Title
Regulating Untaxable Externalities: Are Vehicle Air Pollution Standards Effective and Efficient?

Author
Shapiro, Joseph S.

Publication Date
2022-12-16

Regulating Untaxable Externalities: Are Vehicle Air Pollution Standards Effective and Efficient?*

Mark R. Jacobsen James M. Sallee

Joseph S. Shapiro Arthur A. van Benthem

November 2022

Abstract

What is a feasible and efficient policy to regulate air pollution from vehicles? A Pigouvian tax is technologically infeasible. Most countries instead rely on exhaust standards that limit air pollution emissions per mile for new vehicles. We assess the effectiveness and efficiency of these standards, which are the centerpiece of US Clean Air Act regulation of transportation, and counterfactual policies. We show that the air pollution emissions per mile of new US vehicles has fallen spectacularly, by over 99 percent, since standards began in 1967. Several research designs with a half century of data suggest that exhaust standards have caused most of this decline. Yet exhaust standards are not cost-effective in part because they fail to encourage scrap of older vehicles, which account for the majority of emissions. To study counterfactual policies, we develop an analytical and a quantitative model of the vehicle fleet. Analysis of these models suggests that tighter exhaust standards increase social welfare and that increasing registration fees on dirty vehicles yields even larger gains by accelerating scrap, though both reforms have complex effects on inequality.

JEL Codes: H21, H23, H70, Q50, R40

*We thank Akshaya Jha, Ashley Langer, Kyle Meng, Nick Ryan, Boyoung Seo, Hannes Spinnewijn, and Ann Wolverton for useful discussions, and participants in seminars at AERE, ASSA, CREATE, ITS Davis, ETH Zurich, Georgetown, Grenoble, Harvard, Helsinki, Illinois, Mannheim, LSE, NBER Environmental/Energy, NBER Urban, NBER Transportation in the 21st Century, Oxford, the Paris School of Economics, Stanford, the Tinbergen Institute, Toulouse, U Auckland, UC Berkeley, UC Davis, UC Santa Barbara-Occasional Workshop, U Chicago, UEA, U Hamburg, UPenn, USC, Yale, the World Bank, and ZEW for useful comments; Kipper Berven, Adrian Fernandez, Lillian Holmes, Minji Kim, Ray Kim, Kechen Liu, Vivek Sampathkumar, and Katherine Wong for excellent research assistance; Alberto Ayala, Rasto Brezny, Andy Burnham, Rob Dawson, John German, Dave Good, Tim Johnson, Jarod Kelly, Jim Sidebottom, Allen Lyons, Michael Wang, Tom Wenzel, and Ning Wigraisakda for sharing data and explaining institutions; and NSF SES-1850790 and SES-2117158, USDA NIFA-1021396, USDA NIFA-1010175, and the Kleinman Center for Energy Policy, the Mack Institute, the Wharton Dean's Research Fund, and Analytics at Wharton for generous support.

Affiliations: Jacobsen: UC San Diego and NBER, m3jacobsen@ucsd.edu; Sallee and Shapiro: UC Berkeley and NBER, sallee@berkeley.edu and joseph.shapiro@berkeley.edu; van Benthem: University of Pennsylvania and NBER, arthurv@wharton.upenn.edu

1 Introduction

The world has 1.4 billion passenger vehicles (IHS Markit 2022). How should governments regulate their air pollution? This paper studies the effectiveness and efficiency of air pollution exhaust standards and counterfactual policies.

Vehicle transportation is one of the world's largest sources of air pollution. It accounts for 40 percent of total US emissions of two major air pollutants, carbon monoxide and nitrogen oxides, creates $70 billion in annual pollution-related health and other damages, and causes 37,000 annual premature deaths (National Research Council 2010; Fann et al. 2013; U.S. EPA 2014b). Globally, air pollution from transportation causes a quarter million deaths each year (World Bank 2014; Chambliss et al. 2014).

Textbooks describe optimal policy to address pollution—a corrective or Pigouvian tax equal to the marginal external cost of emissions, or a comparable quantity mechanism (e.g., cap and trade). But taxing vehicle air pollution emissions is infeasible because direct measurement of pollution from individual vehicles is imperfect and prohibitively expensive (Venigalla 2013). We believe no government has ever directly taxed air pollution from vehicles.[1]

Instead, the US, EU, Japan, China, Russia, India, Brazil, and most other countries rely heavily on new vehicle exhaust standards. Exhaust standards set a maximum emission rate per mile for every vehicle. Some standards also impose fleet-wide average requirements.

Exhaust standards have been controversial for decades due to their large costs and ambiguous effectiveness. In the 1970s, Ford executive Lee Iacoca claimed these standards could stop US vehicle production (Kaiser 2003). Congress has issued three requests to the National Academies of Science to provide advice involving exhaust standards (National Research Council 2001, 2004, 2006). Manufacturers have cheated on these standards, including the Volkswagen scandal that involved $22 billion in payments – the largest auto settlement in US history – leading to questions about standards' effectiveness (Yacobucci 2015).

Little economic research, however, scrutinizes exhaust standards. They are separate from fuel economy standards, which target gasoline consumption and have been the focus of much prior literature, reviewed below. We highlight the different patterns and challenges of air pollution and fuel economy. Thus, existing insights and methods from the fuel economy literature do not answer the questions we pose here for vehicle air pollution.

This paper helps to fill this literature gap by investigating several questions. How have vehicle air pollution emission rates changed over time? To what extent have exhaust stan-

[1] Roadside pollution sensing via infrared beams has substantial measurement error for individual vehicles. Scheduled emissions tests ("smog check") when paired with high-stakes incentives can lead to avoidance behaviors, making taxes based on such tests inaccurate (Stedman et al. 1998; Merel et al. 2014; Oliva 2015). Gasoline taxes target greenhouse gas emissions but weakly proxy air pollution (Knittel and Sandler 2018).

dards caused these declines? Are these standards cost-effective? Finally, how might reforms improve policy, either via targeting the stringency of exhaust standards or introducing complementary policies that accelerate vehicle scrap?

We find striking answers to each question. First, the air pollution emissions per mile of the US new vehicle fleet has fallen by more than 99 percent since regulation began in the 1960s. This spectacular decrease may exceed that of any other major sector. Used vehicles follow similar patterns. We conclude that these trends represent genuine, long-term, large declines in exhaust emission rates of US vehicles. We find much smaller declines for carbon dioxide (CO_2) emissions that fuel economy regulations target.

Second, to assess the impact of exhaust standards on emission rates, we exploit variation in exhaust standards between California and federal standards and across classes of vehicles, model years, and pollutants. We find that exhaust standards have caused 50 to 100 percent of the time-series declines in air pollution emission rates. Equivalently, we find an elasticity of vehicle emission rates with respect to exhaust standards of 0.5 to 1.0. Several pieces of evidence support these estimates' internal validity. Event study graphs show that changes in emissions align in time with changes in exhaust standards. We obtain qualitatively similar results when controlling for potential confounding policies—gasoline prices including taxes; and standards for smog check ("inspection and maintenance"), fuel economy, gasoline hydrocarbons, gasoline sulfur content, and ethanol blending. We obtain similar results when separately analyzing each set of standards, generally called Tier 0 (model years 1968-1993), Tier 1 (1994-2003), and Tier 2 (2004-2016). While we find that exhaust standards do not change basic vehicle attributes (horsepower, fuel economy, etc.), they do lead manufacturers to install cleaner engines. This statistical evidence echoes informal assertions by engineers and policymakers that exhaust standards, not secular technological innovation or other forces, account for most decreases in air pollution emission rates from US vehicles.

Third, while the aforementioned regressions suggest exhaust standards are effective, stylized facts suggest that exhaust standards are not cost-effective.[2] They do not equate the marginal cost of abating pollution across vehicles, a necessary condition for cost-effectiveness, because they only weakly regulate pollution from older vehicles. Emission rates of air pollutants (but not CO_2) increase rapidly with age. A majority of air pollution emissions in a calendar year come from vehicles more than 10-15 years old, which are largely exempt from exhaust standards.[3] Registration fees on the oldest and dirtiest used vehicles could in

[2] A cost effective pollution policy minimizes the cost of achieving a given pollution reduction, or maximizes the pollution reduction for a given cost. A pollution policy may increase social welfare yet not be cost effective—the social willingness to pay for its pollution reduction may exceed its costs, even though other policies could have achieved that pollution reduction at even lower cost.

[3] Smog check programs regulate emissions of old dirty vehicles. Most of our data are from areas with

principle discourage ownership of these vehicles. We build a database containing tax rates we collected from US state and local governments describing their vehicle registration fees, motor vehicle taxes, and vehicle property taxes (which we collectively call, "registration fees"). We find that registration fees are higher for newer, cleaner vehicles, and thus encourage ownership of older, dirtier vehicles, thereby exacerbating inefficiencies in fleet turnover. This echoes the broader idea that a commodity tax system which imposes higher tax rates on cleaner goods can cause important environmental damages (Shapiro 2021).

Fourth, we develop an analytical and a quantitative model to evaluate counterfactual policies. The early parts of the paper show regressions analyzing differences in emission rates; the latter parts of the paper combine those data with formal theoretical models to clarify remedies for and implications of the patterns in emission rates. An analytical model with few functional form assumptions provides comparative statics on how counterfactual policies affect social welfare. We analyze the steady state of a continuum of agents who can buy new vehicles from competitive manufacturers or repair new vehicles to drive them as used. Equilibrium used vehicle prices depend on exhaust standards and registration fees, and also determine scrap rates. Our first result shows that tightening new vehicle exhaust standards extends the lifetime of used vehicles, which exacerbates inefficiency from consumers scrapping used vehicles later than is socially optimal. This formalizes the "Gruenspecht Effect," which has been informally noted for many environmental policies. Our second analytical result shows that increasing registration fees on used vehicles can improve social welfare and complement exhaust standards by correcting the low scrap rate for used vehicles.

The quantitative model estimates gains from counterfactual policies. The quantification has a similar basic structure as the analytical model but allows for substitution across over 500 vehicle types differentiated by manufacturer, age, class, and size. The quantitative model also accounts for the engineering cost of meeting exhaust standards and fuel economy standards, Bertrand competition among new vehicle manufacturers, firm expectations, supply-chain (life cycle) emissions from manufacturing vehicles, and transitional dynamics. We study counterfactual changes to exhaust standards or registration fees. For each, we determine the equilibrium that results, then calculate the change in pollution emissions, producer and consumer surplus, environmental damages, and social welfare. The quantification uses data and estimates from earlier parts of the paper.

The quantitative model provides several results. Accelerating the roll-out of tighter (Tier 2) exhaust standards by one year increases social welfare by \$20 to \$30 billion. Policymakers are debating the importance of delays in stringent global climate policy; while we study

smog check programs, suggesting that older vehicles could account for an even larger share of pollution in the absence of smog check programs.

air pollution rather than climate change, we find large consequences of the timing of an environmental policy. Additionally, we find that the benefits of Tier 2 exhaust standards (which operated in the 2000s and 2010s) are 10 to 15 times its costs, and that Tier 2's measured benefits due to avoided premature mortality are 35 percent larger than those of a prominent cap-and-trade market for industrial plants from the same period, the NO_x Budget Program (Deschenes et al. 2018). We find larger gains, around $300 billion in present value, from reforming annual registration fees to reflect the environmental damage of a vehicle's age×type. Changing registration fees creates these benefits primarily by encouraging scrap of old and dirty vehicles. This counterfactual causes scrap of nearly all vehicles aged 25 years old or more. Echoing the Gruenspecht Effect analytical result, levying such environmental registration fees only on new vehicles actually creates welfare losses because new vehicle fees discourage scrap of old vehicles, extending their lifetimes and emissions.

These counterfactuals have complex effects on inequality. Because households in low-income communities drive older and dirtier vehicles, increasing registration fees for dirtier vehicles may trade off equity and efficiency. Exposure to aggregate vehicle emissions, for example because of proximity to highways, is also greater for low-income communities, leading to a potentially progressive environmental incidence. Transportation is a large source of pollution in vulnerable communities (Carlson 2018; Apte et al. 2019). Additionally, recycling revenues from automobile policy substantially influences its regressivity (Bento et al. 2009). We carefully discuss these channels and their political economy implications.

Three ties connect the paper's empirical and theoretical sections. First, they answer complementary parts of the paper's research questions. The empirical analysis studies effectiveness, while the models analyze cost-effectiveness and efficiency. Second, the regressions guide model assumptions. For example, the empirical finding that exhaust standards are effective and (subject to overcompliance) binding motivates corresponding assumptions in the models, and also motivates the counterfactual analysis of tightening standards. Similarly, the empirical finding that age plays a central role in explaining emissions and that existing registration fees exacerbate these patterns motivates both the models' analyses of age-based registration fees and the models' overall focus on fleet composition and scrap. Third, the empirical parts of the paper help assess the properties of the emissions inspections data that the quantitative model uses extensively and that the analytical model uses in a back-of-the-envelope quantification.

This paper utilizes the most comprehensive data on vehicle pollution emission rates ever constructed. It includes a half century of comparable pollution data using the same high-quality measurement method. These data cover nearly every new US light-duty vehicle and light-duty truck sold between 1972 and 2020 and many over the period 1957-1971. We believe this is the longest-lasting comparable microdata on pollution emission rates from any country or sector.[4] We supplement these new vehicle records with 65 million used vehicle test records from three types of tests—used vehicle inspections, official regulatory "in-use" tests, and roadside remote sensing. Our new vehicle data are national. Our main used vehicle data are from the state with the most high quality and extensive used vehicle tests in the US, Colorado, though we corroborate some patterns with additional data from eleven other states and six other countries. Finally, we use the Leontief Inverse of the US input-output table combined with plant-level industrial emissions data to account for the emissions embodied in the manufacturing of new vehicles and the associated supply chain.

This paper builds on several literatures. We provide the first comprehensive analysis of exhaust standards, which are the centerpiece of US Clean Air Act regulation of transportation. Landmark papers study Clean Air Act regulation of industry (e.g., Henderson 1996; Carlson et al. 2000; Greenstone 2002; Walker 2013). Another important literature studies fuel economy standards, which are separate from exhaust standards (Goldberg 1998; West and Williams 2005; Goulder et al. 2012; Jacobsen 2013; Anderson and Sallee 2016; Langer et al. 2017). Analysis of fuel economy standards has developed methods to use the R-squared from a regression to study imperfect targeting of environmental policy (Jacobsen et al. 2020), but the primary challenge we highlight for exhaust standards involves fleet composition and scrap. Existing work largely does not directly analyze exhaust standards' effects.[5]

Additionally, this paper provides the first simple sufficient conditions for stricter environmental policy on new capital to create inefficiency by decreasing scrap. Known as the Gruenspecht Effect (Gruenspecht 1982), this pattern has been informally lamented for decades. Many prominent environmental regulations differ by capital vintage, such as the US Clean Air Act's New Source Review or energy efficiency construction codes (Gruenspecht and Stavins 2002; Stavins 2006). Existing work uses regressions to analyze effects of vintage-differentiated regulations (Bushnell and Wolfram 2012; Bai et al. 2021), or analyzes new-

[4]For example, emissions data from US manufacturing only have firm-level records generally available back to 1990, in many cases come from engineering predictions rather than direct measurement, and can fail data quality tests (Currie et al. 2015). Similarly, regular emissions monitoring from US power plants began in 1980, is quinquennial through 1995, and in many years covers only the largest electricity generating units.

[5]Prior papers describe standards (Bishop and Stedman 2008) or abatement technologies (Bresnahan and Yao 1985); summarize engineering estimates of abatement costs (Fowlie et al. 2012; Cropper et al. 2014); describe model year trends from before versus after standards change using one cross-section of vehicle tests (Kahn 1996a,b), which does not separate effects of age, model year, and standards; undertake simulations of vehicle emissions with a few types of vehicles (Mills and White 1978; Innes 1996; Kohn 1996; Harrington 1997; Walls and Hanson 1999; Fullerton and West 2010; Feng et al. 2013); or compare emissions from electric and gasoline vehicles (Holland et al. 2016). Several papers analyze used vehicle emissions from smog check tests, primarily from California, which measure pollution emission rates from used vehicles and require repairs of the dirtiest vehicles, but those papers do not evaluate exhaust standards (Merel et al. 2014; Knittel and Sandler 2018; Sanders and Sandler 2020).

vehicle purchase fees proportional to CO_2 emissions (Adamou et al. 2013; D'Haultfoeuille et al. 2013). Some papers evaluate programs that encourage retirement of polluting vehicles, including "Cash for Clunkers" (Busse et al. 2012; Sandler 2012; Li et al. 2013; Hoekstra et al. 2017). More broadly, Barahona et al. (2019) and Gillingham et al. (2022) find that policies spurring scrap of old vehicles substantially increase social welfare.

We also create the first national data on, and economic analysis of, vehicle property taxes. Research analyzes property taxes for real estate (e.g., Poterba and Sinai 2008; Cabral and Hoxby 2015) but many property taxes also apply to vehicles. We create a dataset of vehicle property taxes and registration fees from US states, cities, counties, and special districts.

In addition, this research provides the first equilibrium model of vehicle markets and scrap that accounts for air pollution abatement and emissions. Existing frameworks to analyze fuel economy, economy-wide greenhouse gas emissions, or polluting industrial activity do not apply directly to air pollution from vehicles (Goldberg 1998; Goulder et al. 2012; Busse et al. 2013; Jacobsen and van Benthem 2015). The model relates to recent industrial organization papers studying equilibrium trade in used car markets in settings with more general forms of market power and frictions (Biglaiser et al. 2020; Gillingham et al. 2022).

Finally, this research helps answer the question of why pollution in industrialized countries is declining. We describe a setting where a specific regulation accounts for most of a long-term national decrease in pollution emission rates.[6] While many countries and sectors have had large decreases in pollution over time, and most of this decrease reflects cleaner production within an industry rather than reallocation across industries, studies have struggled to assess which economic forces or policies have caused that decline.

The paper proceeds as follows. Section 2 describes policy and technology. Section 3 discusses the data. Section 4 describes emissions trends. Section 5 estimates effects of exhaust standards. Section 6 establishes stylized facts on cost-effectiveness. Section 7 describes the analytical model, Section 8 describes the quantitative model, and Section 9 concludes.

[6]Following Copeland and Taylor (1994) and Grossman and Krueger (1995), researchers have allocated economy-wide changes in pollution into changes in total output ("scale"); changes in the share of output from different industries ("composition"); and changes in pollution emitted per unit of output within a given industry ("technique"). In many regions, technique accounts for most decreases in pollution from manufacturing (Levinson 2009; Cherniwchan et al. 2017; Shapiro and Walker 2018; Copeland et al. 2022).

2 Background on Exhaust Standards

2.1 History of Exhaust Standards

In 1952, chemist A. J. Haagen-Smit discovered that hydrocarbons (HC) and nitrogen oxides (NO_x) emissions from vehicles contribute to smog. By 1959, engineers had developed technology to abate emissions by running exhaust fumes over a catalyst.

Federal regulators have since imposed standards regulating these pollutants and carbon monoxide (CO). We call these regulations, "exhaust standards." Others sometimes call them tailpipe or emission standards. These standards limit the emissions per mile of these pollutants. We refer to the grams of pollution emitted per mile driven as a vehicle's emission rate and the total grams of pollution emitted as emissions. We refer to CO, HC, and NO_x as air pollution, though they are sometimes also called local or criteria pollution, to distinguish them from global pollutants like CO_2. Table 1 summarizes the standards. Appendix A.1 discusses details of standards less directly relevant to our paper.

The 1965 Motor Vehicle Air Pollution Control Act created national standards, called "Tier 0." The 1970 Clean Air Act Amendments substantially expanded them.[7] Standards began for CO and HC in 1968 and for NO_x in 1972.[8] Tier 0 standards periodically tightened through 1993. These standards essentially required every vehicle to have a catalytic converter by the mid-1970s, though catalytic converters were not broadly viable in the 1960s. Automakers developed and deployed catalytic converters to comply with exhaust standards. We focus on federal exhaust standards but the Clean Air Act lets California set its own, tighter exhaust standards. Other countries and US standards have similar structure.

The 1990 Clean Air Act Amendments required Tier 1 standards, which phased in beginning in 1994 and became binding in 1996.[9] A few light-duty trucks could wait until 1997 to comply. Exhaust standards regulate "light-duty vehicles" and "light-duty trucks"; we refer to these as cars and trucks. Tier 1 decreased CO and HC standards more for categories of trucks than for cars, though required similar NO_x decreases in emission rates for cars and trucks. Thus, our analysis of Tier 1 does not focus on NO_x since we exploit differences in stringency between vehicle classes. Tier 2 standards phased in over the years 2004-2009 and continued through 2016. Tier 3 is being phased in from 2017 through 2025.

These standards have the same general approach but different details. Tier 0 and Tier

[7]Corporate Average Fuel Economy Standards are enabled by the Energy Policy and Conservation Act of 1975, a separate law from the Clean Air Act.

[8]All years in this section refer to vehicle model years.

[9]Only 40 percent of vehicles had to comply with Tier 1 in the 1994 model year and 80 percent in 1995. Because many vehicles already met Tier 1 standards in 1993, Tier 1 was most binding for the dirtiest vehicles, which could remain at existing emission levels until model year 1996.

1 define maximum standards. Each standard requires every vehicle in a class (e.g., trucks in a certain weight range) to emit less than the standard. Tier 2 and Tier 3 impose fleet-wide mean standards and tightened the maximum standards. The pollutant used for the fleet-wide average standard differs across regulations.

These standards use the same test to measure a vehicle's emission rate, the Federal Test Procedure. This test specifies the chemical composition of the fuel used in the test, the speed at every second of a 30 minute test, and is run on a dynamometer, a large treadmill-like device; Appendix A.2 discusses details.

Before a vehicle may legally be sold, the EPA must certify that the vehicle meets exhaust standards. In addition to conducting a test, the EPA or manufacturer estimates a "deterioration factor" predicting how emission rates will change during the vehicle's "useful life," which ranges from 50,000 miles and 5 years (whichever comes first) to 150,000 miles or 15 years, depending on the standard. The EPA regulates how manufacturers may determine deterioration factors. Exhaust standards apply to a new vehicle's "certification level," which equals the test result scaled up by the deterioration factor.

Several years after a vehicle is manufactured, the EPA assesses "in-use" compliance. Manufacturers conduct emissions tests on samples of vehicles at up to 150,000 miles and the EPA audits some. If these tests find emission rates above the standard, the vehicle is recalled and the emissions control system repaired or replaced. Between 1975 and 2008, 80 million vehicles, or about 16 percent of all vehicles sold, had recalls, though some of these involved minor reclassifications (U.S. EPA 2008; Department of Energy 2016). Accurately predicting a new vehicle's emission rate at 50,000 or 150,000 miles is challenging. In-use tests and the costs of recalls give manufacturers an incentive to over-comply with exhaust standards. Industry engineers and regulators we interviewed describe over-compliance, sometimes called headroom or a safety margin, as typical for this reason.

2.2 Pollution Abatement Technologies

Explaining technologies used to meet these standards helps interpret results; Appendix A.3 provides details. The approach has changed little since the 1970s: expose exhaust to precious metals inside a catalytic converter, which converts pollution into harmless gases. Because these metals are catalysts, pollution can react with them without consuming or changing them. The precious metal palladium primarily abates CO and HC, which have complementary abatement technologies; rhodium primarily abates NO_x; and platinum abates all three. Under ideal conditions, these reactions eliminate 100 percent of CO, HC, and NO_x.

Lead and sulfur render catalytic converters ineffective by coating the catalyst. Our used vehicle data begin after model year 1975, when vehicles required unleaded gasoline (Mondt 2000). Nonetheless, catalytic converters decrease in effectiveness over time due to remaining low levels of sulfur in gasoline, wear of precious metals, or breakdown of complementary technologies like oxygen sensors.

Would emission rates decline without regulation, due to secular innovation? Engineers and regulators we interviewed argued that technologies that improve vehicle drivability do not affect pollution, so automakers would only decrease emission rates due to regulation. Crandall et al. (1986, pp. 92-93) summarize this view: "There is little evidence to support the view that emission rates would have fallen significantly without the emissions standards program." Innovation may still decrease the marginal cost of controlling vehicle emission rates over time. Because emissions-related recalls are common and costly, even when policy is constant, decreasing marginal abatement costs over time give auto manufacturers an incentive to decrease emission rates even further (additional "overcontrol"), even without tightening standards, to decrease the rate of unexpected recalls.

One may also wonder whether trends in "green" or "warm glow" preferences for environmentally-friendly goods could explain changing vehicle emission rates. We believe this is not a major contributor, in part due to limited consumer information. We have not found anecdotal or statistical evidence that consumers value or even know their vehicle's air pollution emissions, though consumers may have information on fuel economy. Unlike fuel economy, information on a vehicle's air pollution is not easy to find and interpret.[10]

Many environmental policies, including exhaust standards, encourage innovation in abatement technology (Vollebergh 2010; Rozendaal and Vollebergh 2021). The EPA calls exhaust standards "technology forcing" because they can require technologies which have been proven in focused settings but may not have had mass development or adoption. Innovation research finds that the announcement of Tier 0 and Tier 1 standards increased patenting and publishing of technical papers on relevant abatement technologies (Lee et al. 2010, 2011).

2.3 Other Policies Relevant to Emission Rates

Other environmental policies are relevant to our analysis. Our regressions and quantitative model account for them. Corporate Average Fuel Economy standards regulate the mean fuel economy of new vehicles. Fuel economy standards did not change in the periods we study most closely (Department of Transportation 2014). Federal gasoline excise taxes, state retail gasoline taxes, and gasoline prices could affect miles traveled or driving behavior.

[10] Air pollution emission rates are not shown on most leading consumer automotive websites. The EPA calculates a 1 to 10 "smog rating" for vehicles, which now appears in small font on a vehicle's fuel economy sticker. But this rating is not thoroughly explained and was absent for most of our sample period.

Around ten percent of US counties operate smog check programs, where registration requires used vehicles to pass emissions inspections. Our data mostly come from areas with smog check, so our findings that vehicle emission rates rise sharply with age, and our estimates of the benefits of scrapping old dirty vehicles, might be even larger without smog check. Some states and cities regulate the chemical content of gasoline to decrease HC, though not other pollutants (Auffhammer and Kellogg 2011). Colorado, the source of our main data, has not used gasoline with regulated chemical content (U.S. EPA 2019). Ethanol accounts for an increasing share of fuel, in part due to policy. Evidence on how ethanol affects exhaust emission rates is mixed (Hubbard et al. 2014). Governments in 28 states, listed in Appendix F.1, have registration fees that vary with vehicle characteristics, especially value.

3 Data

3.1 New Vehicle Pollution Data

We obtain test results for each new vehicle type from the Annual Certification Test Results Report, also called the Federal Register Test Results Report.[11] We obtain electronic records for model years 1979 to 2019 from the EPA and keyed in records for years 1972-1978 from the Federal Register (1978); see Appendix B.2 for details. Although these data determine compliance with the Clean Air Act, we are not aware of any economics research using them.

For model years 1957-1971, we obtain data on used vehicles tested in AES (1973), which applied the Federal Test Procedure to about 1,000 vehicles aged 1 to 14 years old from five cities. The sample statistically represented the national distribution of vehicle characteristics. In model years before exhaust standards, emission rates of these vehicles do not appear to increase with age and are similar to estimates of uncontrolled emission rates. This is sensible because before exhaust standards, vehicles did not have emissions control systems that could break down. Hence, for these pre-regulation years, new and used vehicles likely had similar emission rates. We identify vehicles meeting California standards in AES (1973) as those in California and vehicles meeting federal standards as those in other states.

3.2 Used Vehicle Pollution Data

Our main used vehicle emission data come from smog check tests in Colorado, which we use for several reasons. While many states test vehicle emissions, recently only Colorado has used the highest-quality test, called IM240 (the inspection and maintenance test that lasts 240

[11] We use "class" to denote cars versus trucks, or weight categories of trucks, and "type" to denote more detailed classification of vehicles such as manufacturer, size, trim, or engine specifications.

seconds). This test provides a short version of the Federal Test Procedure and is considered the "gold standard" of smog check tests for its quality and comparability to the Federal Test Procedure (Sierra Research 1997; Joy et al. 2004; U.S. EPA 2006); Appendix B.1 discusses this comparability.[12] Most other states only obtain a computer description of the status of a vehicle's emissions control system (an "on-board diagnostic test") and do not measure exhaust emission rates for most vehicles. Colorado includes about 12basic difference-in million tests and extensive remote sensing and registration data.[13] Appendix B.3 shows that the Colorado counties have similar driving and emissions patterns to other polluted urban US counties.

The Colorado data cover calendar years 1997 through 2014. In these years, all Colorado gasoline vehicles model year 1982 or later are tested biennially, beginning at age four, so the data cover model year 1982 through 2010. Appendix B.3 describes additional sample restrictions, such as excluding observations missing key variables.

We take a few steps to limit concerns about avoidance and short-term evasion behavior. We restrict the Colorado sample to the first test in a sequence, which is less subject to short-term manipulation concerns. A sequence is a test series for a specific registration, ending in a vehicle passing (and then able to register) or being sold, traded, or driven unregistered. Manipulation is arguably more likely after a vehicle fails the first test. We also include estimates that control for the stringency of the relevant smog check standard. Additionally, we report sensitivity analyses using remote sensing estimates from a Colorado database with over 50 million remote sensing readings; from smaller samples taken in 11 states; from 4 other countries; and from heavy duty trucks (e.g., 18 wheelers).

We show sensitivity analyses from remote sensing, which uses roadside infrared or ultraviolet beams connected to devices that measure pollution concentrations in an exhaust plume. Remote sensing provides data that is believed to be impervious to manufacturer "defeat devices" and that is not generally used in economics papers.[14] Remote sensing, however, has substantial measurement error and imperfectly comparable units versus new or used vehicle tests (Borken-Kleefeld 2013). Appendix Table A1 compares remote sensing and smog check readings from the same vehicle in essentially the same week. If remote sensing and smog check data were perfectly comparable, Appendix Table A1 would obtain regression coefficients and elasticities of one. While matched remote sensing and smog check readings are

[12] The EPA describes the IM240 test as "the most accurate short test available for use in I/M programs" (U.S. EPA 1995). Colorado describes it as "arguably the most accurate emissions test currently in use for replicating the Federal Test Procedure (FTP) that is used to certify new model year vehicles" (AIR 2015).

[13] Most economic research using data on US used vehicle emission rates uses data from California, but its data have lower quality; Appendix A.2 provides details.

[14] Defeat devices typically turn on parts of an emissions control system only when they detect that a vehicle is undergoing a laboratory driving test. Remote sensing observes vehicles during typical on-road driving.

strongly correlated, the magnitude of that regression coefficient ranges from 0.000015 to 435, and the magnitude of the elasticity ranges from 0.01 to 2.98, depending on the pollutant and specification. None of the 95% confidence regions includes zero or one. We interpret remote sensing as an important check on the sign and precision of changes in emission rates, but interpret magnitudes from remote sensing cautiously due to its differences in measurement.

Finally, we report sensitivity analyses from "in-use" tests in California (see Appendix B.5), which have no direct incentives for vehicle owners so are unlikely to suffer from owner manipulation. In-use tests apply the Federal Test Procedure to a sample of vehicles several years old to assess compliance with exhaust standards.

3.3 Other Data Notes

Appendix Table A2 summarizes the samples and coverage of the paper's datasets. We use all years to describe emission rate trends and subsets of years to analyze Tiers 0, 1, and 2. In addition, we use vehicles from model year 1993 and calendar year 2000 to describe fleet-wide emissions, and test year 2000-2014 data to calibrate the quantitative model. Appendices B.6 and B.7 discuss details including concordances, use of the US input-output table to measure the emissions from manufacturing vehicles, and the marginal damages of pollution.

Here we summarize emissions from manufacturing vehicles. We use the Leontief Inverse of the US input-output table, which helps measure the entire supply chain of all goods used to produce a vehicle. We measure emissions from each industry in the vehicle supply chain by using plant-level air pollution emissions data from the National Emissions Inventory. Aggregated, this calculation suggests that manufacturing a new car or truck creates about $600 in environmental damages due to air pollution in the year 2000, including emissions from the entire supply chain, which is in the ballpark of numbers that engineers have estimated from life cycle analyses. These damages fall over time as manufacturing becomes cleaner.

4 Trends in Emission Rates

We first quantify trends in new and used vehicle emission rates. Figure 1 plots mean emission rates in grams per mile from new US vehicles over model years 1957-2020. The figure shows the three air pollutants exhaust standards target—CO, HC, and NO_x. It also shows CO_2, which fuel economy standards target. The graphs show the mean certification level for 50,000 miles, i.e., the emission rate of a new vehicle scaled up by an engineering calculation reflecting 50,000 miles. Each y-axis has log scale. Vertical lines show the year before exhaust standards. The lines with blue squares show the unweighted mean across vehicle types. For model years 2000-2015, the lines with hollow red circles show means weighted by fleet size.

Figure 1 shows that the emissions per mile for each air pollutant have fallen by more than 99 percent since regulation began. CO has fallen by 99.4 percent, HC by 99.7 percent, and NO_x by 99.5 percent. For example, the mean CO emission rate of new US vehicles fell from 83 grams per mile in the 1960s to 0.5 grams per mile in 2020. Even between 1990 and 2018, these emission rates fell by 75 to 95 percent. Unweighted trends and trends weighted by fleet size are similar.[15] We do not believe previous research has directly used these new vehicle test results to measure long-term pollution trends.[16] The long lifetime of vehicles in a setting where emissions are rapidly declining implies that at any given moment, older vehicles are operated alongside newer, cleaner vehicles. This motivates our consideration of policies targeted to accelerate scrap in Section 7. The changes in emission rates between model years we document here underpin the quantitative model of Section 8.

For context, between 1990 and 2018, ambient pollution levels (which depend on emissions from all sources) of CO, NO_2, and ozone fell by 20 to 75 percent (U.S. EPA 2018), suggesting that new vehicles cleaned up faster than other pollution sources. The decrease in emission rates from new vehicles is more rapid than declines in manufacturing emissions or ambient water pollution over this period (Shapiro and Walker 2018; Keiser and Shapiro 2019).

Comparing emission rates in Figure 1 and standards in Table 1 shows that emission rates fall particularly in years when policy tightens. Emission rates are flat before standards begin. Rates then decline rapidly. Figure 1 reflects the large decreases that standards required in 1975. The CO and HC graphs show flatter lines between 1984 and 1993, when standards were flat. Emission rates and standards were also flatter between 2007 and 2017.

Figure 1 also shows that CO_2 fell less than air pollution. CO_2 only fell by 55 percent between 1957 and 2017 and by 25 percent between 1990 and 2017. The changes in CO_2 rates largely occurred in the late 1970s and 2010s, when fuel economy standards tightened. Between 1982 and 2007, both the CO_2 line and fuel economy standards were flat.

Used vehicle emission rates have similar patterns, though they are available for fewer years and are subject to the challenge of disentangling model year, test year, and age effects. Appendix C.2 explains how we analyze Colorado smog check data. Appendix Figure A2

[15] Appendix C.1 discusses data limitations but shows qualitatively similar results for weighted trends before 2000.

[16] Existing evidence does not definitively show these trends. The EPA uses a simulation model, the MOtor Vehicle Emission Simulator (MOVES), to calculate annual vehicle emissions. Dividing predicted national emissions from MOVES by national vehicle miles travelled shows that air pollution emissions per mile have fallen by 98-99 percent since the 1970s (U.S. EPA 2022). MOVES, however, relies on numerous parameters, data sources, and calculation modules; its design changes frequently; and its internal processing can be somewhat non-transparent. Kahn (1996a,b) uses a cross-section of smog check data to calculate decreases in emission rates of 50-90 percent between the early 1970s and late 1980s, though it is difficult to separate age and model year effects in the cross section.

shows that mean used vehicle emission rates for each air pollutant fell by roughly 90 percent between 1982 and 2010; new vehicle emission rates from Figure 1 fell by similar amounts. Mean CO_2 emission rates of the used vehicle fleet actually increased between model years 1990 and 2005, partly due to the increasing market share of light-duty trucks.

5 Effects of Exhaust Standards on Emission Rates

This section describes effects of Tier 0, 1, and 2 exhaust standards on emission rates. We use different approaches for each Tier, reflecting relevant regulations and data. One goal is to understand to what extent exhaust standards caused the trends documented in Section 4. We focus on estimates in logs, though also report estimates in levels, to facilitate comparisons across pollutants and datasets, address outliers, and help interpretation even when manufacturers over-comply with standards. Appendix D discusses sensitivity analyses.

5.1 Econometrics: Effects of Exhaust Standards on Emission Rates

Tier 0. The following equation analyzes how Tier 0 affected emission rates:

$$lnE_{pry} = \beta_1 lnS_{pry} + \eta_{pr} + \lambda_y + \epsilon_{pry} \tag{1}$$

We analyze model years 1957-1971. Each observation represents the mean emission rate of vehicles for pollutant p (CO, HC, NO_x, or CO_2), in region r (California or federal), from model year y. CO and HC faced regulation in the 1960s; NO_x and CO_2 did not. The variables E and S represent emission rates and standards. The term β_1 represents the elasticity of emission rates with respect to exhaust standards. The pollutant×region fixed effects, η_{pr}, address potential confounding from time-invariant differences between vehicles facing California's standards versus those facing federal standards, separately by pollutant. Model year fixed effects, λ_y, address time-varying emission rates common to vehicles nationally.

Tier 1. For Tier 1, we estimate the following equation:

$$lnE_{picy} = \beta_2 lnS_{picy} + X'_{picy}\pi + \mu_{pc} + \nu_{py} + \xi_{pa} + \epsilon_{picy} \tag{2}$$

We analyze model years 1982-2000. We report separate estimates where E represents new or used vehicle emission rates. An observation represents a reading of pollutant p for vehicle i in model year y. The main estimates distinguish vehicle class $c \in car, truck$, which are the most comparable measures of standards. Sensitivity analyses explore more detailed subclasses. For estimates of used vehicle emission rates, we include controls X for age fixed effects, odometer, and other environmental policies that could affect emission rates—fuel economy, fuel content, or smog check standards. The regression includes fixed effects for pollutant×vehicle class, pollutant×model year, and pollutant×age (μ_{pc}, ν_{py}, and ξ_{pa}). The coefficient β_2 represents the elasticity of emission rates with respect to exhaust standards. We cluster standard errors by model year×truck type.

Tier 2. After model year 2000, regulations imposed fleet-wide average standards. Hence, instead of using difference-in-differences across vehicle classes, we analyze the extent to which new vehicle emission rates predict used vehicle emission rates of the same vehicle:

$$lnE^u_{picy} = \beta_3 lnE^n_{picy} + X^{'}_{picy}\zeta + \nu_{py} + \xi_{pa} + \epsilon_{picy} \tag{3}$$

We analyze model years 2000-2010 because the concordance file linking new vehicle engine families and used vehicle Vehicle Identification Number prefixes begins in model year 2000 and our Colorado smog check data conclude in model year 2010. Here E^u is the used vehicle test result of vehicle i, E^n is the new vehicle emissions test result corresponding to used vehicle i, and c, y, and X are defined above. The coefficient β_3 represents the elasticity of used vehicle emission rates with respect to new vehicle emission rates. The regression includes age and model year fixed effects (μ_{pa}, ν_{py}), which vary by pollutant.

5.2 Results: Effects of Exhaust Standards on Emission Rates

We start by graphing raw trend data by class. Figure 2 shows the national time series of exhaust standards (Panels A, C, and E) and new vehicle emissions (Panels B, D, and F). They cover model years 1982-2010. In each graph, the blue solid line describes cars and the dashed red line describes trucks. The vertical dashed lines show when car standards changed; the vertical solid lines show when car and truck standards changed. Each panel shows a different pollutant. Values are measured in grams of pollution emitted per mile.

Figure 2 reveals close correspondence between standards and emissions, which shows that exhaust standards cause large decreases in emission rates. For example, in 1984, truck standards for CO and HC fall abruptly and emission rates do also. In 1996, when Tier 1 rolled out, standards and emissions again move in tandem. A similar pattern occurs for Tier 2 in the mid-2000s.

The main exception here is the decline in NO_x truck emissions in model years 1982-1987 that Panel F shows. California gradually tightened truck standards in these years, while the EPA tightened standards only in 1987. The 1980s new vehicle data do not distinguish California from federal vehicles, so the 1980 trend in NO_x emission rates for trucks may reflect compliance with California's standards.

These graphs also show over-compliance. New vehicle emissions are about half of exhaust standards. The y-axis scale in Panels D-F is nearly half the scale in Panels A-C. For example, in 1990, cars and trucks faced CO standards of 10 and 4, but emission rates for these groups were around 4 and 2. As discussed in Section 2.1, manufacturers over-comply because compliance is ultimately assessed against used vehicles 5 to 10 years later.

We now turn to regressions focused on each Tier of exhaust standards separately.

Effects of Tier 0 Exhaust Standards (Model Years 1957-1971)

Figure 3 shows annual emission rates over model years 1957-1971 1. Panel A shows vehicles facing federal standards and Panel B shows vehicles facing California standards. Each line shows a different pollutant. Federal standards regulated CO and HC in 1968. California standards regulated CO and HC in 1966. Standards only regulated NO_x or CO_2 in 1972 and 1978, respectively. The vertical line in each graph shows the year before regulation began.

Figure 3 suggests that exhaust standards decreased emission rates of regulated pollutants. Before regulation, emission rates of all pollutants were fairly flat. This is consistent with a limited effect of productivity growth on emission rates. When California's exhaust standards began in 1966, CO and HC from California vehicles fell. CO and HC emission rates from federal vehicles only decreased in 1968, when federal regulation began. The other pollutants, CO_2 and NO_x, did not fall when CO and HC standards began, and slightly increased. These other pollutants may have increased because catalytic converters were not viable in the 1960s, so manufacturers then responded to exhaust standards with technologies like combustion modification that can increase NO_x and CO_2 (National Research Council 1988, 2006).

Table 2 shows regressions corresponding to equation (1). Panel A pools pollutants. Panels B and C show one pollutant each. Column (1) is a time series estimate comparing across model years and within each pollutant and region. Columns (2) through (7) provide difference-in-differences estimates comparing across regions and model years.

Table 2 shows that Tier 0 exhaust standards decreased emission rates. The time series estimate in column (1) obtains an elasticity of emission rates with respect to exhaust standards of 0.61 (0.07). Our preferred elasticity estimate is 0.80 (0.08), from the difference-in-differences estimate of column (2). Other estimates in levels or restricted to California or federal vehicles are qualitatively similar (columns (3) through (6)).

Effects of Tier 1 Exhaust Standards (Model Years 1982-2000)

Figure 4 shows event study graphs analyzing the roll out of Tier 1 standards between model years 1990 and 2000. Panels A and B show the change in exhaust standards, Panels C and D show the change in new vehicle emission rates, and Panels E and F show the change in used vehicle emission rates. All these graphs plot differences between trucks and cars by model year, with values for 1993 normalized to zero.

Figure 4 shows that Tier 1 exhaust standards decreased new and used vehicle emission rates. Panels A, B, E and F show that used vehicle emission rates and standards change by similar amounts. Panels C and D show that new vehicle emission rates change less, consistent with initial firm over-compliance. The new vehicle graphs show some differences between cars and trucks in model years 1990-1992. This pattern does not appear for used vehicle emission rates, which matters because used vehicle rates are likely closer to actual on-road emissions.

Table 3 reports regressions corresponding to equation (2). The pooled time-series estimate in column (1) compares across model years and within categories of cars and trucks. The difference-in-differences estimate in column (2) adds model year controls, so exploits changes within class and across model years. Column (3) controls for other policies—fuel economy standards, smog check standards, each vehicle's gasoline cost per mile (equal to the relevant tax-inclusive retail gasoline price divided by the vehicle's fuel economy), the ethanol fuel share, and the fuel sulfur content. Column (4) adds model year×truck linear trends. Column (5) limits the sample to vehicles aged 4 to 6 years. Column (6) restricts the sample to begin in model year 1990. Column (7) estimates the regression in levels rather than logs. Panels A through C analyze used vehicles; Panels D through F analyze new vehicles.

Table 3 shows that Tier 1 exhaust standards decrease used and new vehicle emission rates. The basic difference-in-differences estimate in column (2) is 0.86 (0.08) for used vehicles and 0.54 (0.05) for new vehicles. Controlling for other environmental policies in column (3) does not change the estimate.[17] The other specifications in columns (4) through (7) obtain broadly comparable results, though some point estimates are moderately larger or smaller. Most estimates are precise. Appendix D discusses sensitivity analyses, which obtain qualitatively similar results.

Effects of Tier 2 Exhaust Standards (2000-2010)

Table 4 evaluates the effects of Tier 2 standards on emission rates, using regressions corresponding to equation (3). Columns (1) through (6) repeat the specifications of Table 3.

[17] One interpretation of these estimates is that even if CAFE standards had not been implemented, tightening exhaust standards would have decreased emission rates per mile substantially. But because a vehicle's air pollution emission rates change almost one-for-one with its gasoline consumption, if exhaust standards had not been implemented, tightening CAFE standards would have decreased emissions per mile to some extent. In this sense, each policy alone would have been sufficient to decrease emission rates, though the decrease due to exhaust standards is larger and would have occurred even without CAFE standards.

Columns (7)-(8) add back the abbreviated tests.

Table 4 shows that new vehicle emission rates strongly predict used vehicle emission rates. The pooled elasticities in Panel A are generally around 0.5. Most estimates reject elasticities of both zero and one with 99 percent confidence. Rejecting the null hypothesis of zero implies that new vehicle emissions tests predict a vehicle's actual emission rate. This suggests that even if defeat devices or short-term manipulation occur, enforcement is imperfect, or abatement technologies deteriorate unexpectedly, new vehicle emissions tests strongly predict used vehicle emission rates.

Why are many elasticity estimates below one? Panel E of Table 4 for CO_2 suggests that measurement provides an important answer. A vehicle's fuel economy and associated CO_2 emission rate, unlike its air pollution emission rate, does not typically depreciate with age. Hence, the primary reason why the elasticities in Panel E are below one is measurement error both within and between new and used vehicle tests. The CO_2 elasticities in Panel E range 0.72 to 0.95; all these estimates are significantly less than one, though most are larger than the estimates for air pollution in Panels A through D. Because air pollution emission rates depend on fuel economy and emissions control systems, measurement error may be more important for air pollution than for CO_2.

Binned scatterplot comparisons of new and used vehicle emission rates in Figure 5 show the tight relationship between new and used emission rates of a vehicle type. Each graph groups all new vehicles into twenty equal-sized bins, then plots the mean used vehicle emission rate for each bin plus the linear trend. For all three air pollutants and for CO_2, the points have linear slope, suggesting a constant elasticity of used to new vehicle emissions.

5.3 Discussion: Effects of Exhaust Standards on Emission Rates

This section has described different approaches which find elasticities of emission rates with respect to standards generally between 0.5 and 1.0, suggesting that exhaust standards have caused between half and all of the time series decline measured in Section 4. In this sense, exhaust standards are effective.

How would regulation-induced innovation in abatement technology, discussed in Section 2.2, affect interpretation of our estimates? We interpret our regressions as externally valid to exhaust standards that are not too far beyond the technology frontier. We study cases where technology developed or proved to be sufficient for compliance. Our estimates have some external validity to counterfactual delays in the standards that were implemented, because technology developed to meet these standards. Our estimates may be less externally valid for substantially more rapid tightening of standards. If standards had tightened by 99.5 percent in 1970, for example, the elasticity of emissions with respect to standards would likely have been lower than we estimated.

Regulation-induced innovation, however, does not change the causal interpretation of our estimates. One reason is the evidence from Section 2.1 that regulation is the main incentive to clean up air pollution. Another reason comes from Tier 0. Figure 3 shows that when Tier 0 begins in the 1960s, California regulated CO and HC in 1966, two years before the federal government did. California vehicles decreased emission rates in 1966, but vehicles outside California only decreased emission rates in 1968. California's regulation shows that technology was available for vehicles outside California in 1966, but auto manufacturers waited to install this technology for vehicles outside California until standards required it.

A similar point across vehicle classes applies to subsequent years. Figure 2, for example, shows that in 1984, CO standards tightened sharply for trucks but not cars, and emission rates fell sharply for trucks but not cars. If technology alone drove the 1984 improvements in emission rates, both cars and trucks would have installed it. The 1984 CO truck standards may have led to innovation in pollution control technology, but auto manufacturers installed it because standards required it. Finally, California had more stringent standards during the Tier 2 era that we study using our quantitative model. For example, the California HC standard for light-duty vehicles was 25-40% below the federal standard, and the NO_x standard was 64-88% lower, suggesting that tighter federal standards would have been technologically feasible.

The rest of the paper builds on these results. Section 6 uses the data to describe stylized facts. The analytical and quantitative models of Sections 7 and 8 take from this section that exhaust standards are effective, assess their efficiency, and analyze counterfactuals.

6 Stylized Facts on Cost-Effectiveness and Age

6.1 Emission Rates Increase with Age

Figure 6 plots mean emission rates and annual driving by model year and age. Panels A through C show air pollution, Panel D shows CO_2, and Panel E shows annual miles traveled. The y-axes have logarithmic scale. These visually show the extent to which deterioration of emissions control systems has changed across model years.

The upward-sloping lines in Panels A through C of Figure 6 demonstrate that emission rates for vehicles form a given model year increase with age. This is unsurprising because emissions control systems deteriorate with age. The upward shift of the lines for earlier model years in Panels A through C implies that earlier model years have higher emission rates. The age-emissions profile is similar for most groups of model years, though NO_x controls may be deteriorating more gradually. The y-axis scale implies that these effects are proportional to age. Panel D shows that none of these patterns occur for CO_2. The downward slopes in Panel E imply that older vehicles drive fewer annual miles. This may occur because most households prefer to drive the newer of two vehicles (Archsmith et al. 2020) or because the households that own older vehicles have lower driving demand.

Several additional analyses in the Appendix show similar conclusions but with different contexts or methods. Appendix Figure A4 shows similar patterns in other states and countries and for heavy duty trucks. Appendix Figure A5 shows similar patterns but from regressions including age fixed effects, odometer readings, and vehicle identification number fixed effects. It shows that a vehicle's CO_2 rates and associated fuel economy do not change with age, but a vehicle's air pollution exhaust emission rate increases rapidly with vehicle age. This difference makes sense—as vehicles age, catalytic converters and other pollution abatement technologies break down, increasing emissions. But because end-of-pipe pollution control technologies are not commercially viable for CO_2, vehicles have no CO_2 control systems that would break down with age, so a vehicle's CO_2 emission rate does not change with age.

Does age or odometer account for these patterns? Appendix Figure A5 does control for odometer, and finds that age independently increases deterioration. Appendix Table A6 shows regression analogues to these graphs, suggesting that both age and odometer readings independently increase emissions. Deterioration due to mileage occurs in part because even the low sulfur content in fuel decreases catalytic converter effectiveness. Age may independently cause deterioration because variable weather, aging seals and electronics, and failure of complementary technologies like oxygen sensors and direct injection can decrease catalyst efficiency. We focus on age since existing registration fees already depend on it. We are not aware of US fees that directly depend on odometer readings; because age, unlike odometer, is not susceptible to manipulation; and because taxes that vary only with vehicle age and type simplify modeling the intensive margin of driving choice.

6.2 Older Vehicles Account for a Large Share of Emissions

Exhaust standards limit used vehicle emission rates through in-use testing, but in-use tests only apply to vehicles up to 10-15 years old. Exhaust standards are therefore unlikely to equalize abatement costs across vehicles of different ages, which is a necessary condition for cost-effectiveness (the equimarginal principle). Intuitively if older vehicles cause a large share of emissions, exhaust standards will be less cost-effective.

Figure 7 plots the cumulative distribution of emissions versus vehicle age. The graph shows a cross-section of vehicles in calendar year 2014 from Colorado smog check data. Appendix Figure A6 shows similar patterns from a cohort of model year 1993 vehicles and from Colorado and multi-state remote sensing data.[18] This graph shows a smaller emissions share for the old vehicles in the 1993 cohort, consistent with the idea that model year rather than aging accounts for the majority of this pattern. The vertical red lines show ages 10 and 15. Each graph shows separate curves for each pollutant.

Figure 7 shows that a large share of air pollution emissions come from vehicles older than 10 to 15 years. In these data, 70 to 80 percent of air pollution emissions come from vehicles older than 10 years. Vehicles older than 15 years account for 30 to 50 percent of air pollution emission but only 10 percent of CO_2 emissions. Less CO_2 comes from older vehicles because fuel economy, unlike air pollution, does not change with vehicle age and because fuel economy standards have changed less than exhaust standards across model years. Although older vehicles are driven fewer miles per year and are more likely to be scrapped, their air pollution emission rates are high enough to offset the lower mileage.

Secular trends in vehicle longevity in the US fleet amplify these pollution differences. Appendix Figure A7 shows large linear trends in the mean age of US vehicles over the last half century. In 1970, the mean US vehicle was 6 years old; in 2018, mean vehicle age had doubled to 12 years. This aging likely reflects both improved durability technology for automakers and increasing new vehicle prices via the Gruenspecht Effect.

6.3 Annual Registration Fees are Higher on Cleaner Vehicles

Exhaust standards mandate clean new vehicles. They do not give consumers an incentive to scrap dirty old vehicles and do not give manufacturers an incentive to decrease pollution from aging vehicles. Annual ownership fees that increase with the pollution from a vehicle would give drivers and auto manufacturers incentives to decrease pollution.

Many states and local governments already impose annual registration fees for vehicles that vary with a vehicle's attributes. How do these existing fees vary with emissions?

Figure 8 plots the national mean annual registration fee in dollars for vehicles aged 4 to 18 years. The solid blue line shows the mean annual registration fee; the dashed red line shows the annual air pollution externality from vehicles on the road in calendar year 2000, all in 2019 dollars.

Figure 8 shows that dirtier vehicles face lower registration fees. In other words, these

[18] We show cross-sectional data for 2014 since it is the most recent year when Colorado required smog check test of vehicles aged 4 and older. The Appendix shows cohort data from 1993 since this is the earliest model year where we observe tests of four-year old vehicles.

registration fees implicitly subsidize rather than tax emissions. Owners of 18-year old vehicles pay \$40 less in annual registration fees than the owners of 4-year old vehicles do. But 18-year old vehicles create about \$700 more in air pollution damages than 4-year old vehicles do. Registration fees decrease in age, while annual externalities increase in age. Modifying this incentive is a key consideration of the next two sections.

7 Analytical Model

The previous sections show that exhaust standards decrease emission rates and that registration fees are higher on cleaner vehicles. We now develop a model with few functional form assumptions to clarify how these standards and fees affect scrap and welfare.

Motivated by the trends, regressions, and stylized facts of Sections 4 through 6, we focus on differences in policy and emissions between vehicles of different ages and model years. The quantitative model in Section 8 has heterogeneity within vehicle ages and transition dynamics; here we focus on the steady state. These models seek to clarify mechanisms by which exhaust standards affect emissions and to address questions that the previous sections cannot, such as how different types of exhaust standards and registration fees affect social welfare.

7.1 Analytical Model Setup

We consider a single vehicle type that can last up to two time periods t. A vehicle is initially new (n) and becomes used (u) in the next period. Driving new and used vehicles emits pollution. Manufacturing new vehicles also emits pollution. A measure one continuum of risk-neutral consumers demands vehicles. Pollution is a pure externality, so consumers ignore it in making expenditure decisions. Denote the size of the new and used vehicle market as N and U, respectively, where $N + U = 1$ in a period, so that there is no outside good.[19]

Demand reflects consumers' different taste for new versus used vehicles. We normalize the value of a used vehicle to 0 and let w denote willingness to pay for a new vehicle, distributed $G(w)$, which we assume is non-degenerate and continuous with no mass points. All w are weakly positive, i.e., no consumer prefers a used over a new vehicle at the same price. We assume the distribution $G(\cdot)$ is the same for all consumers and time periods and thus abstract from income effects.

[19]Appendix E.2 derives results allowing for an outside good. The key insights of the model derived here carry over to that model, with the exception of one comparative static related to the size of the used vehicle market, which is ambiguous in the case with an outside good.

New and used vehicle supply have different properties. New vehicle supply comes from competitive, constant returns manufacturing with marginal cost and thus producer price ψ^s. We write the final price to consumers of a new vehicle as $\psi = \psi^s + \tau$, where τ is any tax on new vehicles, explained below. The supply of used vehicles reflects consumer scrap, as follows. A consumer who buys a new vehicle receives a repair cost draw k from the distribution $H(k)$, which we assume is non-degenerate and continuous with no mass points. We assume this distribution is the same for all consumers and time periods. In the next period, this consumer either scraps the vehicle or resells it as used in a competitive, frictionless resale market at price p. We assume the value of scrap is zero.[20]

7.2 Analytical Model Equilibrium

A steady-state equilibrium is a used vehicle price p in all time periods such that consumers choose new versus used vehicle purchases and scrap versus repair to maximize utility, and supply equals demand for both new and used vehicles.

Utility maximization lets us describe used vehicle supply in more detail. A consumer who purchases a new vehicle in one period will repair it in the next period if the used vehicle price exceeds the owner's repair cost draw (i.e., if $p > k$) and will scrap it otherwise. Hence, the share of new vehicles that are repaired and survive as used vehicles equals the cumulative distribution of repair costs, evaluated at the used vehicle price: $H(p)$. Correspondingly, the number of used vehicles supplied equals $U^s = H(p)N$. In equilibrium, $N = 1 - U$, so we can write used vehicle supply as $U^s = H(p)/(1 + H(p))$.

We can also describe used vehicle demand in more detail. The value of a new vehicle to a consumer is its benefit minus its price, $w - \psi$ plus its expected resale value net of repair costs. When deciding whether to scrap or repair a vehicle, the owner receives a repair cost draw. They will repair the vehicle so long as the repair cost k is less than the used vehicle price; otherwise the vehicle is scrapped. Anticipating this, the ex ante expected resale value net of costs is $H(p)$ (the probability that a vehicle will be repaired) times $(p - \bar{k})$, where $\bar{k}$ is the expected cost of repair, conditional on repair being optimal.[21] Thus, a consumer will buy a new vehicle at the start of the period if and only if the surplus from a new vehicle exceeds that of a used, i.e., $w - \psi + H(p)(p - \bar{k}) > -p$. Equivalently, the demand for used vehicles is the probability a consumer does not buy a new vehicle, which is $U^d = G(\psi - p - H(p)(p - \bar{k}))$.

[20] A uniform scrap value would be capitalized into used vehicle prices, which would shift up the price of all used vehicles in equilibrium, but this would not impact the sign of our comparative statics. Adding a scrap value would be equivalent to shifting the distribution of w by a constant, as the scrap value is folded into the normalized value of a used vehicle.

[21] The truncated mean $\bar{k}$ of the repair cost distribution is a function of p: $\bar{k} = 1/H(p) \times \int_{-\infty}^{p} k dH(k)$.

Equating supply and demand for used vehicles provides the key equilibrium condition, where p denotes the equilibrium price:

$$\frac{H(p^*)}{1+H(p^*)} = G(\psi - p^* - H(p^*)(p^* - \bar{k})). \tag{4}$$

The left-hand side of equation (4) describes used vehicles supplied as a function of used vehicle prices p; the right-hand side describes used vehicles demanded as a function of p. Our main results are comparative statics that describe changes in this equilibrium that result from changing primitives. Because supply $(H(p)/(1 + H(p))$ is increasing in p and demand $(G(\psi - p - H(p)(p - \bar{k})))$ is decreasing in p, there will be a unique steady state p^*.[22]

7.3 Analytical Model: Pollution and Policy

We assume the following about pollution, echoing empirical findings from Sections 5 and 6. A new vehicle creates pollution Φ from production and ϕ^n from exhaust. A used vehicle creates exhaust emissions ϕ^u. The difference in externalities between a new and a used vehicle is $\Delta \equiv \Phi + \phi^n - \phi^u$. Exhaust emissions for a used vehicle exceed exhaust emissions for a new vehicle at a given time ($\phi^u > \phi^n$), because tightening exhaust standards cleaned up new vehicles over time or because emissions control systems deteriorate. If $\Delta > 0$, a new vehicle emits more than a used vehicle, after accounting for production and retirement emissions.

We consider two policies. Exhaust standards ω constrain new vehicle exhaust emissions: $\phi^n \leq \omega$. Tighter exhaust standards increase manufacturing costs, so $\psi^{s'}(\omega) \leq 0$.[23] Registration fees for new or used vehicles are τ_n and τ_u. Revenues are recycled lump-sum to consumers. With no outside good, only the new-used difference in tax rates $\tau \equiv \tau_n - \tau_u$ is needed for our analysis. We can write the consumer's price of a new vehicle as $\psi = \psi^s(\omega) + \tau$.

Welfare in the model is private consumer welfare minus costs minus the externality. Costs include used vehicle repair and new vehicle production. As is standard, the potential for welfare improvement from policy comes from correcting the market choice (in this case the share of new vehicles) that prevails when agents ignore externalities. Appendix E.2 shows the model with an outside good, where the outside good share also influences welfare. Our

[22]Uniqueness follows from our assumption that the H and G distributions have no mass points, so there are no flat portions of the supply or demand curve. One exception is if the primitives imply a corner solution where all vehicles are new. This would occur, for example, if the minimum repair costs are sufficiently high. These extremes seem to have little practical interest, so we focus on interior solutions.

[23]Because we describe a steady-state equilibrium, we focus on exhaust standards that cause a constant shift in vehicle manufacturing costs. If the industry learns over time how to reduce emissions at lower cost, then a steady-state standard is tightening over time such that the marginal cost remains constant.

baseline empirical model assumes perfect competition; Appendix F.9 shows results under imperfect (Bertrand) competition, where welfare also reflects profits.

7.4 Analytical Model Results

Proposition 1. *A policy that increases ψ will decrease the scrap rate and increase the market share of used vehicles. The derivative of scrap with respect to new vehicle prices is*

$$\frac{d(1 - H(p^*))}{d\psi} = -h(p^*) \left(\frac{1 + H(p^*)}{\frac{h(p^*)}{g(w^*)(1+H(p^*))} + (1 + H(p^*))^2} \right) < 0 \tag{5}$$

where $w^* = \psi - p^* - H(p^*)(p^* - \bar{k})$ is the marginal type indifferent between used and new vehicles in equilibrium.

Appendix E.1 shows proofs. On the left-hand side of equation (5), the numerator of the derivative is the scrap rate and the denominator is the new vehicle price. The right-hand side of equation (5) evaluates this derivative. Proposition 1 shows that tighter exhaust standards extend vehicle lifetimes by decreasing scrap. Tighter exhaust standards – a lower ω – increase production costs ψ. The negative sign of equation (5) shows that higher production costs decrease equilibrium scrap and thus extend vehicle lifetimes. The mechanism is intuitive. Increasing new vehicle prices causes higher demand for and thus price p^* of used vehicles. For any repair cost draw k, higher used vehicle prices make a consumer less likely to scrap vehicles.

A simple example may clarify. Imagine a driver who crashes an old car, has it towed to a repair shop, and must decide whether to repair or scrap it. If exhaust standards are weak, vehicle production costs and used vehicle values will be relatively low. The cost of repairing the crashed vehicle is more likely to exceed the vehicle's value, so the driver is more likely to scrap the vehicle. If exhaust standards are stringent so that production costs and used vehicle prices are high, the driver is more likely to find that the vehicle's value exceeds the repair cost, and so more likely to repair the vehicle, extending its lifetime.

Proposition 1 also shows that making registration fees higher for new than used vehicles, as Figure 8 shows happens on average in the US, extends vehicle lifetimes. The same holds for any new-vehicle tax—higher relative registration fees on new vehicles are equivalent to a higher τ. The negative sign on the right-hand side of equation (5) shows that this increase in new vehicle purchase prices decreases scrap and extends vehicle lifetimes.

The Gruenspecht Effect posits that policies increasing the prices of new durable goods will extend the life of used durables, which often pollute more. We believe Proposition 1 provides the first formal derivation of it. Gruenspecht (1982) originally considered policy exempting old power plants from pollution standards imposed on new plants, but the Gruenspecht Effect is cited more broadly in discussions of policies affecting power plants, vehicles, home and building construction, and other durables. (Keohane et al. 1998; Stavins 2006; Bushnell and Wolfram 2012; Jacobsen and van Benthem 2015; Anderson and Sallee 2016).

Proposition 1 also implies that vehicles survive longer than is socially optimal if and only if $\tau > \Delta$. In other words, the market share of used vehicles is larger than is optimal if new vehicles are taxed more than their relative pollution damages. The reason is that if consumers internalized pollution externalities, they would perceive a price difference between new and used vehicles equal to $(\psi+\Delta)-(p-H(p)(p-\bar{k}))$. Because we abstract from outside goods here, this is equivalent to treating the new vehicle price as $\psi+\Delta$.[24] This leads to the second result.

Proposition 2. *Welfare in a time period is maximized when $\tau = \Delta$. If $\tau > \Delta$, then moving to τ' where $\tau > \tau' \geq \Delta$ will increase welfare; if $\tau < \Delta$, then moving to τ' where $\tau < \tau' \leq \Delta$ will increase welfare.*

This result is intuitive. In this model, registration fees that differ between new and used vehicles by $\tau = (\Phi + \phi^n) - \phi^u$ can fully correct the pollution externality.[25] Welfare in a time period is improved if we move the tax rate closer to the fully-corrected benchmark.

Figure 8 shows that existing registration fees are higher for newer and cleaner vehicles. Section 6 shows that used vehicles have higher emission rates than new vehicles. If emissions from manufacturing new vehicles are not too large, Proposition 2 implies that flattening registration fees or changing the sign of the correlation between registration fees and age would increase welfare.

Intuitively, exhaust standards and registration fees are complementary. If a counterfactual policy makes exhaust standards tighten more rapidly across model years, the gap Δ between emissions of used and new vehicles grows, and the scrap rate deviates further from the optimum. Registration fees correcting the scrap rate then remedy a larger distortion, implying a greater return to taxing the emissions of used versus new vehicles.

To roughly quantify how pollution rates differ by age, we divide the year 2000 fleet into two categories, new through 9 years old ("new"), and 10 years or older ("used"). Vehicles 9 years and younger accounted for 57% of the fleet and 65% of miles driven. Including estimated production emissions, the typical "new" vehicle causes \$486 of damages per year,

[24] With an outside good, the same results carry over with one exception. Raising the relative price of new vehicles induces a Gruenspecht effect in the same way. The only difference is that, while used vehicles represent a larger share of total vehicle market (i.e., the fleet is older), the total number of used vehicles may rise or fall because the total vehicle market contracts.

[25] In this model, this is the optimal fee policy for a given exhaust standard. In a more detailed setting, miles driven and maintenance could respond to policy, so registration fees would not restore the first-best.

while "used" vehicles cause $1,364 of damages per year. The difference in damages between used and new vehicles, δ, is then $878. These calculations are affected by age and model year because they come from a 2000 cross-section of the fleet, and they take as given the mileage by model year and empirical scrap rates. Thus, despite being driven less, the typical used vehicle produces 2.8 times as much pollution as new vehicles. An efficient relative tax rate would tax used vehicles, whereas existing policy puts a relative tax on new vehicles. This binary division of the fleet hides variation in damages and taxes through a vehicle's life, which the next section explores in detail.

8 Quantitative Model

This quantitative model connects to the paper's other sections in several ways—its analysis of cost-effectiveness and efficiency complements the regressions' analysis of effectiveness; its assumptions and choice of counterfactuals reflect empirical findings that exhaust standards are effective and that emissions rates increase with age; the Colorado smog check pollution data described in Section 3 provide key model inputs; and Propositions 1 and 2 in the analytical model help guide the discussion of counterfactuals. Some key elasticities here come from existing evidence—for example, the scrap elasticity comes from our prior work (Jacobsen and van Benthem, 2015) and the pollution control cost function comes from engineering estimates (U.S. EPA, 1999, 2014a).

8.1 Quantitative Model Details

The model setup is as follows. A representative agent serves several roles. She demands purchase of new vehicles and rental of used vehicles. She also chooses whether to scrap or repair used vehicles available from the previous time period, and therefore she also serves as a competitive "supplier" of used vehicles.[26] Firms produce new vehicles and engage in Bertrand or perfect competition. Motivated by the differences in exhaust standards and emission rates between vehicle classes and ages found in Sections 5 and 6, we allow vehicles to be differentiated by over 500 combinations of class, size, age, and manufacturer. The model accounts for evolution of the vehicle fleet over time.[27]

[26] We would obtain analytically equivalent results, at the cost of additional notation, from modeling a representative consumer and used vehicle supplier as separate agents.

[27] For tractability and data availability, we leave spatial modeling across US counties for future research.

Agent Utility and Demand

Demand for vehicles (new and used) is derived by assuming that the representative agent maximizes a constant elasticity of substitution (CES) utility function $U(v, x)$ in period t (t subscript suppressed) over a composite vehicle v and other goods x, given income M:

$$\max_{v,x} U(v, x) = (\alpha_v v^{\rho_u} + \alpha_x x^{\rho_u})^{\frac{1}{\rho_u}} - \Omega \tag{6}$$

$$s.t. \quad e_v v + e_x x \leq M. \tag{7}$$

Here α_v and α_x are scale parameters that determine demand at baseline prices, and ρ_u represents the elasticity of substitution between vehicles and other goods. Pollution damages Ω are a pure externality, which the agent takes as given. The agent does not have "green preferences" leading her to buy cleaner vehicles out of environmental concern. The per-period prices of the composite vehicle and the composite good are e_v and e_x.

Demand for the composite vehicle v comes from five sequential CES utility nests: vehicles versus other goods, class c, size s, age a, and manufacturer m. Within a nest, demand depends on the per-period cost $e_{c,s,a,m}$ of a differentiated vehicle:

$$e_{c,s,a,m} = r_{c,s,a,m} + \tau_{c,s,a,m} + \sigma_{c,s,a,m}. \tag{8}$$

This cost includes a vehicle rental rate r, which reflects depreciation and repair; vehicle registration fees τ, with revenues rebated lump-sum; and fuel, insurance, and other operating costs σ. In equilibrium, rental rates, taxes and other ownership costs are capitalized in vehicle values. This is a "rental" model of vehicles, so the consumer problem can be solved in isolation each period. Beliefs about next period vehicle prices influence rental costs r, which we discuss further below when describing scrap decisions.

Optimizing this problem implies a standard CES demand system where $q^d_{c,s,a,m}$ denotes demand for each vehicle type conditional on prices. When policy changes per-period costs, the agent reoptimizes vehicle quantities within each nest. Appendix F.3 details this derivation.

We allow miles driven to vary by vehicle class and age based on data but treat mileage within vehicle type×age as exogenous. Our counterfactual policies change the cost of owning a vehicle but not the per-mile operating cost, so we expect their main impact to be on changes in fleet composition rather than miles traveled.

New Vehicle Manufacturers

We present results for both where new vehicle manufacturers engage in either Bertrand or perfect competition. For each class×size, manufacturer m chooses prices p, emissions ϕ, and fuel economy f to maximize profits in time period t, subject to exhaust and fuel economy standards (subscripts m, $a = 0$ suppressed):

$$\max_{p_{c,s,t},\phi_{c,s,t},f_{c,s,t}} \quad \sum_{c,s=1,2}\left[\left(p_{c,s,t} - C^b_{c,s} - C^{\phi}_{c,s,t}(\phi_{c,s,t}) - C^f_{c,s,t}(f_{c,s,t})\right) * q^d_{c,s,t}(\mathbf{p},\mathbf{f})\right] \tag{9}$$

$$C^{\phi}_{c,s,t}(\phi_{c,s,t}) = \chi^t \zeta_{c,s}\left(\frac{\phi_{c,s,0}}{\phi_{c,s,t}} - 1\right) + \xi_{c,s,t} \tag{10}$$

$$s.t. \quad \phi_{c,s,t} \leq \bar{\phi}_{c,s,t} \tag{11}$$

$$s.t. \quad \frac{\sum_s q^d_{c,s,t}(\mathbf{p},\mathbf{f})}{\sum_s (q^d_{c,s,t}(\mathbf{p},\mathbf{f})/f_{c,s,t})} \geq \overline{f}_{c,t}, c \in 1,2. \tag{12}$$

In the profit equation (9), $C^b_{c,s}$ represents per-vehicle production cost at time period $t = 0$ with emissions and fuel economy levels as observed in the baseline, $C^{\phi}_{c,s,t}$ is the per-vehicle cost of controlling exhaust emissions away from the baseline, and $C^f_{c,s,t}$ is the per-vehicle cost of improving fuel economy relative to the baseline.

Demand $q^d_{c,s,t}$ depends on the vector of prices and fuel economies for all vehicles $(\mathbf{p},\mathbf{f})$. Any profits are rebated lump-sum to consumers. We model perfect competition using the limit as $\frac{\partial q^d_{c,s,t}(\mathbf{p},\mathbf{f})}{\partial p_{c,s,t}}$ and $\frac{\partial q^d_{c,s,t}(\mathbf{p},\mathbf{f})}{\partial f_{c,s,t}}$ go to infinity. In this case the first-order conditions in (9) reduce to zero profit conditions that also satisfy the exhaust emissions and fuel economy constraints in (11) and (12).[28] In equilibrium, competitive new vehicle prices translate into per-period costs r and fuel economy translates into per-period operating costs σ.

Equation (10) describes the cost function for controlling exhaust emissions, in years 2002 and beyond, above a baseline level of control applied to vehicles in model year 2000. It builds on the general convex form in Bovenberg et al. (2008). The term $\chi < 1$ describes the rate of innovation in pollution control technology. The term $\zeta_{c,s}$ varies the relative control cost by vehicle class and size. The residual $\xi_{c,s,t}$ comes from the least squares calibration of χ and $\zeta_{c,s}$ to match the EPA's engineering cost estimates for Tier 2 and Tier 3 exhaust standards (Appendix F.6 provides details). This form and calibration has useful properties—adding no control above that in the 2000 model year adds no cost beyond that in the 2000 model year; a given level of emissions control becomes cheaper over time; marginal pollution

[28] Under perfect competition, vehicles are priced so $p_{c,s,t} = C^b_{c,s} + C^{\phi}_{c,s,t}(\phi_{c,s,t}) + C^f_{c,s,t}(f_{c,s,t})$ plus the shadow cost of vehicle c, s with respect to the fuel economy constraint in time t, and so $\phi_{c,s,t} \leq \bar{\phi}_{c,s,t}$ for each vehicle.

control costs rise smoothly; it exactly matches EPA's projected costs in a world where emissions standards are introduced at the historical rate; and it adapts engineering data from the EPA's analyses when applying arbitrary counterfactual exhaust standards. Sensitivity analyses examine alternative control costs. Motivated by the regressions in Section 5 and the idea that manufacturers primarily or only change exhaust rates due to standards, we assume that exhaust standards bind for all manufacturers.

Exhaust standards $\overline{\phi}$ in equation (11) cap exhaust emissions per vehicle, separately by vehicle class. We calibrate $\overline{\phi}$ to historical data which already includes any over-compliance. We assume the same over-compliance persists in counterfactuals. Fuel economy (CAFE) standards require that the harmonic average of fuel economies within a class $c \in$ (car, truck) must exceed $\overline{f}_{c,t}$, which is the form of CAFE relevant over most years this model analyzes. Because fuel economy standards average within a manufacturer×class, firms equalize marginal compliance costs across vehicles in each class.

Vehicle Scrap Decisions

We refer to the representative agent's capacity as a competitive supplier of used vehicles as "vehicle rental suppliers." Vehicle rental suppliers begin each period with a stock of used vehicles from the previous period and take as given rental rates $r_{a,t}$ for used vehicles (subscripts c, s and m suppressed). At the period's start, each vehicle receives a repair cost draw $k_{a,t}$ that must be paid to survive, or the vehicle is scrapped. To generate a constant-elasticity scrap decision, we assume the cumulative distribution of repair cost shocks is $H(k_{a,t}) = 1 - b_a(k_{a,t})^{\gamma_a}$, where b_a is a scale parameter (we calibrate) and γ_a (we take from the literature) controls the elasticity of the scrap rate with respect to vehicle value. This cumulative density corresponds to a probability density $h(k_{a,t}) = -b_a\gamma_a(k_{a,t})^{\gamma_a - 1}$ defined over the support $k_{a,t} \geq (1/b_a)^{(1/\gamma_a)}$. Vehicle rental suppliers maximize current and expected rental receipts minus the cost of repairs and new vehicle purchases.

Vehicle rental suppliers expect that rental rates follow $\mathbb{E}[r_{c,s,a,m,t+1}] = r_{c,s,a,m,t}$.[29] With these expectations, the sequence of used vehicle resale values is (derived in Appendix F.4):

$$\begin{aligned} p_{a_{max},t} &= r_{a_{max},t} \\ p_{a,t} &= r_{a,t} + (1 - y_{a+1,t})\left(\frac{p_{a+1,t} - \bar{k}_{a+1,t}}{1+\delta}\right), \qquad a = 1, \ldots, a_{max} - 1. \end{aligned} \tag{13}$$

[29] We do not assume rational expectations about future vehicle rental rates but we do adjust expectations based on upcoming changes in fuel economy and registration fees. This adjustment happens at a slower rate than if suppliers had fully forward looking expectations; see Appendix F.7. "Surprises" are possible along transitions after a policy shock, but once the system reaches a new steady state, this form of naive expectations will, by definition of the steady state, match fully forward looking expectations.

Here δ is the per-period discount rate, $y_{a,t}$ is the scrap rate, and $\bar{k}_{a,t}$ is expected expenditure on repair per vehicle of a given age, which follows from the repair cost density $h(k_{a,t})$:

$$\begin{aligned}\bar{k}_{a,t} &\equiv \mathbb{E}(k_{a,t}|k_{a,t} < p_{a,t}) \\ &= \frac{b_a^{-1/\gamma_a}\gamma_a - b_a\gamma_a p_{a,t}^{1+\gamma_a}}{(1+\gamma_a)\left(1 - b_a p_{a,t}^{\gamma_a}\right)}.\end{aligned} \tag{14}$$

Applying the used vehicle values from (13), vehicle rental suppliers choose the following set of scrap rates and thus used vehicle supply:

$$\begin{aligned}y_{a,t} &= b_a(p_{a,t})^{\gamma_a} \\ q^s_{a,t} &= q_{a-1,t-1} * (1 - y_{a,t}).\end{aligned} \tag{15}$$

We let γ_a vary with class and size and choose b_a to match scrap rates in the baseline data.

Vehicle rental suppliers also choose how many new vehicles to purchase. Vehicle manufacturers sell new vehicles at price $p_{0,t}$ (0 refers to age; t to the time period). Because vehicle rental suppliers earn zero expected and realized profits in steady state, they purchase new vehicles until their profits are zero; $r_{0,t}$ equals depreciation between new and one-period old vehicles adjusted for repair and scrap:[30]

$$r_{0,t} = p_{0,t} - (1 - y_{1,t})\Big(\frac{p_{1,t} - \bar{k}_{1,t}}{1+\delta}\Big). \tag{16}$$

Because equation 13 shows that $p_{1,t}$ is a function of rental prices and the repair cost density, new vehicle rental price becomes a function of new vehicle purchase price, used vehicle rental prices, and the repair cost density.

Equilibrium and Welfare

A competitive equilibrium of this model is a series of vectors of new vehicle prices, used vehicle rental rates, new vehicle emission rates, and new vehicle fuel economy levels $(p_{c,s,0,m,t}, r_{c,s,a,m,t}, \phi_{c,s,0,m,t}, f_{c,s,0,m,t})$ such that the representative agent maximizes utility (6) subject to the budget constraint (7); scrap decisions follow equation (15); vehicle manufacturers maximize profits as in (9) subject to exhaust and fuel economy standards in (11) and

[30] Along transition paths additional accounting flows need to be tracked. In particular, the supplier can experience rental flows that are greater or less than the depreciation it assigns in any given year along a transition. The timing of changes in accounting profits depends on the depreciation method the supplier uses to value its capital. Appendix F.9 finds that, over the long run, welfare does not depend importantly on this choice; the depreciation method influences only the timing of perceived gains and losses.

(12); and supply of each vehicle equals demand ($q^s_{c,s,a,m,t} = q^d_{c,s,a,m,t}$). We solve for equilibrium in each time period in sequence, by iteratively applying the exhaust and fuel economy constraints, and using a globally convergent quasi-Newton algorithm (Broyden's method; Appendix F.5 provides details).

We measure the effect of counterfactual policy on social welfare from the equivalent variation of utility. Exhaust standards and registration fees affect social welfare by changing vehicle manufacturing, demand decisions, and environmental externalities.

8.2 Data and Parameters

This model analyzes two vehicle classes (car and truck), two sizes (small or large), nineteen age categories (ages 0 to 37, grouped in two-year bins to reduce the computation) and seven manufacturers (Ford, General Motors, Chrysler, Toyota, Honda, Other Asian, and European). There are thus 28 vehicle types per age and 532 (=28*19) vehicle types.

We summarize data and parameters for the quantitative model here; Appendix F.1 and Appendix Table A7 provide details. We calibrate the model to leading industry data on vehicle prices and composition for the 2000 U.S. vehicle fleet and follow vehicles through 2020;[31] Appendix F.2 discusses how baseline model outputs compare to the data. This period lets us observe the evolution of emission rates over the following 20 years. We use our life cycle measure of the emissions from the supply chain of manufacturing a new vehicle. The model also incorporates age, class, and size specific averages for vehicle miles traveled. We take the elasticity of the scrap rate with respect to vehicle value from Jacobsen and van Benthem (2015). We calculate the value of external damages Ω outside the equilibrium algorithm since it is additively separable.[32] We measure pollution damages from the AP3 model (Tschofen et al. 2019), which accounts for emissions from each US county, atmospheric transport (i.e., wind speed and direction), functions relating ambient pollution concentrations to outcomes like mortality, and the value of a statistical life. Our baseline quantification analyzes perfect competition among new vehicle manufacturers, though a sensitivity analysis accounts for market power.[33] We discuss sensitivity analyses varying many of these parameters.

[31] We begin in the year 2000 because it lets us follow vehicle types as they age. This primarily encompasses the roll out of Tier 2 exhaust standards.

[32] It is $\Omega_t = \sum_{c,s,a,m} \phi_{c,s,a,m,t} vmt_{c,s,a} \theta q_{c,s,a,m,t} + \sum_{c,s,m} \Phi_{c,s,m,t} q_{c,a,0,m,t}$, where $\phi_{c,s,a,m,t}$ indicates per-mile exhaust emissions, $vmt_{c,s,a}$ denotes vehicle miles traveled, θ are damages per ton of emissions, and $\Phi_{c,s,m,t}$ reflects damages from emissions associated with the manufacturing of a new vehicle.

[33] The baseline quantification assumes perfectly competitive manufacturers because then pollution externalities provide the only distortion, letting us focus on the welfare effects of alternative policies that are second-best along a single dimension.

8.3 Counterfactual Policies

We evaluate two classes of policy.[34] The first changes exhaust standards. Actual Tier 2 exhaust standards rolled out over the period 2004 through 2006 then applied through model year 2016. Data from Section 5 indicates that annual damages from new vehicles decreased by 77 percent during the roll out of Tier 2 standards. We consider counterfactual policies that delay or accelerate these improvements by four or eight years. We also consider a uniform tightening of exhaust standards by 10 percent. We implement these counterfactuals by changing exhaust standards $\overline{\phi}_{c,s,t}$.

We choose these exhaust standard counterfactuals for several reasons. Tier 2 is the main set of exhaust standards which changed over the years 2000-2020 where we have best data coverage. Studying acceleration or delay of these standards lets us measure the annual value of Tier 2. Policymakers also frequently debate the timing of important environmental policies. Studying a 10 percent change in exhaust standards helps think about broad general changes in exhaust standards.

One could think of accelerating Tier 2 as encouraging earlier adoption of abatement technologies in a scenario where they were available. Some evidence suggests this scenario is plausible. Increasing catalyst mass (precious metals—palladium, platinum, and rhodium) is available in any year at additional cost and represents a large component of abatement costs (U.S. EPA 2014a). Appendix Table A8 shows that 70 to 90 percent of new vehicles met Tier 2 standards four years early, and 50 percent met Tier 2 standards eight years early. The share emitting less than half of Tier 2 standards early (i.e., that overcomplied) was lower. Pinning down the precise technological feasibility of implementing standards four to eight years early is beyond the scope of this paper, but we believe these counterfactuals are realistic enough to be relevant.

The second class of counterfactuals covers four possible changes to annual registration fees. The first adds fees equal to the annual pollution damages of each age$\times$vehicle type. The second scales these fees to be revenue-neutral. The third imposes fees on new vehicles only, reflecting lifetime environmental damages. The fourth makes registration fees flat. We implement these counterfactuals by changing registration fees τ in equation (8). These counterfactuals hold the path of exhaust standards fixed at their actual, historical value,

[34] The quantitative model is flexible enough to analyze many other possible types of policies, such as a tax on vehicle miles travelled (VMT). We have chosen to save VMT taxes and other classes of counterfactuals for future work, however, for several reasons. Focusing on counterfactual registration fees and property taxes maximizes coherence and consistency with the rest of this paper. These counterfactuals change vehicle purchase prices but not per-mile driving costs, which lets us focus the model accordingly. Additionally, vehicle registration fees and exhaust standards are common and vary substantially across space, time, and vehicle type/attributes, which suggests that reforms of these policies in the direction of an externality-based fee system may be politically feasible.

and recycle registration fee revenue to the representative agent. Welfare gains mirror those in Proposition 2 in the analytical model.

We study these registration fee counterfactuals for several reasons. Our empirical results show strong age deterioration, so we focus on a policy targeting vehicle age. State and local governments charge registration fees that vary with vehicle attributes. The technical, and perhaps political, ability to consider such policies makes reforms in the direction of externality-based fees interesting and plausible. A full damage-based type×age fee is the natural baseline to evaluate even if states are more likely to implement partial versions. Adding revenue-neutrality to the fee system may further improve political feasibility. Finally, many existing policies target new vehicles, so restricting fees to those vehicles may be politically feasible.

8.4 Results

Table 5 shows how counterfactual policies affect several outcomes. Column (1) describes market surplus, equal to consumer surplus under perfect competition. Column (2) shows the change in pollution damages. Column (3) shows the change in social welfare, and column (4) shows the change in tax revenues, all in cumulative billions of 2019 dollars. Columns (5) through (7) show the percent change in cumulative pollution emissions over the same 20-year horizon, relative to baseline. Each row considers one counterfactual. Panel A examines changes in exhaust standards and Panel B examines changes in registration fees.

Counterfactual Exhaust Standards

Table 5, row 1, shows that delaying implementation of Tier 2 exhaust standards by four years decreases social welfare by \$107 billion, or \$27 billion per year. Delaying standards slightly increases market surplus and massively increases pollution damages. A four-year delay in Tier 2 increases total pollution emissions by five to ten percent. Exhaust standards generate no tax revenue. Row 2 shows slightly smaller per-year effects for an eight-year delay in Tier 2. Columns (5) through (7) show that an eight-year delay produces nearly double the total pollution increase as a four-year delay. Rows 3-4 show that accelerating Tier 2 by four or eight years increases social welfare by \$113 or \$175 billion in present value. While accelerating Tier 2 decreases surplus in the vehicle market somewhat, it decreases pollution damages by far more. Row 5 of Table 5 describes a more modest 10% improvement in standards relative to the baseline. This increases welfare by \$25 billion over 20 years.

Several benchmarks suggest these magnitudes are economically important. If the benefits of Tier 2 were measured against a value of a statistical life of \$10 million, they would represent around 2,700 fewer deaths per year. This is an appropriate benchmark because almost all the monetized benefits of decreasing NO_x and VOC emissions are due to avoided premature mortality (Tschofen et al. 2019). Another benchmark is other recent environmental policies. An important cap-and-trade market for industrial NO_x implemented over this period, the NO_x Budget Program (NBP), prevented an estimated 2,000 premature deaths per year (Deschenes et al. 2018). Thus, Tier 2 exhaust standards create about 35 percent larger annual health benefits due to avoided premature mortality than this prominent cap-and-trade market. Comparing columns (1) and (2) of Table 5 suggests Tier 2 has a benefit/cost ratio of ten to fifteen; this ratio is in line with those of other recent federal air quality regulations (Keiser et al. 2019). If one took the pollution changes documented for Tier 0 and Tier 1 in Section 5 and extrapolated the types of numbers estimated here for Tier 2, however, they would likely imply welfare gains from Tier 0 and Tier 1 exhaust standards in the trillions of dollars.

Counterfactual Registration Fees

We also consider counterfactuals that vary registration fees. Table 5, row 6, shows that making registration fees proportional to environmental damages produces a present-value social welfare gain of \$322 billion and produces \$1.2 trillion in additional revenue over 20 years, or \$60 billion annually. These counterfactual registration fees have double the welfare gains from accelerating counterfactual Tier 2 exhaust standards. The environmental registration fees decrease cumulative vehicle emissions by a third.

This reform heavily taxes the oldest vehicles. Figure 9, Panel A, shows the fee that this counterfactual imposes for vehicles of each age. These graphs average across vehicle types within an age. The fee for 0-year old vehicles reflects both exhaust emissions and air pollution damages from vehicle manufacturing. Vehicles more than 20 years old face an annual registration fee of over \$2,000, which exceeds the resale value of these vehicles.[35]

The solid line in Figure 9, Panel B, shows that this policy leads households to scrap a third of 15-year old vehicles, half of 20-year old vehicles, and 90 percent of 25 year old vehicles. This is an extraordinary change in the fleet of older vehicles. Put another way, most vehicles aged over 25 and older here have environmental damages exceeding their annual ownership

[35] Such reforms might affect unregistered driving, though we conjecture that such effects would be modest. The only estimate of unregistered driving rates we could find describes California (U.S Bureau of Transportation Statistics 2011). Only 1% of all vehicles were unregistered more than 3 months after the registration deadline, and practically none (0.03%) after more than two years. For comparison, estimates suggest that over 10 percent of U.S. drivers are uninsured; thus, most uninsured drivers' vehicles are registered. While unregistered driving may increase in response to higher registration fees, regulators can also increase enforcement of vehicle registration requirements, which only requires observing a vehicle's license plate.

cost. The dashed line in Panel B shows the environmental gains due to vehicles of each age, which has a hump shape that peaks at vehicles of age 24. Younger vehicles have lower emissions rates. Vehicles age 25 and older pollute more per mile, but there are few such vehicles in the baseline and they are driven few miles per year.

Row 7 of Table 5 shows a revenue-neutral version of the age×vehicle type registration fee, which taxes dirty vehicles and subsidizes clean vehicles (a "feebate"). It increases welfare somewhat less, about $230 billion, because it produces composition changes but not a downsizing of the entire fleet, as vehicles remain under-priced on average. Rows 6-7 shed light on the role of composition versus scale effects. Roughly, the revenue-neutral fee system creates welfare gains through improved composition. The externality tax improves composition and also reduces the scale of the market in line with the externality.

Table 5, row 8, shows that charging registration fees for new vehicles only, with the fee equal to lifetime external damages from the vehicle, modestly *decrease* social welfare, by $20 billion in present value. This perverse result reflects the power of the Gruenspecht Effect highlighted in Proposition 1. Although these fees encourage new vehicle buyers to choose cleaner vehicles, they also increase the price of all vehicles, which decreases scrap and keeps dirty used vehicles on the road longer. This phenomenon also underscores why the difference-in-differences regressions of Section 5 imply a mixed review of exhaust standards. While Section 5 shows exhaust standards decrease emission rates, this model quantification implies exhaust standards also extend the lifetime of dirtier used vehicles.

Figure 9 shows this example of the Gruenspecht Effect in action. Panel C shows that the average new vehicle has lifetime pollution damages of about $4,500, though new vehicle registration fees in this counterfactual vary by vehicle type and this graph shows the average across types. Charging that externality only to new vehicles decreases purchase of new vehicles, by over 25 percent.[36] Panel D shows that the number of surviving used vehicles increases, especially vehicles 15-30 years old. The new vehicle fee substantially extends used vehicle lifetimes, for precisely the dirtiest vehicles.

Table 5, row 9, shows the effect of changing current registration fees to be identical for all vehicle ages and types. Figure 9, Panel E, shows that this counterfactual decreases registration fees by up to $50 for vehicles younger than 5 years old and increases them by up to $30 for older vehicles. This reform increases social welfare by $18 billion in present value and decreases pollution emissions by around 2 percent. The smaller impact for this counterfactual versus the externality-based fee in rows 6 and 7 reflects the idea that the inefficiency of current registration fees is less due to an implicit subsidy to pollution (which

[36]This relatively elastic response in the first year of policy diminishes in later years as used vintages become in shorter supply.

row 9 remedies) and more due to the failure to price externalities (which rows 5 and 6 address).

Appendix F.9 discusses variations in parameters, data, and assumptions about market power, which produce qualitatively similar results. It also describes how spatially-varying emissions rates boost the benefit-cost ratio of age-based fees in MSAs, and that bans on vehicles become cost effective at age 14 in MSAs but only at age 26 in non-MSA areas.

8.5 Inequality, Environmental Justice, and Political Economy

This analysis provides a menu describing the consequences of different policies' impacts on pollution, surplus, and social welfare. A full analysis should also consider these policies' effects on different socioeconomic groups. Incidence is important directly and for assessing political feasibility. Concern about an equal distribution of environmental quality is a top priority in some jurisdictions.

The counterfactuals we study affect inequality through several channels. Lower-income households tend to own older and more polluting vehicles, so increasing registration fees on dirtier used vehicles could have regressive initial incidence. Panel A of Appendix Figure A11 uses data from the National Highway Travel Survey (U.S. Federal Highway Administration 2001) to show that vehicle owners with household income below $10,000 have a mean vehicle age close to 12 years, while owners with income above $80,000 have a mean vehicle age of 7 years. Similarly, Panel B shows that vehicle owners with less than a high school degree have a mean vehicle age of 10.5 years, while owners with a graduate degree have a mean vehicle age of 7 years. Panel C of the figure displays the distribution of vehicle ages for high and low incomes in more detail.

Appendix Table A9 shows the change in discounted annualized fees paid across the income distribution for our various policy counterfactuals, accounting for scrap of the oldest vehicles.[37] Under age×type vehicle registration fees, fees go up somewhat more for higher-income households—they own newer cars so the per-car fee is less, but they own more cars in total. Overall, however, this fee system is still regressive in that lower-income households pay more as a fraction of income. A revenue-neutral age×type fee system that returns revenues on a per-vehicle basis becomes even more regressive, again because higher-income households own more cars. New-vehicle fees, in contrast, place much of the burden on wealthier groups, but as shown in Table 5, fail to produce pollution improvements.

Several other factors determine the full incidence of emissions policies. First, reforming registration fees and exhaust standards affects the resale value of used vehicles. Making

[37] Our main simulation uses a representative consumer. To account for differential scrap rates by income group, we augment the model. Details are provided in Appendix F.8.

registration fees more proportional to pollution will decrease the value of older polluting vehicles that lower-income households disproportionately own. In the longer run, the lower rental or ownership costs of older and dirtier vehicles will offset some of the effect of changed registration fees.

Second, as shown in Table 5, registration fees raise substantial revenue. Overall incidence depends on how those revenues are redistributed. Registration fees proportional to environmental damages generate $60 billion in annual revenues. Dispersing revenue equally to each household or through the income tax system could produce progressive outcomes. This is not relevant for the revenue-neutral registration fees or exhaust standards.

Third, the health impacts of vehicle pollution reduction may disproportionately benefit low-income households. Similar patterns occur with other corrective taxes (Allcott et al. 2019). Older and dirtier vehicles are disproportionately owned by households that reside in low-income communities. If these vehicles are disproportionately driven near those communities, or pollute them, increasing registration fees on dirty used vehicles could create outsize environmental benefits to those communities. Transportation is a leading source of pollution in vulnerable communities, some of which border major roads (Stuart et al. 2009; Rowangould 2013; Carlson 2018; Apte et al. 2019). Quantifying where vehicles are driven, separately by demographic of owner and vehicle attribute, is a complex task we leave for future research. The net effect of the regressive fee channel and the possibly progressive pollution channel is ambiguous and may vary with the specific counterfactual.

One other impact on political feasibility is worth noting. The registration fee policies we analyze increase the cost of owning used vehicles, which can increase new vehicle demand. Hence, auto manufacturers, a powerful interest group, may support such reforms, particularly if revenue-neutral. At the same time, exhaust standards increase new vehicle prices and encourage substitution to used vehicles, so may be expected to receive less support from auto manufacturers.

What is the broad political feasibility of reforming vehicle registration fees? Only some states impose registration fees that vary with vehicle value or age. The pattern of these states does not obviously reflect geography or politics. While it is hard to generalize globally, Japan does have a national "shaken" registration fee which increases with vehicle age. In general, we believe that mass increases to registration fees are politically sensitive, but moderate reforms to fee patterns, particularly revenue-neutral reforms, have political feasibility in some areas. Our goal is related to that of the optimal taxation literature—to identify the efficiency and equity of potential reforms, while recognizing that the political feasibility of these reforms varies.

9 Conclusions

Vehicle air pollution exhaust standards are arguably among the world's most important environmental policies, particularly for transportation. They have been the subject of little economics research. This contrasts with fuel economy standards, a separate set of regulations that influential economics research has studied carefully. It likewise contrasts with the influential research on the US Clean Air Act's regulation of industry.

This paper examines US exhaust standards over the last half century. We first document vast declines of over 99 percent in air pollution emissions per mile from new US vehicles since exhaust standards began in the 1960s. Panel data regressions using various time periods, datasets, and research designs find that exhaust standards have caused most of that downward emissions trend. Several stylized facts, however, suggest that these standards are not cost-effective because they do not tightly regulate emissions from older vehicles. Additionally, registration fees and property taxes are lower on older and dirtier used vehicles. An analytical model highlights the "Gruenspecht Effect," which policy debates have informally mentioned for decades but has not been rigorously derived before—environmental standards and other policies raising the price of new, clean capital counterproductively extend the lifetime of used, dirty capital. The analytical model also suggests potential efficiency gains from increasing registration fees on old dirty vehicles. A quantitative model finds present-value net benefits in the hundreds of billions of dollars from setting annual registration fees equal to the pollution damages of a vehicle age$\times$type. Using externality-based registration fees appears to have larger benefits than further tightening standards, though both produce substantial gains. In sum, we conclude that vehicle exhaust standards have been remarkably effective, but they have left room for improvement in cost effectiveness, and feasible policy reforms can thus generate large welfare gains.

Given the enormous decreases in pollution from passenger transportation this paper documents, do additional reforms have economically important magnitudes? Although pollution used to be an even worse problem, the 37,000 annual US deaths mentioned at the beginning highlight that pollution is still costly.

We conclude with several areas we believe are important for future work. First, how important are issues in this paper for ongoing fleet composition trends? Although electric vehicles represent less than 1 percent of the US fleet today, industry forecasts suggest electric vehicles may constitute half the fleet in the year 2050 (Cage 2022). Thus, while the transition to electric vehicles will require most of the twenty-first century, policymakers in regions with a clean electric grid will face a trade-off between clean new electric vehicles and polluting older gasoline vehicles. The question of how policy should deal with legacy pollution at that stage will mirror the questions we analyze here. Anticipating that transition may inform policy for electric vehicles today. In addition, this paper shows steady downward trends in emission rates even for gasoline vehicles. While we quantify effects of varying past policy reforms, what are potential welfare gains from current or future additional reforms? Continuing deterioration of emissions control systems with age suggests that in the future when vehicles are cleaner, older used gasoline vehicles may continue emitting the majority of pollution. Such analysis would require projection or imputation of many of the data used in the quantitative model, but are relevant to future policy.

Second, are the environmental benefits of removing the most polluting older vehicles concentrated in low-income communities? While making annual registration fees better reflect pollution damages can create large environmental benefits, it can also create concerns about environmental justice because vulnerable communities may pay a larger share of those fees. At the same time, if vehicle air pollution disproportionately affects vulnerable communities, cleaning it up can improve the equity of overall environmental outcomes.

Third, to what extent should the kinds of policies we study differ across space? Driving in exurbs, suburbs, and city centers creates different levels of externalities, including congestion and pollution damages. Many European cities have addressed these issues with low-emission zones that restrict driving to relatively clean vehicles. Appendix F.9 highlights these policies in a simple framework, but studying such questions in more detail requires models emphasizing spatial differentiation.

Fourth, do the ideas and findings here generalize to other countries? Because most middle- and low-income countries use exhaust standards with stringency set years behind the US, the ideas advanced here are potentially relevant to China, India, Mexico, and many other countries. Testing whether our findings generalize to other countries would be valuable.

Fifth, what are the magnitude, environmental, and welfare consequences of "leakage" due to policies encouraging scrap of polluting old vehicles? For example, suppose the US implemented some reforms we analyze; how would these reforms affect exports of old US vehicles to Mexico, and how would such exports affect welfare in both countries? If Mexico implemented such reforms, one could ask a similar question for Mexico's used vehicle exports to Central America. Davis and Kahn (2010) study these questions for NAFTA and California's smog check policies but one could ask similar questions for exhaust standards and registration fees in broader settings.

Finally, how externally valid are our findings to other types of environmental policy? For example, we find that pollution emission rates have declined precipitously and that environmental policy is the leading cause. Aspects of those findings also appear to apply to electricity generation, industrial air pollution, and municipal water pollution (Shapiro 2022).

The Gruenspecht Effect is relevant for drinking water treatment, coal-fired electricity generation, and industrial water pollution regulation (Stavins 2006). Our quantitative model finds that while tightening pollution standards can produce welfare gains, revising tax instruments to reflect environmental damages can produce larger welfare gains; this broad conclusion of the relative efficiency of taxes over standards is a common theme in environmental economics.

References

Adamou, A., S. Clerides, and T. Zachariadis (2013). Welfare implications of car feebates: A simulation analysis. *Economic Journal 124(578)*, F420–F443.

AES (1973). A study of emissions from light duty vehicles in six cities. Technical report, Automotive Environmental Systems.

AIR (2015). Automobile inspection and readjustment (AIR) program annual report calendar year 2014. Technical report, Colorado Air Pollution Control Division, Department of Public Health & Environment.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019). Regressive sin taxes, with an application to the optimal soda tax. *Quarterly Journal of Economics 134(3)*, 1557–1626.

Anderson, S. T. and J. M. Sallee (2016). Designing policies to make cars greener: A review of the literature. *Annual Review of Resource Economics 8*, 157–180.

Apte, J. S., S. E. Chambliss, C. W. Tessum, and J. D. Marshall (2019). A method to prioritize sources for reducing high pm2.5 exposures in environmental justice communities in california. Technical report, UT Austin.

Archsmith, J., K. Gillingham, C. R. Knittel, and D. S. Rapson (2020). Attribute substitution in household vehicle portfolios. *RAND Journal of Economics 51*(4), 1162–1196.

Auffhammer, M. and R. Kellogg (2011). Clearing the air? The effects of gasoline content regulation on air quality. *American Economic Review 101(6)*, 2687–2722.

Bai, J., S. Li, D. Xie, and H. Zhou (2021). Environmental protection or environmental protectionism? Evidence from tail pipe emission standards in china. *AEA Papers and Proceedings 111*, 381–85.

Barahona, N., F. Gallego, and J.-P. Montero (2019). Vintage-specific driving restrictions. *Review of Economic Studies 87*(4), 1646–1682.

Bento, A. M., L. H. Goulder, M. R. Jacobsen, and R. H. von Haefen (2009). Distributional and efficiency impacts of increased US gasoline taxes. *American Economic Review 99(3)*, 667–699.

Biglaiser, G., F. Li, C. Murry, and Y. Zhou (2020). Intermediaries and product quality in used car markets. *The RAND Journal of Economics 51(3)*, 905–933.

Bishop, G. A. and D. H. Stedman (2008). A decade of on-road emissions measurements. *Environmental Science & Technology 42(5)*, 1651–1656.

Borken-Kleefeld, J. (2013). Guidance note about on-road vehicle emissions remote sensing. Technical report, IPCC.

Bovenberg, A. L., L. H. Goulder, and M. R. Jacobsen (2008). Costs of alternative environmental policy instruments in the presence of industry compensation requirements. *Journal of Public Economics 92*, 1236–1253.

Bresnahan, T. F. and D. A. Yao (1985). The nonpecuniary costs of automobile emissions standards. *RAND Journal of Economics 16(4)*, 437–455.

Bushnell, J. B. and C. D. Wolfram (2012). Enforcement of vintage differentiated regulations: The case of new source review. *Journal of Environmental Economics and Management 64(2)*, 137–152.

Busse, M. R., C. R. Knittel, J. Silva-Risso, and F. Zettelmeyer (2012). Did "cash for clunkers" deliver? The consumer effects of the car allowance rebate system. Unpublished Mimeograph, MIT.

Busse, M. R., C. R. Knittel, and F. Zettelmeyer (2013). Are consumers myopic? Evidence from new and used car purchases. *American Economic Review 103(1)*, 220–56.

Cabral, M. and C. Hoxby (2015). The hated property tax: Salience, tax rates, and tax revolts. NBER Working Paper No. 18514.

Cage, F. (2022). The long road to electric cars. https://graphics.reuters.com/AUTOS-ELECTRIC/USA/mopanyqxwva/.

Carlson, A. E. (2018). The Clean Air Act's blind spot: Microclimates and hotspot pollution. *U.C.L.A. Law Review 65*, 1036–1088.

Carlson, C., D. Burtraw, M. Cropper, and K. L. Palmer (2000). Sulfur dioxide control by electric utilities: What are the gains from trade? *Journal of Political Economy 108(6)*, 1292–1326.

Chambliss, S. E., R. Silva, J. J. West, M. Zeinali, and R. Minjares (2014). Estimating source-attributable health impacts of ambient fine particulate matter exposure: Global premature mortality from surface transportation emissions in 2005. *Environmental Research Letters 9*(10), 104009.

Cherniwchan, J., B. R. Copeland, and M. S. Taylor (2017). Trade and the environment: New methods, measurement, and results. *Annual Review of Economics 9*, 59–85.

Copeland, B. R., J. S. Shapiro, and M. S. Taylor (2022). *Handbook of International Economics*, Chapter Globalization and the Environment. Elsevier Science Publishing.

Copeland, B. R. and M. S. Taylor (1994). North-south trade and the environment. *Quarterly Journal of Economics 109(3)*, 755–787.

Crandall, R. W., H. K. Gruenspecht, T. E. Keeler, and L. B. Lave (1986). *Regulating the Automobile*. The Brookings Institution.

Cropper, M. L., Y. Jiang, A. Alberini, and P. Baur (2014). Getting cars off the road: The cost-effectiveness of an episodic pollution control program. *Environmental and Resource Economics 57*(1), 117–143.

Currie, J., L. Davis, M. Greenstone, and R. Walker (2015). Environmental health risks and housing values: Evidence from 1,600 toxic plant openings and closings. *American Economic Review 105(2)*, 678–709.

Davis, L. W. and M. E. Kahn (2010). International trade in used vehicles: The environmental consequences of nafta. *American Economic Journal: Economic Policy 2*(4), 58–82.

Davis, S. C. (1997). Transportation energy data book. Technical report. Edition 17, Oak Ridge National Laboratory.

Department of Energy (2016). Fact #907: January 11, 2016 light vehicle sales at a record high in 2015. https://www.energy.gov/eere/vehicles/fact-907-january-11-2016-light-vehicle-sales-record-high-2015.

Department of Transportation (2014). Summary of fuel economy performance 2014. Technical report, NHTSA. https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/performance-summary-report-12152014-v2.pdf. Accessed July 15, 2021.

Deschenes, O., M. Greenstone, and J. S. Shapiro (2018). Defensive investments and the demand for air quality: Evidence from the NOx budget program. *American Economic Review 108*(12), 3814–54.

D'Haultfoeuille, X., P. Givord, and X. Boutin (2013). The environmental effect of green taxation: The case of the French bonus/malus. *Economic Journal 124(578)*, F444–F480.

Fann, N., C. M. Fulcher, and K. Baker (2013). The recent and future health burden of air pollution apportioned across U.S. sectors. *Environmental Science & Technology 47*(8), 3580–3589.

Federal Register (1972-1978). Data on annual certification test results. Accessed July 15, 2021.

Feng, Y., D. Fullerton, and L. Gan (2013). Vehicle choices, miles driven, and pollution policies. *Journal of Regulatory Economics 44(1)*, 4–29.

Fowlie, M., C. R. Knittel, and C. Wolfram (2012). Sacred cars? Cost-effective regulation of stationary and nonstationary pollution sources. *American Economic Journal: Economic Policy 4(1)*, 98–126.

Fullerton, D. and S. E. West (2010). Tax and subsidy combinations for the control of car pollution. *B. E. Journal of Economic Analysis & Policy 10(1)*, 1–33.

Gillingham, K., F. Iskhakov, A. Munk-Nielsen, J. Rust, and B. Schjerning (2022). Equilibrium trade in automobiles. *Journal of Political Economy*.

Goldberg, P. K. (1998). The effects of the corporate average fuel efficiency standards in the US. *Journal of Industrial Economics 46(1)*, 1–33.

Goulder, L. H., M. R. Jacobsen, and A. A. van Benthem (2012). Unintended consequences from nested state and federal regulations: The case of the Pavley greenhouse-gas-per-mile limits. *Journal of Environmental Economics and Management 63(2)*, 187–207.

Greenstone, M. (2002). The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 Clean Air Act amendments and the Census of Manufacturers. *Journal of Political Economy 110(6)*, 1175–1219.

Grossman, G. M. and A. B. Krueger (1995). Economic growth and the environment. *Quarterly Journal of Economics 110(2)*, 353–377.

Gruenspecht, H. K. (1982). Differentiated regulation: The case of auto emissions standards. *American Economic Review 72*(2), 328–31.

Gruenspecht, H. K. and R. N. Stavins (2002). New source review under the Clean Air Act:

Ripe for reform. *Resources 147*, 19–23.

Harrington, W. (1997). Fuel economy and motor vehicle emissions. *Journal of Environmental Economics and Management 33*(3), 240–252.

Henderson, V. J. (1996). Effects of air quality regulation. *American Economic Review 86(4)*, 789–813.

Hoekstra, M., S. L. Puller, and J. West (2017). Cash for Corollas: When stimulus reduces spending. *American Economic Journal: Applied Economics 9*(3), 1–35.

Holland, S. P., E. T. Mansur, N. Z. Muller, and A. J. Yates (2016). Are there environmental benefits from driving electric vehicles? The importance of local factors. *American Economic Review 106*(12), 3700–3729.

Hubbard, C. P., J. E. Anderson, and T. J. Wallington (2014). Ethanol and air quality: Influence of fuel ethanol content on emissions and fuel economy of flexible fuel vehicles. *Environmental Science and Technology 48*(1), 861–867.

IHS Markit (2022). How well do you know the global vehicle market? Technical report.

Innes, R. (1996). Regulating automobile pollution under certain, competition, and imperfect information. *Journal of Environmental Economics and Management 31*(2), 219–239.

Jacobsen, M. R. (2013). Evaluating US fuel economy standards in a model with producer and household heterogeneity. *American Economic Journal: Economic Policy 5(2)*, 148–187.

Jacobsen, M. R., C. R. Knittel, J. M. Sallee, and A. A. van Benthem (2020). The use of regression statistics to analyze imperfect pricing policies. *Journal of Political Economy 128*(5), 1826–1876.

Jacobsen, M. R. and A. A. van Benthem (2015). Vehicle scrappage and gasoline policy. *American Economic Review 105(3)*, 1312–38.

Joy, R. W., P. L. Heirigs, G. D. Torgerson, M. S. Denis, T. C. Austin, J. Gordon, B. Tefft, and J. Lindner (2004). Development of the IM147: An alternative inspection/maintenance mass-emission transient test to address vehicle preconditioning concerns. *Journal of the Air and Waste Management Association 54*(3), 269–285.

Kahn, M. E. (1996a). The efficiency and equity of vehicle emissions regulation: Evidence from California's random audits. *Eastern Economic Journal 22(4)*, 457–465.

Kahn, M. E. (1996b). New evidence on trends in vehicle emissions. *RAND Journal of Economics 27(1)*, 183–196.

Kaiser, W. (2003). Clean Air Act and American automobile industry. *Icon 9*, 31–43.

Keiser, D. A., C. L. Kling, and J. S. Shapiro (2019). The low but uncertain measured benefits of US water quality policy. *Proceedings of the National Academy of Sciences 116*(12), 5262–5269.

Keiser, D. A. and J. S. Shapiro (2019). Burning rivers to crystal springs? US water pollution regulation over the last half century. *Journal of Economic Perspectives 33*(4), 51–75.

Keohane, N. O., R. L. Revesz, and R. N. Stavins (1998). The choice of regulatory instruments in environmental policy. *Harvard Environmental Law Review 22*, 313–368.

Knittel, C. R. and R. Sandler (2018). The welfare impact of second-best uniform-Pigouvian taxation: Evidence from transportation. *American Economic Journal: Economic Pol-*

icy 10(4), 211–242.

Kohn, R. (1996). An additive tax and subsidy for controlling automobile pollution. *Applied Economics Letters 3(7)*, 459–462.

Langer, A., V. Maheshri, and C. Winston (2017). From gallons to miles: A disaggregate analysis of automobile travel and externality taxes. *Journal of Public Economics 152*, 34–46.

Lee, J., F. M. Veloso, and D. A. Hounshell (2011). Linking induced technological change, and environmental regulation: Evidence from patenting in the u.s. auto industry.

Lee, J., F. M. Veloso, D. A. Hounshell, and E. S. Rubin (2010). Forcing technological change: A case of automobile emissions control technology development in the us. *Technovation 30*, 249–264.

Levinson, A. (2009). Technology, international trade, and pollution from US manufacturing. *American Economic Review 99(5)*, 2177–2192.

Li, S., J. Linn, and E. Spiller (2013). Evaluating 'cash-for-clunkers': Program effects on auto sales and the environment. *Journal of Environmental Economics and Management 65*(20), 175–193.

Merel, P., A. Smith, J. Williams, and E. Wimberger (2014). Cars on crutches: How much abatement do smog check repairs actually provide? *Journal of Environmental Economics and Management 67(3)*, 371–395.

Mills, E. S. and L. J. White (1978). *Approaches to Controlling Air Pollution*, Chapter Government Policies toward Automotive Emissions Control, pp. 348–409. Cambridge, MA: MIT Press.

Mondt, J. R. (2000). *Clean Cars: The History and Technology of Emission Control Since the 1960s.* Society of Automotive Engineers.

National Commission on Air Quality (1981). To breathe clean air. Technical report, National Commission on Air Quality.

National Research Council (1988). *Air Pollution, the Automobile, and Public Health.* Washington, DC: The National Academies Press.

National Research Council (2001). Evaluating vehicle emissions inspection and maintenance programs. Technical report, NRC.

National Research Council (2004). Air quality management in the United States. Technical report, NRC.

National Research Council (2006). *State and Federal Standard for Mobile-Source Emissions.* Washington, DC: The National Academies Press.

National Research Council (2010). *Hidden Costs of Energy: Unpriced Consequences of Energy Production and Use.* Washington, DC: The National Academies Press.

Oliva, P. (2015). Environmental regulations and corruption: Automobile emissions in Mexico City. *Journal of Political Economy 123(3)*, 686–724.

Poterba, J. and T. Sinai (2008). Tax expenditures for owner-occupied housing: Deductions for property taxes and mortgage interest and the exclusion of imputed rental income. *American Economic Review: Papers & Proceedings 98*(2), 84–89.

Rowangould, G. M. (2013). A census of the US near-roadway population: Public ehalth and environmental justice considerations. *Transportation Research Part D: Transport and Environment 25*, 59–67.

Rozendaal, R. L. and H. R. Vollebergh (2021). Policy-induced innovation in clean technologies: Evidence from the car market. *CESifo Working Paper No. 9422*.

Sanders, N. J. and R. Sandler (2020). Technology and the effectiveness of regulatory programs over time: Vehicle emissions and smog checks with a changing fleet. *Journal of the Association of Environmental and Resource Economists 7*(3), 587–618.

Sandler, R. (2012). Clunkers or junkers? Adverse selection in a vehicle retirement program. *American Economic Journal: Economic Policy 4(4)*, 253–281.

Shapiro, J. S. (2021). The environmental bias of trade policy. *Quarterly Journal of Economics 136*(2), 831–886.

Shapiro, J. S. (2022). Pollution trends and US environmental policy: Lessons from the past half century. *Review of Environmental Economics and Policy 16*(1), 42–61.

Shapiro, J. S. and R. Walker (2018). Why is pollution from U.S. manufacturing declining? The roles of environmental regulation, productivity, and trade. *American Economic Review 108*(12), 3814–54.

Sierra Research (1997). Development of a proposed procedure for determining the equivalency of alternative inspection and maintenance programs. Technical report, Sierra Research, Inc.

Stavins, R. N. (2006). Vintage-differentiated environmental regulation. *Stanford Law Review 25*, 29–63.

Stedman, D. H., G. A. Bishop, and R. S. Slott (1998). Repair avoidance and evaluating inspection and maintenance programs. *Environmental Science & Technology 32(10)*, 1544–1545.

Stuart, A. L., S. Mudhasakul, and W. Sriwatanapongse (2009). The social distribution of neighborhood-scale air pollution and monitoring protection. *Journal of the Air & Waste Management Association 59*(5), 591–602.

Tschofen, P., I. L. Azevedo, and N. Z. Muller (2019). Fine particulate matter damages and value added in the US economy. *Proceedings of the National Academy of Sciences 116*(40), 19857–19862.

U.S Bureau of Transportation Statistics (2011). Unregistration rates for on-road vehicles in california. Technical report, Washington, DC: U.S. BTS. https://www.bts.gov/archive/publications/journal_of_transportation_and_statistics/volume_07_number_23/paper_01/index.

U.S. EPA (1973). A report on automotive fuel economy. Technical report, U.S. EPA.

U.S. EPA (1995). *EPA I/M Briefing Book*. Office of Air and Radiation, U.S. EPA. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=9101LWU3.TXT.

U.S. EPA (1999). Regulatory impact analysis - control of air pollution from new motor vehicles: Tier 2 motor vehicle emissions standards and gasoline sulfur control requirements. Technical report, U.S. EPA, EPA420-R-99-023.

U.S. EPA (2006). EPA's oversight of the vehicle inspection and maintenance program needs improvement. Technical report, U.S. EPA. Report No. 2007-P-00001.

U.S. EPA (2008). 2008 progress report: Vehicle and engine compliance activities. Technical report, U.S. EPA. https://nepis.epa.gov/Exe/ZyPDF.cgi/P1008K2D.PDF?Dockey=P1008K2D.PDF.

U.S. EPA (2014a). Control of air pollution from motor vehicles: Tier 3 motor vehicle emission and fuel standards final rule. Technical report, U.S. EPA, EPA-420-R-14-005.

U.S. EPA (2014b). National emissions inventory 2014. https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data. Accessed January 30, 2017.

U.S. EPA (2016). EPA emission standards for light-duty vehicles and trucks. Technical report, U.S. EPA, https://www.epa.gov/emission-standards-reference-guide/epa-emission-standards-light-duty-vehicles-and-trucks. Accessed October 10, 2019.

U.S. EPA (2018). Our nation's air: Status and trends through 2018. Technical report, U.S. EPA. https://gispub.epa.gov/air/trendsreport/2018/.

U.S. EPA (2019). *Reformulated Gasoline*. https://www.epa.gov/gasoline-standards/reformulated-gasoline.

U.S. EPA (2021). The 2021 epa automotive trends report: Greenhouse gas emissions, fuel economy, and technology since 1975. Technical report, U.S. EPA.

U.S. EPA (2022). History of reducing air pollution from transportation in the united states. https://www.epa.gov/transportation-air-pollution-and-climate-change/accomplishments-and-success-air-pollution-transportation#:~:text=Compared%20to%201970%20vehicle%20models,nitrogen%20oxides%20and%20particle%20emissions).

Venigalla, M. (2013). *Handbook of Measurement in Science and Engineering*, Chapter Mobile Source Emissions Testing, pp. 233–240. John Wiley & Sons, Inc.

Vollebergh, H. R. (2010). Fuel taxes, motor vehicle emission standards and patents related to the fuel-efficiency and emissions of motor vehicles. joint meetings of tax and environment experts. Report from the OECD's Joint Meetings of Tax and Environment Experts.

Walker, R. (2013). The transitional costs of sectoral reallocation: Evidence from the Clean Air Act and the workforce. *Quarterly Journal of Economics 128(4)*, 1787–1835.

Walls, M. and J. Hanson (1999). Distributional aspects of an environmental tax shift: The case of motor vehicle emissions taxes. *National Tax Journal 52(1)*, 53–65.

West, S. E. and R. C. Williams (2005). The cost of reducing gasoline consumption. *American Economic Review Papers and Proceedings 95(2)*, 294–299.

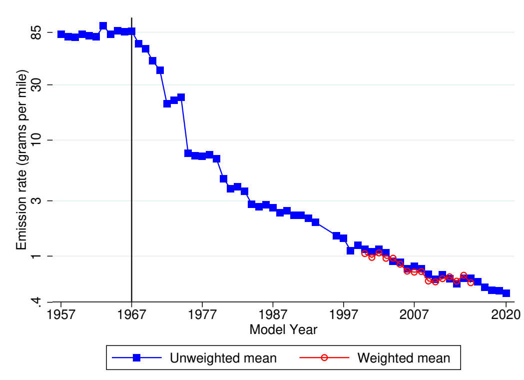
World Bank (2014). Transport for health: The global burden of disease from motor road transport. Technical report, World Bank.

Yacobucci, B. D. (2015). EPA's in-use emissions testing and Volkswagen's "defeat device". *CRS Insight*. https://www.hsdl.org/?view&did=787654.

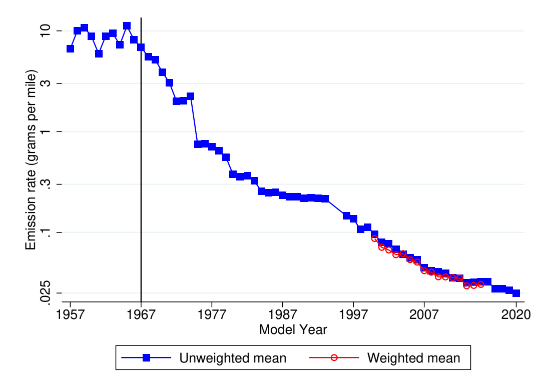
Figures and Tables

Figure 1: Mean Pollution Emission Rates of New US Vehicles, 1957-2020

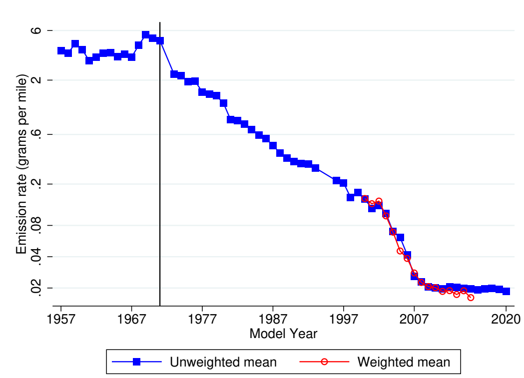
(a) Carbon monoxide (CO)



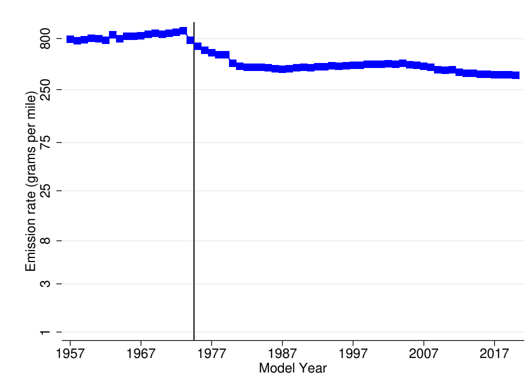
(b) Hydrocarbons (HC)



(c) Nitrogen oxides (NO_x)

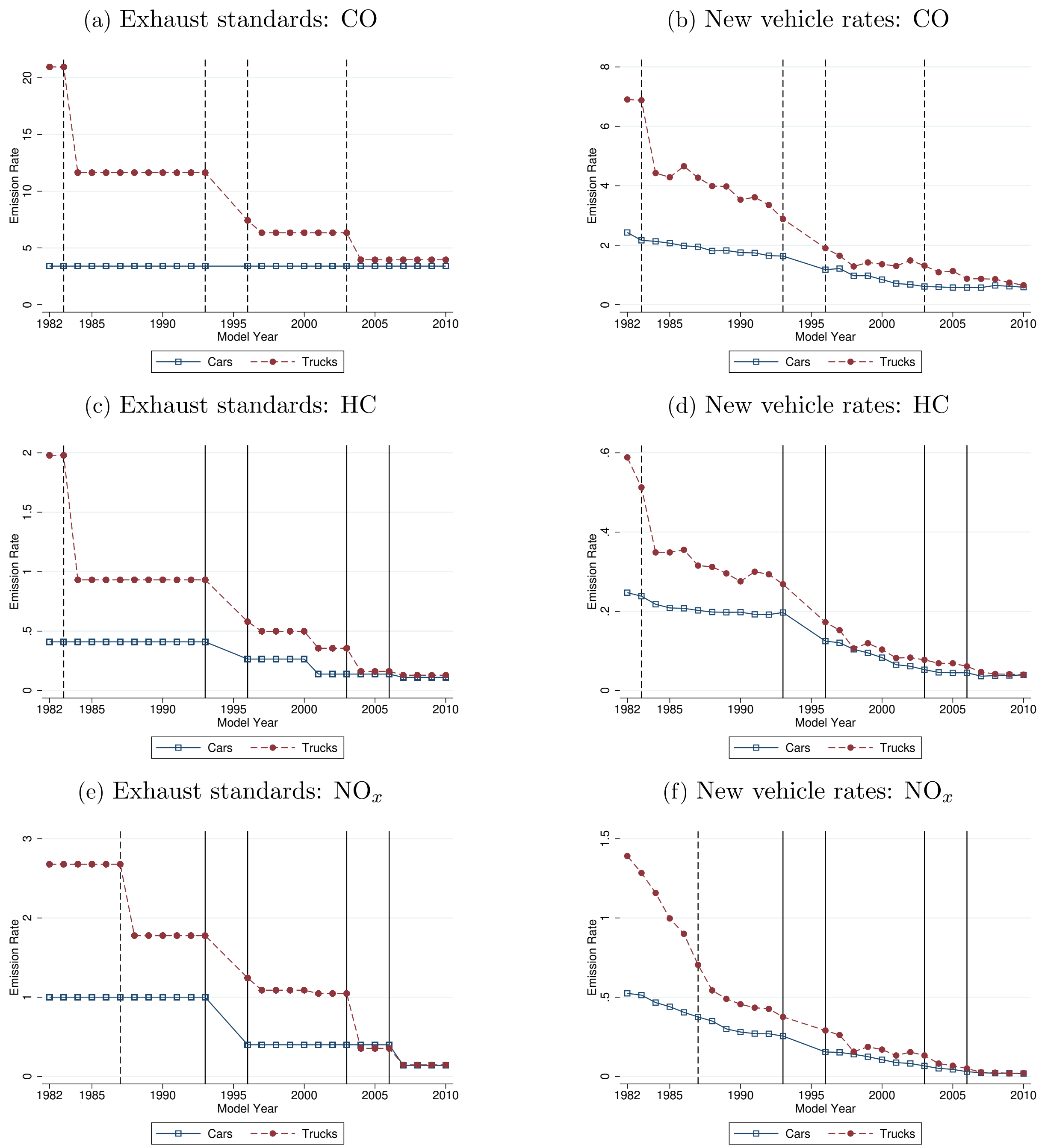


(d) Carbon dioxide (CO_2)



NOTES: Y-axes have logarithmic scale. Graphs use full sample of new vehicle test data and AES (1973). For Panels A-C, model years 1957-1971 are means of a sample of used vehicles given an FTP test. Model years 1972-2020 are from certification test records for 50,000 miles. Model years 1972-4 received an earlier version of the FTP test ("FTP72"). We concord FTP72 to FTP values, separately by pollutant, using ratios for all vehicles in AES (1973). Vertical line depicts year before exhaust standards began. CO_2 data are sales-weighted fleet-wide averages. CO_2 data converted from mile per gallon data, from U.S. EPA (1973) for 1957-1975 and U.S. EPA (2021) for 1975-2020. We splice the two CO_2 series to have the same mean in 1975. Weights for CO, HC, and NO_x in the red lines with circles are the frequency of each vehicle in Colorado remote sensing data.

Figure 2: Exhaust Standards and Emission Rates, Cars Versus Trucks



(a) Exhaust standards: CO

(b) New vehicle rates: CO

(c) Exhaust standards: HC

(d) New vehicle rates: HC

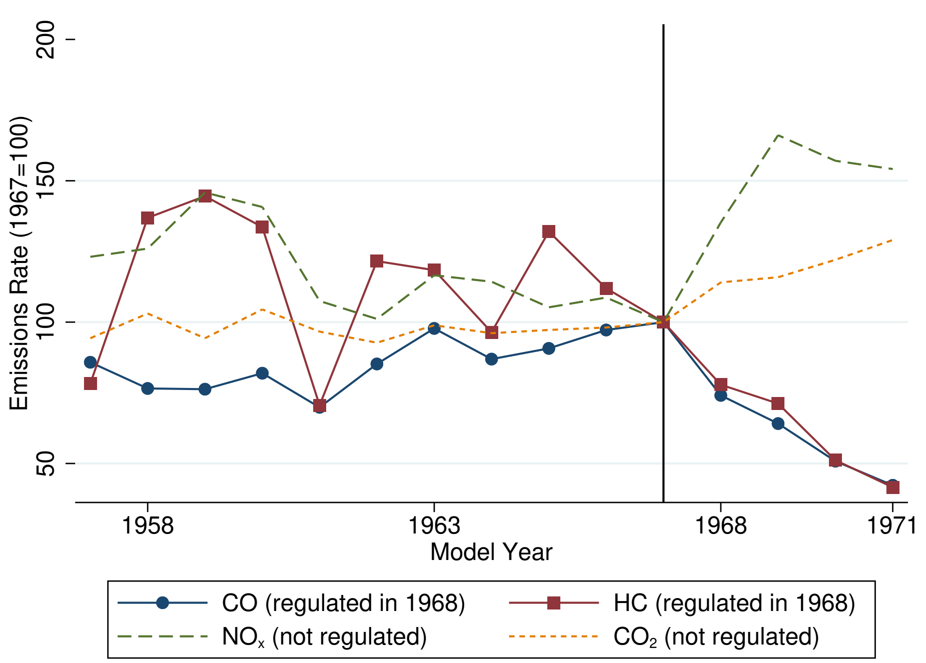
(e) Exhaust standards: NO_x

(f) New vehicle rates: NO_x

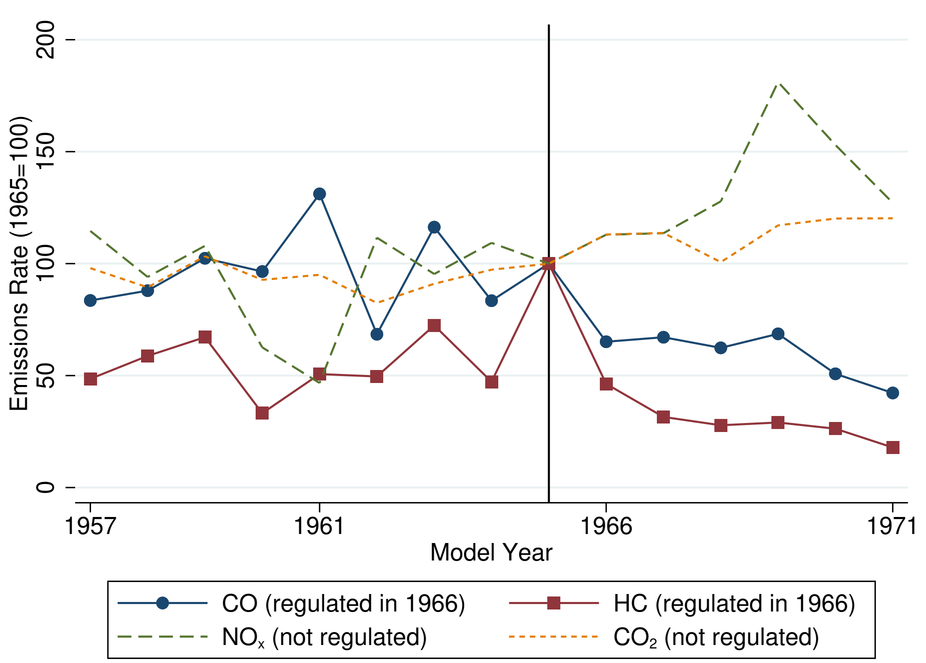
Notes: Dashed vertical lines show years when standards change for cars only; solid vertical lines show years when standards change for both cars and trucks. Each panel uses full sample, restricted to model years 1982-2010. Panels D through F show certification levels, equal to raw test results scaled up by deterioration factors for 50,000 miles. Appendix A.1 explains details. Beginning in 1988 for NO_x and 1994 for other pollutants, standards distinguish sub-groups of trucks based on weight; graphs show weighted means of standards across these groups, with weights equal to the proportion of each vehicle from model year 1993 in Colorado smog check test data.

Figure 3: Event Study Graphs for Tier 0 Exhaust Standards, 1957-1971

(a) Vehicles outside California, model years 1957-1971



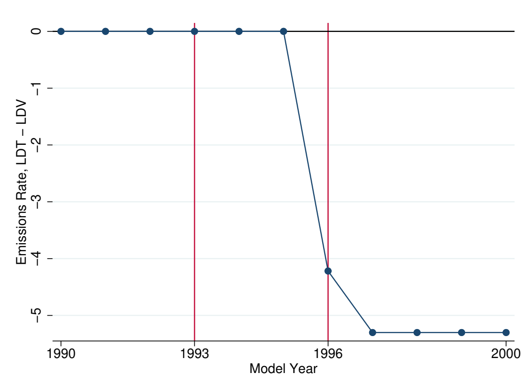
(b) Vehicles in California, model years 1957-1971



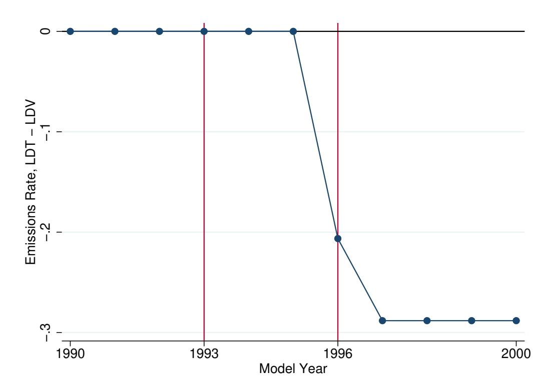
NOTES: Graphs use full sample from AES (1973). All emission rates are in grams per mile, scaled to equal 100 in 1967 (Panel A) or 1965 (Panel B).

Figure 4: Event Study Graphs for Tier 1 Exhaust Standards, 1990-2000

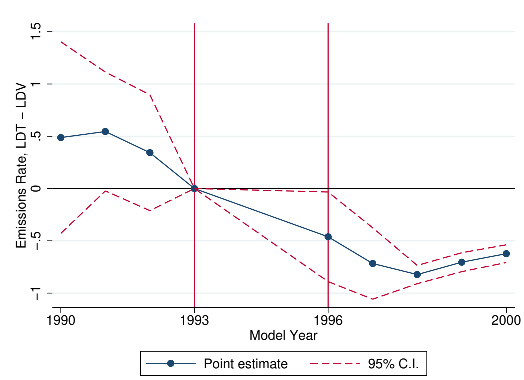
(a) Exhaust standards: CO



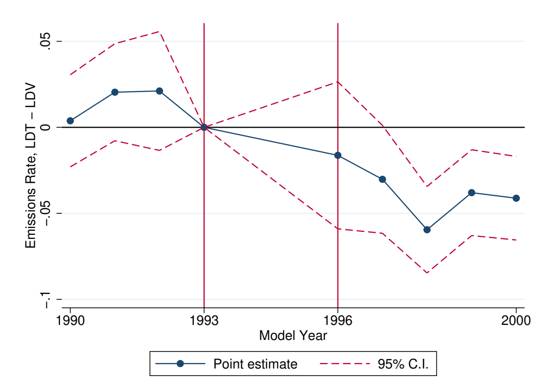
(b) Exhaust standards: HC



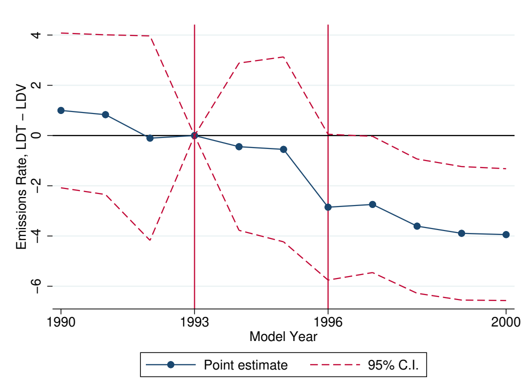
(c) New vehicle emissions: CO



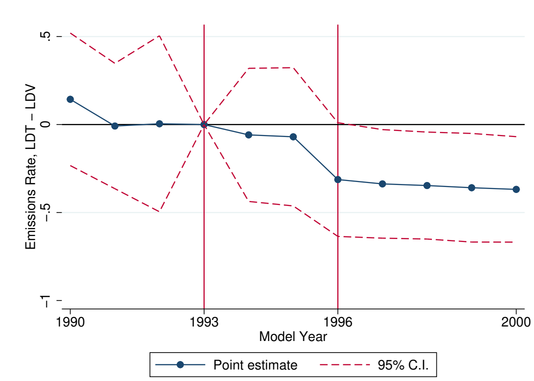
(d) New vehicle emissions: HC



(e) Used vehicle emissions: CO



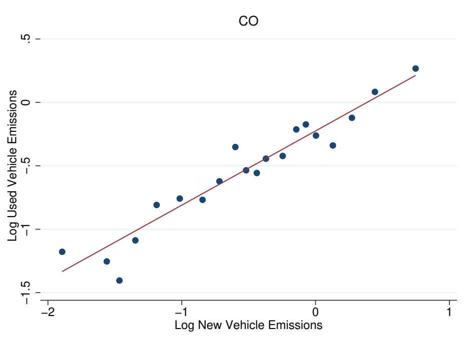
(f) Used vehicle emissions: HC



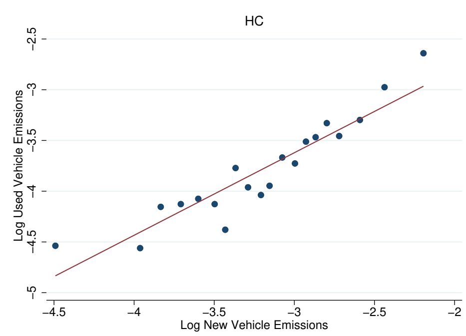
NOTES: Graphs use model years 1990-2000 from new vehicle tests (Panels C and D) or Colorado smog check data (Panels E and F). Emissions are measured in grams per mile. In panels A and B, each class×model year is weighted by its share in the 1993 Colorado smog check test data. Panels C and D show certification levels for 50,000 miles. New vehicle emission rate data are unusable for 1994-1995 (see Appendix B.2). Reference year is 1993. Standard errors are clustered by model year×truck type.

Figure 5: Used Versus New Emission Rates for Tier 2 Exhaust Standards, 2000-2010

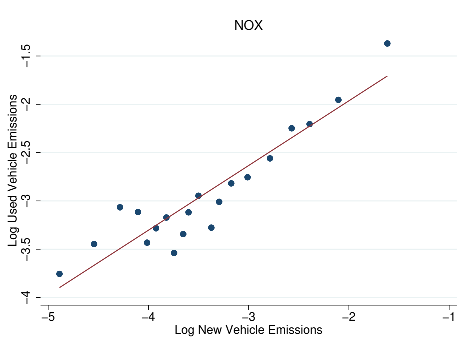
(a) Carbon monoxide (CO)



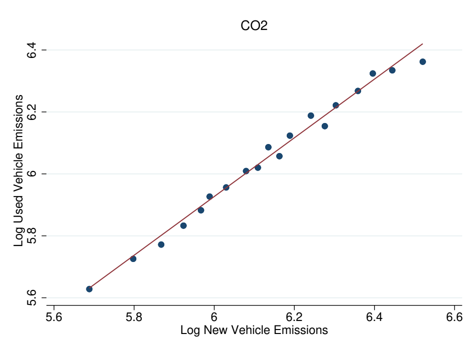
(b) Hydrocarbons (HC)



(c) Nitrogen oxides (NO_x)



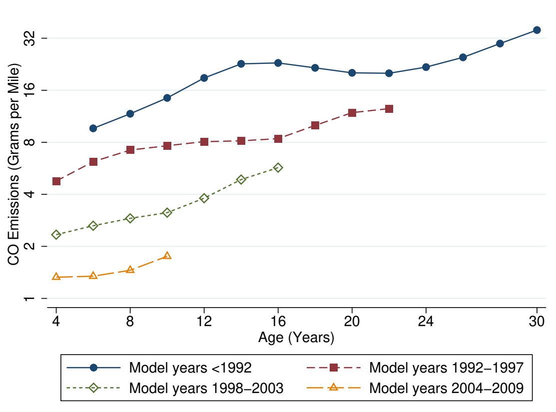
(d) Carbon dioxide (CO_2)



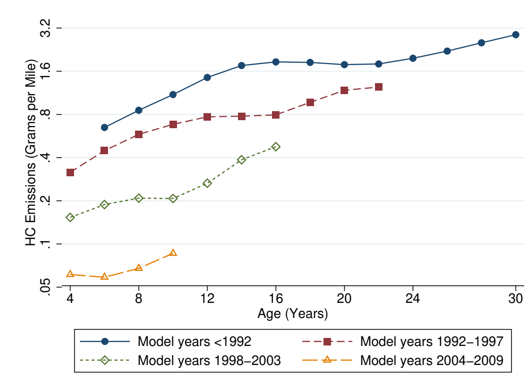
Notes: Graphs show binned scatter plots. Graphs use new vehicle tests and Colorado smog check data. Graphs exclude small share of observations with zero pollution readings.

Figure 6: Used Vehicle Emission Rates and Miles Traveled, by Model Year and Age

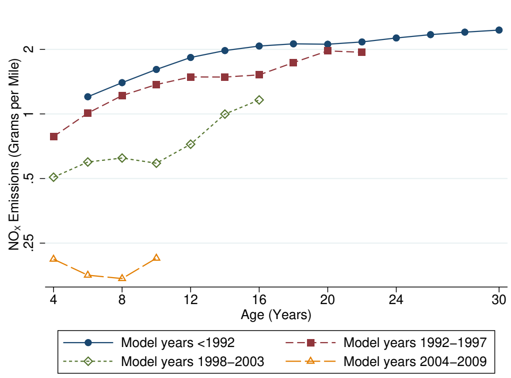
(a) Carbon monoxide (CO)



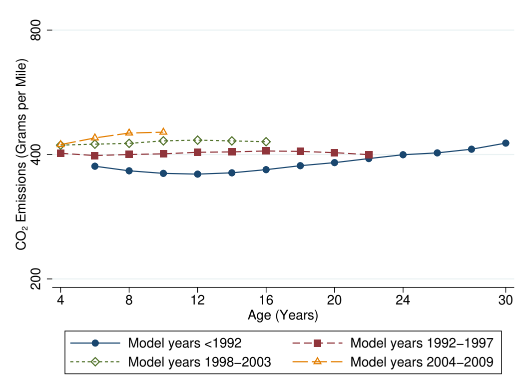
(b) Hydrocarbons (HC)



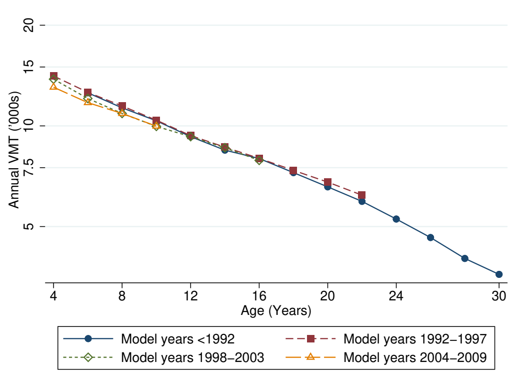
(c) Nitrogen oxides (NO_x)



(d) Carbon dioxide (CO_2)

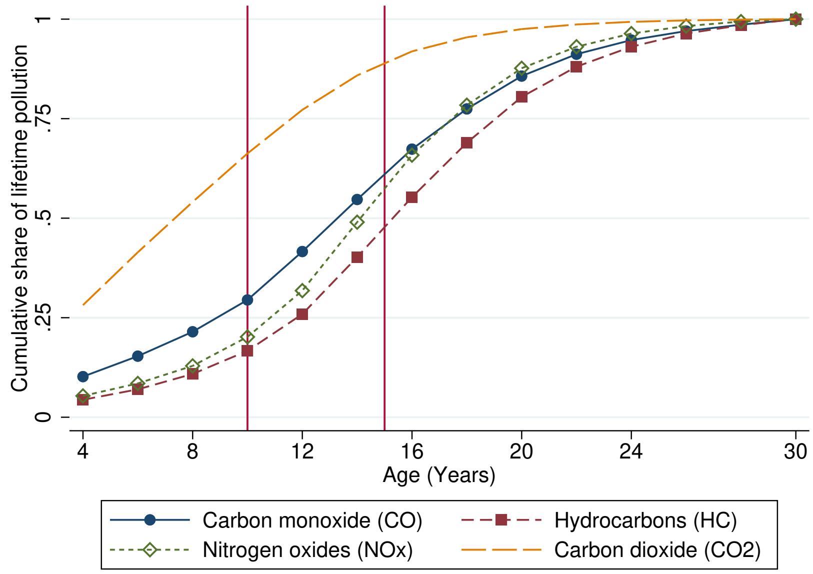


(e) Annual vehicle miles traveled



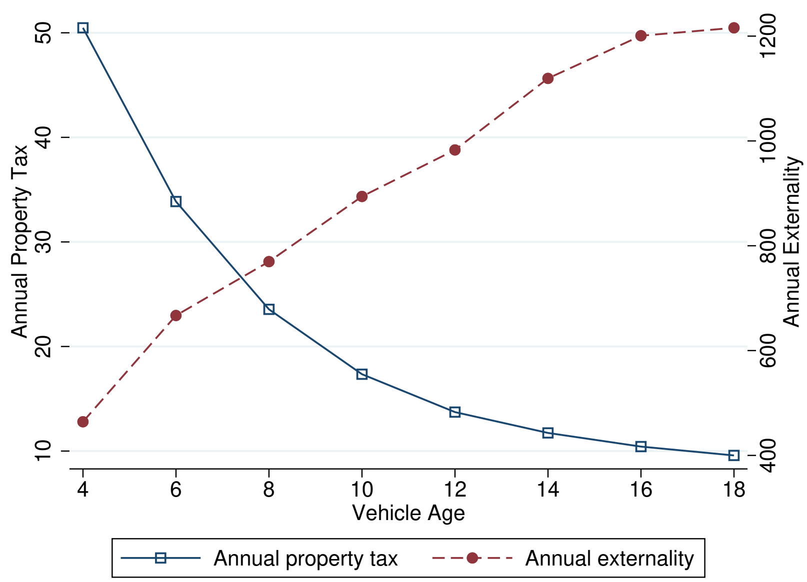
NOTES: Figures use full sample from Colorado smog check data. Points represent mean emission rates in a given model year×age cell, averaged across all vehicles in the data. Y-axes have logarithmic scale.

Figure 7: Cumulative Share of Fleet Emissions from Each Vehicle Age



NOTES: Each line shows the cumulative distribution for total pollution emissions from each age. Vertical lines at ages 10 & 15 show when exhaust standards stop applying. Pollution for a vehicle equals the emission rate times miles driven. Miles equals change in vehicle odometer since the previous test, divided by years since the previous test. For a vehicle's first test, decimal years equals age. Data from 2014 Colorado inspections.

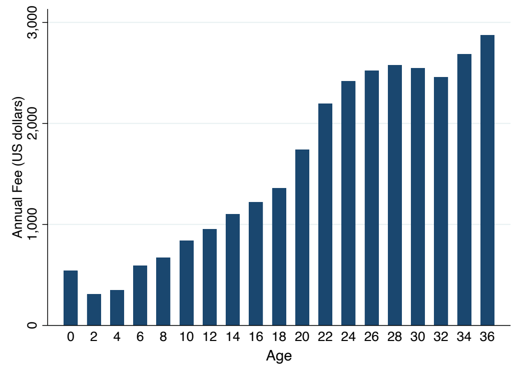
Figure 8: Annual Pollution Externalities, Property Taxes, and Vehicle Age



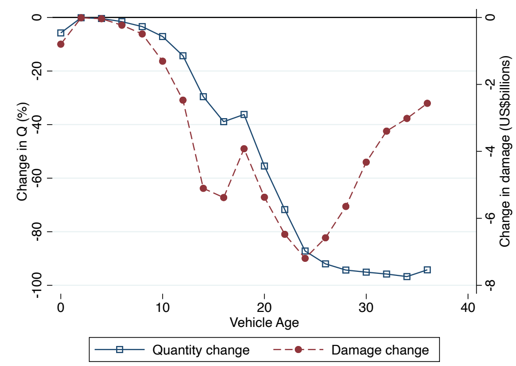
NOTES: Graph measures market shares of VIN prefixes using calendar year 2000 Colorado inspections to calculate the mean externality and tax by age. Vehicle values are from the National Automobile Dealers Association used retail prices. Currency in 2019$. Property taxes are weighted across regions by population.

Figure 9: Model-Based Estimates: Levels of Counterfactual Registration Fees and Effects on Fleet Composition and Pollution Damages

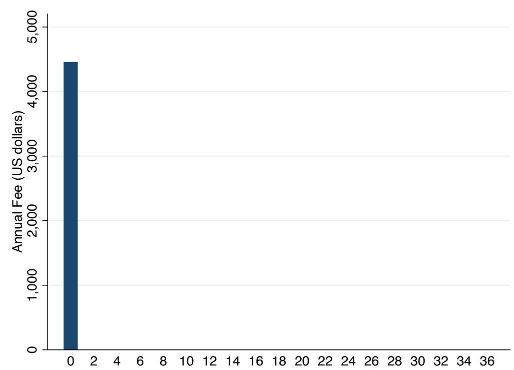
(a) Age×type fee



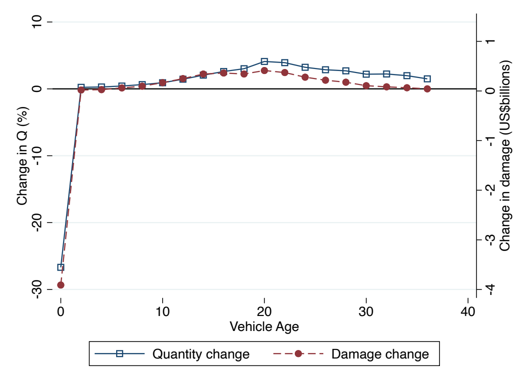
(b) Age×type fee: Effect on fleet



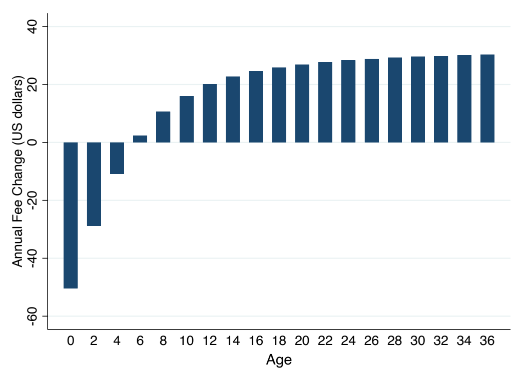
(c) New vehicle fee



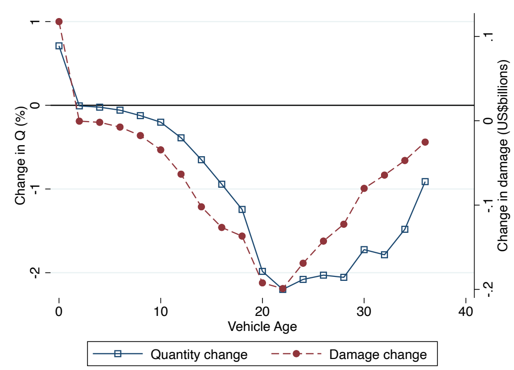
(d) New vehicle fee: Effect on fleet



(e) Change in fees due to flat fee



(f) Flat fee: Effect on fleet



Notes: Panels B, D, and F show the model-based estimates of the impact of counterfactual policies on the calendar year 2000 fleet and environmental damages. Currency values are in 2019$, deflated using the Consumer Price Index for urban consumers.

Table 1: Federal Exhaust Standards

Policy	Model years	Light-duty vehicles CO	Light-duty vehicles HC	Light-duty vehicles NO_x	Light-duty trucks CO	Light-duty trucks HC	Light-duty trucks NO_x	Mean Limit	Mean Pollutant
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Uncontrolled	-1967	90.0	8.200	3.40	90.0	8.200	3.40	—	—
Tier 0	1968-1971	34.0	4.100	—	34.0	4.100	—	—	—
	1972-1974	28.0	3.000	3.10	28.0	3.000	3.10	—	—
	1975-1976	15.0	1.500	3.10	20.0	2.000	3.10	—	—
	1977-1978	15.0	1.500	2.00	20.0	2.000	3.10	—	—
	1979	15.0	1.500	2.00	18.0	1.700	2.30	—	—
	1980	7.0	0.410	2.00	18.0	1.700	2.30	—	—
	1981-1983	3.4	0.410	1.00	18.0	1.700	2.30	—	—
	1984-1987	3.4	0.410	1.00	10.0	0.800	2.30	—	—
	1988-1993	3.4	0.410	1.00	10.0	0.800	1.50	—	—
Tier 1	1994-1996	3.4	0.250	0.40	10.0	0.250	0.85	—	—
	1997-2000	3.4	0.250	0.40	5.2	0.250	0.85	—	—
NLEV (8 states)	1999-2000	3.4	0.250	0.40	5.2	0.250	0.85	0.075	NMOG
NLEV	2001-2003	3.4	0.139	0.40	5.2	0.250	0.80	0.075	NMOG
Tier 2	2004-2006	3.4	0.125	0.40	3.4	0.139	0.40	0.070	NO_x
	2007-2016	3.4	0.100	0.14	3.4	0.100	0.14	0.070	NO_x
Tier 3	2017-2025	4.2^+	0.16^+		4.2^+	0.16^+		0.030	NMOG+NO_x

NOTES: CO is carbon monoxide, HC is hydrocarbons, NO_x is nitrogen oxides, NMOG is non-methane organic gases. All numbers are for gasoline vehicles, measured in grams per mile by the Federal Test Procedure. See Appendix A.1 for details. Columns (5) through (7) show mean standards across truck types, with weights equal to the proportion of each vehicle from model year 1993 in Colorado smog check data. For policies that impose a fleet-wide mean limit, columns (2) through (7) show the limit for the highest bin. $^+$Tier 3 standards apply at 150,000 miles, whereas earlier policies apply at lower mileage. Tier 3 has a combined NMOG+NO_x standard, which is phased in and reaches 0.030 in model year 2025. Uncontrolled emissions are calculated based on emission rates and estimates from vehicles before emissions controls. Sources: National Commission on Air Quality (1981); Bresnahan and Yao (1985); Davis (1997); U.S. EPA (2016).

Table 2: Effects of Tier 0 Exhaust Standards on Vehicle Emissions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel A. Carbon monoxide and hydrocarbons (CO and HC)						
Exhaust standard	0.61***	0.80***	0.97***	0.62***	0.90***	0.59***	0.82***
	(0.07)	(0.08)	(0.18)	(0.08)	(0.09)	(0.12)	(0.18)
N	105	105	105	60	60	45	45
	Panel B. Carbon monoxide (CO)						
Exhaust standard	0.48***	0.46**	0.76***	0.52***	—	0.52***	—
	(0.07)	(0.18)	(0.18)	(0.07)	—	(0.07)	—
N	30	30	30	15	—	15	—
	Panel C. Hydrocarbons (HC)						
Exhaust standard	0.76***	0.22	0.52*	0.71***	—	0.71***	—
	(0.11)	(0.20)	(0.28)	(0.13)	—	(0.13)	—
N	30	30	30	15	—	15	—
Fixed effects:							
Pollutant × region	X	X	X	X	X	X	X
Model year	—	X	X	—	X	—	X
Levels	—	—	X	—	—	—	—
California only	—	—	—	X	X	—	—
Federal only	—	—	—	—	—	X	X

NOTES: Dependent variable is the emission rate in grams/mile from AES (1973). Regressions are in logs except where otherwise noted. Robust standard errors are in parentheses. Before standards began, "exhaust standards" are defined to equal the unconstrained emission rate from Table 1. Asterisks denote p-value < 0.10 (*), <0.05 (**), or <0.01 (***).

Table 3: Effects of Tier 1 Exhaust Standards on Used and New Vehicle Emission Rates

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Carbon monoxide and hydrocarbons (CO and HC), used vehicles							
Exhaust standard	2.02***	0.86***	1.01***	0.39***	1.15***	0.88***	1.34***
	(0.12)	(0.08)	(0.09)	(0.13)	(0.13)	(0.11)	(0.13)
N	17,165,695	17,165,695	16,874,083	17,165,695	3,352,360	13,379,341	17,165,695
Panel B. Carbon monoxide (CO), used vehicles							
Exhaust standard	2.03***	0.81***	0.77***	0.45***	1.02***	0.77***	1.34***
	(0.14)	(0.07)	(0.09)	(0.17)	(0.12)	(0.09)	(0.13)
N	8,568,269	8,568,269	8,422,458	8,568,269	1,670,269	6,675,107	8,568,269
Panel C. Hydrocarbons (HC), used vehicles							
Exhaust standard	2.02***	1.09***	2.20***	0.23	2.44***	1.97***	1.47***
	(0.14)	(0.22)	(0.25)	(0.22)	(0.35)	(0.33)	(0.13)
N	8,597,426	8,597,426	8,451,625	8,597,426	1,682,091	6,704,234	8,597,426
Panel D. Carbon monoxide and hydrocarbons (CO and HC), new vehicles							
Exhaust standard	1.29***	0.54***	0.52***	0.36***	—	0.35***	0.29***
	(0.10)	(0.05)	(0.06)	(0.07)	—	(0.12)	(0.01)
N	17,039	17,039	17,039	17,039	—	11,111	17,039
Panel E. Carbon monoxide (CO), new vehicles							
Exhaust standard	1.36***	0.54***	0.54***	0.33**	—	0.35***	0.29***
	(0.09)	(0.06)	(0.07)	(0.14)	—	(0.12)	(0.01)
N	8,522	8,522	8,522	8,522	—	5,557	8,522
Panel F. Hydrocarbons (HC), new vehicles							
Exhaust standard	1.25***	0.53***	0.49***	0.34***	—	0.35	0.22***
	(0.11)	(0.05)	(0.05)	(0.07)	—	(0.30)	(0.01)
N	8,517	8,517	8,517	8,517	—	5,554	8,517
Fixed effects							
Truck × pollutant	X	X	X	X	X	X	X
Model yr. × pollutant	—	X	X	X	X	X	X
Age × pollutant	X	X	X	X	X	X	X
Odometer	X	X	X	X	X	X	X
CAFE standards	—	—	X	—	—	—	—
Smog check stds.	—	—	X	—	—	—	—
Gasoline cost per mile	—	—	X	—	—	—	—
Ethanol share	—	—	X	—	—	—	—
Sulfur content	—	—	X	—	—	—	—
Model yr. × truck trend	—	—	—	X	—	—	—
Ages 4-6	—	—	—	—	X	—	—
Model yrs. 1990-2000	—	—	—	—	—	X	—
Levels	—	—	—	—	—	—	X

Notes: Dependent variable is the emission rate in grams/mile. Independent and dependent variables are in logs except where otherwise noted. Estimates use model years 1982-2000, except for column (6). Panels A-C use the Colorado inspection data; Panels D through F use the new vehicle inspection data. Odometer includes linear and squared odometer and odometer terms. New vehicle data in Panels D through F lacks age, odometer, and controls for other policies besides CAFE. Standard errors are clustered by model year×truck type. Asterisks denote p-value <0.10 (*), <0.05 (**), <0.01 (***).

Table 4: Assessment of Tier 2 Exhaust Standards: Do New Predict Used Vehicle Emission Rates?

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Carbon monoxide (CO) and hydrocarbons (HC) and nitrogen oxides (NO_x)								
New vehicle emission rate	0.67***	0.50***	0.49***	0.49***	0.48***	0.74***	0.21***	0.39***
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.06)	(0.01)	(0.05)
N	216,933	216,918	216,918	216,918	106,965	216,918	9,757,515	9,757,515
Panel B. Carbon monoxide (CO)								
New vehicle emission rate	0.59***	0.58***	0.60***	0.58***	0.58***	0.76***	0.16***	0.51***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	(0.06)	(0.01)	(0.06)
N	72,311	72,306	72,306	72,306	35,655	72,306	3,252,505	3,252,505
Panel C. Hydrocarbons (HC)								
New vehicle emission rate	0.81***	0.61***	0.51***	0.60***	0.49***	0.96***	0.35***	1.25***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.08)	(0.01)	(0.06)
N	72,311	72,306	72,306	72,306	35,655	72,306	3,252,505	3,252,505
Panel D. Nitrogen oxides (NO_x)								
New vehicle emission rate	0.67***	0.34***	0.36***	0.34***	0.33***	1.16***	0.20***	1.36***
	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)	(0.10)	(0.01)	(0.09)
N	72,311	72,306	72,306	72,306	35,655	72,306	3,252,505	3,252,505
Panel E. Carbon dioxide (CO_2)								
New vehicle emission rate	0.95***	0.87***	0.84***	0.87***	0.84***	0.79***	0.77***	0.72***
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N	72,311	72,306	72,306	72,306	35,655	72,306	3,252,505	3,252,505
Age, model year FE	—	X	X	X	X	X	—	—
Truck indicator	—	X	X	X	X	X	—	—
Odometer	—	X	X	X	X	X	—	—
CAFE standards	—	—	X	—	—	—	—	—
Smog check standards	—	—	X	—	—	—	—	—
Gasoline cost per mile	—	—	X	—	—	—	—	—
Ethanol share	—	—	X	—	—	—	—	—
Sulfur content	—	—	X	—	—	—	—	—
Model year × truck trend	—	—	—	X	—	—	—	—
Ages 4-6	—	—	—	—	X	—	—	—
Levels	—	—	—	—	—	X	—	X
Include abbreviated tests	—	—	—	—	—	—	X	X

NOTES: Dependent variable is the used vehicle emission rate in grams/mile. Regressions are in logs except where otherwise noted. Regressions use model years 1982-2000 of new vehicle tests and Colorado smog check data. Columns (1) through (7) use the observations which completed all 240 seconds of the smog check test (Appendix B.3 describes details). New vehicle emission rate is certification level for 50,000 miles. Estimates correspond to the specification of Table 3, column (1), except where otherwise noted. Smog check standard is not defined for CO_2. Standard errors are clustered by VIN prefix. Asterisks denote p-value <0.10 (*), <0.05 (**), <0.01 (***).

Table 5: Model-Based Estimates: Effects of Counterfactual Exhaust Standards and Registration Fees

	Change in market surplus	Change in pollution damages	Total change in social welfare = (1) - (2)	New tax revenue	Percent change in cumulative emissions		
					CO	HC	NOx
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Counterfactual Exhaust Standards							
1. Delay Tier 2 by four years	8.2	115.3	-107.2	0.0	8.1	4.6	10.3
2. Delay Tier 2 by eight years	13.3	198.2	-184.9	0.0	15.8	8.1	17.8
3. Accelerate Tier 2 by four years	-9.9	-122.8	112.9	0.0	-6.3	-4.7	-10.8
4. Accelerate Tier 2 by eight years	-20.7	-195.2	174.5	0.0	-9.7	-7.5	-17.1
5. Tighten standards 10 percent	-2.3	-27.0	24.7	0.0	-1.4	-1.0	-2.4
Panel B. Counterfactual Registration Fees							
6. Age×type fee	-170.6	-492.5	321.9	1,167.5	-42.3	-42.7	-24.6
7. Age×type fee, revenue neutral	-113.9	-343.7	229.7	0.0	-33.2	-33.5	-15.7
8. New vehicle fee	-16.5	3.2	-19.7	399.6	1.7	1.8	-0.3
9. Flat registration fee	-3.2	-20.7	17.5	0.0	-1.9	-1.9	-1.1

NOTES: Policies start in calendar year 2000 and effects are calculated over 20 years. Values in columns (1) through (4) are in billions of \$2019. Values in columns (5) through (7) are percent changes. Social welfare is defined as consumer + producer surplus − pollution damages, which equals welfare for a social welfare function that abstracts from distribution. As we assume perfect competition among vehicle manufacturers, market surplus equals consumer surplus. The main text describes each counterfactual policy.

Appendix

A Background Material on Standards, Tests and Technology (Section 2)

A.1 Exhaust Standards: Additional Details

Evaporative Emissions. Tier 1 introduced evaporative emission standards, which come from gasoline evaporation due to ambient temperature fluctuations, vaporized gasoline during regular driving, evaporation from a hot vehicle in the hour after it is turned off, permeation once some components of the engine system are saturated with fuel, and refueling while gasoline is pumped (Manufacturers of Emissions Controls Association 2010). Because evaporative test methods have changed over our sample and only target HC from evaporation (not from exhaust, and not for CO or NO_x), and since our remote sensing and smog check data do not record evaporative emissions, we do not analyze them. Pollution control systems for evaporative emissions are separate from control systems for exhaust emissions.

Engine Families. An "engine family" describes the exact configuration of engine and abatement technology in a vehicle, and does not map one-to-one to make and model. Each vehicle also has an "evaporative family" describing the abatement technologies used to control evaporative emissions in the vehicle. If an engine family violates exhaust standards, vehicles with that engine family are recalled. Between model years 1990 and 2015, the number of engine families in a given year ranged from about 200 to 700, and the average model year had over 400 different engine families. Although the precise definition of a "model" depends on how different trims and extensions are included, between model years 1990 and 2015, the number of manufacturer×model pairs in a model year ranged from 150 to 400, with over 200 in the average model year. So on average, each model has two different engine families.

Other Pollutants. Exhaust standards use several different measures of HC. Tier 0 regulated total HC. Tier 1 added limits on non-methane HC, since methane is a HC that does not easily form ozone pollution, and since a key reason to regulate HC is to decrease ozone. The National Low Emissions Vehicle Program (NLEV), Tier 2, and Tier 3 regulate non-methane organic gases (NMOG), which includes non-methane HC emb compounds – alcohols and aldehydes – which form ozone but which the traditional method of measuring HC excludes. Based on mean levels, non-methane HC are 90 percent of total HC and non-methane organic gases are 94 percent of total HC (Mondt 2000; U.S. EPA 2003). NLEV, Tier 2, and Tier 3 added restrictions on emissions of particulate matter and formaldehyde, which we do not analyze since most of our data do not measure them.

Fleet-Wide Average. In 1999, eight Northeast states voluntarily applied tighter standards, called the National Low Emissions Vehicle (NLEV) program. Other states joined in 2001. NLEV, Tier 2, and Tier 3 limit the sales-weighted fleet-wide mean of each manufacturers' vehicles' emissions. These standards limit different pollutants— NLEV limits mean NMOG; Tier 2 limits fleet mean NO_x; and Tier 3 limits fleet mean NO_x plus NMOG (Table 1).

If a fleet-wide standard limits only one pollutant, why do auto manufacturers limit all pollutants? The answer reflects how regulators calculate the fleet-wide average. The EPA first certifies the CO, HC, and NO_x emission rates for each engine family. NLEV, Tier 2, and Tier 3 then define several bins. Each bin specifies the maximum emission rate for each pollutant that a vehicle in that bin may emit. For example, under Tier 2, a vehicle may only qualify for bin 2 if its NMOG emission rate is below 0.01, if its CO rate is below 2.1, and if its NO_x rate is below 0.02. The EPA calculates the fleet-wide average based on the threshold values of the bin in which a vehicle is categorized, not over the certified emissions level for an individual vehicle (Federal Register 1995, p. 52748). Under these regulations, for each pollutant, the standard of the highest bin is the maximum standard for the regulation. This is the value shown in Table 1.

Mileage and Age Values for Standards. Each exhaust standard specifies regulated mileage and age levels. Standards refer to these values as "intermediate life" or "full useful life." These values are used for calculating deterioration factors, conducting in-use tests, and determining recalls. In-use tests exclude vehicles with broken emissions control systems, though systematic failure of such systems can justify recalls.

The exact mileage and age values differ across standards. Under Tier 0, cars only faced standards for 10 years or 50,000 miles (whichever comes first), and trucks only faced standards for 10 years or 100,000 miles. Under Tier 1 and NLEV, cars and trucks faced an intermediate life standard at 50,000 miles and a full useful life standard at 100,000 miles. The full useful life standards were twenty to fifty percent higher than the intermediate life standards in order to reflect the greater mileage.

Some Tier 2 bins impose intermediate standards at age 5 years or 50,000 miles. All bins face standards at the full useful life. By default, the Tier 2 full useful life applies to 10 years or 120,000 miles. Tier 2 gives manufacturers the option to be exempt from the intermediate useful life standards, but then for the full useful life to apply at 150,000 miles. Under Tier 3, all cars and trucks face standards at 15 years or 150,000 miles, whichever comes first.

Defining Categories of Trucks. Many exhaust standards differ by vehicle type. Light-duty cars and trucks include vehicles with gross vehicle weight below 8,500 pounds and payload capacity up to 4,000 pounds. The difference between car and truck depends on the vehicle's purpose and design. Sport utility vehicles, minivans, and passenger trucks qualify as trucks. Several standards distinguish categories of trucks, sometimes called LDT1, LDT2, LDT3, and LDT4, based on a vehicle's curb weight or gross vehicle weight.[1]

This paper focuses on US cars and trucks. Heavy-duty vehicles and other countries have their own exhaust standards. Tier 2 began regulating medium-duty passenger vehicles, which have a gross weight value rating of 8,500 pounds or greater. To increase comparability over time and across data sets, we exclude medium-duty vehicles from the analysis.

California Exhaust Standards. The Clean Air Act allows the California Air Resources Board (CARB) to set its own exhaust standards and allows other states to adopt California's standards. Between 2004 and 2014, thirteen other states adopted California standards.[2] They are sometimes known as "Section 177 states," since Section 177 of the Clean Air Act

[1] Curb weight is the weight of a vehicle with standard equipment but no passengers or cargo. Gross vehicle weight is a vehicle's weight with standard equipment, passengers, and cargo.

[2] The thirteen states are Connecticut, Delaware, Maine, Maryland, Massachusetts, New Jersey, New Mexico, New York, Oregon, Pennsylvania, Rhode Island, Vermont, and Washington.

allows other states to adopt California's standards, or "CARB states."

Apart from the analysis of Tier 0, we do not separately analyze standards in California due to data limitations. Appendix A.2 discusses the lower quality of used vehicle emission rate data for California. Used vehicles driven in California may also be either vehicles originally certified to California standards, or vehicles originally certified to federal standards but then imported into California. Vehicle Identification Numbers do not always identify whether a vehicle was originally certified to California or federal standards. Additionally, new vehicles may be certified to either or both federal and California standards.

International Standards. To assess the global relevance of this paper, we reviewed exhaust standards for the 20 largest countries measured by GDP or population (30 countries total). All the industrialized and middle-income countries in this group apply US, EU, or national exhaust standards. Only three of the thirty countries, all low-income, appear to have no standards: Bangladesh, Ethiopia, and the Democratic Republic of Congo. Globally, some low-income countries have exhaust standards, though many do not. Some countries impose standards with a lag. For example, Europe currently uses Euro 6 standards, while Nigeria, Pakistan, and the Philippines all impose earlier versions of the European standards.

A.2 Test Details

The FTP test also goes by the acronyms FTP75, EPA75, and CVS75. It is also used to determine a vehicle's official city fuel economy (California Air Resources Board 2002; Thompson et al. 2014). The mechanics of an FTP test are fairly straightforward. Levels of each pollutant are measured through a hose attached to the vehicle's exhaust pipe. The vehicle is tested on a chassis dynamometer, with mean speed of 30 miles per hour and maximum speed of 57 miles per hour.

Before an FTP test, a vehicle is shut off for 24 hours, so the first part of an FTP test measures emissions from an unheated engine (a "cold start"). The cold start is increasingly important. Catalysts are most effective once heated to several hundred degrees Fahrenheit, which takes time after a vehicle is turned on.

In-use tests are FTP tests conducted 5-15 years after a vehicle is manufactured, to assess compliance with exhaust standards. Two government agencies conduct in-use tests—the EPA, to determine compliance with federal standards; and CARB, to determine compliance with California standards. To obtain samples of vehicles for in-use tests, the EPA randomly chooses vehicle owners, emails them a letter requesting to use their vehicle for a specified amount of time, and offers compensation.

Regulators have added three other new vehicle tests for highway driving, aggressive driving, and air conditioning. The three new tests are called "Supplemental FTP" tests. Tier 1 in 1994 added a "cold CO" test conducted at 20°F. These are not applied across most of our sample and are less comparable with our used vehicle data, so we do not analyze them.

The text mentions that Colorado has higher-quality used vehicle emissions data than other states. Most states use on-board diagnostic (OBD) tests, in which a monitor is connected to a vehicle's computer that checks the function of its control systems, but does not actually measure the physical concentration of exhaust emissions. Besides Colorado, only California has required exhaust tests for most vehicles over most of the period 2000-2020, including vehicles manufactured after model year 1996, when vehicle OBD systems became mandatory. California uses the Acceleration Simulation Mode, in which an engine is operated at a constant speed (e.g., 2500 rotations per minute). This test misses acceleration, deceleration, and breaking. California's test is also reported in a different scale than FTP or IM240 (parts per million rather than grams per mile), decreasing comparability. Available methods for comparing between these scales require several calibrated parameters and nonlinear equations, which introduce measurement error.

A.3 Abatement Technologies

Several innovations since the 1970s have increased catalytic converters' effectiveness (Palucka 2004). One involves maintaining a more precise ratio of fuel to oxygen in the engine, which has been accomplished using carburetors (1970s), single-point fuel injection (1980s), multi-point fuel injection (1990s), sequential fuel injection (2000s), and direct fuel injection (2010s).

A second innovation makes catalytic converters heat more quickly when a vehicle starts, since catalysts must have temperatures above 500°F to function fully. A third set of innovations has applied precious metals onto increasingly thin and durable materials, including to spherical beads (1970s), honeycomb-like ceramic materials (1980s), roughly-textured materials (1990s), and most recently using precision manufacturing techniques to apply layers of precious metal in ways that prevent the metals from agglomerating over time. Increasing the volume and mass of the precious metals also improves abatement.

One counterproductive innovation has been the sale of after-market defeat devices. These are illegal but appear to be more common for diesel pickup trucks bypassing diesel particulate filters. The EPA increased enforcement against these after-market devices in 2020.

In the paper, we model the cost of reducing emissions at a given point in time as an increase in marginal cost. This is consistent with costs being due primarily to increases in the quantities of precious metals, which is a good depiction of the Tier 2 era. That said, there certainly are fixed costs, but we note that these are not tied to a single model and thus would be unlikely to affect product choice or entry/exit of particular models. Rather, the fixed costs are at the firm, or even industry level, because when engineers figure out better catalytic converter designs, the innovations can be deployed broadly across all vehicles that share a manufacturing and design platform.

B Data for Reduced Form Analysis (Section 3)

B.1 Comparability of New and Used Vehicle Tests

Several studies and reviews provided a new vehicle FTP test and a used vehicle IM240 test to the same vehicle on the same day or otherwise very close in time (U.S. EPA 1992; CARB 1993; Kelly 1994; U.S. EPA 1995; Bishop et al. 1996; Sierra Research 1997; AIR 1999). Each study reports the relationship between FTP and IM240 emission rates, separately for the three main pollutants (CO, HC, and NO_x), both as a regression coefficient (with standard error) and R-squared. Most of the individual studies use around 100 vehicles, though in total the studies we reviewed include about 2,000 vehicles. Pooling across studies and weighting each by its sample size, the coefficient and R-squared for comparing FTP versus IM240 are 0.97 and 0.73 for CO; 1.23 and 0.86 for HC; and 0.82 and 0.82 for NO_x.

B.2 New Vehicle Emissions Data

In most years, the new vehicle emissions data list vehicle types tested, plus other vehicle types with the same engine and abatement technology but which are not directly tested. Because the EPA certifies emission rates and deterioration factors to be the same within an engine family×emission control system, we link the certification levels for all vehicles to the mean reporting certification levels within the same group.

Different years' data distinguish these groups with slightly different criteria. In model years 1971 to 2000, vehicles are grouped by manufacturer. In model years 1971-1977, vehicles are also grouped by engine family. In model years 1978-1984, vehicles are further subdivided by evaporative family. In model years 1985-1995, vehicles are divided by engine family×emission control system combination. In model years 1996-2000, vehicles are also divided by evaporative family×emission control system combination. Beginning in 2001, manufacturers group engine, fuel, and abatement technology into "durability" and "test" groups. "Engine family" is here synonymous with "test group." When possible, we create separate observations for California- and federal-certified test results for each vehicle, though this categorization changes over years. Model names listed within groups sometimes contain unabbreviated test vehicle model names, in which case we impute a duplicate observation for a model. Imputed models are assigned the other characteristics (e.g., engine displacement and horsepower) equal to the mean of those of the test vehicles in their groups.

The data include several test procedure categories that are minor variants of the FTP (e.g., California-certified versus standard gasoline, or a test performed at a specified temperature versus ambient temperature).[3] For years when the test procedure is listed, we further separate our imputed emissions data by procedure. A small share of tests report non-methane hydrocarbon and non-methane organic gases; we convert these to total hydrocarbons using the conversion ratios discussed in the main text. If a vehicle does not report a certification level for 50,000 miles, we impute it as the reported certification level for another distance times the median ratio of that certification level to the 50,000 certification level, where the median is calculated from all vehicles reporting both certification levels.

The new vehicle and associated data report many vehicle attributes. Unfortunately, however, most innovations discussed in Appendix A.3 that improved the performance of catalytic converters, like rapid catalyst heating and catalyst application methods, are not reported in any systematic vehicle-level data that we know exists. Most of the new vehicle data report whether a vehicle is a car or truck. We generally group vehicle class by these categories, though some analyses use these further sub-categories. For model years before 1988, we estimate the truck type of a vehicle using test weight thresholds. For model years 1988 and beyond, we determine truck types by linking it to Colorado smog check data. We

[3]We include the test procedures CVS 75 and later (without canister load) plus Federal or California 2 or 3 day test procedures. Formally, this includes EPA test procedure numbers 2, 21, 25, 31, 35, 51, and 52. This is only relevant after model year 1998, when other procedures (such as the Supplemental Federal Test Procedures) appear.

match across these data using make, model, drive type, trim, displacement, horsepower, and similar variables.

Model years 1972-4 used an earlier design of the FTP test (FTP72). We inflate emissions for these years by the mean ratio of the standard FTP test to values for the earlier (FTP72) test (AES 1973). Model year 1971 used a different version of the FTP test (FTP71), which we do not have a way to concord to the alternative versions of the test (FTP72 and FTP75), so we exclude the 1971 new vehicle emissions data.

The FTP data for test years 1994 and 1995 reports raw test results but not deterioration factors or certification levels. They also have a smaller sample than surrounding years, different mean raw test results, and a disclaimer that the EPA is "not able to provide the normal report format." Hence, we largely exclude new vehicle data from these years.

The five cities covered in the older 1957-1971 model year data are Chicago, Houston, Los Angeles, St. Louis, and Washington, DC. We use the data in the AES national sample.

In several parts of the paper, to convert miles per gallon data to grams of CO_2 per mile, we use the standard emission rate of 19.37 pounds CO_2 per gallon gasoline from the Energy Information Agency, and the conversion rate of 453 grams per pound.

B.3 Used Vehicle Emissions Data: Colorado Smog Check

We obtain these data from the Colorado Department of Public Health and the Environment.

The analysis sample imposes several restrictions. We exclude observations with missing odometer, missing values for CO, HC, or NO_x emissions, or a vehicle identification number (VIN) that is not the standard 17 digits. We clean reported odometer readings following Knittel and Sandler (2018). We winsorize pollution readings at the 99.9th percentile.

A vehicle which appears to be especially clean in the first part of an IM240 test is allowed to complete the test before the full 240 second test is complete, a process Colorado calls "Fast Pass." The Colorado data then report an imputed value for the emission rate that Colorado regulators estimate would have been recorded in a complete 240-second test. In recent years, a randomly-chosen set of vehicles are required to complete the full 240 second test, a sub-sample we use in estimates with recent years of data.

We exclude some Colorado data with lower quality or limited comparability. Colorado vehicles model year 1981 and earlier undergo a low-quality test (two-speed idle), which we do not analyze. Colorado also provided smog check data for calendar years 1995-6 and 2015-6, but we do not use them. The 1995-6 data appear to use different methods for measuring CO. In 2015-6, vehicles aged 4 through 9 years are exempt from tests. Colorado's current contractor began managing the program in 1995, which is the first year Colorado began using a more stringent ("enhanced") smog check program (Air Pollution Control Division 2013). The measure of annual externalities in Figure 8 winsorizes the annual externality at the 99.9th percentile.

In these and other used vehicle data, we define a vehicle's age as the year when pollution was measured (i.e., its test year) minus its model year.

How representative are the Colorado data? They come from counties with similar driving and emissions patterns to other polluted urban counties. All Colorado counties with smog check data are in "nonattainment" Clean Air Act status for ozone pollution, meaning they have high ambient ozone levels and strict regulations, including this vehicle smog check program. Out of 3,000 US counties, only 265 were in ozone nonattainment in the year 2014, though those 265 counties account for 44 percent of the US population. Compared to other ozone nonattainment counties, these Colorado counties have moderately higher vehicle NO_x and VOC emissions per square mile, but lower vehicle miles traveled (VMT) and population density.[4] Statistically, these Colorado counties are indistinguishable from other ozone nonattainment counties along each of these dimensions individually, and marginally indistinguishable in a joint test (F-statistic of 1.98, p-value of 0.099). Perhaps unsurprisingly, compared to attainment counties, which are cleaner and less urban, the Colorado smog check counties have significantly higher emissions, driving, and population density.

Colorado requires tests of vehicles registered to addresses in nine counties in and north of Denver—Boulder, Broomfield, Denver, Douglas, and Jefferson counties, and some parts of Adams, Arapahoe, Larimer, and Weld counties. Colorado began testing vehicles registered in the Northern Colorado counties of Larimer and Weld only in November 2010.

B.4 Used Vehicle Emissions Data: Remote Sensing

Colorado's remote sensing program, called RapidScreen or CleanScreen, began in calendar year 2004 and is managed by Colorado's Department of Public Health and the Environment (Hawkins et al. 2010; Opus Inspection 2016; Klausmeier 2017). Colorado's remote sensing records cover calendar years 2009 to 2016 and include over 50 million observations. If a vehicle receives two or more clean remote sensing readings, it is exempted from the standard smog check test. About a third of Colorado vehicles are thereby exempted from smog check tests. We include estimates correcting for potential selection caused by the exempt vehicles.

The Colorado data include each vehicle's VIN, which the remote sensing system identifies by photographing a vehicle's license plate and using administrative records to link to the VIN. The data report CO and CO_2 in percentages and HC and NO_x in parts per million (ppm). We convert these values to grams of pollution per mile using average conversion rates from Bishop and Haugen (2018), Table 3, and EPA fuel economy ratings.

We also use several additional remote sensing samples collected from the Fuel Efficiency Automobile Test (FEAT) (Bishop et al. 1989). We obtain measurements from the FEAT Reports data (`http://www.feat.biochem.du.edu/light_duty_vehicles.html`, accessed on 3/2/2017 for US data and 12/8/2020 for multi-country and heavy duty truck data). FEAT emits an infrared beam from a device on one side of the road, which is then read by a receiver on the opposite side of the road. In the leading method, an infrared beam detects CO, CO_2, and HC, and an ultraviolet beam detects NO_x (Bishop and Haugen 2018). Some detectors measure only nitric oxide, a component of NO_x, which we include with NO_x data for comparability.

We report sensitivity analyses using a multi-state remote sensing sample that includes many FEAT collection events.[5] Appendix B.4 describes details. A collection event is a city where researchers collected data in a particular year. FEAT's data on vehicles come from devices located ten inches above ground, so they measure emissions from light-duty

[4]Comparisons in this paragraph use data from the year 2014, a year for which many of these data are available, using data from the EPA's National Emissions Inventory.

[5]The multi-state remote sensing sample includes data from Arizona, California, Illinois, Maryland, Nebraska, Nevada, Oklahoma, Pennsylvania, Texas, Utah, and Washington.

vehicles and light-duty trucks but not heavy-duty trucks (which have higher tailpipes). In analyses of control system deterioration, we exclude remote sensing observations of very old age categories with less than 50 observations per age category.

We also report patterns of vehicle emissions, age, and deterioration from a multi-country remote sensing sample using FEAT data from Monterrey, Mexico; Auckland, New Zealand; Rotterdam, The Netherlands; Toronto, Canada; Melbourne, Australia; and Milan, Italy. This represents all FEAT data from countries which include information on the vehicle's model year, which is needed to measure the vehicle's age.

Most remote sensing data measure each pollutant as the percent of total gas. For comparisons with FTP or IM240 values, we convert these units to grams per mile, using conversion rates from Bishop and Haugen (2018). A reasonable share of raw remote sensing data have negative values (e.g., due to measurement error, they resemble a normal distribution around a small number, which has some mass below zero), and correspondingly a reasonable share of the translated data have negative grams/mile values. To avoid excluding these values from analysis, we generally work with remote sensing data in inverse hypersine values, rather than in logs. In either raw percent or transformed, we winsorize the remote sensing data at the 0.5th and 99.5th percentile to address outliers.

Finally, our sensitivity analysis uses measurements of emissions from heavy duty trucks collected from the Port of Los Angeles and from a highway weigh station in Northern California from the On-Road Heavy-Duty Vehicle Emissions Monitoring System (OHMS) (Bishop et al. 2015). OHMS is a tent with a collection pipe where heavy-duty trucks drive slowly, then sensors process the exhaust plume.

Appendix Table A1 compares remote sensing to Colorado smog check data. In 65,000 cases, we observe an individual vehicle (a 17-digit Vehicle Identification Number) in the remote sensing data, and then observe the same vehicle in the smog check data the following week. We allow a one-week lag between the data to avoid the possibility that a smog check caused a vehicle to be repaired, which could make the remote sensing value differ from the smog check reading. If the remote sensing and smog check data gave identical results, these matched pairs of observations would have the same value.

Appendix Table A1 finds that remote sensing and smog check values are very strongly correlated. The t statistics from regressing one measure on the other range from 8 to over 100. In this sense, remote sensing strongly co-moves with smog check tests.

At the same time, the units have different scales. Panel A regresses the remote sensing reading on the smog check reading for the same pollutant. The correlation between the two readings in inverse hypersines is 0.10 (0.003) for CO, though is 0.53 and 2.98 for HC and NO_x. Panel B shows that when we reverse the variables, so smog check is the dependent variable and remote sensing the explanatory variable, these correlations range from 0.01 to 0.16. These values are all far from one. If we analyze these relationships in levels (g/mile) rather than inverse hypersine, these relationships become even further from one. Ultimately, these comparisons suggest that remote sensing data are strongly associated with smog check inspection data, but comparing units between the two types of measurement is difficult.

B.5 Used Vehicle Emissions Data: In-Use Tests

Our "in-use" test data cover model years 2004 through 2014, were conducted in calendar years 2008 through 2017, cover vehicles 0 to 6 years old, and include about 10,000 observations. We obtain these data from the California Air Resources Board (CARB). We keep observations which we can match to fuel economy data, and set the fuel economy of each vehicle in a test group equal to the test group mean. We winsorize the pollution values at the 99.9th percentile. While we only use these in one sensitivity analysis, they may represent a combination of vehicles certified to California and federal standards, so could have more measurement error than other data.

B.6 Emissions from Manufacturing

The quantitative model incorporates estimates of the pollution from manufacturing vehicles. We calculate these rates using input-output tables, following Lyubich et al. (2018).

We use the 2002 benchmark tables after redefinitions from the Bureau of Economic Analysis (BEA) (U.S. Bureau of Economic Analysis 2020).[6] We utilize three BEA tables: the make table, use table, and import matrix. The use table describes the dollar inputs of each commodity, including imports, required to produce a dollar output of each industry. We subtract imports from the use table to produce a domestic use table, describing the domestic inputs required to produce an output. We combine the make and use tables to produce an input-output table. We then calculate the Leontief Inverse, equal to $(I - A)^{-1}$, where I is the identity matrix and A is the input-output table. This Leontief Inverse represents the dollars of domestic inputs required to produce a dollar of output in each industry, including direct inputs to an industry, inputs to that input, inputs-to-inputs-to-inputs, etc., including the entire value chain (also sometimes called the entire life cycle, supply chain, or footprint).

To measure emissions from each input industry, we use data from the 2002 National Emissions Inventory (U.S. EPA 2014b). We calculate emissions from each NAICS industry, and concord this to BEA industries using a concordance file for the year 2002 from the BEA. We measure the emissions per dollar of output for each industry by dividing the NEI industry emissions totals by the gross output of each industry from the input-output files.

To measure emissions from each output industry, we multiply the emission rate of each input industry by its input share in the Leontief Inverse. Summing these values across all input industries gives a measure of the tons of pollution emitted per dollar of output in each industry. Multiplying this by the gross output of each industry gives an estimate of the short tons of each pollutant emitted to produce the entire year 2002 output of each industry, including emissions from the entire domestic value chain of each industry. We focus on these values for two industries—light-duty vehicles (NAICS and BEA code 336111) and light trucks and utility vehicles (NAICS and BEA code 336112).

To measure emissions per vehicle manufactured, we link these emissions data to the number of vehicles manufactured in the year 2002, obtained from the US. Federal Reserve (Federal Reserve Bank of St. Louis 2021). Dividing an industry's total emissions by the

[6]The BEA publishes many versions of the table; the version "after redefinitions" has processing to improve the quality of the relationship between input use and product categories, and this version is commonly used for life cycle analysis and calculating the Leontief Inverse.

number of vehicles manufactured gives an estimate of the tons of pollution emitted per vehicle manufactured. We multiply emissions of each pollutant by the damage rates used in the rest of the paper (in \$2019) to measure pollution damages per vehicle manufactured (details in Section F.1 below). This calculation obtains an estimate of \$605 in damages per light-duty vehicle manufactured and \$595 in damages per light-duty truck manufactured. We summarize this as an estimate of \$600 in pollution damages per vehicle manufactured.

An alternative is to use engineering calculations of the emissions required for different materials used to produce a vehicle. The GREET Model, managed by Argonne National Laboratory (part of the U.S. Department of Energy), is probably the leading such engineering model. Using GREET, we calculate damages of \$827 per vehicle in the year 2019 (\$2019). While this number is not identical to the \$600 value we obtain from the input-output table, given the numerous differences between engineering models and input-output tables, it is notable that the two approaches give estimates with the same order of magnitude.[7]

The quantitative model also uses an estimate of the trend in emissions per vehicle manufactured. We calculate this trend by comparing emissions and real revenue from US industry between 2002 and 2017, then measuring the two-year time step. We use these years since the National Emissions Inventory occurs every few years and since the economic census takes place in years ending in 2 and 7. To measure these trends, we obtain data on industrial emissions from the National Emissions Inventory. We measure the 2002-2017 trend for each of the three pollutants on which we focus (CO, NO_x, HC). We weight the trend across pollutants by each pollutant's marginal damages. We measure the trend in the value of sales from industry (defined here to include manufacturing, utilities, and mining) from the Economic Census, and deflate it by the GDP deflator. The resulting ratio of industrial pollution in 2017 to 2002 is 0.5694, and of real industrial output is 1.0695. Finally, we calculate the implied decrease in pollution per dollar output per two-year time period in the model as 8.06 percent, calculated from the equation $(0.5694/1.0695)^{(1/(15/2))} - 1$.

B.7 Other Data

We obtain data on the attributes of each VIN prefix using a set of files purchased from an industry vendor, typically called a VIN decoder. These files indicate the fuel economy, retail price, weight, horsepower, torque, and unique engine identifier associated with each VIN prefix, among other characteristics.[8] For model years 2000 and later, the VIN decoder identifies the engine families for each VIN prefix, which is distinct from and does not map

[7] Specifically, GREET and our calculation using input-output tables reflect many differences. GREET describes emissions in 2019, while the input-output table reflects emissions in the year 2002. The input-output table includes all inputs while GREET includes only the most important inputs. The input-output table includes the entire supply chain while GREET focuses only on a few steps down the chain. The input-output calculation uses observed emissions for the year 2002, while GREET uses a combination of modeled engineering emission factors from different years and observed emissions from the National Emissions Inventory. GREET performs most calculations in physical units, while input-output tables perform most calculations in monetary units. GREET calculates emissions from vehicle scrap and manufacturing, while input-output tables calculate emissions from manufacturing only, and include scrap only to the extent that it occurs in the value chain supporting vehicle manufacturing.

[8] A VIN prefix is the first eight digits of the VIN plus digits ten and eleven. This uniquely identifies the manufacturer, vehicle attributes, model year, and code of the plant that manufactured it.

one-to-one with the vendor's proprietary engine identifier. We link these engine families to new vehicle FTP test results. We calculate the new vehicle FTP emissions for each VIN prefix as the mean for all engine families and tests linked to it.

We also need to identify the exhaust standards that apply to vehicles in most of these datasets. Colorado's smog check data identify the class of each vehicle (car or weight categories of trucks). For other data, we link each vehicle's VIN prefix to the Colorado smog check data to determine each vehicle's class.[9]

The quantitative model calculates the environmental externality from each vehicle by year. To measure a vehicle's emissions per mile, we use its Colorado smog check results. To calculate annual miles traveled, we use the change in odometer readings since the previous smog check divided by decimal years elapsed or, for a vehicle's first test, since its model year. To measure damages, we use county-specific estimates of the marginal damages of NO_x and VOC emissions from ground level sources, estimated from the AP3 model (Tschofen et al. 2019).[10] They exclude CO, so we take a national value of CO marginal damages from Matthews and Lave (2000). We use the Bureau of Labor Statistics Consumer Price Index for urban consumers to express all currency values in 2019 real dollars.

In the sensitivity analysis using selection correction models, we use Colorado vehicle registration data, which we obtained for calendar years 2005-2013. We merge this data with Colorado smog check and remote sensing results that are usable for each registration date. Smog check results can be used for registration for up to 24 months, while remote sensing results are valid for up to 12 months. [11]

Some regressions use additional data to control for potential confounding variables. We measure gasoline prices per million British Thermal Units (BTU) from the State Energy Database System (SEDS), then convert to price per gallon of gasoline using annual data on BTU per barrel from the Energy Information Administration's Monthly Energy Review. We measure the ethanol share from SEDS as fuel ethanol (excluding denaturant) for transportation divided by the sum of ethanol and motor gasoline, all measured in BTUs. We measure the annual sulfur content of fuel as sulfur dioxide emissions from highway vehicles, according to the EPA's summaries of the National Emissions Inventory (U.S. EPA, 2014b), divided by the FHWA's Highway Statistics report of highway use of gasoline by state×year in gallons. The gasoline price, ethanol, and sulfur data represent the calendar year of the test.

C Additional Empirical Results: Trends (Section 4)

C.1 Fleet-Weighted New Vehicle Emission Trends

Figure 1 in the main text shows mean emissions of all new US vehicles, averaged across vehicle types. For model years 2000-2015, where we have a concordance between emissions data identifier codes and VIN prefixes, those graphs also show fleet-weighted emission rates,

[9] In the Colorado data, the regulatory class for a VIN prefix varies infrequently across individual vehicles. We take the modal regulatory class for a VIN prefix. Some of our analyses compare between car and truck rather than across weight categories of truck, since the car/truck distinction is more consistently measured.

[10] They do not report values for Broomfield County, Colorado, which was created from four other Colorado counties; we calculate this value as the mean of the other four counties it was created from.

[11] 5 CCR 1001-13, Reg 11, Motor Vehicle Emissions Inspection Program

where the weights are the share of each VIN prefix in the Colorado remote sensing data. The CO_2 data underlying Figure 1 are the weighted average across models, where weights equal the sales of models.

The air pollution microdata underlying Figure 1 report the emissions rate for an "engine family" or similar aggregate; Appendix A.1 provides additional background on engine families. Engine families have complex links with vehicle models or Vehicle Identification Number (VIN) prefixes. Some models and VIN prefixes have many different possible engine families, and some engine families correspond to many different possible VIN prefixes, and for model years before 2000 we are not aware of any concordances linking these different identifiers. The similarity of the weighted and unweighted series in Figure 1 in years after 2000 provides one piece of evidence that weighting does not substantially change these trends.

We spent considerable effort constructing our own concordance for years before 2000 between the new vehicle emissions data and Wards Automotive Yearbooks sales data. It is difficult to construct this link accurately. Vehicle types in the new vehicle sales data and the emissions have different descriptions. An identification code in the emissions data typically requires a many-to-many match with vehicle types in the sales data. The differences occur for many reasons. For example, the data may refer to different vehicle trims of a given make and model without any clear relationship; vehicles with the same underlying engineering (sometimes called "badge engineered") may have different model names across datasets; vehicles with multiple fractional corporate parents (e.g., a brand might have multiple owners, which change over time) may be listed with different make; and the distinction between model and trim is fuzzy and differs across datasets and years.

Despite these strong caveats, using our effort at constructing this concordance, Appendix Figure A1 shows fleet-weighted averages for new vehicle emissions over the model years 1981 to 2015. The average is weighted by Wards sales data. We identified the emission rate for 50 to 80 percent of sold vehicles in a given model year. The sales-weighted data show more year-to-year volatility than the unweighted data, which partly reflects differences in match rates across model years. At the same time, the weighted and unweighted series have extremely similar patterns overall, and in most years are within a few percent of each other. Broadly, these results echo those of the main text and suggest that weighted and unweighted fleet averages have similar levels and trends.

C.2 Used Vehicle Emission Trends

Section 4 of the main text summarizes trends in emission rates from used vehicles. This appendix provides details.

Estimating emissions trends for used vehicles involves some challenges. The evolution of emission rates with age may vary by model year, which makes it complex to distinguish age from model year, even if the sample includes age fixed effects. The used vehicle data begin in model year 1982 since vehicles from before model year 1982 are exempt from the higher-quality (IM240) emission test. We report an additional trend using only the vehicles ages 4 to 6 years old and with odometer between 40,000 and 60,000 miles, which are more comparable to the new vehicle data. We also separately report the 25th percentile of emissions by model year, which helps address both the effects of outliers and Colorado's remote sensing policies (CleanScreen). If CleanScreen exempts a third of vehicles from inspections, then the median

Colorado vehicle would be about the 25th percentile of Colorado inspections.

Appendix Figure A2 shows that used vehicles followed qualitatively similar patterns to new vehicles in Figure 1. Mean used vehicle emission rates for each air pollutant fell by 90 percent between model years 1982 and 2010; new vehicle emissions from Figure 1 fell by similar amounts. Trends for vehicles ages 4-6, or for the 25th percentile of emissions, are similar, though levels are lower in all model years. The used vehicle data suggest that CO_2 emission rates actually increased slightly. Used vehicles also suggest some patterns coincident with exhaust standards, for example, a steeper slope in 1995-1996 when Tier 1 became binding, and a flatter slope after 2007. In the 1980s and 1990s (though less in the late 2000s), the used vehicle data show more of a steady downward trend in years when exhaust standards did not change. This time series makes it unclear whether this is due to confounding of age, model year, and test year effects, to measurement error, or to true improvements in abatement technology and its durability.

Comparing Table 1 and Appendix Figure A2 also shows that in model years before 1995, average used vehicle emission rates are close to standards, and in some years above them. In model year 1990, for example, Table 1 shows that the standards for CO were 3.4 and 10 for cars and trucks, while Appendix Figure A2 shows that the associated mean emissions for all vehicles was around 10. Used vehicle smog check tests include vehicles with broken emissions control systems, while standards and in-use tests exclude them. For example, although the model year 1990 exhaust standards for CO are 3.4 and 10, a model year 1990 vehicle can pass a Colorado smog check inspection with an emission rate up to 15. Nonetheless, these patterns are consistent with EPA and related reports from the late 1970s and 1980s showing mean emission rates of used vehicles that are close to or exceed the relevant exhaust standards in that earlier time period (Mills and White 1978; Jones 1980; Lorang 1984; Crandall et al. 1986; Manufacturers of Emissions Controls Association 1995).[12]

D Additional Empirical Results: Effects of Exhaust Standards on Emission Rates (Section 5)

This section discusses alternative estimates of how Tier 0, Tier 1, and Tier 2 exhaust standards have affected emission rates.

The beginning of Section 5.2 from the main text discusses graphs in Figure 2 showing new vehicle exhaust standards and new vehicle emission rates spanning Tier 0, Tier 1, and Tier 2. Appendix Figure A3 shows versions of these graphs for smog check and remote sensing data. Those used vehicle data suggest qualitatively similar patterns, though with smoother adjustment potentially in part due to the greater measurement error in used vehicle tests and the complication of separating test year, model year, and age effects.

For the analysis of Tier 1, Appendix Table A3 shows estimates using different specifications—excluding the phase-in model years 1994 and 1995, specifying the dependent variable as

[12] In the largest such study, which included 2,000 FTP tests on vehicles from model years 1975-1980, over half of vehicles in all cities outside California violated the relevant federal standard for one or more pollutants (U.S. EPA 1980). In this time period, carburetors designed to abate CO led to rough idling, and adjustments by mechanics, dealers, or owners to address the rough idling dramatically increased CO (UPI 1976).

emissions per gallon (one way to control for fuel economy standards), distinguishing separate categories of trucks, accounting for selection into the sample due to remote sensing, and directly estimating effects on remote sensing emissions from both the Colorado and the multi-state remote sensing samples. All these estimates are precise. Most magnitudes of the estimates using the smog check samples in columns (1) through (6) are between 0.5 and 1.0.

Appendix Table A3, columns (5) and (6), shows estimates that address selection into inspection tests due to Colorado's remote sensing program. In recent calendar years, a third of vehicles receive clean remote sensing readings (CleanScreen) and are exempted from smog check tests. Column (5) reports OLS estimates where each observation is a vehicle that registered in the calendar years for which we have Colorado state registration data. The dependent variable is the mean smog check test reading for a registered vehicle which has an associated smog check reading for that registration. Column (6) reports a Heckman selection model using the sample of column (5), and including vehicles that received a CleanScreen pass but did not have smog check tests. As an instrument in the selection equation, we use the number of times a vehicle passed the CleanScreen road-side monitoring devices. This is arguably a good instrument for selection, since selection is based on whether a vehicle has two or more clean CleanScreen readings, and not on any kind of average. Hence, the more often a vehicle happens to pass the remote sensing devices, the more likely the vehicle is to pass CleanScreen and the less likely the vehicle undergoes a smog check inspection. The selection results in column (6) are close to the OLS results in column (5) using the same sample, which is one piece of evidence that CleanScreen selection does not bias our main estimates.

As noted earlier, the magnitudes for remote sensing are unfortunately not comparable to the numbers for the new and used vehicle smog check tests, but most remote sensing estimates are precise.

Appendix Table A4 analyzes mechanisms for exhaust standards to affect emission rates. Each table entry summarizes a separate regression corresponding to equation (2). Comparing Rows 1 and 2 shows that controlling for engine family fixed effects attenuates estimates by a third. This suggests that two thirds of the effects of standards on emission rates is within an engine family—auto manufacturers improved pollution abatement technology across model years while keeping the identical engine brand, stroke, etc. One third of the effects of standards on engines come from replacing dirtier with cleaner engines. Rows 3-8 show little evidence that standards affect vehicle attributes like horsepower or torque. We find some effects on vehicle prices, though they diminish in magnitude and precision with trend controls.

Section 5.2 from the main text analyzes how Tier 2 affects emission rates. Appendix Table A5 obtains qualitatively similar estimates from sensitivity analyses using in-use tests, Colorado remote sensing data, and the multi-state remote sensing sample. As discussed in Section 3.2, the units of the remote sensing tests are less comparable and obtain varying magnitudes, but the signs are in the expected direction and the remote sensing estimates are precise.

E Analytical Model: Proofs and Outside Good (Section 7)

E.1 Proofs

Proposition 1. We first demonstrate that an increase in ψ will lead to an increase in the equilibrium price of used vehicles p^*. This increase in equilibrium price is associated with an increase in equilibrium quantity of used vehicles because the change to ψ shifts the demand for used vehicles and affects a movement along the used vehicle supply curve.

To obtain the result, we use implicit differentiation on the equilibrium condition (equation (4)), using $w^* = \psi - p^* - H(p^*)(p^* - \bar{k})$ for brevity:

$$\frac{\partial}{\partial p^*}\left(\frac{H(p^*)}{1+H(p^*)}\right) dp^* = \frac{\partial}{\partial p^*} g(w^*)\frac{\partial w^*}{\partial p^*} dp^* + \frac{\partial}{\partial p^*} g(w^*)\frac{\partial w^*}{\partial \psi} d\psi \tag{F.1}$$

For the left-hand side of equation (F.1), see that:

$$\frac{\partial}{\partial p^*}\left(\frac{H(p^*)}{1+H(p^*)}\right) = \frac{h(p^*)(1+H(p^*)) - h(p^*)H(p^*)}{(1+H(p^*))^2} = \frac{h(p^*)}{(1+H(p^*))^2} > 0. \tag{F.2}$$

To analyze the right-hand side of equation (F.1), we first calculate intermediate results. The truncated mean of repair costs $\bar{k}$ depends on the used vehicle equilibrium price. Its derivative is:

$$\begin{aligned}\frac{\partial \bar{k}}{\partial p^*} &= \frac{\partial}{\partial p^*}\frac{\int_{-\infty}^{p^*} kh(k)dk}{H(p^*)} = \frac{p^* h(p^*)H(p^*) - h(p^*)\int_{-\infty}^{p^*} kh(k)dk}{H(p^*)^2} \\ &= \frac{h(p^*)}{H(p^*)}(p - \bar{k}).\end{aligned} \tag{F.3}$$

The derivative of the expected used vehicle resale value net of repair costs $H(p^*)(p^* - \bar{k})$ with respect to price is:

$$\begin{aligned}\frac{\partial}{\partial p^*} H(p^*)(p^* - \bar{k}) &= h(p^*)(p^* - \bar{k}) + H(p^*)\left(1 - \frac{d\bar{k}}{dp^*}\right) \\ &= h(p^*)(p^* - \bar{k}) + H(p^*)\left(1 - \frac{h(p^*)}{H(p^*)}(p^* - \bar{k})\right) \\ &= h(p^*)(p^* - \bar{k}) + H(p^*) - h(p^*)(p^* - \bar{k}) \\ &= H(p^*),\end{aligned} \tag{F.4}$$

The second line substitutes in equation (F.3).

These results allow us to calculate the derivative of w^* with respect to p^*:

$$\frac{\partial w^*}{\partial p^*} = \frac{\partial}{\partial p^*}(\psi - p^* - H(p^*)(p^* - \bar{k}(p^*))) = -1 - H(p^*).$$

Substituting equation (F.2) for the left-hand side of (F.1), substituting (F.4) into the right-hand side and noting that $\partial w^*/\partial\psi = 1$, yields the desired comparative static:

$$\frac{h(p^*)}{(1+H(p^*))^2}dp^* = -g(w^*)(1+H(p^*))dp^* + g(w^*)d\psi \tag{F.5}$$
$$\frac{dp^*}{d\psi} = \frac{1+H(p^*)}{\frac{h(p^*)}{g(w^*)(1+H(p^*))} + (1+H(p^*))^2} > 0. \tag{F.6}$$

The sign follows because the functions are all distributions and hence weakly positive.

We have now shown the first part of the result, p^* rises in ψ. The repair rate is $H(p^*)$ and the scrap rate is $1 - H(p^*)$. The repair rate rises in ψ because $H'(p) = h(p) > 0$. Conversely, the repair rate declines in ψ, at rate $-h(p)dp/d\psi$. The used vehicle market share U rises in ψ because $U = H(p)/(1+H(p))$, the derivative of which is positive, as shown in equation (F.2).

Proposition 2. Our steady state framework describes situations in which prices and scrap rates are constant over time. The proposition is specific about welfare in a time period because welfare may vary across time periods if the externalities change over time. To clarify this in the derivation here, we use t to denote the time period, allowing only the externality to potentially vary over time, consistent with the steady-state interpretation of the model.

Social welfare from vehicles W_t in a period t is private benefits from the new vehicle (the integral of w over those who choose a new vehicle), minus the cost of new vehicles (new vehicle market share times ψ), minus the cost of used vehicles (used vehicle market share times average repair costs, conditional on the vehicle being repaired), minus the externality (the used vehicle share times emissions of used vehicles plus the new vehicle share times emissions of new vehicles). In steady state, the stock of used vehicles is the same each period, so changes in the stock do not appear in the welfare expression.

We define social welfare along these lines as a function of w', the cutoff value above which an agent ends up with a new vehicle, and below which the agent ends up with a used vehicle. To describe the optimum, we assume the planner can directly choose w'. For any w', there is an implied repair rate that determines a cutoff repair cost below which all vehicles are repaired, denoted k'. Thus, w' determines k' and hence average repair costs $\bar{k}$. Specifically:

$$(1 - G(w'))H(k') = G(w').$$

Rearranging yields:

$$H(k') = \frac{G(w')}{1 - G(w')}$$
$$1 + H(k') = \frac{1}{1 - G(w')} \tag{F.7}$$

Implicit differentiation shows:

$$h(k')dk' = \frac{g(w')}{(1 - G(w'))^2}dw'$$
$$\frac{dk'}{dw'} = \frac{g(w')}{h(k')}\frac{1}{(1 - G(w'))^2} > 0. \tag{F.8}$$

The total repair costs of used vehicles can be written in two distinct but equivalent ways. Here we write it as the integral over the repair cost distribution from the minimum (0) up to the endogenously determined cutoff (k') times the size of the new vehicle market:

$$K = (1 - G(w')) \int_0^{k'} k dH(k).$$

Social welfare in period t is thus written:

$$\begin{aligned} W_t = & \int_{w'}^{\infty} w dG(w) - (1 - G(w'))\psi - (1 - G(w')) \int_0^{k'} k dH(k) \\ & - G(w')\phi_t^u - (1 - G(w')(\Phi + \phi_t^n). \end{aligned}$$

The first order condition is:

$$\begin{aligned} \frac{dW_t}{dw'} = & - g(w')w' + g(w')\psi + g(w') \int_0^{k'} k dH(k) - (1 - G(w'))k'h(k') \frac{dk'}{dw'} \\ & - g(w')\phi_t^u + g(w')(\Phi + \phi_t^n) = 0. \end{aligned} \tag{F.9}$$

Simplify the first-order condition (F.9) by dividing through by $g(w')$ and moving w' to the left-hand side and denoting the solution value of w' by w^s for "social" optimum:

$$w^s = \psi + \int_0^{k'} k dH(k) - \frac{(1 - G(w'))}{g(w')} k'h(k') \frac{dk'}{dw'} + \Phi + \phi_t^n - \phi_t^u. \tag{F.10}$$

Substitute equation (F.8) and use the definition of $\bar{k}$ to rewrite equation (F.10) as:

$$w^s = \psi + H(k')\bar{k} - \frac{(1 - G(w'))}{g(w')} k'h(k') \frac{g(w')}{h(k')} \frac{1}{(1 - G(w'))^2} + \Phi + \phi_t^n - \phi_t^u. \tag{F.11}$$

Simplify and substitute equation (F.7) to yield:

$$w^s = \psi + H(k')\bar{k} - (1 + H(k'))k' + \Phi + \phi_t^n - \phi_t^u. \tag{F.12}$$

Rearranging equation (F.12) yields:

$$w^s = \psi - k' - H(k')(k' - \bar{k}) + \Phi + \phi_t^n - \phi_t^u.$$

The private market outcome is described by a cutoff value w^*, with $w > w^*$ choosing a new vehicle and others choosing used, where $w^* = p - \psi + \tau - H(p)(p - \bar{k})$, and the cutoff repair cost k' is equal to p, which satisfies the equilibrium quantity condition.

Thus, when $\tau = \Phi + \phi_t^n - \phi_t^u$, the private market will solve:

$$\begin{aligned} w^* &= \psi + \tau - p - H(p)(p - \bar{k}) \\ &= \psi - p - H(p)(p - \bar{k}) + \Phi + \phi_t^n - \phi_t^u. \end{aligned}$$

The benchmark fee thus causes the private market to choose an equilibrium cutoff value w^* that equals w^s, the social optimum in period t. The remainder of the result states that moving towards this welfare maximizing point raises welfare, which is true by concavity of the welfare function.

E.2 Analytical Model with an Outside Good

The main model presented in the paper allows for no substitution to an outside good. This abstracts from some potential patterns of substitution. In this section, we describe and analyze a version of the analytical model that includes an outside good. Our main results persist in this version of the model.

Setup. Some changes are required to adapt the model to account for an outside good. The original model normalizes the utility of the used vehicle to zero. Here we normalize the utility of the outside good to zero.

We also describe a simple set of preferences to reflect this addition and maintain tractability. We assume that an agent whose utility of the new vehicle is w has utility from the used vehicle equal to $z(w) = zw$, where $0 < z < 1$. This is a strong assumption, but intuitive. The main implication is monotonic sorting—agents with the highest level of w purchase a new vehicle, those with a middle range purchase a used vehicle, and those with the lowest w choose the outside good.

We also need to specify taxes separately for used and new vehicles, because not only the relative tax matters. We denote τ_u the tax paid by the buyer on used vehicles, and τ_n the tax paid by the buyer on new vehicles.

The assumption of a competitive, constant marginal cost new vehicle supply means that the buyer's price of a new vehicle is $\psi + \tau_n$. Because ψ is fixed, an increase in the tax rate on new vehicles fully passes on to buyers. The equilibrium buyer's price of used vehicles is denoted $p + \tau_u$, with p being the price received by sellers. Because p is an equilibrium object, pass-through of a tax on used vehicles depends on the shape of supply and demand.

Supply. The owner of a new vehicle repairs the vehicle if and only if their repair cost k is below the equilibrium (seller) price p, the probability of which is $H(p)$. The used vehicle supply is thus the size of the new vehicle market in equilibrium times the repair rate, or $NH(p)$. Supply is perfectly elastic in the new vehicle market and the outside good.

Demand. Consumers are indexed by their preference w. A consumer with preference w chooses between a new vehicle, a used vehicle, or the outside good. Our assumption that a consumer with new vehicle utility w has used vehicle utility zw ensures a monotonic sorting, where the w distribution will be partitioned with the highest values choosing a new vehicle, a middle range choosing used, and a bottom range choosing the outside good. We restrict our attention to cases where all three choices have some market share.

We can thus describe the equilibrium by the cutoff values that form the boundaries of the partition. Denote the lowest type who buys a used vehicle as $\underline{w}$. This consumer is indifferent between the outside good and a used vehicle, so $\underline{w}$ is defined by:

$$\begin{aligned} 0 &= z\underline{w} - p - \tau_u \\ \underline{w} &= \frac{p + \tau_u}{z}. \end{aligned}$$

Denote the highest type that buys a used vehicle as $\overline{w}$. This consumer is indifferent between a new and used vehicle:

$$\begin{aligned} \overline{w} - \psi - \tau_n + H(p)(p - \bar{k}) &= z\overline{w} - p - \tau_u \\ \Rightarrow \overline{w} &= \frac{(\tau_n - \tau_u) + (\psi - p) - H(p)(p - \bar{k})}{1 - z}. \end{aligned}$$

The partition is summarized in Appendix Figure A8, which shows the payoffs from each choice as a function of w, for a given p. The values of a new car and a used car are shown as two lines with w on the x-axis and payoffs (in dollars) on the vertical axis. The used car value has a slope of z and a y-intercept at $-p - \tau_u$, which would be the payoff for an agent with zero valuation of the used good. The new car line starts off at a lower intercept, $-\psi - \tau_n + H(p)(p - \overline{h})$ but rises at a faster slope, equal to 1.[13]

Agents make the vehicle choice, including the outside good, with the highest payoff. For an agent with $w < \underline{w}$, the best choice will be the outside good (payoff of 0), because both used and new vehicles have a negative payoff. Individuals with $\underline{w} < w < \overline{w}$ will prefer a used car. Because the slope of the new car payoff is steeper, for a sufficiently high w individuals will have $w > \overline{w}$ and will thus prefer a new car.

Market shares. The size of the new vehicle market is $1 - G(\overline{w})$. The size of the used vehicle market is $G(\overline{w}) - G(\underline{w})$.

Equilibrium. The equilibrium requires that p is such that used vehicle supply $((1 - G(\overline{w}))H(p))$ equals used vehicle demand $G(\overline{w}) - G(\underline{w})$. This equilibrium condition can be written equivalently as follows, where $\overline{w}$ and $\underline{w}$ are on opposite sides of the equation, which facilitates differentiation below:

$$H(p) - G(\underline{w}) = (1 + H(p)G(\overline{w}). \tag{F.13}$$

Comparative statics. There remains one endogenous price in the model, p. In equilibrium, all agents make the optimal choice of new, used, or outside good, and repairs are made whenever $k < p$. The new vehicle market clears at price $\psi + \tau_n$, and the outside good market clears at price 0. The equilibrium price p clears the market for used vehicles.

We are interested in how changes in τ_n and τ_u affect the market, noting that a change in ψ has the same effects on the market as a change in τ_n. The results from the model are summarized in Appendix Table E.1. Derivations of the results are included below.

Table E.1: Comparative Statics Summary for Model with Outside Good

Variable	Outcome				
	O	U	N	p	$H(p)$
τ_n (or ψ)	+	?	-	+	+
τ_u	+	-	+	-	-

In this model, an increase in the new vehicle price (from either ψ or τ_n) causes the overall vehicle market to shrink. Equivalently, the outside good share rises. The quantity of new vehicles shrinks. The price of used vehicles rises, which means that the repair rate increases. An increase in the new vehicle price has an ambiguous effect on the size of the used vehicle market. Intuitively, used vehicles are a larger share of a smaller market, and this can lead to an increase or a decrease in total size, depending on which of those factors dominates.

[13] For a sufficiently high tax on used vehicles, or a sufficiently expensive minimum repair cost, an equilibrium exists with no used vehicles and the diagram would be qualitatively different. Our attention is limited to cases where there are some used vehicles and some selection to the outside good.

An increase in the used vehicle tax causes the overall size of the vehicle market to shrink (equivalently, the outside good share rises). The used vehicle market shrinks. The new vehicle market expands. The price of a used vehicle falls, so the repair rate declines.

How do these results compare to the model with no outside good? In that model, Proposition 1 says that an increase in the new vehicle price ψ increases the used vehicle market and hence decreases the new vehicle market, increases the repair rate, and decreases the equilibrium used vehicle price. These results are the same with the outside good, but the effect on the absolute size of the used vehicle market is ambiguous. Used vehicles are a larger share of the total vehicle market, but the market shrinks so the absolute size is ambiguous. This adds nuance to the Gruenspecht effect discussed in the main text. With an outside good, raising the price of new durables lowers the equilibrium scrap rate and makes used durables a larger fraction of the market. But the total number of used durables in the market could nevertheless decline, if the market shrinks enough.

Similarly, an increase in the relative tax on new vehicles $\tau = \tau_n - \tau_u$ decreases the scrap rate and increases the market share of used vehicles. With an outside good, increasing the relative tax on new vehicles can come from either an increase in τ_n or a decrease in τ_u. Either case decreases the scrap rate and increases the relative share of used vehicles as a fraction of the total vehicle market. As noted above, the effect of an increase in τ_n on the absolute magnitude of used vehicles is ambiguous, whereas the effect of a decrease in τ_u is not. This is the only difference in comparative statics between the two versions of the model.

Proposition 2 from the main text pertains to welfare. In the case with two tax rates and an outside good, a broader set of welfare results are possible. If one tax rate is set equal to marginal damages (say $\tau_n = \phi_t^n$), then the welfare maximizing value of the other tax equals marginal damages, and moving that tax toward marginal damages improves welfare.

To sign the comparative statics above, we first show that the used vehicle price rises with an increase in the new vehicle tax ($dp/d\tau_n > 0$) and that the price will fall with an increase in the used vehicle tax ($dp/d\tau_u < 0$). We then totally differentiate the cutoff values $\underline{w}$ and $\overline{w}$. Given the signs of $dp/d\tau_n$ and $dp/d\tau_u$, we rearrange those derivatives to yield clear signs.

For purposes of notation, write the equilibrium condition in equation (F.13) as $A = B$, where $A = H(p) + G(\underline{w})$ and $B = (1 - H(p))G(\overline{w})$. Then, implicit differentiation yields:

$$\frac{\partial A}{\partial p}dp + \frac{\partial A}{\partial \tau_u}d\tau_u = \frac{\partial B}{\partial p}dp + \frac{\partial B}{\partial \tau_u}d\tau_u.$$

Rearranging:

$$\frac{dp}{d\tau_u} = \frac{\frac{\partial A}{\partial \tau_u} - \frac{\partial B}{\partial \tau_u}}{\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}}. \tag{F.14}$$

Likewise, for τ_n the same steps yield:

$$\frac{dp}{d\tau_n} = \frac{\frac{\partial A}{\partial \tau_n} - \frac{\partial B}{\partial \tau_n}}{\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p}}. \tag{F.15}$$

Equations (F.14) and (F.15) have the same denominator. The two terms in the denominator are:

$$\begin{aligned}
\frac{\partial B}{\partial p} &= (1-H(p))g(\overline{w})\frac{\partial \overline{w}}{\partial p} - h(p)G(\overline{w}) \\
&= (1-H(p))g(\overline{w})\left(\frac{1}{1-z}(-1-H(p))\right) - h(p)G(\overline{w}) \\
&= -(1-H(p))(1+H(p))\frac{g(\overline{w})}{1-z} - h(p)G(\overline{w}) \\
\frac{\partial A}{\partial p} &= h(p) + g(\underline{w})\frac{\partial \underline{w}}{\partial p} = h(p) + \frac{g(\underline{w})}{z}
\end{aligned}$$

Combining yields the denominator for either comparative static, which is negative:

$$\begin{aligned}
\frac{\partial B}{\partial p} - \frac{\partial A}{\partial p} &= -(1-H(p))(1+H(p))\frac{g(\overline{w})}{1-z} - h(p)G(\overline{w}) - h(p) - \frac{g\left(\underline{w}\right)}{z} \\
&= -\underbrace{(1-H(p))(1+H(p))\frac{g(\overline{w})}{1-z}}_{(+)} - \underbrace{h(p)(1+G(\overline{w}))}_{(+)} - \underbrace{\frac{g(\underline{w})}{z}}_{(+)} \\
&< 0.
\end{aligned}$$

The numerator for equation F.14 (the τ_u case) is:

$$\frac{\partial A}{\partial \tau_u} - \frac{\partial B}{\partial \tau_u} = \frac{g(\underline{w})}{z} + (1-H(p))\frac{g(\overline{w})}{1-z} > 0.$$

Thus, $dp/d\tau_u < 0$ because the numerator and denominator of (F.14) are both negative. The numerator for equation F.15 (the τ_u case) is:

$$\frac{\partial A}{\partial \tau_n} - \frac{\partial B}{\partial \tau_n} = 0 - (1-H(p))\frac{g(\overline{w})}{1-z} < 0.$$

Thus, $dp/d\tau_n < 0$ because the numerator and denominator of F.15 have opposite signs.

With these effects on price signed, we can derive the market size effects by differentiating the market size expressions, recognizing that the cutoff values $\overline{w}$ and $\underline{w}$ will change, both because of direct effects and because of the impact of the tax on p.

Used vehicle taxes. An increase in the tax on used vehicles will increase the outside good share (shrink the total vehicle market):

$$\frac{dO}{d\tau_u} = \frac{dG(\underline{w})}{d\tau_u} = \frac{g(\underline{w})}{z}\left(1 + \frac{dp}{d\tau_u}\right) > 0$$

It also positively affects the new vehicle market share:

$$\frac{dN}{d\tau_u} = \frac{d(1-G(\overline{w}))}{d\tau_u} = \frac{g(\overline{w})}{1-z}\left(1 + \frac{dp}{d\tau_u} + H(p)\frac{dp}{d\tau_u}\right) > 0$$

The effect on the used vehicle market size is just the difference of these two effects, so the used vehicle market size must decline.

New vehicle taxes: An increase in the tax on new vehicles will increase the size of the outside good (shrink the total vehicle market):

$$\frac{dO}{d\tau_n} = \frac{dG(\underline{w})}{d\tau_n} = \frac{g(\overline{w})}{z}\left(\frac{\partial p}{\partial \tau_n}\right) > 0$$

An increase in the tax on new vehicles will decrease the size of the new vehicle market. This must be true because the overall vehicle market shrinks, and the repair rate $H(p)$ rises because $dp/d\tau_n > 0$. The effect of a tax on new vehicles on the overall size of the used vehicle market is ambiguous. The tax shrinks the overall size of the vehicle market, but increases the fraction of vehicles that are used.

F Quantitative Model: Additional Details (Section 8)

This appendix provides detail supporting the quantitative model and shows results from a range of additional counterfactuals. The appendix begins by providing a list of data sources and parameters needed for the analysis (F.1) and how baseline model outputs line up with the data (F.2). It then provides detail on several aspects of model specification, namely: how the utility function described in the text resolves into the demand specification (F.3), explains how our assumption about agent beliefs about price changes translates into used vehicle prices (F.7), our solution algorithm (F.5), calibration of the model to the initial period (F.6), and the model mechanics regarding vehicle depreciation (F.7). We then offer an extension of the representative agent model, calibrated using vehicle ownership divided over income groups, to characterize the likely distributional implications of policy counterfactuals (F.8). Lastly, we report a variety of sensitivity analyses (F.9).

F.1 Data for Quantitative Model

This c (summarized in Appendix Table A7), assumptions, and extrapolations used to construct the quantitative model. The final input data for the model is for two-year age bins $a = 0, 1, \ldots, 18$, where age bin 0 corresponds to 0-1 year old vehicles, age bin 1 corresponds to 2-3 year old vehicles, etc.

Many data and parameters described here are primitives (i.e., they do not change in equilibrium) that we assign to a vehicle based on some combination of vintage, age and class. These include annual vehicle miles traveled, fuel economy, scrap elasticities and damages per ton. Scrap rates, vehicle prices and vehicle quantities are equilibrium objects. We use data on vehicle prices and quantities to calibrate the initial fleet.

Vehicle miles traveled (VMT). The data source for vehicle miles traveled is the Colorado emissions smog check dataset (Colorado Department of Public Health and Environment 2016). The quantitative analysis uses data from test year 2014, which gives us raw VMT data for the widest age range—vehicles aged 4-32 years. We tabulate average VMT by age by class and size. The cutoff used for size is the median curb weight for each vehicle class (3,000 lbs. for cars and 4,000 lbs. for trucks). We assume that VMT for 0-3 year old vehicles equals VMT at age 4, as these newer vehicles are exempt from emissions testing and therefore not observed in the Colorado data. Likewise, we assume that VMT for vehicles ages 33-37 equals VMT at age 32.

Vehicle prices. Data on vehicle prices are from NADA (National Automobile Dealers Association 2012). This dataset contains used vehicle resale values for vehicles between 1 and 19 years old. We extrapolate prices for new vehicles (assuming that the depreciation rate between 0 and 1 year old vehicles equals that between 1 and 2 year old vehicles) as well as for 20-37 year old vehicles. The latter extrapolation is performed as follows. We use pricing data for 19 year old and 27 year old vehicles from the 2019 Kelley Blue Book (KBB) (Kelley Blue Book Co. 2019). For each of the 28 make-class-size combinations in the quantitative model, we select the model that appears in most model-year by age by calendar-year combinations, except if it was a sports car (such vehicles are not representative). We exclude vehicles for which the KBB does not go back to model year 1992, unless there is no vehicle in the KBB data that goes back that far in time (this applies to 3 out of 28 categories, for which the earliest model year is 1993 or 1995). We use the "buy from private party" option as this seems most relevant for old vehicles. We use "fair value" for a middle-of-the-road trim, without added options, in "good" (the most common) condition. After collecting resale values for the 19 year old and 27 year old vehicles, we took the ratio of the average prices, which indicates 37.7% depreciation between ages 19 and 27. We then extrapolate the NADA price data by setting a fixed (calibrated) annual depreciation percentage for ages 27-37 and a linear interpolation of the depreciation percentage between age 19 (for which we observe prices in the NADA data) and the assumed percentage for age 27 (based on the KBB data). The calibrated depreciation for ages 27-37 is -4.2% annually.

Vehicle quantities. We use Wards Automotive Yearbooks data on the composition of new vehicle sales by size, class and manufacturer (Wards Intelligence 2002). We then calibrate the quantity of new vehicles sold (i.e., apply a scaling factor) such that the resulting magnitude of the total (new and used) fleet equals the total fleet size from the Wards Automotive Yearbook 2002 (which reports vehicle quantities for the year 2000). The total fleet size in the year 2000 is 221 million vehicles. Finally, holding the total quantity of vehicles of each class, size, manufacturer and age fixed, we adjust the light-duty truck share to match the average car v. truck profile over the period 2000-2014 using data from the Federal Reserve Bank of St. Louis (2014). This adjustment adds realism to model estimates, as the share of trucks has risen sharply over the last several decades and not adjusting for this trend would lead to an overstated used truck fleet (and, therefore, overstated emissions damages).

Inflation. We use the Consumer Price Index for all items in U.S. city average, all urban consumers, not seasonally adjusted (U.S. Bureau of Labor Statistics 2021).

Scrap elasticities. We take elasticties of vehicle scrap with respect to the used vehicle resale value from Jacobsen and van Benthem (2015). We take their estimate by class, size and age category (1-8 years old vs. 9+ years old). These elasticities for the younger age category are -0.758, -0.979, -0.816 and -0.617 for small cars, large cars, small trucks and large trucks, respectively. For the older age categories, the elasticities are -0.514, -0.500, -0.811 and -1.018. These elasticities are treated as fixed parameters. Combined with equilibrium prices, they result in scrap rates that are endogenous outcomes of the model.

Scrap rates. We calculate scrap rates by age, class and size from vehicle registration data from R.L. Polk & Company (2009), used in Jacobsen and van Benthem (2015). Scrap rates for one year old vehicles (not observed in the data) are assumed equal to scrap rates for two year old vehicles of the same class and size. We also do not observe scrap rates for vehicles ages 32-37 years, so we assume their scrap rates are equal to the scrap rates of 27-31 year old vehicles of the same class and size. These scrap rates are taken as initial starting points for the baseline simulation; changes to scrap rates depend on the scrap elasticity and equilibrium prices.

Fuel economy. Fuel economy data for new vehicles in the year 2000 come from the U.S. Department of Energy and aggregated to the make by class by size level (U.S. Department of Energy 2022). We use data on realized fuel economy of the fleet by model year to calculate fuel economy ratings to model years older than 2000 (National Highway Traffic Safety Administration 1978, 2014). From this, we observe that the fuel economy of new vehicles was almost flat for the period 1982-2000. For model year 1963 (corresponding to the oldest possible age in our model for the 2000 fleet, 37 years old) to 1982, we compute a trend in annual fuel economy, separately for cars and trucks (0.9735 and 0.9797, respectively). We then assign a vehicle's fuel economy rating based on when it was produced, using the estimated trend only for vehicles produces before 1982.

We measure Corporate Average Fuel Economy Standards for cars and trucks from the National Highway Traffic Safety Administration (2011) (CAFE Standards 1978-2010) and the U.S. EPA (2010) (CAFE Standards 2011-2016). For model years 2017 and beyond, we assume CAFE standards for cars and trucks increase linearly at the rate observed over 2000-2014.[14]

Finally, we use curvature parameters to calibrate the fuel economy cost functions as described in Appendix F.6. To represent baseline values beginning in model year 2000, we use an estimate of the costs of fuel economy using engineering data from the National Research Council (National Research Council 2002). Their costs can be approximated closely with a quadratic function in fuel economy.

Pollution per mile. The pollution data in the model are averages of CO, HC and NO_x emissions per mile by age, class and size from the Colorado smog check data. We use data for the vehicle fleet observed in calendar years 2000, 2002, ..., 2014 consistent with our two-year age bins.

Because the Colorado smog check data end in calendar year 2014, we extrapolate emission rates for calendar years 2015 and beyond. We do this using a combination of age deterioration factors (i.e., a given model year becomes dirtier as it ages) and model year improvement factors (i.e., every subsequent model year has lower new vehicle emission rates). We estimate the age deterioration factor using a regression of log pollution rates on age and VIN prefix fixed effects. We plot the age fixed effects and fit a linear relationship for ages 4-19. For vehicles age 20 and older, the relationship is flat, and we assume no further deterioration as a result. See Appendix F.6 for detail on calibration of emissions functions beyond the 2000-2014 time period. We estimate the vintage improvement factor for new vehicles as the average rate of decline in new vehicle emission rates over the period 2014-2020; see Appendix B.2 for details.

We also extrapolate emissions for unobserved model years in the 2000, 2002, ..., 2014 emissions data. The raw pollution data for the 2000 fleet describe vehicles aged 4-18 years.[15]

[14] These differ from the actual fuel economy standards over the years 2017-2021 but we make a long-term assumption that standards will progress steadily at historical rates.

[15] Colorado smog check inspections are required for 4-year old vehicles and for vehicles with model year

We extrapolate down to ages 2-3 using emission rates for 4-5 year-old vehicles in the 2002 fleet, and down to ages 0-1 using emission rates for 4-5 year-old vehicles in the 2004 fleet. We extrapolate up to ages 19-25 using the exponential annual deterioration factor calculated as (emissions for 18-year-old vehicles/emissions for 4-year-old vehicles)$^{1/14}$, which combines age deterioration effects and vintage improvement effects. For vehicles aged 26 years and older, we use a more conservative linear extrapolation based on the pollution deterioration between age 18 and age 25.

Pollution damages per ton. The pollution damages for HC and NO_x are taken from Tschofen et al. (2019), weighted across counties by population. For CO, the damages are from Knittel and Sandler (2018).

Vehicle property taxes. We created a database with vehicle property tax rates using a variety of state and local sources. Most of these tax rates come from state government websites, though the relevant division of the government varies by state (typically the state department of revenue, department of motor vehicles, or state law), and in some cases we corresponded directly with staff to clarify rates. Tax rates can vary at the county, special district, school district, or city level. We take an unweighted mean of the tax rates over the geographies within a county to aggregate up to the county level. Names of these registration fees vary by state and county—they can be called vehicle excise fees, vehicle personal property tax, vehicle ad valorem tax, or just a motor vehicle tax. Some states apply percentages that vary with vehicle age. In total, twenty-eight states have such registration fees.[16]

Other parameters. The quantitative model has several other parameters. We use an annual discount rate of 3.0%, which is one of the two standard discount rates used by the EPA and National Highway Traffic Safety Administration in their impact analysis for environmental regulation. We take GDP for the year 2000 from the U.S. Bureau of Economic Analysis: \$10.25 trillion (\$15.22 trillion in \$2019) (U.S. Bureau of Economic Analysis 2020). We assume a GDP growth rate of 0.5% per year, chosen to match the growth rate in total vehicle miles traveled between 2000 and 2014 reported in the Highway Statistics (Highway Statistics 2017). We use the gasoline price in the year 2000, obtained from the U.S. EIA: \$1.51 in \$2000 (\$2.24 in \$2019) (U.S. Energy Information Administration 2015). We assume an autonomous rate of improvement in fuel economy technology of 1.8% per year (Knittel 2011). Vehicle demand elasticities are taken from Jacobsen and van Benthem (2015). The values are $\rho_{t,s,a} = 0.5$ for all manufacturer nests, $\rho_{t,s} = 0.575$ for all age nests, $\rho_t = 0.55$ for both size nests, and $\rho_v = 0.5$ for the car/truck nest.[17] The highest-level utility parameter determines the substitution between vehicles and other goods. Our central case value for this parameter implies an aggregate elasticity of demand for vehicles (including gasoline cost) of 0.75. The corresponding value used for ρ_u is -0.33.

In the counterfactuals that accelerate Tier 2 by eight years, the first year of the policy change is 2000, making 2008 standards apply in 2000, 2010 standards apply in 2002, etc.

$\geq$ 1982. So in the calendar year 2000 fleet, we mostly only observe emissions for vehicles aged 4-18 years.

[16]The 28 states are Alabama, Arizona, Arkansas, California, Colorado, Connecticut, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Massachusetts, Michigan, Minnesota, Missouri, Mississippi, Montana, Nebraska, Nevada, New Hampshire, North Carolina, Rhode Island, South Carolina, Virginia, Washington, West Virgina, and Wyoming.

[17]These values are calibrated to an average own-price elasticity of -2.4 (Austin and Dinan 2005) and approximate the substitution patterns seen in the new vehicle source data in that study.

F.2 Model Fit

The quantitative model incorporates vehicle emissions rates for CO, HC, and NO_x taken from the full sample of Colorado smog check data, corresponding to that displayed in Figure 6 in the main text. Transformations of the emissions data to create a comprehensive model of age- and time-based depreciation factors to project out of sample, and to match vehicle definitions and age bins in the quantitative model, mean that assessing fit is an important check. Panels A and B of Appendix Figure A9 plot the emissions per mile of CO and HC in the quantitative model against the original Colorado smog check data. Each point is an age, class (passenger car or light truck) and vintage bin. The data line up closely along the diagonal with small average deviations. Vehicle quantities in the national sample used in the quantitative model appear slightly skewed toward more polluting vehicle sizes than in Colorado (somewhat more points appear to the right of the diagonals than to the left) though the effect is small in magnitude.

Next, we compare cumulative emissions by age in the quantitative model to those in the Colorado data. Here we expect some, though perhaps not major, differences: vehicle class and age shares in the quantitative model come from data on national sales and scrap rates of all vehicles, while Colorado may have a distinct fleet composition. Panel A of Appendix Figure A10 shows the cumulative distribution of emissions by pollutant and vehicle age in the quantitative model while panel B reproduces Figure 7 for Colorado from the main text and is provided for comparison. The patterns nationally and in the Colorado sample are similar; with the quantitative model calibration reflecting a national fleet with a somewhat greater fraction of driving and pollution occurring in vehicles less than 10 years old.

F.3 Demand

We assume that agent utility follows a standard constant elasticity of substitution (CES) form. We show here that this functional form for utility produces demand curves for each vehicle type that depend on per-period costs e, as well as a set of scale parameters and CES substitution parameters.

Maximizing utility (6) subject to the budget constraint (7) from the main text yields demand:

$$v(e_v, e_x, M) = \left(\frac{\alpha_v}{e_v}\right)^{\frac{1}{1-\rho_u}} \frac{M}{\alpha_v^{\frac{1}{1-\rho_u}} e_v^{\frac{\rho_u}{\rho_u-1}} + \alpha_x^{\frac{1}{1-\rho_u}} e_x^{\frac{\rho_u}{\rho_u-1}}} \tag{G.1}$$

$$x(e_v, e_x, M) = \left(\frac{\alpha_x}{e_x}\right)^{\frac{1}{1-\rho_u}} \frac{M}{\alpha_v^{\frac{1}{1-\rho_u}} e_v^{\frac{\rho_u}{\rho_u-1}} + \alpha_x^{\frac{1}{1-\rho_u}} e_x^{\frac{\rho_u}{\rho_u-1}}}$$

The associated price of one unit of utility (the CES composite or ideal price index) is

$$e^* = \left(\alpha_v^{\frac{1}{1-\rho_u}} e_v^{\frac{\rho_u}{\rho_u-1}} + \alpha_x^{\frac{1}{1-\rho_u}} e_x^{\frac{\rho_u}{\rho_u-1}}\right)^{\frac{\rho_u-1}{\rho_u}} \tag{G.2}$$

The consumer buys M/e^* units of composite good. From (G.1) and (G.2), demand for composite vehicles, or demand for other goods, relative to demand for composite good, is

$$\frac{v(e_v, e_x, M)}{M/e^*} = \left(\frac{\alpha_v e^*}{e_v}\right)^{\frac{1}{1-\rho_u}}$$
$$\frac{x(e_v, e_x, M)}{M/e^*} = \left(\frac{\alpha_x e^*}{e_x}\right)^{\frac{1}{1-\rho_u}}$$

At each level, the representative agent minimizes the cost of the given amount of the composite good:

$$\min_{c_i} \sum_{i=1}^{n} e_i c_i \quad \text{s.t. } Q = \left(\sum_{i=1}^{n} \alpha_i q_i^{\rho}\right)^{\frac{1}{\rho}}$$

for $i = 1, \ldots, n$, where Q is the (given) amount of the composite good demanded. Solving this yields the following solution for nest 5 (and analogous solutions for nests 4, 3 and 2):

$$\frac{v_{c,s,a,m}}{v_{c,s,a}} = \left(\frac{\alpha_{c,s,a,m} e_{c,s,a}}{e_{c,s,a,m}}\right)^{\frac{1}{1-\rho_{c,s,a}}}, \quad m = 1, \ldots, 7 \tag{G.3}$$
$$e_{c,s,a} = \left(\sum_{m=1}^{7} \alpha_{c,s,a,m}^{\frac{1}{1-\rho_{c,s,a}}} e_{c,s,a,m}^{\frac{\rho_{c,s,a}}{\rho_{c,s,a}-1}}\right)^{\frac{\rho_{c,s,a}-1}{\rho_{c,s,a}}}$$

The solution to the problem in nest 1 is described in equation (G.1).

We now solve for demands in all nests, given prices, parameters, and income. We solve for the demand ratios and $e_{c,s,a}$ at nest 5, then for nest 4, etc. Using the e_v, e_x obtained for nest 1 above and total income M, one can now solve for the level of nest 1 demand v and x. Finally, the solutions for the levels of demand at the sub-nests can be calculated using the earlier obtained demand ratios. We denote demand at the finest nest by $q^d_{c,s,a,m} \equiv v_{c,s,a,m}$.

We calibrate scale parameters α as functions of prices, quantities, and ρ_u, as follows:

1. Set $e_{c,s,a} = 1$ for all c, s, a.[18]

2. Determine $v_{c,s,a}$ given $e_{c,s,a}$, observed vehicle demands $v_{c,s,a,m}$, and the relationship $\sum_m e_{c,s,a,m} v_{c,s,a,m} = e_{c,s,a} v_{c,s,a}$.

3. Calculate $\alpha_{c,s,a,m}$ by rearranging equation (G.3).

F.4 Derivation of Used Vehicle Pricing Equation (13)

This appendix explains how our assumption of "no change" beliefs about rental rates translates into a set of used vehicle values. First consider the suppliers' problem for vehicles that are entering age a_{max}. The suppliers enter period t with $q_{a_{max}-1,t-1}$ of these vehicles and solve:

[18] Total expenditure on the composite good $v_{c,s,a}$ is uniquely determined by the demands and prices of the specific goods, but units for $v_{c,s,a}$ are arbitrary. Hence we can define units so $e_{c,s,a} = 1$.

$$\max_{p_{a_{max},t}} H_{a_{max}}(p_{a_{max},t}) \left(r_{a_{max},t} - \bar{k}_{a_{max},t}\right)$$

Where $H_{a_{max}}$ is the survival probability applied to the endowment of $q_{a_{max}-1,t-1}$ vehicles and $\bar{k}_{a_{max},t}$ is expected repair expenditure. Future periods do not enter this maximization problem since vehicles are scrapped with certainty after age a_{max}. The solution is to choose a cutoff value for repair $p_{a_{max},t} = r_{a_{max},t}$. Applying the repair cost density $h(\cdot)$ the quantity of vehicles supplied is:

$$q^s_{a_{max},t} = q_{a_{max}-1,t-1} * (1 - b_{a_{max}}(p_{a_{max},t})^{\gamma_{a_{max}}})$$

The second term inside the parentheses is the scrap rate, $y_{a_{max},t} = b_{a_{max}}(p_{a_{max},t})^{\gamma_{a_{max}}}$

Now consider vehicles entering age $a_{max} - 1$ at time t. The suppliers enter period t with $q_{a_{max}-2,t-1}$ of these vehicles and choose a cutoff $p_{a_{max}-1,t}$ such that they only repair vehicles with a repair cost draw below this cutoff. They take rental rates $r_{a_{max}-1,t}$ and $\mathbb{E}[r_{a_{max},t+1}]$ as given and maximize rental income this period, plus potential rental next period, less repair expenditures. When $\mathbb{E}[r_{a_{max},t+1}] = r_{a_{max},t}$, the cutoff from above ($p_{a_{max},t}$) also serves as a continuation value for the decision problem on vehicles of age $a_{max} - 1$. Similarly, the cutoff for age $a_{max} - 1$ vehicles is the continuation value in the age $a_{max} - 2$ decision. This makes us use "p" to represent the repair decision cutoffs since in equilibrium they equal the price or asset value of used vehicles.

F.5 Equilibrium Solution Algorithm

Because producer decisions about new vehicles depend on used-vehicle prices, which in turn depend on the new market, we take a nested iterative approach. All conditions must be satisfied in equilibrium. The model is exactly identified—the unknowns include 532 vehicle prices (the outside good price is normalized to one), 28 fuel economies, and 28 exhaust emission rates, with one equation per unknown.[19] The following steps are computationally efficient:

1. Given new vehicle prices $p_{c,s,0,m,t}$ and fuel economy levels $f_{c,s,0,m,t}$, solve for rental rates $r_{c,s,a,m,t}$ so $q^s_{c,s,a,m,t} = q^d_{c,s,a,m,t}$ for all $a > 0$. This involves iterating over the demand system and scrap versus repair decisions. The solution is a vector of 504 used vehicle rental rates, one per vehicle type.

2. Given used vehicle rental rates and list of which fuel economy constraints bind, solve the profit maximization problem in equation (9). This yields 28 new-vehicle prices $p_{c,s,0,m,t}$ and fuel economy levels $f_{c,s,0,m,t}$. We assume the exhaust constraints (11) bind on each vehicle and enter the cost function in equation (10).

[19] Analytical results for equilibrium existence and uniqueness have not been established for this class of models, so the analysis assesses equilibrium uniqueness numerically, using a broad range of starting values and alternative algorithms including a simple Newton's method.

3. Given the new vehicle demand quantities for each manufacturer and fleet, update the vector of which 14 fuel economy constraints bind. If a constraint was non-binding but is being violated, make it bind. If a constraint was binding but its Lagrange multiplier is negative, make it non-binding.

Solution begins by guessing a vector of binding fuel economy constraints and iterating between solving nests 0 and 1 until convergence. We use Broyden's method, a globally convergent quasi-Newton algorithm, to solve for the prices equating supply and demand in each nest. Once new and used vehicle prices are found so both markets clear, constraints in nest 2 are evaluated. If changes are made to the vector of binding constraints, the model re-solves the lower nests. This process continues until all equilibrium conditions are satisfied and no changes occur in nest 2. Equilibria are calculated for every two-year time period in sequence.

F.6 Calibration

Scale parameters. We calibrate the scale parameters α as described in Appendix F.3. We choose the scale parameter b_a to match baseline scrap rates in the data and to set γ_a.

Fuel Economy Cost Function. Equation (12) describes CAFE standards for each manufacturer's cars and trucks. The consumer also cares about fuel economy through gasoline costs in (8). We use CAFE standards from 2000, our base year. Year 2000 CAFE standards applied separately to each manufacturer and vehicle class, without possibility of trading between classes or manufacturers. Hence, we express CAFE standards as a threshold for the harmonic average fuel economies of each manufacturer's car and truck fleets.

The cost function for fuel economy in equation (9) is:

$$C^f_{c,s,t}(f_{c,s,t}) = \kappa^1_{c,s,t}(f_{c,s,t} - \tilde{f}_{c,s,t}) + \kappa^2_{c,s}(f_{c,s,t} - \tilde{f}_{c,s,t})^2$$

where $\tilde{f}_{c,s,t}$ is baseline fuel economy observed in our data. We calibrate values of the parameters $\kappa^1_{c,s,t}$ and $\kappa^2_{c,s}$. The calibration of $\kappa^1_{c,s,t}$ follows from the first-order conditions of the producer problem. Specifically, at the profit-maximizing point the value of an additional unit of fuel economy to producers (which we assume they set equal to the slope of cost given in $\kappa^1_{c,s,t}$) equals the willingness to pay for fuel economy by consumers plus the shadow value of fuel economy under pre-existing CAFE standards. For demand, we assume willingness to pay for a marginal improvement in fuel economy reflects the discounted stream of savings on gasoline. For the shadow value of CAFE standards we use estimates from Jacobsen (2013). To calibrate the second derivative of the fuel economy cost function, $\kappa^2_{c,s}$, we use the coefficient on fuel economy squared from a regression of engineering cost on fuel economy and fuel economy squared with the vehicle design data reported in National Research Council (2002).

Exhaust Emissions Cost Function. We calibrate the exhaust cost function (10) to minimize the sum of squared differences between the costs of exhaust standards in our model and those described in the Tier 2 and Tier 3 Regulatory Impact Analyses (U.S. EPA 1999, 2014a). The analyses report additional costs from Tier 2 and Tier 3 (combined and fully phased in) between \$90 for small cars and \$414 for large trucks.

Calibrated values of $\zeta_{c,s}$ from equation (10) reflect costs ranging from \$5.26 (small cars) to \$23.69 (large trucks) for a ten percent reduction in emission rates. The calibrated value of χ is 0.985. Allowing χ to vary with (c, s) does not substantially improve the fit.

The term $\xi_{c,s,t}$ in equation (10) reflects the calibration residual. Including it means that our baseline exactly matches the costs from the regulatory impact analyses; $\zeta_{c,s}$ and χ determine deviations in cost when exhaust policy deviates from the baseline.

It may be informative to compare the abatement technology assumptions in equation (10) against other approaches in the literature. The structure here follows that in Bovenberg et al. (2008). This is related to an approach used in much macro-climate change research, which assumes greenhouse gas emissions equal output, times a trend in emissions intensity, times the secular long-term trend in emissions intensity (Nordhaus 2013). Our approach fits historic data on emissions and costs, and thus also includes the residual term $\xi_{c,s,t}$ and the actual emissions data $\phi_{c,s,t}$ rather than simply the trend.[20]

Pollution. Baseline pollution emission rates evolve as follows. We use raw emissions data from Colorado smog check described in Section 3.2 to measure emission rates for calendar years 2000 through 2014. Emissions for time steps beyond 2014 ($t = 8$ in our notation) are calibrated as:

$$\phi_{p,a,t}|t > 8 = agefactor_{p,a}\phi_{p,0,t-a} \tag{G.4}$$

where $\phi_{p,0,t-a}$ are emissions of the vehicle when it was new and $agefactor_{p,a}$ captures deterioration of emissions with age. Calibrated values of $agefactor_{p,a}$ reflect annual rates of deterioration (increase) in CO, HC and NO_x of 3.6%, 5.6% and 4.0% through age 19, and zero thereafter.

When $a \geq t$ performing this computation requires inferring new-vehicle emissions before 2000. To do this we apply:

$$\phi_{p,0,t}|t < 1 = \frac{\phi_{p,1-t,1}}{agefactor_{p,1-t}} \tag{G.5}$$

Finally, for vehicles produced after 2014 we use new-vehicle emissions data through 2020 and apply $agefactor_{p,a}$ as above. In some sensitivity analyses we run the model past 2020, and there we extrapolate new vehicle emissions using observed improvements between 2014 and 2020.

F.7 Other Model Mechanics

Calculating model dynamics. When the model algorithm moves between time periods, it calculates a new equilibrium as described in Section 8.1, with updated exogenous parameters (e.g., income growth) but also given the fleet from the previous period's equilibrium. The fleet evolves so $q_{a,t} = (1 - y_{a,t})q_{a-1,t-1}$.

[20] Economy-wide models of air pollution can use one of several alternative models—production may generate potential pollution, and then abatement decreases actual emissions relative to the potential; or pollution abatement takes an endogenously-chosen share of productive factors, while goods production uses the rest; or firms have a separate production function for pollution, which uses abatement investments as an input. If goods production is Cobb-Douglas in standard inputs and in pollution, then these alternative interpretations of pollution abatement are analytically equivalent (Copeland and Taylor 2003; Shapiro and Walker 2018).

Depreciation. Counterfactual policies affect the value of existing used vehicles in ways that the vehicle rental suppliers do not expect. The timing of when changes in capital value enter the supplier's accounting method (and so are returned to households) influences the pattern of welfare effects. In the long run, any deferred changes in asset value must eventually appear, but discounting means the choice of accounting method could affect social welfare conclusions.

We assume new vehicle purchases and repairs are immediately fully depreciated:

$$\pi_t = \sum_{a=0}^{18} \Big((r_{a,t} - \tilde{k}_{a,t}) q_{a,t} \Big) - p_{0,t} q_{0,t}$$

Here, accounting profits for the vehicle rental supplier equal rental income minus spending on repairs and replacements. Profits will then be positive when the fleet is shrinking and negative when the fleet is growing. With a shrinking fleet, for example, vehicles from previous periods still bring in rental income, but some baseline expenditures to repair and replace them are no longer being made. Appendix F.9 discusses alternative approaches to computing depreciation.

Accounting for Expected Changes in Fuel Economy and Emission Rates

While the core of the model reflects simple steady-state expectations about used vehicle prices (i.e., vehicle suppliers assume future used vehicle values will match current ones), we can allow some sophistication in the form of adjustments to expectations based on attributes. Specifically we account for expected increases in future rental rates due to improving fuel economy and emission rates over time:

$$\begin{aligned} &\mathbb{E}[r_{c,s,a,m,fut.}] = \\ &\quad r_{c,s,a,m,cur.} + \upsilon * (\tau * vmt * (\phi_{j,cur.} - \phi_{j,fut.}) + p_{gas} * vmt * (1/f_{j,cur.} - 1/f_{j,fut.})) \end{aligned} \tag{G.6}$$

Here $fut.$ refers to future, $cur.$ to current, and $\upsilon \in [0, 1]$ controls how much of the difference between current and future attributes of vehicle j the supplier expects to be reflected in future rental values. The true value (if the supplier had rational expectations) is intermediate since both demand and supply will shift.

The value of υ affects the time path of accounting profits for the supplier. A low value of υ means that the supplier has positive surprises in the future when vehicles rent for more than expected, since they have better fuel economy and lower emissions than current versions of the same vehicle. In an accounting sense, too much depreciation is charged early on and so offsetting rents appear later. At the same time, low values of υ also mean that more scrap will occur in the short run, since suppliers expect future used values to stay low, so additional pollution gains occur. The main analysis uses a value of $\upsilon = 0.5$ and Appendix F.9 shows that welfare results aggregated over time are not sensitive to this choice.

F.8 Quantitative Model: Effects by Income Group

We apply data on the distribution of vehicles by income group to consider the likely incidence of our counterfactual registration fee policies across the income spectrum. Vehicle choice data by age and income from the 2001 NHTS (U.S. Federal Highway Administration 2001), chosen to line up with our central policy counterfactuals, appear in Panel C of Appendix Figure A11.[21] The highest income bin in the sample (annual income greater than $80,000 per household) appears in green, with a distribution of choices sharply skewed toward newer vehicles. Households from the lowest income bins (less than $20,000 annual income) are shown in red, and own vehicles from a much older section of the age distribution.

Appendix Table A9 presents registration fee payments at the baseline and under our central set of policy counterfactuals. Row 1 shows how baseline fees assessed in proportion to vehicle value (the fixed component of registration fees is assumed unchanged throughout) increase with income. Higher income households own newer, more valuable, vehicles, and more of them. We note that the incidence of existing fees (even when considering only the portion proportional to vehicle value) is regressive as a fraction of income.

Next, we compute the hypothetical fees that households in each income bin would pay if a pollution based fee varying with vehicle age and type were assessed and no re-optimization occurred: that is, the incidence if all households were to keep their baseline vehicle choice as in the 2001 NHTS. Reading row 2 from left to right, there are two competing effects: higher income households own more vehicles and so pay more fees, but they also own newer vehicles and so the pollution-based fees per vehicle are smaller. The effect of the increasing number of vehicles (average vehicles per household appears at the bottom of the table for reference) dominates through the 40-50k income bin, meaning higher income households pay slightly more pollution-based fees even though their per-car fees are lower. At higher levels of income the two effects cancel nearly exactly. The values shown are the annualized cost of all fees expected over 20 years of the counterfactual: payments in any individual year decline over time with improvements in pollution control and also reflect the transition path as older vehicles are removed from the economy. Overall, the fees assessed across income groups are similar in absolute terms, ranging from $170 to $205, making them sharply regressive as a fraction of income.

Row 3 makes use of our equilibrium counterfactual, where households scrap the majority of vehicles older than age 24 and thus avoid paying many of the highest fees. To consider the incidence of fees by income group we need to reassign vehicles that remain in the counterfactual equilibrium back into income bins, such that aggregate vehicle choice matches the modeled outcome and such that the fraction of households in each bin remains fixed. Among the set of allocations that satisfy these requirements we take the simple, and we think neutral, approach of reallocating vehicles such that the changes for each income group are kept in proportion to the baseline choices for that group.[22] For example, a group that tends to split its demand between middle-aged and old vehicles will switch most of their demand to middle-aged vehicles after the policy shock. A group that tends to own new, middle-aged, and old vehicles relatively equally would shift their demand from old to a combination of new and middle-aged.

The incidence of policy shown in row 3 shifts with scrap: Because lower income groups own more of the oldest cars to start with, they also do most of the vehicle scrap in response

[21] We further disaggregate by vehicle class and size in the analysis that follows.

[22] Mathematically this amounts to solving for two vectors, weights on vehicles and weights on income groups, such that when the weights are multiplied by baseline choices we arrive at a new matrix of choices satisfying the constraints on income bins and aggregate vehicle choices.

to fees: total fees paid fall 35% when accounting for equilibrium effects (i.e. between rows 2 and 3 for the lowest income group). Higher income groups also see fees fall: they substitute from middle-aged cars (which are now mostly owned by the lower income groups) into the newest vehicles and see fees paid fall 26%. Incidence remains regressive as in row 2, but not quite as sharply regressive after accounting for differential scrap rates.

Rows 4 through 6 investigate the remaining registration fee counterfactuals we examine in Section 8. With a revenue-neutral structure (where pollution-based fee revenue is dispersed equally to each vehicle registration) the wealthiest households gain relative to the baseline due to their large number of vehicles per household. The pollution-based fee raises large amounts of revenue and so alternative recycling structures, for example dispersing revenue equally to each household or through the income tax system, would produce very different and potentially progressive outcomes. New-vehicle fees in row 5 place much of the burden on wealthier groups, but as we discuss in the main text, they fail to produce pollution improvements. A simple flattening of fees in row 6 amounts to a more modest version of the revenue-neutral system in row 4 in terms of distributional outcomes.

F.9 Quantitative Model: Sensitivity Analyses

Alternative Elasticities, Baselines, and Policies

Appendix Table A10 reports a range of sensitivity analyses. Panel A repeats baseline results for the eight year delay of Tier 2 and the age×type registration fee from Table 5.

Panel B evaluates the Tier 2 delay counterfactual under four alternative elasticities. Rows 3 and 4 assume 50% lower and higher elasticities of scrap with respect to vehicle resale value. Rows 5 and 6 assume 50% lower and higher elasticities of substitution between vehicle vintages, i.e., how easily consumers substitute between vehicles of different vintages. Results are similar in all four cases; the exhaust standard delay changes the age profile of the fleet only slightly, and so changing parameters that control flexibility along this dimension has little effect.

Panel C investigates alternative baselines. Rows 7 and 8 assume CAFE standards are more stringent and that income grows more rapidly, while row 9 assumes that the ratio of miles traveled for new versus old vehicles is 5 (our data in the main analysis assume a ratio of 3.4). More stringent CAFE standards imply somewhat longer vehicle lifetimes, slightly slowing the damage done from a delay in exhaust standards and therefore reducing the (discounted) total damage change. Faster income growth and the alternative VMT schedule imply a slightly newer VMT-weighted fleet, somewhat increasing the present value of pollution damage from the counterfactual delay.

Panel C, row 10 allows for Bertrand competition among new vehicle producers. This adds a pre-existing distortion to the economy: market power reduces new vehicle sales and the overall number of vehicles, and it lengthens vehicle lifetimes. We calibrate elasticities such that markups from the producer problem are 25% in the baseline. Our first order conditions for firms include equilibrium effects in the current period. This allows the used vehicle market to adjust in response to new-vehicle price decisions, but it abstracts from the effects of current-period price decisions on future-period used markets. This myopia parallels the consumer problem. Implicitly, dynamic competition and other features of competition outside our model are assumed to be captured in the 25% markup and insensitive to the policy counterfactuals. In the context of the counterfactual exhaust standard delay in row 10, the longer lifetimes associated with imperfect competition imply that vehicles will take longer to work through the fleet, slightly reducing discounted harms. Row 11 considers a higher gasoline price, and row 12 considers an internal discount rate of 7% instead of 3%. The welfare results in row 12 are still discounted at 3% to provide a useful comparison for the table; it is the way market participants and asset values are constructed inside the model that differs. Welfare effects remain relatively stable across these scenarios: the welfare cost of delaying Tier 2 is -$185 billion in the main estimate, and in the sensitivity analyses this ranges from -$175 billion to -$202 billion.

Panel D evaluates the age×type registration fee counterfactual under alternative elasticities controlling flexibility in the age profile. Since this policy counterfactual operates directly on vehicle age we expect more sensitivity to the elasticities than in Panel B above: more elastic scrap in row 14 increases the utility of counterfactual policies since it lowers the cost of altering the fleet. Similarly, larger elasticities of substitution in row 16 predict larger welfare gains from age× type registration fees; when people more easily substitute across vehicle ages the gains from a registration fee policy are larger. Sensitivity appears largest to vintage substitution, with the high case implying about 60% greater net welfare gains.

Panel E investigates alternative baselines, now comparing effects of the age×type registration fee in different settings. Alternative trajectories of CAFE standards or income growth (rows 17 and 18) have little effect on the welfare gains, most of which are coming early in the simulated period. Under the alternative VMT schedule in row 19 (which makes older vehicles driven relatively less, and so less important to overall pollution) the system of age×type registration fees produces 11% smaller welfare gains. Bertrand competition among new vehicle producers (row 20) creates a pre-existing distortion which is now partially corrected by the age×type registration fee policy; it performs somewhat better in this environment since it now addresses both a market power distortion and the pollution externality. In row 21, the 50% higher baseline gasoline price reduces available welfare gains because the age×type tax is smaller relative to baseline ownership costs. Put another way, the higher gasoline price means that some of the switching away from used vehicles (which have slightly worse fuel economy) and especially used light trucks has already happened. Finally, row 22 shows internal discount rates have relatively small impact; this is due to the disproportionate share of fees that fall early in the simulated time period.

Panel F considers alternative counterfactual policies. Row 23 assumes that the marginal cost of emissions reductions is five times as high as in the baseline. This could reflect increased prices of precious metals used as catalysts. Net benefits fall from $25 billion in the baseline (Table 5) to $15 billion. Row 24 considers a scaled-down version of the age×type fee in row 2, now set at only 10 percent of damages. The smaller fee in row 24 produces a benefit-cost ratio of 19, primarily since marginal costs of distorting the vehicle age profile are increasing in the size of the distortion and row 24 describes a scaled-down policy. This relatively small policy could also be regarded as a better-targeted version of "flattening" existing registration fee structures and it generates economically large benefits. Row 25 examines a used vehicle fee that conditions on age only. The age-based registration fee in row 25 obtains a welfare gain that is 95% as large as the age×type registration fee in row 2. Put another way, almost all the benefits of this fee are due to differentiation between vehicle ages, rather than vehicle types. Finally, row 26 considers a flattening of registration fees starting from a higher baseline tax rate (0.68% versus 0.31% in the central case). Larger baseline fees mean that flattening the structure will have a larger effect: We find it scales approximately linearly with the starting fee rate and now leads to an approximately 4% reduction in emissions.

Alternative Depreciation Approaches and Expectations

We also investigated two approaches to computing depreciation which differ from that of Section F.7. One alternative immediately credits capital gains and charges capital losses. In this alternative, profits in time period t equal rental income less expected depreciation, which includes expected scrap and repairs and is equal to zero by equation (13), plus unexpected appreciation or depreciation between time periods due to the policy. The other alternative uses a schedule of depreciation for the original capital that is determined at vehicle purchase and then held fixed. Repair spending is depreciated immediately. The fixed depreciation schedule could reflect, for example, a pre-determined set of payments to a bank made to cover the original vehicle purchase. In this setting a reduction in rental rates (e.g., associated with a pollution tax) results in a sequence of losses since rental income falls short of the pre-determined payments each year as a given vehicle continues to age. The loss resulting from the policy shock will be more spread out than in case 2 above.

Experimentation found that the main depreciation approach and the first alternative produced similar results, while the second alternative increases the discounted welfare gain (over 20 years) from the age×type registration fee by about a fourth. This is because the third depreciation method allows much of the cost of policy (most of which is added new car purchases) to be deferred. We use the main depreciation method both for its simplicity and because it provides a conservative estimate of potential welfare gains. The welfare gains across the three methods should converge as the time horizon expands: reassuringly we find that over a 60 year time horizon the modeled welfare costs fall within 10% of one another.

We also investigated alternative choices of the υ parameter from equation (G.6) governing expectations around fuel economy and emission rates. We experimented with values between 0 and 0.6; values >0.6 can imply negative price expectations in some periods and prevent the model from converging. Welfare gains from the age×type registration fee over a 20-year horizon range from \$327 to \$316 billion, bracketing the central case estimate of \$322 billion. From this we take that the model-based estimates over time are not especially sensitive to this choice about expectations.

Spatially-Varying Damages and Low-Emissions Zones

The aggregate model we consider offers limited insight into spatial differences in pollution and policy, but we consider some variants of the model here that suggest important patterns. First, we run the model with two re-calibrations where damages are held either at the average for counties that are part of an MSA (denser, more urban counties) or for counties not in any MSA (the remainder of the U.S.). Damages in MSA counties are 3.5 times higher than damages in non-MSA counties following the estimates in Tschofen et al. (2019). Because most of the population resides in MSA counties, the main estimates in Table 5 come much closer to the "MSA" re-calibration.

In our main analysis the counterfactual policy of assessing registration fees equal to age by type specific aggregate damages produces a benefit-cost ratio of 2.9. When assigning the somewhat higher damages in MSA counties the benefit-cost ratio rises to 3.0. If all counties had non-MSA level damages, it falls to 2.4. Note that both the taxes, and benefits, assigned in the non-MSA counterfactual are much smaller.

The differences become more stark when considering coarse policies assigning high fees (independent of damages) to vehicles over a particular age. This counterfactual is similar in spirit to the "low emissions zone" policies present in many cities in Europe.[23] To approximate a discrete policy of this type we consider counterfactuals with a large, fixed registration fee that begins at a set age.[24] In our main analysis, age cutoff policies become cost effective beginning at age 16: that is, a ban on vehicles 16 and over is (just) cost effective. Bans on vehicles age 20 and older have large benefit-cost ratios and are similar to some of our main counterfactuals. When applying the level of damages present in MSA counties, bans on vehicles age 14 and older become cost effective. The even higher damages present in city centers, or the most densely populated counties, would likely take this pattern further. In contrast, using damages from counties not part of any MSA, bans by age are only cost effective when placed for vehicles age 26 and older.

SM References

AIR (1999). 1999 audit of the Colorado AIR program. Technical report, Office of the State Auditor.

Air Pollution Control Division (2013). Automobile inspection and readjustment (AIR) program annual report calendar year 2012. Technical report, Colorado Department of Public Health and the Environment.

Austin, D. and T. Dinan (2005). Clearing the air: The costs and consequences of higher cafe standards and increased gasoline taxes. *Journal of Environmental Economics and Management 50*(3), 562–582.

Bishop, G. A. and M. J. Haugen (2018). On-road remote sensing of automobile emissions in the Denver area: Winter 2017. Technical report, Coordinating Research Council.

Bishop, G. A., R. Hottor-Raguindin, D. H. Stedman, P. McClintock, E. Theobald, J. D. Johnson, D.-W. Lee, J. Zietsman, and C. Misra (2015). On-road heavy-duty vehicle emissions monitoring system. *Environmental Science & Technology 49*(3), 1639–1645.

Bishop, G. A., J. R. Starkey, A. Ihlenfeldt, W. J. Williams, and D. H. Stedman (1989). IR long-path photometry: A remote sensing tool for automobile emissions. *Analytical Chemistry 61(10)*, 671A–677A.

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

[23] The policies apply based on expected tailpipe emissions, but in practice this effectively removes vehicles based on age.

[24] Because we model a smooth CES utility function we cannot actually push all vehicles of any category out, so we choose the fee here such that 90% of vehicles at or beyond the age cutoff are eliminated.

Bishop, G. A., D. H. Stedman, J. de la Garza Castro, and F. J. Davalos (1997). On-road remote sensing of vehicle emissions in Mexico. *Environmental Science and Technology 31*(12), 3505–3510.

Bovenberg, A. L., L. H. Goulder, and M. R. Jacobsen (2008). Costs of alternative environmental policy instruments in the presence of industry compensation requirements. *Journal of Public Economics 92*, 1236–1253.

California Air Resources Board (2002). The federal test procedure. *Emissions Inventory Series 1(9)*, 1–2.

CARB (1993). Assessment of acceleration simulation mode (ASM) testing as an alternative to the IM240 transient dynamometer test. Technical report, California Air Resources Board.

Colorado Department of Public Health and Environment (1995-2016). Data on vehicle emissions inspection. Accessed July 15, 2021.

Copeland, B. R. and M. S. Taylor (2003). *Trade and the Environment: Theory and Evidence*. Princeton University Press.

Crandall, R. W., H. K. Gruenspecht, T. E. Keeler, and L. B. Lave (1986). *Regulating the Automobile*. The Brookings Institution.

Davis, S. C. and R. G. Boundy (2021). Transportation energy data book, edition 38. Technical report, Oak Ridge National Laboratory.

Federal Register (1995). Control of air pollution from new motor vehicles and new motor vehicle engines: Voluntary standards for light-duty vehicles.

Federal Reserve Bank of St. Louis (1963-2014). FRED economic data. https://fred.stlouisfed.org/graph/?id=DAUTOSAAR,DLTRUCKSSAAR,FAUTOSAAR,FLTRUCKSSAAR,. Accessed July 15, 2021.

Federal Reserve Bank of St. Louis (1977-2021). Data on motor vehicle assemblies. https://fred.stlouisfed.org/series/MVALTTRCKN and https://fred.stlouisfed.org/series/MVAAUTOASN. Accessed August 6, 2021.

Hawkins, D., C. Hart, C. Fulper, J. Warila, D. Brzezinski, S. Kishan, T. H. DeFries, J. Kemper, and J. Sidebottom (2010). Evaluation of evaporative leaks using RSD and inventory implications. Technical report, U.S. EPA.

Highway Statistics (1900-2017). Public road mileage, lane-miles, and VMT 1900-2017. https://www.fhwa.dot.gov/policyinformation/statistics/2017/vmt421c.cfm. Accessed July 15, 2021.

Jacobsen, M. R. (2013). Evaluating US fuel economy standards in a model with producer and household heterogeneity. *American Economic Journal: Economic Policy 5(2)*, 148–187.

Jacobsen, M. R. and A. A. van Benthem (2015). Vehicle scrappage and gasoline policy. *American Economic Review 105(3)*, 1312–38.

Jones, A. D. (1980). Testing of new technology three-way catalyst equipped vehicles in Los Angeles. Technical report, Automotive Environmental Systems.

Kelley Blue Book Co. (2019). Kelley blue book. https://www.kbb.com/car-values/. Accessed February 10, 2020.

Kelly, K. J. (1994). Correlation of I/M240 and FTP emissions for alternative motor fuel act test vehicles. Technical report, SAE Special Publication 1053 - Progress in Emissions ControlTechnologies (SAE 941901).

Klausmeier, R. (2017). Automobile inspection and readjustment program. Technical report, dKC de la Torre Klausmeier Consulting.

Knittel, C. R. (2011). Automobiles on steroids: Product attribute trade-offs and technological progress in the automobile sector. *American Economic Review 101*(7), 3368–99.

Knittel, C. R. and R. Sandler (2018). The welfare impact of second-best uniform-Pigouvian taxation: Evidence from transportation. *American Economic Journal: Economic Policy 10*(4), 211–242.

Lorang, P. A. (1984). Emissions of in-use vehicles: Update on the U.S. experience. Technical report, U.S. EPA.

Lyubich, E., J. S. Shapiro, and R. Walker (2018). Regulating mismeasured pollution: Implications of firm heterogeneity for environmental policy. *AEA Papers and Proceedings 108*, 136–42.

Manufacturers of Emissions Controls Association (1995). Emission control system upgrades for gasoline-powered light-duty vehicles: An available option for reducing ROG and NOx emissions. Technical report, MECA.

Manufacturers of Emissions Controls Association (2010). Evaporative emission control technologies for gasoline powered vehicles. Technical report, MECA.

Matthews, H. S. and L. B. Lave (2000). Applications of environmental valuation for determining externality costs. *Environmental Science & Technology 34*(8), 1390–1395.

Mills, E. S. and L. J. White (1978). *Approaches to Controlling Air Pollution*, Chapter Government Policies toward Automotive Emissions Control, pp. 348–409. Cambridge, MA: MIT Press.

Mondt, J. R. (2000). *Clean Cars: The History and Technology of Emission Control Since the 1960s*. Society of Automotive Engineers.

National Automobile Dealers Association (1993-2012). Used car guide. Unpublished data. Accessed July 15, 2021.

National Highway Traffic Safety Administration (1955-1978). Historical passenger car fleet average characteristics. https://one.nhtsa.gov/cars/rules/cafe/historicalcarfleet.htm. Accessed July 15, 2021.

National Highway Traffic Safety Administration (2011). Summary of fuel economy performance 2011. http://www.nhtsa.gov/staticfiles/rulemaking/pdf/cafe/2011_Summary_Report.pdf. Accessed July 15, 2021.

National Highway Traffic Safety Administration (2014). Summary of fuel economy performance 2014. hwww.nhtsa.gov/staticfiles/rulemaking/pdf/cafe/Performance-summary-report-12152014-v2.pdf. Accessed July 15, 2021.

National Research Council (2002). *Effectiveness and Impact of Corporate Average Fuel Economy (CAFE) Standards*. Washington, DC: The National Academies Press.

Nordhaus, W. (2013). DICE 2013R: Introduction and user's manual. Technical report, Yale

University.

Opus Inspection (2016). Remote sensing programs. https://web.archive.org/web/20170401230240/opusinspection.com/remote-sensing-device-technology/remote-sensing-programs/.

Palucka, T. (2004). Doing the impossible. *Invention & Technology Magazine 19*(3).

R.L. Polk & Company (1993-2009). Data on vehicles in operation. Unpublished data. Accessed July 30, 2019.

Shapiro, J. S. and R. Walker (2018). Why is pollution from U.S. manufacturing declining? The roles of environmental regulation, productivity, and trade. *American Economic Review 108*(12), 3814–54.

Sierra Research (1997). Development of a proposed procedure for determining the equivalency of alternative inspection and maintenance programs. Technical report, Sierra Research, Inc.

Thompson, G. J., D. K. Carder, M. C. Besch, A. Thiruvengadam, and H. K. Kappanna (2014). In-use emissions testing of light-duty diesel vehicles in the United States. Technical report, West Virginia University.

Tschofen, P., I. L. Azevedo, and N. Z. Muller (2019). Fine particulate matter damages and value added in the US economy. *Proceedings of the National Academy of Sciences 116*(40), 19857–19862.

UPI (1976). E.P.A. tests show ill-set carburetors are polluting air. *New York Times*. July 22, p. 28.

U.S. Bureau of Economic Analysis (1929-2020). Data on national income and product accounts. https://apps.bea.gov/iTable/iTable.cfm?reqid=19&step=2#reqid=19&step=2&isuri=1&1921=survey. Accessed July 16, 2021.

U.S. Bureau of Labor Statistics (1930-2021). Consumer price index. https://data.bls.gov/cgi-bin/surveymost?cu, series number CUUR0000SA0: all urban consumers - all items - US city average. Accessed July 15, 2021.

U.S. Department of Energy (1978-2022). Fuel economy data. https://www.fueleconomy.gov/feg/download.shtml. Accessed October 26, 2021.

U.S. Energy Information Administration (1929-2015). Data on average historical annual gasoline pump price, 1929-2015. https://www.energy.gov/eere/vehicles/fact-915-march-7-2016-average-historical-annual-gasoline-pump-price-1929-2015. Data from Monthly Energy Review, Table 9.4. Accessed July 16, 2021.

U.S. EPA (1980). FY79 study of emissions from passenger cars in six cities: Volume I. Technical report, U.S. EPA.

U.S. EPA (1992). I/M costs, benefits, and impacts analysis. Technical report, U.S. EPA.

U.S. EPA (1995). Development of a methodology for estimating basic emission rates for use in the MOBILE emission factor model. Technical report, U.S. EPA.

U.S. EPA (1999). Regulatory impact analysis - control of air pollution from new motor vehicles: Tier 2 motor vehicle emissions standards and gasoline sulfur control requirements. Technical report, U.S. EPA, EPA420-R-99-023.

U.S. EPA (2003). Conversion factors for hydrocarbon emission components. Technical report, U.S. EPA.

U.S. EPA (2010). Light-duty vehicle greenhouse gas emission standards and corporate average fuel economy standards; final rule. http://edocket.access.gpo.gov/2010/pdf/2010-8159.pdf. Accessed July 15, 2021.

U.S. EPA (2014a). Control of air pollution from motor vehicles: Tier 3 motor vehicle emission and fuel standards final rule. Technical report, U.S. EPA, EPA-420-R-14-005.

U.S. EPA (2014b). National emissions inventory 2014. https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data. Accessed January 30, 2017.

U.S. Federal Highway Administration (2001). National household travel survey. https://nhts.ornl.gov/downloads. Accessed August 4, 2021.

Wards Intelligence (2002). Ward's automotive yearbook 2002.

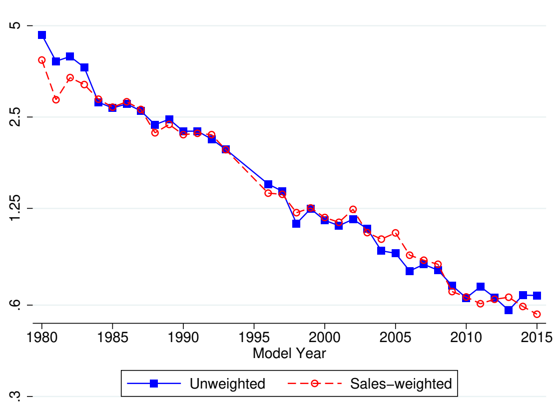
Xie, S., J. Bluett, G. Fisher, G. Kuschel, and D. Stedman (2005). On-road remote sensing of vehicle exhaust emissions in Auckland, New Zealand. *Clean Air and Environmental Quality 39*(4), 37–42.

Zhang, Y., D. H. Stedman, G. A. Bishop, P. L. Guenther, and S. P. Beaton (1995). Worldwide on-road vehicle exhaust emissions study by remote sensing. *Environmental Science & Technology 29*(9), 2286–2294.

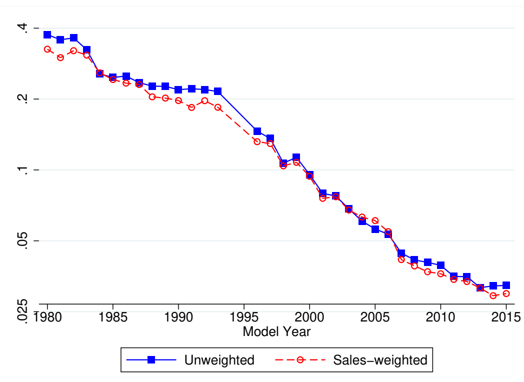
Appendix Figures and Tables

Figure A1: Mean Pollution Emission Rates of New US Vehicles, Weighted and Unweighted

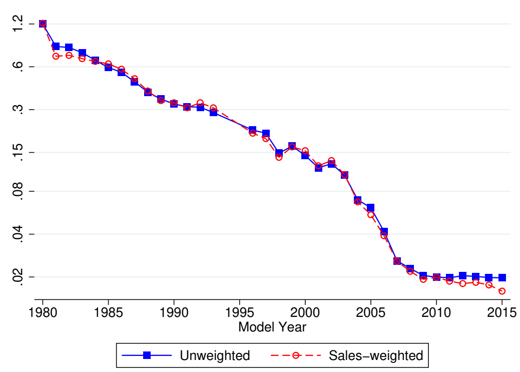
(a) Carbon monoxide (CO)



(b) Hydrocarbons (HC)

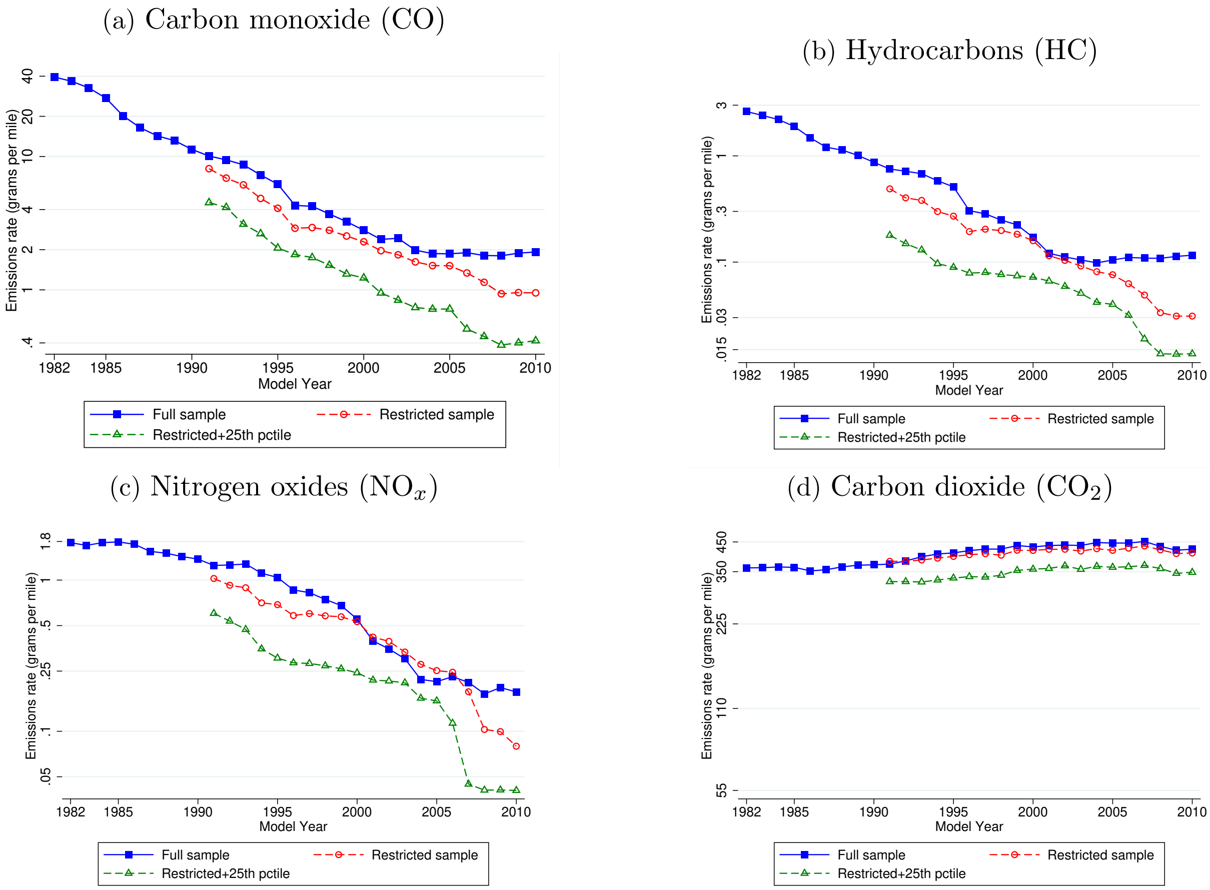


(c) Nitrogen oxides (NO_x)



NOTES: Blue solid line shows unweighted trend reprinted from Figure 1. Red dashed line shows the subset of Wards data for which we could accurately identify the emission rate, weighted by sales.

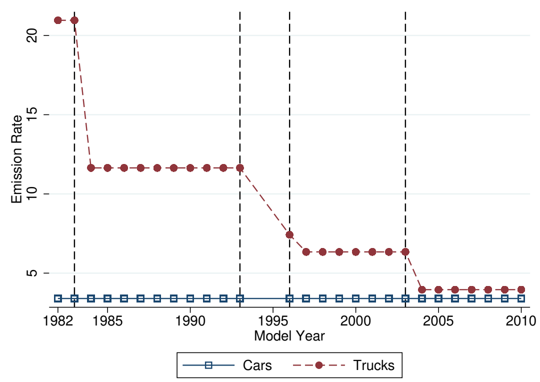
Figure A2: Mean Air Pollution Emission Rates of Colorado Used Vehicles



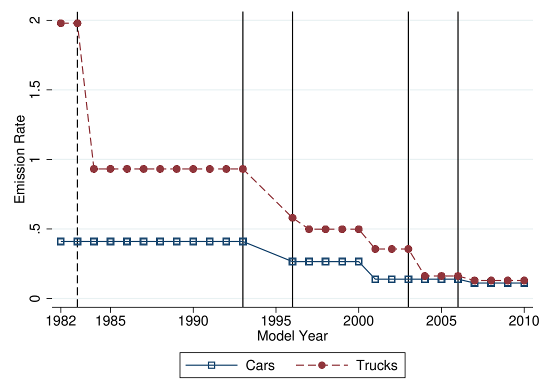
NOTES: Blue solid line shows full sample from Colorado smog check tests. Restricted sample limits the sample to 4 to 6 year old vehicles with 40,000 to 60,000 miles in model years 1991-2010. The 25th percentile line is estimated from quantile regressions. Graphs show fitted values for model year fixed effects plus a constant from regressions. Full sample regressions include age fixed effects and a linear odometer control.

Figure A3: Exhaust Standards and Emission Rates, Cars Versus Trucks

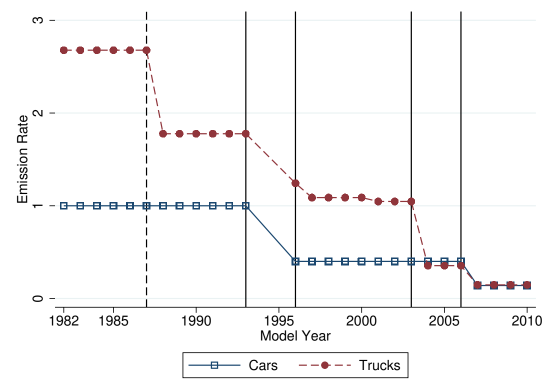
(a) Exhaust standards: CO



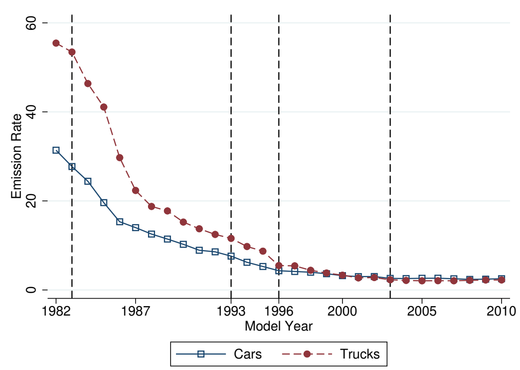
(b) Exhaust standards: HC



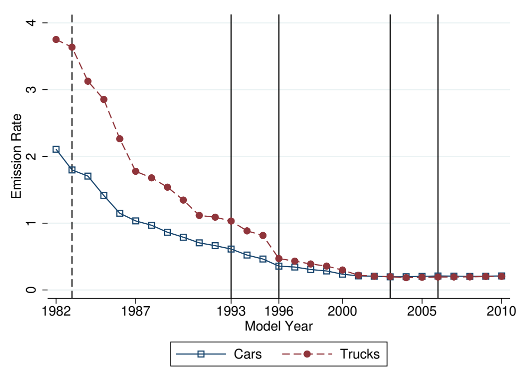
(c) Exhaust standards: NO_x



(d) Colorado smog check: CO



(e) Colorado smog check: HC



(f) Colorado smog check: NO_x

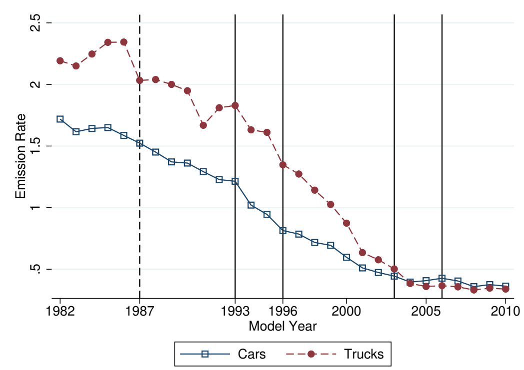
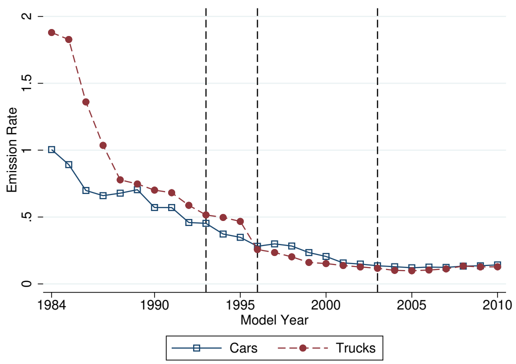
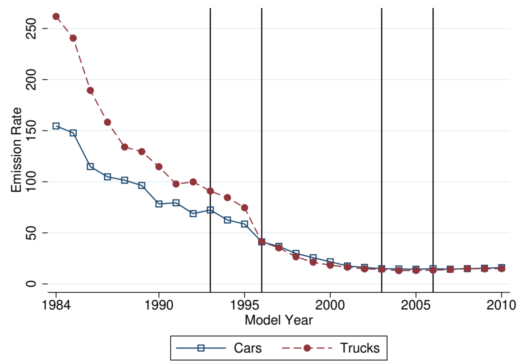


Figure A3: Exhaust Standards and Emission Rates, Cars Versus Trucks

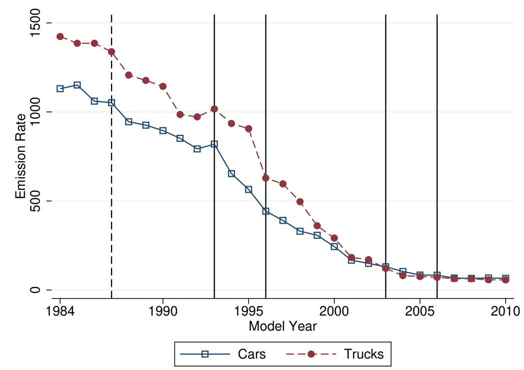
(g) Colorado remote sensing: CO



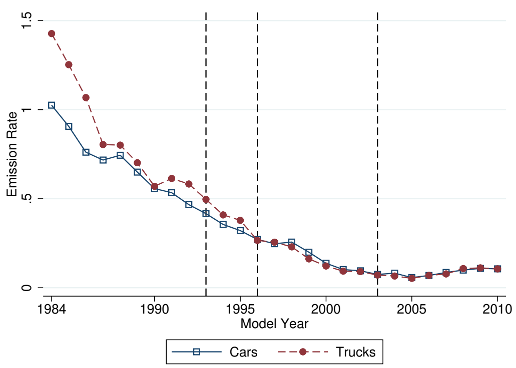
(h) Colorado remote sensing: HC



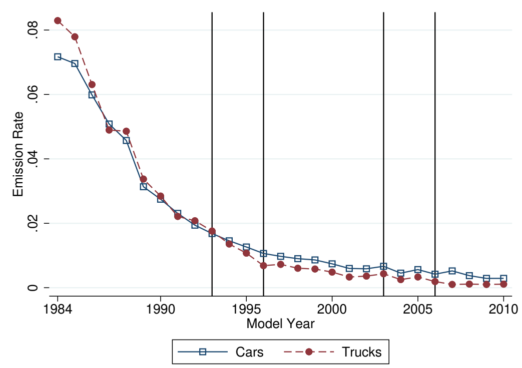
(i) Colorado remote sensing: NO_x



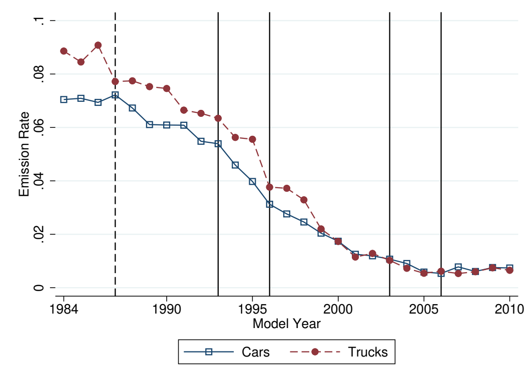
(j) Multi-state remote sensing: CO



(k) Multi-state remote sensing: HC



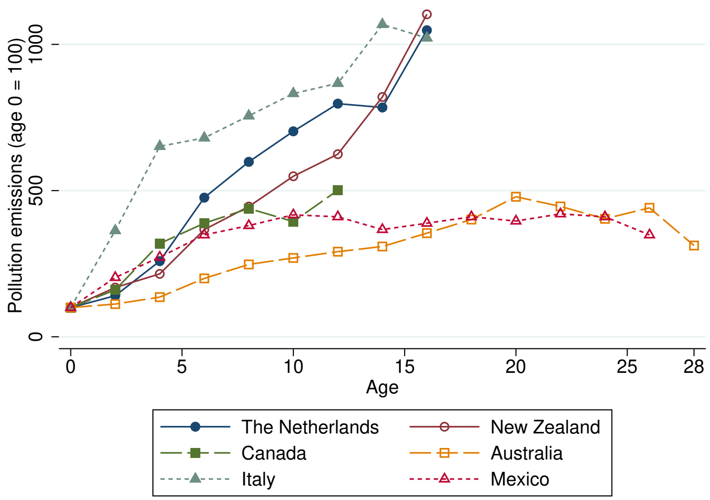
(l) Multi-state remote sensing: NO_x



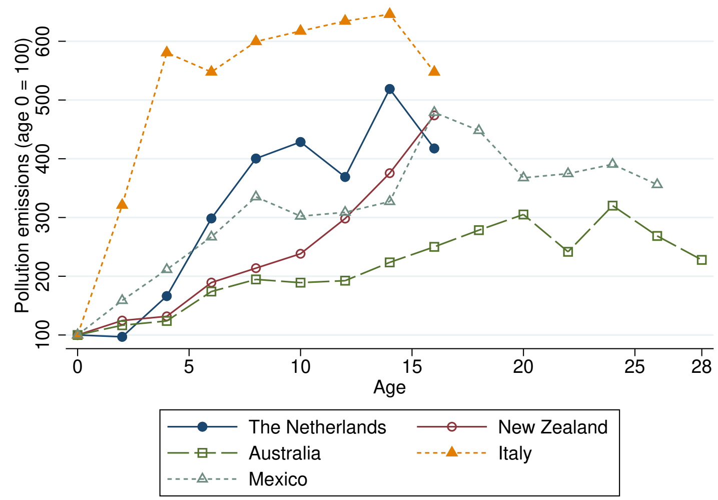
NOTES: Each panel uses full sample, restricted to model years 1982-2010. See text for explanation of mileage and age at which these standards apply, and for comparing different measures of HC over model years. Beginning in 1988 for NO_x and 1994 for other pollutants, graphs show weighted means across truck types, with weights equal to the proportion of each vehicle from model year 1993 in Colorado smog check data. Graphs show fitted values for model year plus a constant (for cars) or plus model year interacted with truck indicator plus a constant (for trucks) from regressions that also control for age fixed effects. Dashed vertical lines show years standards change for cars only; solid vertical liens show years when standards change for both cars and trucks.

Figure A4: Emissions by Age, Separately by Country, State, and for Heavy-Duty Trucks

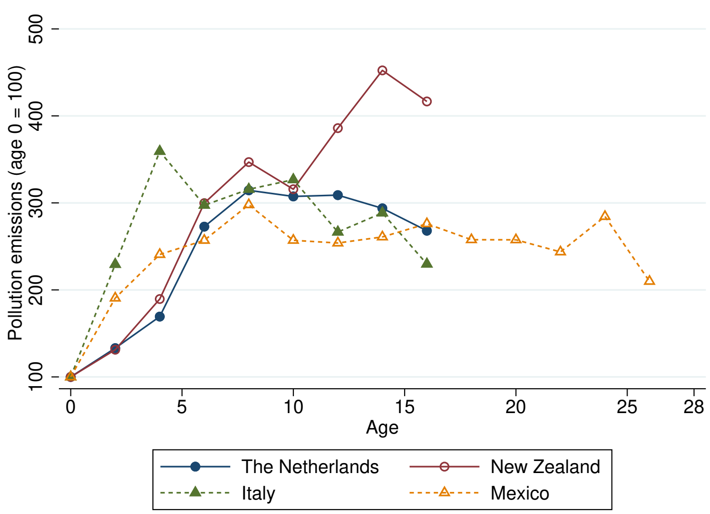
(a) By country: carbon monoxide (CO)



(b) By country: hydrocarbons (HC)



(c) By country: nitrogen oxide (NO)



(d) By country: carbon dioxide (CO_2)

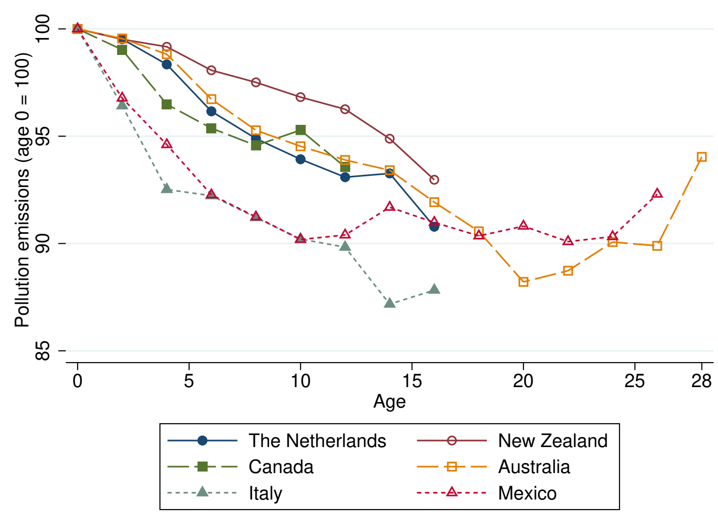
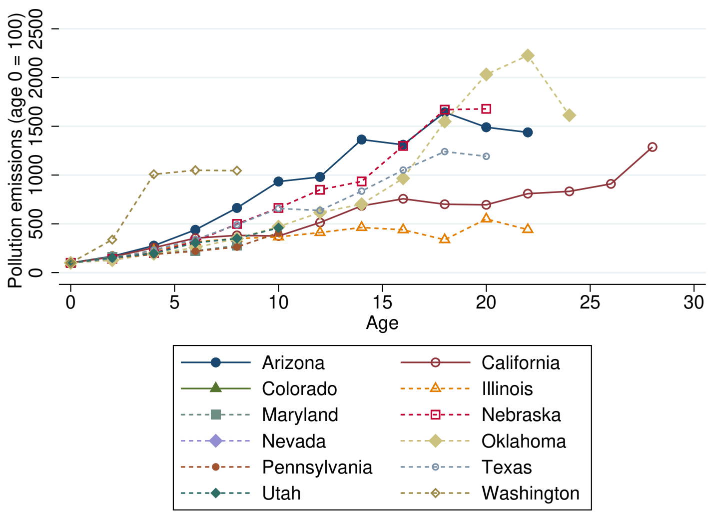
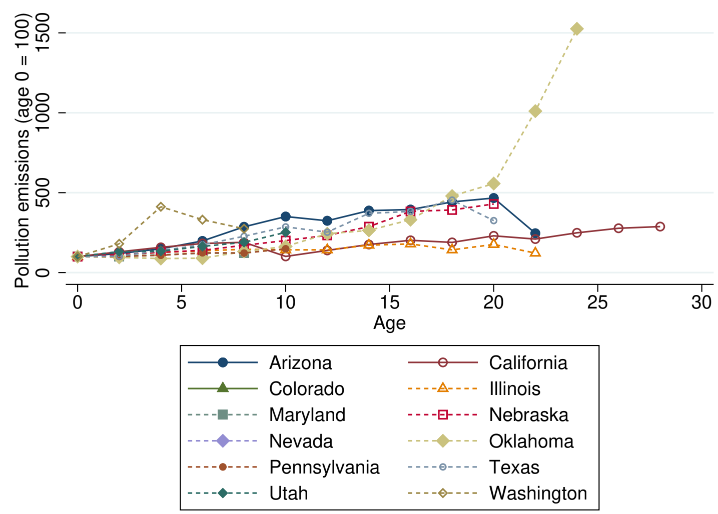


Figure A4: Emissions by Age, Separately by Country, State, and for Heavy-Duty Trucks (Continued)

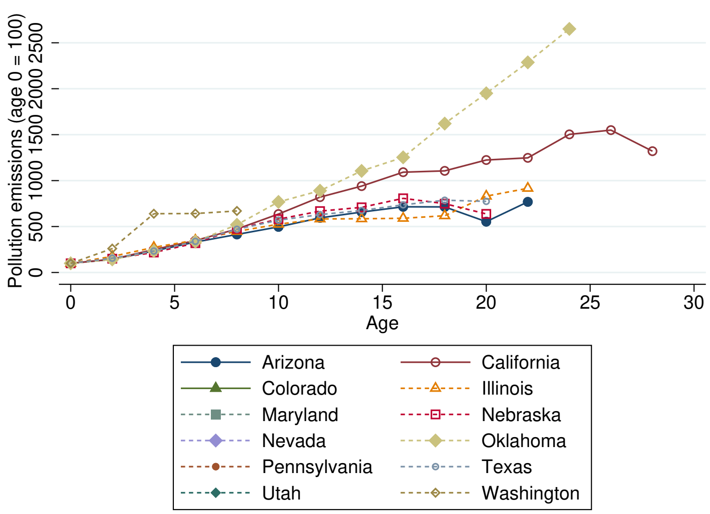
(e) By state: carbon monoxide (CO)



(f) By state: hydrocarbons (HC)



(g) By state: nitrogen oxide (NO)



(h) By state: carbon dioxide (CO_2)

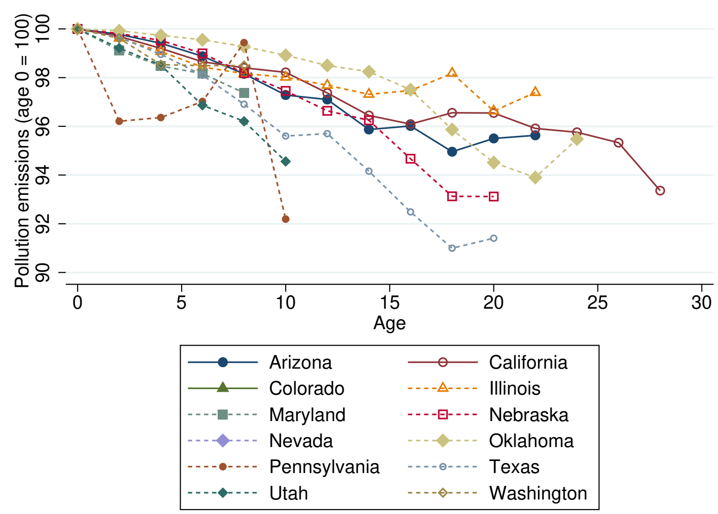
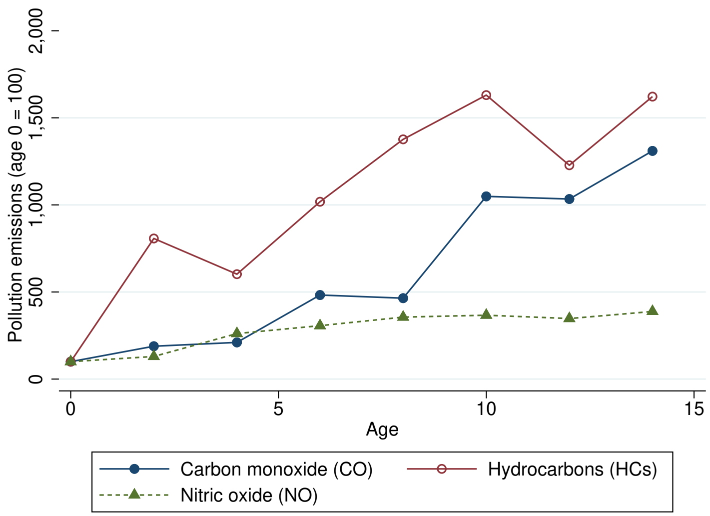


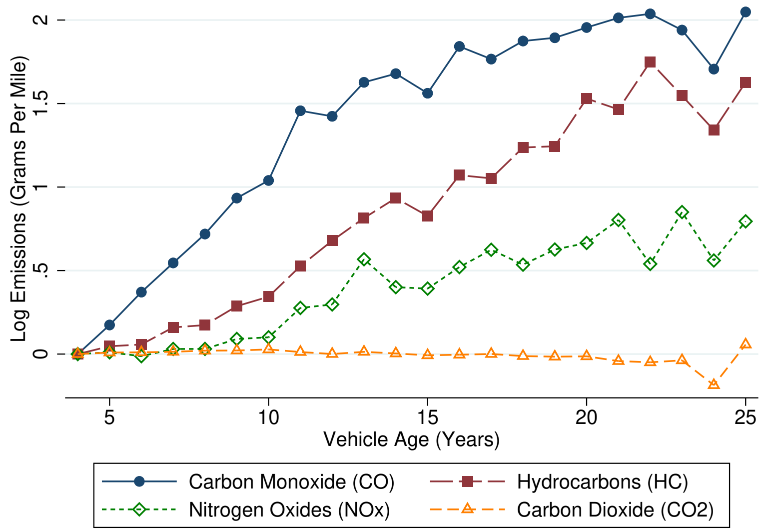
Figure A4: Emissions by Age, Separately by Country, State, and for Heavy-Duty Trucks (Continued)

(i) Heavy duty trucks



Notes: Graphs use roadside remote sensing data from Zhang et al. (1995), Bishop et al. (1997), and Xie et al. (2005). The value for the lowest age group in each category is normalized to 100. Graphs group ages 1 and earlier (including model years after the observed driving year) into age 1.

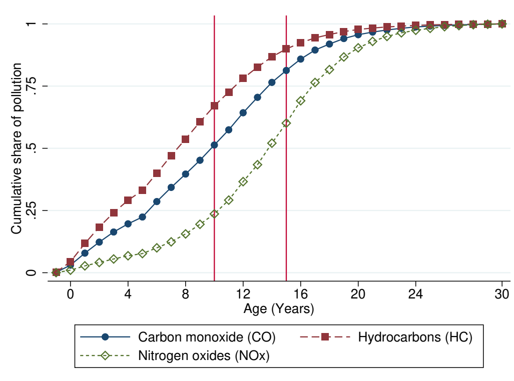
Figure A5: Air Pollution but Not CO_2 Increases with Vehicle Age



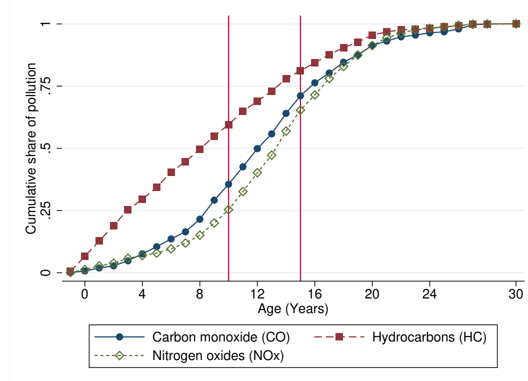
NOTES: Graph shows age fixed effects α_a from a regression including vehicle fixed effects μ_i and controls for odometer and odometer squared o: $E^u_{it} = \sum_j \alpha_j 1[age_{it} = j] + \gamma_1 o_{it} + \gamma_2 o^2_{it} + \mu_i + \epsilon_{it}$. Regression uses Colorado 240-second sample.

Figure A6: Cumulative Share of Fleet Emissions from Each Vehicle Age, Alternative Estimates

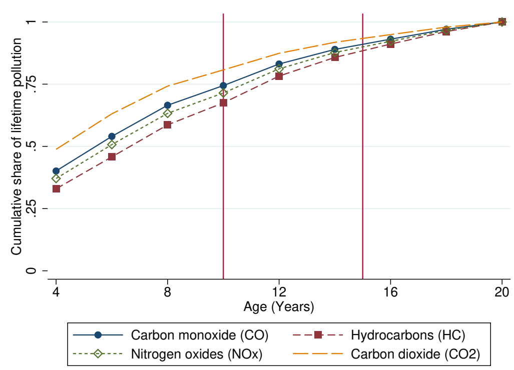
(a) 2014 fleet, Colorado remote sensing



(b) 2014 fleet, multi-state remote sensing

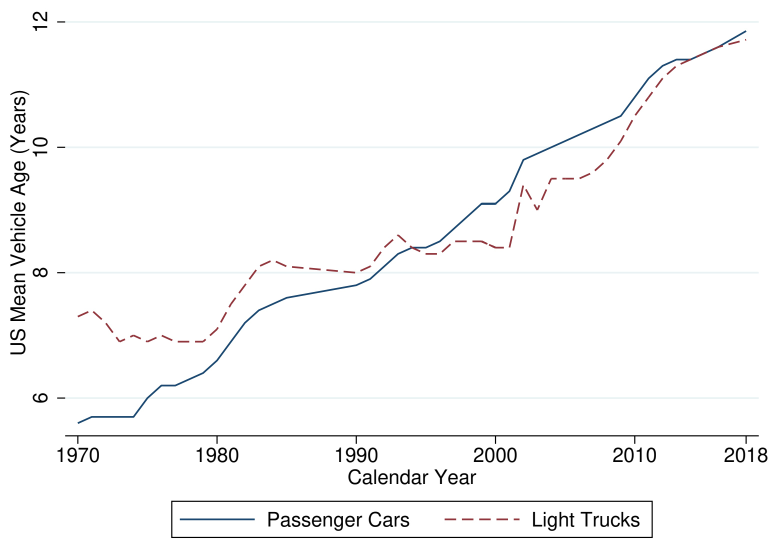


(c) 1993 cohort, Colorado inspections



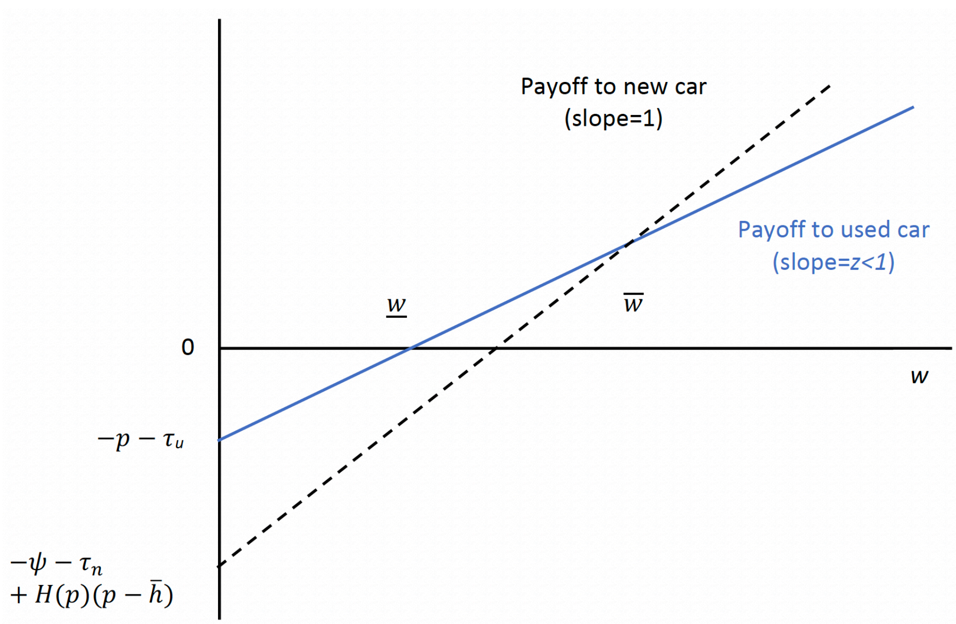
Notes: Each line shows the cumulative distribution for total pollution emissions from each age. Each pollutant is a separate line. Vertical lines at ages 10 and 15 show when exhaust standards stop applying. Pollution for an individual vehicle equals the emission rate measured in an individual test times miles driven. Miles driven is calculated as the change in a vehicle's odometer since the last test for that Vehicle Identification Number divided by the number of decimal years since the last test for that Vehicle Identification Number. For a vehicle's first test, this value of years is assumed to equal the vehicle's age. In Panels A and B, we assume that the number of times each vehicle passes a remote sensing detector is proportional to the vehicle's miles driven, so each value equals the share of total emissions detected by remote sensing that come from each age group.

Figure A7: US Mean Vehicle Age, by Calendar Year



NOTES: Data from Davis and Boundy (2021).

Figure A8: Schematic of Choice with an Outside Good

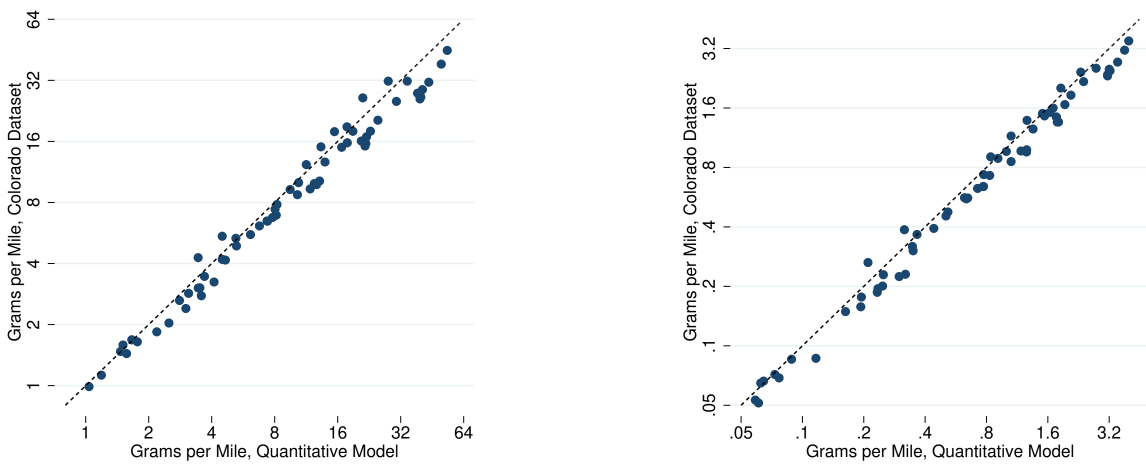


NOTES: Figure depicts the payoff to the three options of outside good, used vehicle, and new vehicle as a function of w. For $w < \underline{w}$, the outside good (value 0) will have the highest payoff. Between $\underline{w}$ and $\overline{w}$, the used vehicle has the highest value. If $w > \overline{w}$, the new vehicle will be chosen.

Figure A9: Quantitative Model Calibration Versus Emissions by Age and Class

(a) Carbon monoxide (CO)

(b) Hydrocarbons (HC)

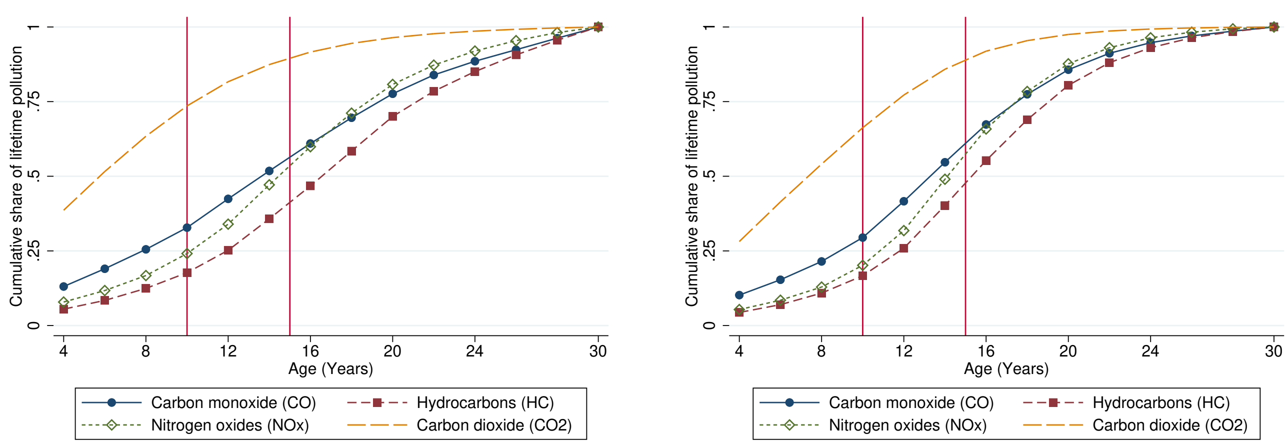


NOTES: Figures compare the quantitative model calibration to the full sample from Colorado smog check data. Points represent mean emission rates in a given model year×age×vehicle class cell, averaged across all vehicles in the data. Axes have logarithmic scale.

Figure A10: Cumulative Emissions: Quantitative Model Versus Colorado Sample

(a) Quantitative Model

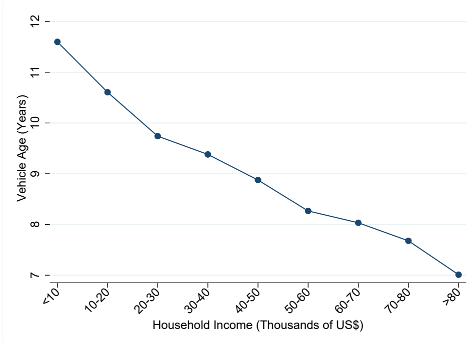
(b) Colorado Smog Check Data



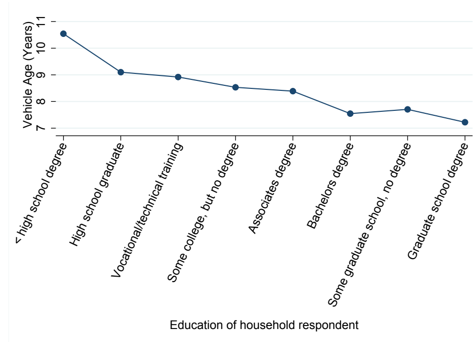
NOTES: Figure compares cumulative emissions in the quantitative model (calibrated to the national vehicle age profile) to that in the Colorado smog check data.

Figure A11: Vehicle Age Across Demographic Groups

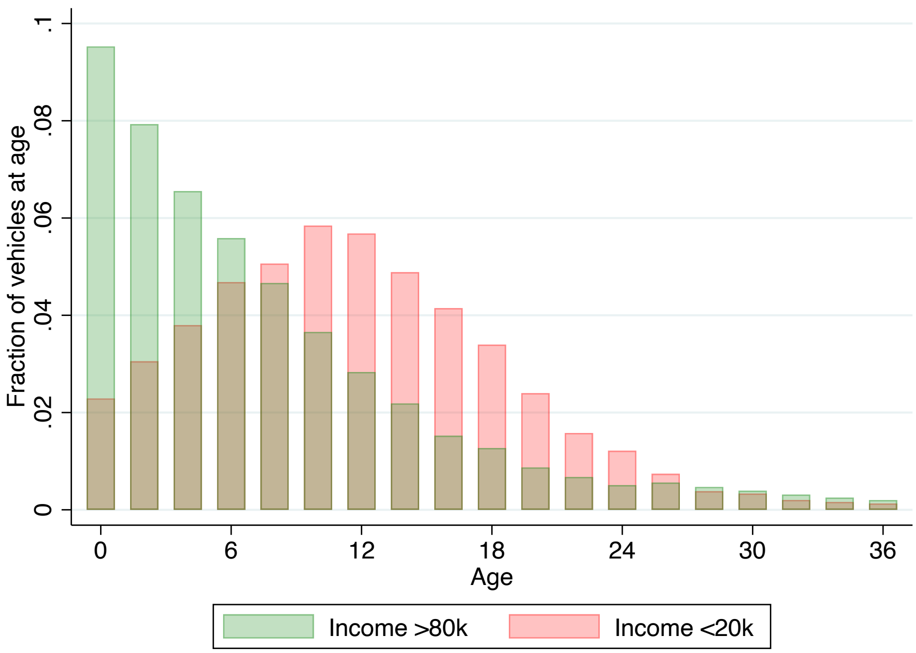
(a) Average Vehicle Age by Income



(b) Average Vehicle Age by Education Level



(c) Vehicle Age Distribution by Income Group



Notes: Data from National Household Travel Survey 2001 (U.S. Federal Highway Administration 2001). Income measure is total income across all household members (HHINCTTL). Education measure represents the education level of the household respondent (HHR_EDUC).

Table A1: Colorado Remote Sensing Versus Smog Check

	Carbon Monoxide (CO)	Hydrocarbons (HC)	Nitrogen Oxides (NO_x)
	(1)	(2)	(3)
Panel A: Regress remote sensing on smog check (inverse hypersine)			
Smog check	0.0989426***	0.5273313***	2.9836975***
	(0.0028322)	(0.0150110)	(0.0338312)
Panel B: Regress smog check on remote sensing (inverse hypersine)			
Remote sensing	0.1600617***	0.0138776***	0.0277832***
	(0.0050490)	(0.0003593)	(0.0003623)
Panel C: Regress remote sensing on smog check (g/mi)			
Smog check	0.0100927***	4.5860504***	434.9084140***
	(0.0007075)	(0.5802705)	(25.1872745)
Panel D: Regress smog check on remote sensing (g/mi)			
Remote sensing	0.1407309***	0.0000597***	0.0000151***
	(0.0107362)	(0.0000052)	(0.0000009)

NOTES: Data includes 65,327 observations. Each observation represents the mean pollution for a 17-digit VIN (an individual vehicle) in a particular week and year. To be in the sample, a VIN must appear in the Colorado remote sensing data in a given week and the Colorado smog check data the following week; this matched observation is used in the analysis. Standard errors clustered by 17-digit VIN. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

Table A2: Datasets and Samples

		Main data			Data used for sensitivity analyses		
	Sample	New vehicle tests	Older tests (AES 1973)	Colorado smog check	Colorado remote sensing	Multi-state remote sensing	In-use
		(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Characteristics of full sample							
Model years	Full	1972-2019	1957-'71	1982-2010	1984-2017	1982-2016	2004-'14
Calendar (test) years	Full	1972-2019	1972	1997-2014	2009-2016	1988-2015	2004-'17
N	Full	32,985	851	11,670,943	49,322,100	1,146,026	10,720
Type of test	Full	FTP	FTP	IM240	Rapid-Screen	FEAT	FTP
Panel B. Number of observations in each sample							
N	1982-2000	9,120	—	8,612,261	11,329,026	823,621	—
N	2000-2010	7,761	—	3,667,890	33,538,516	295,890	7,861
N	1993 cohort	520	—	652,195	432,286	44,151	—
N	2000 fleet	734	—	591,245	0	61,669	—
N	2014 fleet	960	—	854,035	6,324,084	5,875	—

NOTES: FTP is federal test procedure, IM240 is inspection and maintenance test for 240 seconds. The year listed for fleet sample ('90, '00, etc.) refer to calendar (test) year when a dataset measures emissions, not to model years. Some figures and tables use subsets of the indicated sample in cases where the variable(s) of interest are not available in observations (e.g., data distinguishing truck types are only available for model years 1982-2010).

Table A3: How Do Tier 1 Exhaust Standards Affect Vehicle Emissions? Sensitivity Analyses

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A. Carbon monoxide and hydrocarbons (CO and HC)								
Exhaust standard	0.81***	0.83***	0.96***	0.62***	0.71***	0.74***	0.13***	0.32***
	(0.07)	(0.07)	(0.08)	(0.05)	(0.10)	(0.10)	(0.04)	(0.07)
N	17,066,835	14,378,345	16,193,779	17,066,835	1,825,361	2,968,700	20,969,068	649,340
Panel B. Carbon monoxide (CO)								
Exhaust standard	0.75***	0.77***	0.83***	0.64***	0.63***	0.66***	0.09***	0.28***
	(0.07)	(0.06)	(0.07)	(0.05)	(0.08)	(0.08)	(0.02)	(0.07)
N	8,518,949	7,175,846	8,082,646	8,518,949	911,107	1,484,350	10,484,534	324,670
Panel C. Hydrocarbons (HC)								
Exhaust standard	1.05***	1.02***	1.80***	0.57***	1.22***	1.25***	0.61***	0.72***
	(0.17)	(0.16)	(0.18)	(0.09)	(0.24)	(0.25)	(0.20)	(0.10)
N	8,547,886	7,202,499	8,111,133	8,547,886	914,254	1,484,350	10,484,534	324,670
Main estimates	X							
Exclude 1994-5		X						
Emissions per gallon			X					
Truck type disaggregate				X				
Registration data					X			
Selection correction						X		
Colorado remote sensing							X	
Multi-state remote sensing								X

Notes: The dependent variable is the emission rate. Each observation is an individual vehicle. Emission rates and standards are in logs in columns (1) through (6) and inverse hyperbolic sine for columns (7) and (8). Columns (1) through (6) use model years 1982-2000 of Colorado inspections data. Main estimates in column (1) correspond to column (1) of Table 3. See paper text for details of other estimates. Standard errors are clustered by model year $\times$ light duty truck (LDT) type. In column (8), fixed effects differ by remote sensing study. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

Table A4: Effects of Tier 1 Exhaust Standards on Intermediate Outcomes and Mechanisms

	Carbon monoxide (CO)		Hydrocarbons (HC)	
	(1)	(2)	(3)	(4)
Effects of exhaust standards on …				
1. Used vehicle emissions	0.75***	0.52*	1.90***	1.55
	(0.10)	(0.27)	(0.36)	(0.94)
2. Used vehicle emissions: Within-engine changes	0.42***	0.38**	1.17***	1.40***
	(0.08)	(0.16)	(0.23)	(0.39)
3. Miles per gallon	0.03	-0.03	0.07	-0.15
	(0.05)	(0.11)	(0.17)	(0.33)
4. Vehicle retail price	-0.18**	-0.03	-0.52*	0.07
	(0.08)	(0.16)	(0.27)	(0.49)
5. Curb weight	-0.02	0.03	-0.07	0.11
	(0.07)	(0.15)	(0.23)	(0.45)
6. Horsepower	-0.08	-0.13	-0.23	-0.29
	(0.07)	(0.12)	(0.23)	(0.40)
7. Torque	-0.03	-0.05	-0.06	-0.03
	(0.10)	(0.21)	(0.33)	(0.64)
8. Engine displacement	-0.02	-0.02	-0.02	0.05
	(0.10)	(0.22)	(0.33)	(0.67)
Model year × truck trends	—	X	—	X

NOTES: Data cover model years 1990-2000. Rows 1-2 use Colorado inspection data. Rows 4-8 use new vehicle data. Standard errors clustered by model year×truck type have p < 0.10, 0.05, or 0.01 (*, **, ***).

Table A5: Tier 2: Do New Vehicle Emissions Predict Used Vehicle Emissions? Other Data

	In-use tests		Colorado remote sensing		Multi-state remote sensing	
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. Carbon monoxide (CO)					
New vehicle emissions	0.637***	0.630***	0.107***	0.092***	0.504***	0.397***
	(0.015)	(0.015)	(0.003)	(0.003)	(0.011)	(0.011)
N	7,839	7,839	36,313,589	36,313,589	296,657	296,657
	Panel B .Hydrocarbons (HC)					
New vehicle emissions	0.791***	0.771***	4.296***	4.713***	0.849***	0.173***
	(0.055)	(0.057)	(0.098)	(0.113)	(0.054)	(0.061)
N	7,765	7,765	36,354,100	36,354,100	296,999	296,999
	Panel C. Nitrogen oxides (NO_x)					
New vehicle emissions	0.623***	0.582***	9.210***	7.821***	0.886***	0.627***
	(0.043)	(0.046)	(0.267)	(0.305)	(0.020)	(0.026)
N	7,793	7,793	36,328,110	36,328,110	296,795	296,795
Age	—	X	—	X	—	X
Model year FE	—	X	—	X	—	X

NOTES: See Table 4 notes. Columns (1) and (2) use logs, columns (3) through (6) use inverse hypersine. Standard errors clustered by VIN prefix. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

Table A6: Used Vehicle Emissions, by Age and Model Year

	(1)	(2)	(3)	(4)	(5)	(6)
			Panel A. Carbon monoxide (CO)			
Age	0.018**	0.029**	0.023*	0.023***	0.031***	0.022***
	(0.008)	(0.013)	(0.013)	(0.003)	(0.004)	(0.004)
Model Year	-0.138***	—	—	-0.135***	—	—
	(0.006)	—	—	(0.003)	—	—
Age ×Model Year	0.000	0.000	0.000	—	—	—
	(0.001)	(0.001)	(0.001)	—	—	—
Odometer	—	—	0.110***	—	—	0.109***
	—	—	(0.009)	—	—	(0.008)
N	11,474,087	11,474,087	11,474,087	11,474,087	11,474,087	11,474,087
			Panel B .Hydrocarbons (HC)			
Age	0.019**	0.103***	0.093***	0.038***	0.049***	0.036***
	(0.009)	(0.017)	(0.016)	(0.004)	(0.006)	(0.006)
Model Year	-0.174***	—	—	-0.159***	—	—
	(0.006)	—	—	(0.005)	—	—
Age ×Model Year	0.002**	-0.004***	-0.005***	—	—	—
	(0.001)	(0.001)	(0.001)	—	—	—
Odometer	—	—	0.180***	—	—	0.151***
	—	—	(0.011)	—	—	(0.009)
N	11,616,611	11,616,611	11,616,611	11,616,611	11,616,611	11,616,611
			Panel C. Nitrogen oxides (NO_x)			
Age	-0.048***	0.079***	0.071***	0.024***	0.033***	0.024***
	(0.008)	(0.015)	(0.015)	(0.004)	(0.005)	(0.005)
Model Year	-0.176***	—	—	-0.120***	—	—
	(0.007)	—	—	(0.006)	—	—
Age ×Model Year	0.006***	-0.003***	-0.004***	—	—	—
	(0.001)	(0.001)	(0.001)	—	—	—
Odometer	—	—	0.125***	—	—	0.100***
	—	—	(0.009)	—	—	(0.009)
N	11,634,349	11,634,349	11,634,349	11,634,349	11,634,349	11,634,349
			Panel D. Carbon dioxide (CO_2)			
Age	0.000	-0.006***	-0.006***	0.003***	0.000	0.001
	(0.004)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Model Year	0.010**	—	—	0.012***	—	—
	(0.004)	—	—	(0.002)	—	—
Age ×Model Year	0.000	0.000***	0.000***	—	—	—
	(0.000)	(0.000)	(0.000)	—	—	—
Odometer	—	—	-0.008***	—	—	-0.005***
	—	—	(0.001)	—	—	(0.001)
N	11,669,895	11,669,895	11,669,895	11,669,895	11,669,895	11,669,895
VIN FE	—	X	X	—	X	X

NOTES: Estimates include full sample of Colorado smog check data. Model year is re-centered around 1981 (=raw model year − 1981). Standard errors are clustered by model year×truck type. All emission rates in logs. Asterisks denote p-value < 0.10 (*), <0.05 (**), <0.01 (***).

Table A7: Data and Parameters in the Quantitative Model

Parameter input	Symbol	Sources	Value(s)
Scrap elasticities	γ_a	Jacobsen and van Benthem (2015)	(-0.50, -1.02)
Pollution damages	θ	Tschofen et al. (2019)	\$1,045 (CO)
		Knittel and Sandler (2015)	\$15,047 (HC)
			\$35,566 (NO_x)
Discount rate	δ	U.S. Environmental Protection Agency	3.0% per year
GDP growth rate	—	U.S. Environmental Protection Agency	0.5% per year
Autonomous fuel economy improvement rate		Knittel (2011)	1.8% per year
Vehicle demand elasticities	ρ	Jacobsen and van Benthem (2015)	See Section F.1
Pollution reduction cost parameters	$\chi, \zeta_{c,s}$	U.S. Environmental Protection Agency	See Section F.6
Fuel economy cost parameters	$\kappa^1_{c,s}, \kappa^2_{c,s}$	National Research Council (2002)	See Section F.6
Data input	**Sources**	**Value(s)**	
Vehicle miles traveled	$vmt_{c,s,a}$	Colorado Dept. Public Health and Environment	—
Vehicle prices	$p_{c,s,a,m,t}$	National Automobile Dealers Association	—
		Kelley Blue Book	
Vehicle quantities	$q_{c,s,a,m,t}$	Wards Intelligence	—
		Federal Reserve Bank of St. Louis	
Inflation	—	U.S. Bureau of Labor Statistics	—
Scrap rates	$y_{c,s,a,m,t}$	R.L. Polk & Company	—
Fuel economy	$f_{c,s,0,m,t}$	U.S. Department of Energy	—
		National Highway Traffic Safety Administration	
		U.S. Environmental Protection Agency	
Pollution per mile	$\phi_{c,s,a,m,t}$	Colorado Dept. Public Health and Environment	See Section 3
		U.S. Environmental Protection Agency	
Pollution from manufacturing	$\Phi_{c,s,m,t}$	U.S. Bureau of Economic Analysis	See Section B.6
		National Emissions Inventory	
Vehicle registration fees	$\tau_{c,s,a,m,t}$	Jacobsen et al. (2021)	—
Household vehicle characteristics	—	U.S. Federal Highway Administration	See Figure A7
GDP (2000)	M	U.S. Bureau of Economic Analysis	\$15.22 trillion
Gasoline price (2000)	p_{gas}	U.S. Energy Information Administration	\$2.24/gallon

NOTES: In column 2, — indicates values used in the quantitative model but not in equations in this paper. In column 4, — indicates values in the paper's replication files but not easily summarized in one value here. Dollar values in \$2019.

Table A8: Technology and Timing of Exhaust Standards

	Compliance			Overcompliance		
	Year t	Year t-4	Year t-8	Year t	Year t-4	Year t-8
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Tier 2 (2004)						
Cars HC: 0.125	1.00	0.84	0.53	0.88	0.47	0.02
Trucks HC: 0.139	1.00	0.86	0.44	0.75	0.25	0.02
Trucks NO_x: 0.40	1.00	0.96	0.85	0.97	0.66	0.29
Panel B: Tier 2 (2007)						
Cars HC: 0.100	1.00	0.91	0.67	0.91	0.69	0.23
Cars NO_x: 0.14	1.00	0.93	0.65	0.99	0.60	0.21
Trucks HC: 0.10	1.00	0.83	0.40	0.79	0.32	0.06
Trucks NO_x: 0.14	1.00	0.66	0.43	0.98	0.26	0.09

NOTES: "Compliance" describes the share of vehicle types with emission rate less than the standard for the indicated standard year, class, pollutant and model year. Tier 2 began in 2004 but its standards tightened in 2007. Overcompliance describes the share of vehicle types with emission rate less than 50% of the standard for the indicated standard year, class, pollutant, and model year. Year t indicates the year of implementation listed in each row (1996, 2004, or 2007); Years t-4 and t-8 describe four and eight years earlier. Table shows only the pollutants and vehicles where a Tier changes standards (e.g., Tier 1 did not change CO standards for cars).

Table A9: Quantitative Model-Based Estimates: Incidence of Fees by Income Group

Income bin:	<10k	10-20k	20-30k	30-40k	40-50k	50-60k	60-70k	70-80k	>80k
Panel A. Baseline fees per household									
1. Baseline	6.5	7.8	10.0	11.9	13.7	15.5	16.7	18.1	20.2
Panel B. Changes per household when counterfactual registration fees are applied									
At baseline vehicle choice:									
2. Age x type fee	175.0	170.7	185.7	196.0	204.5	201.9	199.6	200.3	201.4
At equilibrium vehicle choice:									
3. Age x type fee	114.0	112.4	122.1	135.1	141.6	143.5	144.7	146.8	149.4
4. Age x type fee, revenue neutral	28.0	19.1	15.1	12.1	4.7	-3.8	-10.0	-15.2	-25.0
5. New vehicle fee	11.2	14.5	24.2	32.8	40.9	50.2	57.0	67.0	87.0
6. Flat registration fee	3.9	3.0	1.9	1.3	0.5	-0.6	-1.3	-2.2	-3.5
Panel C. Notes for interpretation									
Number of vehicles per household	1.36	1.48	1.69	1.90	2.07	2.20	2.29	2.38	2.53
Fraction of households in income bin	0.04	0.10	0.13	0.14	0.12	0.11	0.08	0.07	0.21

NOTES: Annualized fees are in 2019 dollars and expressed as the discounted sum of fees paid over 20 years, divided by 20. Baseline fees shown include only payments proportional to vehicle value; fixed charges per vehicle are not included.

Table A10: Quantitative Model-Based Estimates: Sensitivity Analyses

	Change in market surplus	Change in pollution damages	Total change in social welfare = (1) - (2)	New tax revenue	Percent change in cumulative emissions		
					CO	HC	NO_x
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A. Baseline case							
1. Delay Tier 2 by eight years	13.3	198.2	-184.9	0.0	15.8	8.1	17.8
2. Age×type fee	-170.6	-492.5	321.9	1,167.5	-42.3	-42.7	-24.6
Panel B. Alternative elasticities, delay eight years							
3. 50% lower scrap elasticity	13.1	199.6	-186.5	0.0	15.9	8.2	17.9
4. 50% higher scrap elasticity	13.5	196.8	-183.3	0.0	15.7	8.0	17.7
5. 50% lower vintage substitution	13.3	197.6	-184.2	0.0	15.8	8.1	17.7
6. 50% higher vintage substitution	13.2	198.4	-185.1	0.0	15.8	8.1	17.8
Panel C. Alternative baselines, delay eight years							
7. More stringent CAFE standards	12.6	188.0	-175.4	0.0	15.0	7.6	16.8
8. Faster income growth	13.8	204.4	-190.6	0.0	16.2	8.3	18.2
9. Alternative VMT schedule	13.4	215.5	-202.2	0.0	17.9	9.2	19.2
10. Imperfect competition	15.7	199.5	-183.8	0.0	15.9	8.2	17.9
11. Higher gasoline price	13.6	201.2	-187.6	0.0	16.0	8.2	18.0
12. Higher internal discount rate	14.3	196.2	-181.9	0.0	15.7	8.0	17.6
Panel D. Alternative elasticities, age×type fee							
13. 50% lower scrap elasticity	-170.5	-475.7	305.3	1,185.3	-41.3	-41.5	-23.4
14. 50% higher scrap elasticity	-170.1	-497.4	327.3	1,162.8	-42.4	-43.1	-25.1
15. 50% lower vintage substitution	-143.0	-376.4	233.4	1,286.3	-32.0	-32.1	-18.6
16. 50% higher vintage substitution	-187.7	-558.3	370.6	1,101.1	-47.7	-48.4	-28.4
Panel E. Alternative baselines, age×type fee							
17. More stringent CAFE standards	-169.7	-493.5	323.8	1,168.7	-42.3	-42.6	-24.6
18. Faster income growth	-171.5	-493.3	321.8	1,178.8	-42.1	-42.6	-24.5
19. Alternative VMT schedule	-164.4	-449.9	285.5	1,189.5	-39.2	-39.9	-22.9
20. Imperfect competition	-164.7	-492.4	327.7	1,168.4	-42.1	-42.5	-24.6
21. Higher gasoline price	-161.1	-451.0	289.9	1,212.5	-39.2	-39.5	-22.2
22. Higher internal discount rate	-182.1	-496.1	314.0	1,162.7	-42.8	-43.2	-24.7
Panel F. Alternative policies							
23. 10% exhaust improvement, higher cost	-11.5	-26.8	15.3	0.0	-1.4	-1.0	-2.3
24. Small (10%) age-type registration fee	-5.3	-98.4	93.1	156.2	-9.8	-9.5	-4.2
25. Age-based registration fee	-180.9	-487.9	307.0	1,162.6	-41.7	-42.2	-24.5
26. Flat registration fee (from 0.68% base)	-7.9	-45.9	38.0	0.0	-4.2	-4.2	-2.5

NOTES: Values in are in billions of $2019 discounted to the base year. Social welfare is defined as consumer + producer surplus − pollution damages, which equals welfare for a social welfare function that abstracts from distribution. See paper text for details of each case. The smaller registration fee in row 24 is 10% of that in the baseline case in row 2.