\hat{\beta}_1) \times A_j \times DISTANCE_j$. This situation is depicted as the dashed emissions path in panel (a) of Fig. 4 and corresponds to the smallest possible cumulative benefits. In contrast, if the emissions recuperation relative to the counterfactual occurs at a later point in time before the subsequent inspection, the cumulative emissions reductions should be larger. If one is willing to assume that the failure event happens with equal probability at any point in time between inspections and that the vehicle is driven uniformly during the period, then the expected emissions reductions are the same as those calculated assuming that emissions accumulate linearly in distance between inspections, $(1-\hat{\beta}_1/2) \times A_j \times DISTANCE_j$. This situation is depicted in panel (b) of Fig. 4.

We compute these two measures of emissions reductions for vehicles in our CAP fleet that failed the emissions test. Since we know the total cost of each CAP repair, we also compute the implied cost-effectiveness of emission-related repairs. Summary statistics for repair costs are provided in Table 1. Our cost measure includes direct repair costs (parts and labor) paid by CAP and by the vehicle owner, excluding the cost of the initial failing test. On average, repair costs for the sub-sample of vehicles having failed the emissions test are \$461, and they are slightly higher for vehicles having failed at the "gross polluter" level.

[29] This pattern happens even if we add $E_j^0 \times DISTANCE_j$ as a control. We tried to interact $A_j$ with more flexible functions of distance traveled. The overall relationship between abatement loss and distance remained very flat.

**Table 9**
Cost effectiveness analysis.

| Assumption | (1) Abatement persistence | | (2) Historical deterioration | | (3) Random repair failure | | (4) Short-lived repair | |
|---|---|---|---|---|---|---|---|---|
| Sample | Failed emissions | Gross polluters | Failed emissions | Gross polluters | Failed emissions | Gross polluters | Failed emissions | Gross polluters |
| $E=E_{CO}+E_{NOx}+E_{HC}$ | | | | | | | | |
| Emissions saved (mill. lbs) | 46.60 | 33.72 | 38.73 | 28.16 | 36.16 | 26.24 | 25.73 | 18.76 |
| | (–) | (–) | (0.21) | (0.17) | (0.20) | (0.15) | (0.39) | (0.31) |
| Repair cost per pound ($/lb) | 1.69 | 0.99 | 2.03 | 1.18 | 2.17 | 1.27 | 3.05 | 1.77 |
| $E=E_{NOx}+E_{HC}$ | | | | | | | | |
| Emissions saved (mill. lbs) | 5.89 | 3.59 | 4.61 | 2.81 | 4.39 | 2.67 | 2.88 | 1.74 |
| | (–) | (–) | (0.03) | (0.01) | (0.02) | (0.01) | (0.04) | (0.02) |
| Repair cost per pound ($/lb) | 13.33 | 9.27 | 17.04 | 11.84 | 17.91 | 12.49 | 27.27 | 19.13 |

Standard errors for emissions saved are reported in brackets.

We do not attempt to derive an overall measure of I/M program cost-effectiveness. Such a measure would include costs other than emission-related repairs, such as administrative costs and the opportunity cost of driving to and from the station, as well as co-benefits such as improved fuel economy.[30] The measure would also account for costs associated with inspecting vehicles that do not fail the emissions test and net benefits from early retirement of high emitters.

Because emission-related repairs lie at the core of I/M programs, our cost-effectiveness measures are meaningful by themselves. They capture the direct and irreducible part of the cost of repairing non-conforming vehicles, assuming away indirect costs associated with selecting such vehicles and bringing them to the repair location. In their evaluation of the California Enhanced Smog Check Program, Singer and Wenzel (2003, p. 2588) write

> Repair of vehicles that failed an initial, official Smog Check appears to be the most important mechanism of emission reductions, but pre-inspection maintenance and repair also contributed substantially. Benefits from removal of nonpassing vehicles accounted for a small portion of total benefits.

Because our sample only includes repairs performed at high-quality stations and subsidized by the state, we expect our cost-effectiveness measures to constitute lower bounds to the cost effectiveness of repairs performed at an average station, and to that of preemptive repairs performed outside of a recorded inspection.

We report results in Table 9 for two aggregation rules: our baseline aggregation by weight, and a second one that ignores the contribution of CO. To this end, we use the estimates for model year groupings reported in Table 5 for the baseline aggregation and in Appendix B for NOx+HC. Because the contribution of CO by weight dominates our baseline measure of emissions, the omission of CO results in much lower emissions abatement and much higher cost per pound of emissions avoided. We report results separately for all vehicles in our sample that failed the emissions test and for the subset of gross polluting vehicles.

Column (1) of Table 9 reports the emissions reductions and the cost per pound of emissions saved for our fleet of failing vehicles, based on the persistence hypothesis. Fleet-level emissions reductions amount to about 47 million pounds of aggregate emissions, with an abatement cost of $1.69/lb. Because our results show that emissions deteriorate faster after repairs than they would have in the absence of repairs, this method overestimates emissions reductions and thus underestimates abatement costs.

A more sophisticated approach would assume that emissions accrue smoothly at the historical deterioration rate at the post-repair emissions level. This approach implicitly assumes that repairs are as good as new. Results of cost effectiveness following this approach are reported in column (2). Emissions savings for vehicle $j$ are calculated as $(1+\hat{\beta}_0/2)\times A_j \times DISTANCE_j$, assuming a locally linear emissions accrual between inspections. Because the estimated emissions trajectory is concave, the slope of the counterfactual trajectory is lower than the historical accrual rate, and calculated emissions reductions are lower than under the persistence hypothesis. The corresponding decrease in fleet-level emissions reductions is sizable, about 17%.

Figures reported in column (3) are based on the assumption that a single repair failure occurs at a random moment between inspections, and are equivalent to those obtained assuming that emissions recuperation after repairs is a smooth process occurring as the vehicle is being driven. Under this scenario, fleet-level emissions reductions are 22% lower than under the persistence hypothesis, and 7% lower than under historical deterioration. Column (4) is based on the assumption that repairs are short-lived, in the sense that the loss in repair-induced abatement occurs immediately after repairs. As such,

[30] McConnell (1990) explicitly accounts for such costs in the context of Maryland's centralized I/M program and shows that driver costs represent more than two-thirds of net repair expenditures and administrative costs. Harrington et al. (2000) find in the context of Arizona's centralized I/M program that motorists' waiting time and travel costs account for about 30% of total costs.

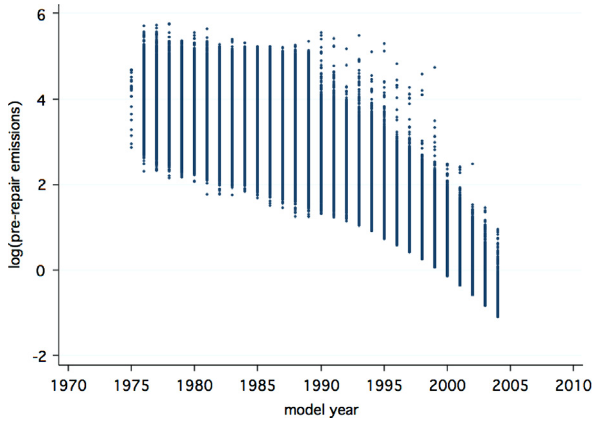


**Fig. 5.** Logarithm of pre-repair emissions by model year.

figures in this column correspond to the lowest fleet-level emissions reductions and the highest cost per pound of emissions avoided. Fleet-level emissions reductions are 45% lower than those obtained under the persistence hypothesis.

Should failure events occur equi-probably at any point in time between inspections, the resulting cost is $2.17 per pound of aggregate emissions avoided. This cost drops to $1.27 per pound of emissions if one only considers repairs performed on gross polluting vehicles. This finding confirms previous results of Mérel and Wimberger (2012) who find that the marginal abatement cost of high emitters is much smaller than that of cleaner vehicles. (Their study does not account for differences in miles driven across vehicles.) If we instead assume that the abatement loss occurs shortly after repairs, the resulting cost per pound of emissions becomes $3.05/lb for vehicles which failed the emissions test, and $1.77/lb for gross polluters only. Therefore, an almost two-fold improvement in cost effectiveness can be achieved by repairing gross polluters only.

An even more promising avenue would be to exempt vehicles that are unlikely to be high emitters, for instance by extending the current exemption for recent model-year vehicles, as this would greatly reduce testing costs. (About 16% of the vehicles subject to Smog Check actually fail the test on average, California Air Resources Board and Department of Consumer Affairs/Bureau of Automotive Repair, 2004.) Fig. 5 plots the logarithm of pre-repair emissions against vehicle model year and clearly suggests that one way to "catch" high-polluting vehicles without requiring universal testing is to focus on older model-year vehicles. We calculated that emissions reductions commensurate with those obtained for gross polluters, and at a commensurate cost per pound, could be obtained by focusing on model years 1989 and older within our sample of CAP vehicles (including vehicles having passed the emissions test).[31]

*Comparison with other measures*

California's Smog Check Program targets three different pollutants, which makes it difficult to compare our cost-effectiveness measures to the cost effectiveness of other methods of air pollution abatement. However, we obtain some perspective if we focus on NOx and HC. A commonly used cost-effectiveness threshold in California is that from the Carl Moyer program, which provides grants to reduce mobile-source emissions of NOx, VOCs, and inhalable particulate matter (California Air Resources Board, 2008). As of 2008, the program specified a threshold of $16,000 per weighted ton of emissions for program funding, which would translate into $8 per pound of HC and NOx combined.[32] Our results for NOx and HC suggest that CAP repairs fail to meet that threshold. Under the assumption of repair persistence, the threshold is almost met by gross polluters, but this is no longer true once we account for the lack of repair persistence, even assuming as-good-as-new repairs ($11.84/lb). If we assume a repair failure at a random moment between inspections, the cost per pound for the set of vehicles that failed the emissions test rises to $17.91/lb, more than twice the Carl Moyer threshold. Given the fact that our calculations only capture expenditures for CAP vehicles that failed the emissions test and reflect state-of-the-art repairs, it seems clear that the threshold would be exceeded at the California fleet level.[33]

A recent report by BAR on the costs and benefits of CAP for the fiscal year 2011–2012 provides some validation for our estimates (California Bureau of Automotive Repair, 2013). According to this report, CAP spent $9,529,675 to repair 23,507 vehicles during that fiscal year. The associated emissions reductions were estimated to be 1772.1 tons of combined NOx, HC and CO on a yearly basis, or 309.6 tons of NOx and HC. Given that the average time between the two I/M cycles is 1.76 years

[31] More specifically, 24.13 million lbs of CO-NOx-HC are abated under the random repair failure scenario, at a cost of $1.34/lb. Focusing on even older vehicles (model year 1984 and older) decreases the cost ratio to $1.04/lb, but it reduces the total abatement to a mere 8.24 million pounds.

[32] The cost-effectiveness threshold was raised in 2011 to $16,640 per ton and again as of April 1, 2012 to $17,080 per ton.

[33] Vehicles failing the inspection for visual or functional reasons may experience small emissions reductions and still incur significant test and repair costs. Vehicles passing the inspection experience no emissions reductions but still incur test costs. In addition, our CAP vehicles are dirtier on average than vehicles failing the California Smog Check, see Table 2. Our results and those of Mérel and Wimberger (2012) suggest that cost effectiveness ratios are lower for dirtier vehicles, so our cost per pound likely understates that of the average repair in California.

for vehicles failing emissions in our sample, the cumulative expected emissions reductions between two cycles for this recent set of CAP repairs would rise to about 3120 tons for HC, NOx and CO and to 545 tons for HC and NOx. The implied cost-effectiveness ratio would be $1.53/lb of HC, NOx and CO and $8.74/lb of HC and NOx. These values are slightly lower than our estimates of cost effectiveness under the persistence assumption, and therefore lower than our preferred estimates taking account of emissions recuperation.

The results of Fowlie et al. (2012b) also provide a point of comparison for our estimates. These authors compute engineering-based estimates of the cost of reducing NOx emissions from power plants and motor vehicles in the U.S. Their preferred estimate for the marginal abatement cost of NOx emissions from newly manufactured motor vehicles is $0.45/lb of NOx, while that from power generation is $0.95/lb, indicating that mobile source NOx reductions are more cost-effective. In contrast, our estimates suggest that reducing emissions of in-use vehicles through I/M programs is costly. When fully accounting for CO reductions, our more optimistic estimate of cost effectiveness is $1.75/lb, a figure already above Fowlie et al.′s (2012b) estimate for point sources. Our estimate increases by one order of magnitude when focusing on NOx and HC reductions.

As such, our findings indicate that I/M repairs are far from being cost effective when compared to other NOx reduction efforts.

## Conclusion

I/M programs require motorists to pay for emission-related repairs whenever the tailpipe emissions of their vehicles, measured at regular time intervals, exceed regulatory thresholds. In this paper, we measure the persistence of repair-induced emissions abatement throughout the I/M cycle using data from California over the period 2000–2010. Because California has a relatively old fleet of vehicles, its Smog Check Program continues to represent an important lever to control emissions from mobile sources, in addition to the secular nationwide tightening in emissions standards for newly manufactured vehicles.

We show that on average 41% of the repair-induced abatement is lost by the subsequent Smog Check inspection. About 28% of the abatement is lost due to the local concavity in the emissions trajectory of vehicles: repairs tend to bring vehicles back to a portion of their trajectory where emission deterioration is occurring at a faster rate. To the extent that vehicles cannot be repaired to be "better than new," there is not much a regulator could do to eliminate this loss. The remaining 13% loss in abatement is directly attributable to a lack of repair durability. In principle, this loss could be reduced by improving the quality of repairs, but the fact that our sample of vehicles underwent repairs at high-performing stations suggests that limited scope may exist for such improvement.

Our results have implications for calculating the benefits of I/M programs. A common assumption is that the rate of emissions deterioration of a vehicle over usage is essentially unaffected by repairs, so that cumulative benefits from repairs can be calculated by multiplying the initial abatement by the expected miles traveled. Another assumption, which has been used in the California Air Resources Board′s EMFAC model, is that emission deterioration after a repair follows historical emission deterioration at the post-repair emissions level. None of these assumptions finds empirical support in our study, although the second certainly provides a better approximation than the first.

In our most optimistic scenario regarding the timing of the repair failure, cumulative emissions reductions for our fleet are 22% lower than those calculated assuming abatement is persistent and 7% lower than those calculated based on historical deterioration. In our less optimistic scenario, emissions reductions are 45% lower than under abatement persistence, and 34% lower than under historical deterioration. Our estimates indicate that Smog Check repairs are expensive in comparison to existing federal and state policies to reduce air pollutants such as NOx, and likely fail to meet commonly accepted cost-effectiveness thresholds.

Our empirical finding that emission-related repairs, despite their ability to reduce the emissions rates of vehicles, are not as good as new—in the sense that the resulting deterioration rate is higher than the historical one—illustrates the broader proposition that fixes made on defective durable goods cannot completely erase the effects of time and use. As we hope to have shown here, this proposition has direct implications regarding the estimation of cumulative benefits from repairing such goods.

## Appendix A. Data set

Repairs recorded in BAR′s CAP database between July 2000 and August 2010 were matched to BAR′s Vehicle Identification Database (VID) to obtain Smog Check records and vehicle characteristics, and to the DMV′s Vehicle Registration Database (VRD) to obtain registration due dates and registration status. The VID is publicly available; the VRD and the CAP database are not. The CAP database is organized by application type (to either the repair or the retirement program). These administrative records include the many applications denied or withdrawn. Applications denied but approved on re-application within 180 days were consolidated into one record, as were applications for repairs that were re-assigned to the retirement program without any state payment for the repair. Repairs at one shop over several days were consolidated into one record for each VIN. Most vehicles appear only once over the sample period. Since CAP subsidies are available only to those vehicles failing their regular biennial inspection, we first checked that the recorded inspection results (in the VID) and the corresponding CAP repairs (in the CAP database) could be linked to a legitimate registration due date.

**Table 10**
Results for the model-year groupings, $E = E_{NOx} + E_{HC}$.

| Dep. var. $\Delta E_j$ | (1) 1975–1980 | (2) 1981–1983 | (3) 1984–1986 | (4) 1987–1992 | (5) 1993–1995 | (6) 1996–2000 | (7) 2001–2003 | (8) 2004 |
|---|---|---|---|---|---|---|---|---|
| $E_j^0$ | −0.250*** | −0.403*** | −0.388*** | −0.477*** | −0.444*** | −0.398*** | −0.218 | −0.340 |
| | (0.036 ) | (0.028) | (0.014) | (0.012) | (0.025) | (0.044) | (0.129) | (0.265) |
| $A_j = E_j^0 - E_j^1$ | 0.399*** | 0.559*** | 0.540*** | 0.531*** | 0.482*** | 0.418*** | 0.497*** | 0.421 |
| | (0.031) | (0.023) | (0.011) | (0.009) | (0.017) | (0.044) | (0.091) | (0.257) |
| Usage history | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| VIN prefix f. e. | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Usage characteristics | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Obs. | 4844 | 9056 | 28,000 | 71,979 | 46,097 | 44,204 | 5053 | 370 |
| $R$-squared | 0.254 | 0.331 | 0.274 | 0.207 | 0.088 | 0.213 | 0.461 | 0.565 |

Reported standard errors are heteroskedasticity-robust.
*** Statistical significance at the 0.1% level.

Our usable sample was constructed as follows from the population of vehicles for which CAP funded repairs during the period, a total of 368,675 vehicle-repairs.

We first used the VID to recover the sequence of tests corresponding to the CAP repair—typically, a fail followed by a pass, though for many vehicles there was a sequence of fails before the pass, in which case we used the first fail in the sequence. We dropped 38,690 observations for which the CAP repairs did not result in a passing certificate.[34] A large number of CAP vehicles (77,752) do not have a subsequent Smog Check on record. Most of those were junked, issued a salvage certificate, registered as non-operational, or moved out of state, as indicated by the VRD; some have disappeared from California's records. To the extent that these vehicles are likely to accumulate emissions at a different rate than other vehicles, omitting them from the sample will introduce bias. This is particularly true of vehicles junked or registered as non-operational, because these likely accrue emissions at a faster rate than other vehicles. As such, our estimates will have to be understood as reflective of the vehicles that survive to the next Smog Check.

Because we track emissions using concentrations measured using the ASM procedure, 21,701 vehicles with incomplete ASM readings had to be excluded. Most of these have some property such as four-wheel drive that makes ASM testing impossible, but some others were tested at one time on an ASM machine and other times not, for no obvious reason.

We excluded a handful of vehicles with model year 1974, as they would have been the only vehicles in one of the model-year groups based on the Smog Check emission standard categories. We also excluded a few vehicles for which the VIN was incorrectly recorded and thus could not be used to construct the VIN Prefix fixed effects. Finally, we excluded vehicles for which the odometer increment between the passing test of the first cycle and the first test of the next cycle (after correcting for odometers having rolled over) was either nonpositive of greater than 150,000 miles. We further excluded vehicles for which the time elapsed between the passing test of the first cycle and the first test of the next cycle was less than 180 days or exceeded 1095 days.

Overall, this set of selections led us to 217,067 potentially usable observations. Among those, we further excluded observations if the lapse of time between the first failing test and the passing test was larger than 180 days. After these subtractions, we ended up with 209,603 observations, each corresponding to a CAP repair.

The VIN Prefix of a vehicle essentially reflects the following characteristics: make, model, engine, transmission, model year and plant where the vehicle is manufactured. We use the VIN Prefix as fixed effects in our main specification. For vehicles with model year 1980 and older, there is no common VIN Prefix among vehicles sharing the same characteristics. For these older vehicles we thus construct "pseudo" VIN Prefix fixed effects based on information directly recorded in the VID regarding vehicle model year, class, make, cylinders, engine and transmission. The pseudo fixed effects are constructed by categorizing and interacting these variables.

## Appendix B. Model-year groupings, no CO

See Table 10 above.

## Appendix C. Biases on $\hat{b}_0$ and $\hat{b}_1$

For tractability, we assume that all variables are demeaned and ignore the covariates $\mathbf{X}_j$ and $\mathbf{Z}_j$. We denote $\mathbf{E}_j = (E_j^0\ E_j^1)$ and denote $L_j$ the emissions lemon variable included in the error term $\varepsilon_j$ in model (3). We posit the following correlation

[34] Many of these vehicles were junked, the repair having been no more than diagnosis.

structure:

$$\begin{cases} E_j^0 = \alpha_0 L_j + u_{0j} \\ E_j^1 = \alpha_1 E_j^0 + u_{1j} = \alpha_1 \alpha_0 L_j + \alpha_1 u_{0j} + u_{1j} \end{cases}$$

with plim $(N^{-1}\mathbf{L}'\mathbf{u}_0) = 0$, plim $(N^{-1}\mathbf{L}'\mathbf{u}_1) = 0$, plim $(N^{-1}\mathbf{u}_0'\mathbf{u}_1) = 0$, plim $(N^{-1}\mathbf{L}'\mathbf{L}) = \sigma_L^2$, plim $(N^{-1}\mathbf{u}_0'\mathbf{u}_0) = \sigma_0^2$, and plim $(N^{-1}\mathbf{u}_1'\mathbf{u}_1) = \sigma_1^2$.

The bias on $\hat{\mathbf{b}} = \begin{pmatrix} \hat{b}_0 \\ \hat{b}_1 \end{pmatrix}$ will be proportional to

$$(\text{plim } (N^{-1}\mathbf{E}'\mathbf{E}))^{-1}(\text{plim } (N^{-1}\mathbf{E}'\mathbf{L})).$$

We have that

$$\text{plim } (N^{-1}\mathbf{E}'\mathbf{E}) = \begin{pmatrix} \alpha_0^2\sigma_L^2 + \sigma_0^2 & \alpha_0^2\alpha_1\sigma_L^2 + \alpha_1\sigma_0^2 \\ \alpha_0^2\alpha_1\sigma_L^2 + \alpha_1\sigma_0^2 & \alpha_0^2\alpha_1^2\sigma_L^2 + \alpha_1^2\sigma_0^2 + \sigma_1^2 \end{pmatrix}$$

and

$$\text{plim } (N^{-1}\mathbf{E}'\mathbf{L}) = \begin{pmatrix} \alpha_0\sigma_L^2 \\ \alpha_0\alpha_1\sigma_L^2 \end{pmatrix}$$

so that, after simplification

$$\text{bias}\left(\hat{\mathbf{b}}\right) \propto \frac{\alpha_0\sigma_L^2}{\alpha_0\sigma_L^2 + \sigma_0^2}\begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

## References


Ando, A., McConnell, V., Harrington, W., 2000. Costs, emissions reductions, and vehicle repair: evidence from Arizona. Journal of the Air and Waste Management Association 50 (4), 509–521.

Aufhammer, M., Kellogg, R., 2011. Clearing the air? the effects of gasoline content regulation on air quality. American Economic Review 101, 2687–2722.

Bishop, G.A., Stedman, D.H., Ashbaugh, L., 1996. Motor vehicle emissions variability. Journal of the Air and Waste Management Association 46, 667–675.

Bishop, G.A., Stedman, D.H., Peterson, J.E., Hosick, T.J., Guenther, P.L., 1993. A cost-effectiveness study of carbon monoxide emissions reduction utilizing remote sensing. Journal of the Air and Waste Management Association 43, 978–988.

California Air Resources Board, 2001a. Section 4: Development of the Basic Emission Rates. EMFAC 2000 On-Road Emission Model Methodology Documentation. ⟨http://www.arb.ca.gov/msei/onroad/doctable_test.htm⟩.

California Air Resources Board, 2001b. Section 8: Methodology Used to Model Inspection and Maintenance (I&M) Programs. EMFAC 2000 On-Road Emission Model Methodology Documentation. ⟨http://www.arb.ca.gov/msei/onroad/doctable_test.htm⟩.

California Air Resources Board, 2003. 2003 State and Federal Strategy for the California State Implementation Plan, Section II Mobile Sources. ⟨http://www.arb.ca.gov/planning/sip/stfed03/revsect2.pdf⟩.

California Air Resources Board, 2004. California Enhanced Inspection and Maintenance (I/M) Program Evaluation, Technical Support Document Part 1. ⟨http://www.arb.ca.gov/msprog/smogcheck/jun04/tsd_part1.pdf⟩.

California Air Resources Board, 2008. The 2008 Carl Moyer Program Guidelines Part IV. ⟨http://www.arb.ca.gov/msprog/moyer/2008guideline_updates.htm⟩.

California Air Resources Board, 2009. Transitioning Away from Smog Check Tailpipe Emission Testing in California for OBD II Equipped Vehicles. ⟨http://www.arb.ca.gov/msprog/smogcheck/march09/transitioning_to_obd_only_im.pdf⟩.

California Air Resources Board, 2012a. 2012 Smog Check Performance Report. ⟨http://www.bar.ca.gov/80_BARResources/ftp/pdfdocs/2012%20Smog%20Check%20Performance%20Report.pdf⟩.

California Air Resources Board, 2012b. Quick and Easy Smog Checks Coming to California. News Release, October 18, 2012.

California Air Resources Board and Department of Consumer Affairs/Bureau of Automotive Repair, 2004. April 2004 Evaluation of the California Enhanced Vehicle Inspection and Maintenance (Smog Check) Program. ⟨http://www.arb.ca.gov/msprog/smogcheck/dec05/final_approved_smogcheck_report_leg_2005.pdf⟩.

California Bureau of Automotive Repair, 2013. CAP Annual Report for Fiscal Year 2011–2012. ⟨http://www.smogcheck.ca.gov/80_BARResources/01_CAP&GoldShield/CAPAnnualReport_11_12.pdf⟩.

Chay, K.Y., Greenstone, M., 2005. Does air quality matter? Evidence from the housing market. Journal of Political Economy 113 (2), 376–424.

Delucchi, M.A., Murphy, J.J., McCubbin, D.R., 2002. The health and visibility cost of air pollution: a comparison of estimation methods. Journal of Environmental Management 64, 139–152.

Downing, P.B., 1973. An economic analysis of periodic vehicle inspection programs. Atmospheric Environment 7, 1237–1246.

Environmental Protection Agency, 2001. Determination of NOx and HC Basic Emission Rates, OBD and I/M Effects for Tier 1 and later LDVs and LDTs. ⟨http://www.epa.gov/otaq/models/mobile6/r01056.pdf⟩.

Environmental Protection Agency, 2002. MOBILE6 Inspection/Maintenance Benefits Methodology for 1981 through 1995 Model Year Light Vehicles. ⟨http://www.epa.gov/otaq/models/mobile6/r02014.pdf⟩.

Environmental Protection Agency, 2003. Major Elements of Operating I/M Programs. ⟨http://epa.gov/oms/epg/420b03012.pdf⟩.

Environmental Protection Agency, 2008. The Plain English Guide to the Clean Air Act—Cars, Trucks, Buses, and "Nonroad" Equipment. ⟨http://www.epa.gov/air/peg/carstrucks.html⟩.

Environmental Protection Agency, 2010a. MOVES2010 Highway Vehicle Temperature, Humidity, Air Conditioning, and Inspection and Maintenance Adjustments. ⟨http://www.epa.gov/otaq/models/moves/420r10027.pdf⟩.

Environmental Protection Agency, 2010b. National Ambient Air Quality Standards (NAAQS). ⟨http://www.epa.gov/air/criteria.html⟩.

Environmental Protection Agency, 2011a. Development of Emission Rates for Light-Duty Vehicles in the Motor Vehicle Emissions Simulator (MOVES2010). ⟨http://www.epa.gov/otaq/models/moves/documents/420r11011.pdf⟩.

Environmental Protection Agency, 2011b. Health Effects. ⟨http://www.epa.gov/airquality/carbonmonoxide/health.html⟩.

Environmental Protection Agency, 2011c. Light-duty exhaust emission rates in MOVES2010. In: Presentation at MOVES Workshop, Ann Arbor, MI. ⟨http://www.epa.gov/otaq/models/moves/conference2011/ld-exhaust-moves-2011.pdf⟩.
Environmental Protection Agency, 2012. Using MOVES to Prepare Emission Inventories in State Implementation Plans and Transportation Conformity: Technical Guidance for MOVES2010, 2010a and 2010b. ⟨http://www.epa.gov/otaq/models/moves/documents/420b12028.pdf⟩.
Fowlie, M., 2010. Emissions trading, industry restructuring, and investment in pollution abatement. American Economic Review 100, 837–869.
Fowlie, M., Holland, S.P., Mansur, E.T., 2012a. What do emissions markets deliver and to whom? Evidence from Southern California′s NOx trading program. American Economic Review 102 (2), 965–993.
Fowlie, M., Knittel, C., Wolfram, C., 2012b. Sacred cars: optimal regulation of stationary and non-stationary pollution sources. American Economic Journal: Economic Policy 4 (1), 98–126.
Freeman, A.M., 2002. Environmental policy since earth day i: what have we gained?. Journal of Economic Perspectives 16 (1), 125–146.
Fullerton, D., West, S.E., 2002. Can taxes on cars and on gasoline mimic an unavailable tax on emissions? Journal of Environmental Economics and Management 43, 135–157.
Fullerton, D., West, S.E., 2010. Tax and subsidy combinations for the control of car pollution. B.E. Journal of Economic Analysis and Policy 10 (1), Article 8.
Harrington, W., McConnell, V., 1994. Modeling in-use vehicle emissions and the effects of inspection and maintenance programs. Journal of the Air and Waste Management Association 44 (6), 791–799.
Harrington, W., McConnell, V., Ando, A., 2000. Are vehicle emission inspection programs living up to expectations? Transportation Research Part D 5, 153–172.
Harrington, W., McConnell, V.D., Cannon, M., 1998. A behavioral analysis of EPA′s MOBILE emission factor model. Discussion Paper 98–47, Resources for the Future. Washington, DC.
Hubbard, T.N., 1997. Using inspection and maintenance programs to regulate vehicle emissions. Contemporary Economic Policy 15, 52–62.
Innes, R., 1996. Regulating automobile pollution under certainty, competition, and imperfect information. Journal of Environmental Economics and Management 31, 219–239.
Knittel, C.R., Sandler, R., 2011. Cleaning the Bathwater *with* the Baby: The Health Co-Benefits of Carbon Pricing in Transportation. NBER Working Paper No. 17390.
Lawson, D.R., 1993. "Passing the Test"—human behavior and California′s smog check program. Journal of the Air and Waste Management Association 43 (12), 1567–1575.
Lawson, D.R., 1995. The costs of "M" in I/M—reflections on inspection/maintenance programs. Journal of the Air and Waste Management Association 45, 465–476.
McConnell, V., 1990. Costs and benefits of vehicle inspection: a case study of the Maryland region. Journal of Environmental Management 30, 1–15.
Mérel, P., Wimberger, E., 2012. Improving air quality in California′s San Joaquin Valley: the role of vehicle heterogeneity in optimal emissions abatement. Journal of Environmental Economics and Management 63, 169–186.
Muller, N.Z., Mendelsohn, R., Nordhaus, W., 2011. Environmental accounting for pollution in the United States economy. American Economic Review 101, 1649–1675.
National Research Council, 2000. Modeling Mobile-Source Emissions. National Academy Press, Washington, DC.
National Research Council, 2001. Evaluating Vehicle Emissions Inspection and Maintenance Programs. National Academy Press, Washington, DC.
Nickell, S., 1981. Biases in dynamic models with fixed effects. Econometrica 49 (6), 1417–1426.
Sierra Research, 2009. Evaluation of the California Smog Check Program Using Random Roadside Data. Report Prepared for the California Air Resources Board and the California Bureau of Automotive Repair. ⟨http://www.arb.ca.gov/msprog/smogcheck/march09/roadsidereport.pdf⟩.
Singer, B.C., Wenzel, T.P., 2003. Estimated emission reductions from California′s enhanced smog check program. Environmental Science and Technology 27, 2588–2595.
Small, K.A., Kazimi, C., 1995. On the costs of air pollution from motor vehicles. Journal of Transport Economics and Policy 29 (1), 7–32.
Wenzel, T., 2001. Evaluating the long-term effectiveness of the Phoenix IM240 program. Environmental Science and Policy 4, 377–389.
Wenzel, T., Singer, B.C., Slott, R., 2000. Some issues in the statistical analysis of vehicle emissions. Journal of Transportation Statistics, 1–14.
Wenzel, T.P., Singer, B.C., Slott, R.R., Stedman, D.H., 2004. Short-term emissions deterioration in the California and Phoenix I/M programs. Transportation Research Part D 9, 107–124.
White, L.J., 1982. U.S. automotive emissions controls: how well are they working?. American Economic Review 72 (2), 332–335.