Appendix

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

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 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 I).

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 III)

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 VIII 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 IV)

C.1 Fleet-Weighted New Vehicle Emission Trends

Figure I 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 IV 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 I. Mean used vehicle emission rates for each air pollutant fell by 90 percent between model years 1982 and 2010; new vehicle emissions from Figure I 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 I 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 V)

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

The beginning of Section V.B from the main text discusses graphs in Figure II 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 V.B 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 III.B, 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 VII)

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 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 a 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). We focus on cases where 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:

$$0 = z\underline{w} - p - \tau_u$$
$$\underline{w} = \frac{p + \tau_u}{z}.$$

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

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

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 denomi-

nator 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:

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

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 VIII)

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 VI 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 VII 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 III.B 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 VIII.A, 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 VIII. 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 V.

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 V) 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 V 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.

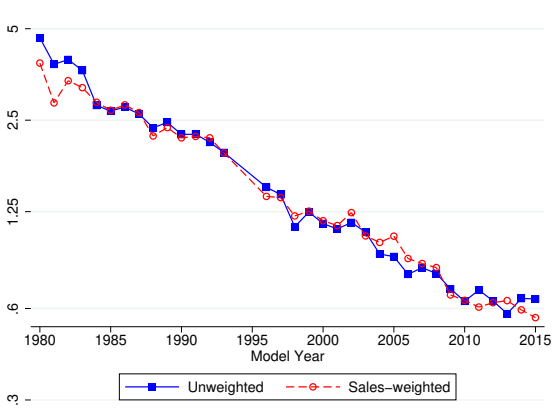
Xie, S., J. Bluett, G. Fisher, G. Kuschel, and D. Stedman (2005). On-road remote sensing of vehicle exhaust emissions in Auckland, New Zealand. *Clean Air and Environmental Quality 39*(4), 37–42.

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.

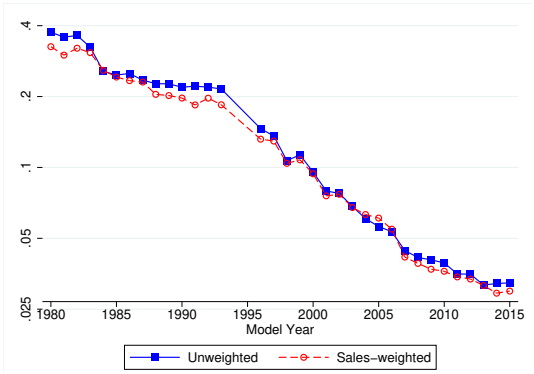
Appendix Figures and Tables

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

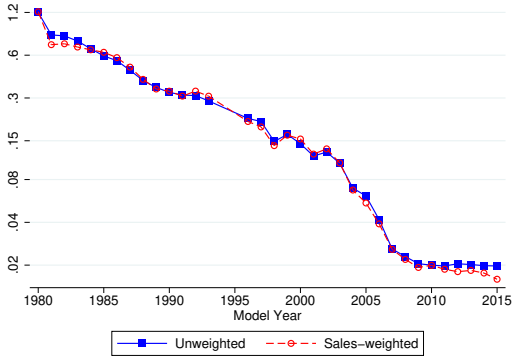
(A) Carbon monoxide (CO)



(B) Hydrocarbons (HC)



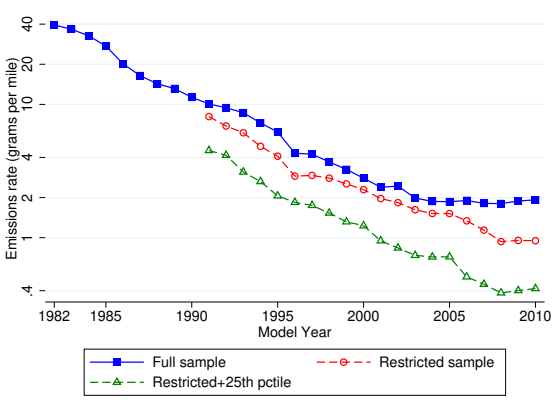
(C) Nitrogen oxides (NO_x)



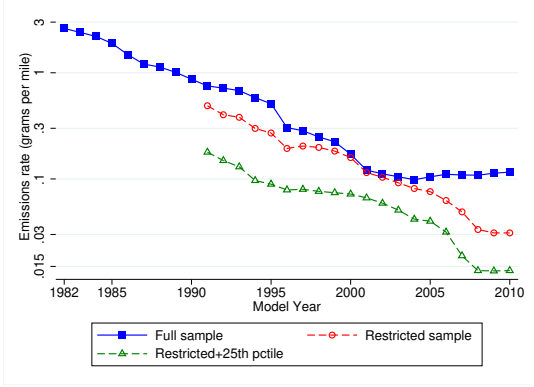
NOTES: Blue solid line shows unweighted trend reprinted from Figure I. Red dashed line shows the subset of Wards data for which we could accurately identify the emission rate, weighted by sales.

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

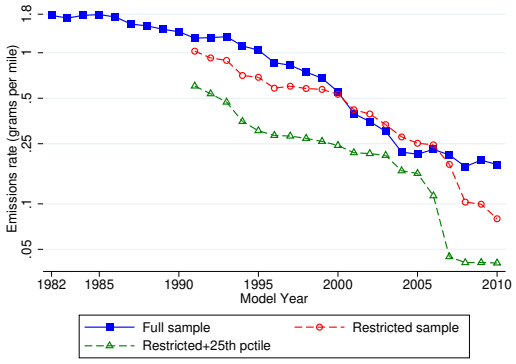
(A) Carbon monoxide (CO)



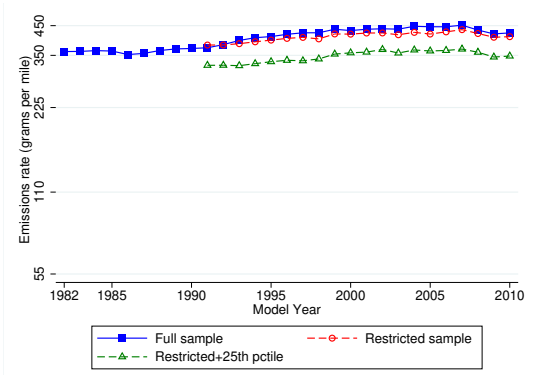
(B) Hydrocarbons (HC)



(C) Nitrogen oxides (NO_x)



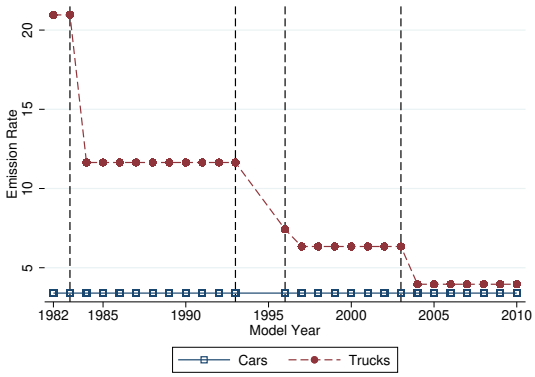
(D) Carbon dioxide (CO_2)



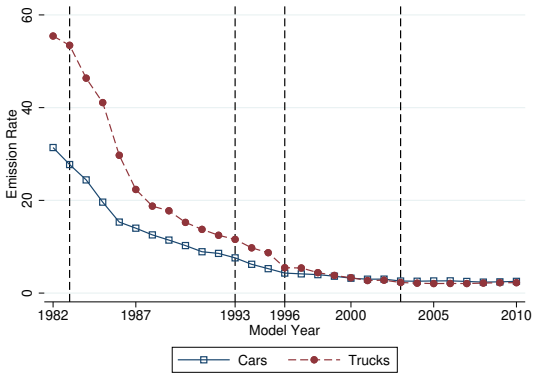
NOTES: Blue solid line shows full sample from Colorado smog check tests. Restricted sample limits the sample to 4 to 6 year old vehicles with 40,000 to 60,000 miles in model years 1991-2010. The 25th percentile line is estimated from quantile regressions. Graphs show fitted values for model year fixed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

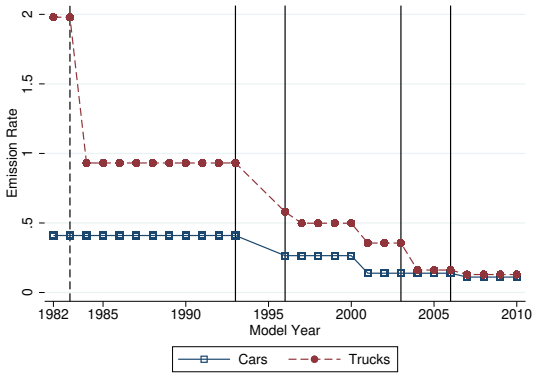
(a) Exhaust standards: CO



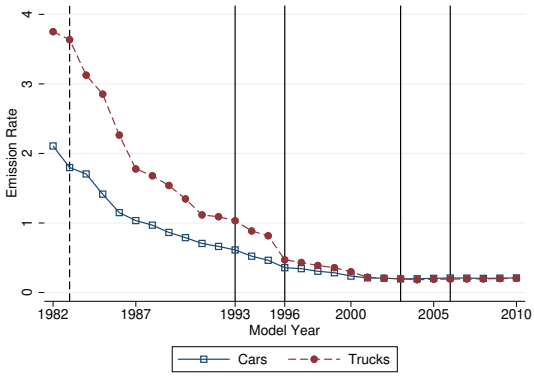
(b) Colorado smog check: CO



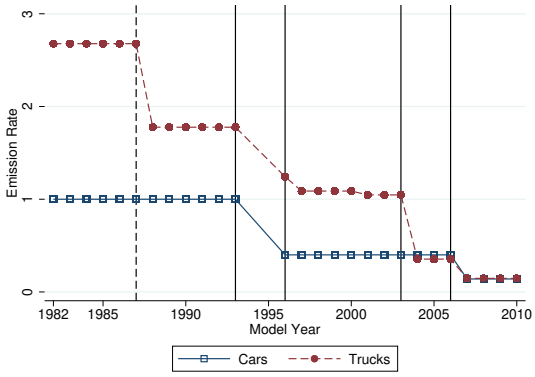
(c) Exhaust standards: HC



(d) Colorado smog check: HC



(e) Exhaust standards: NO_x



(f) Colorado smog check: NO_x

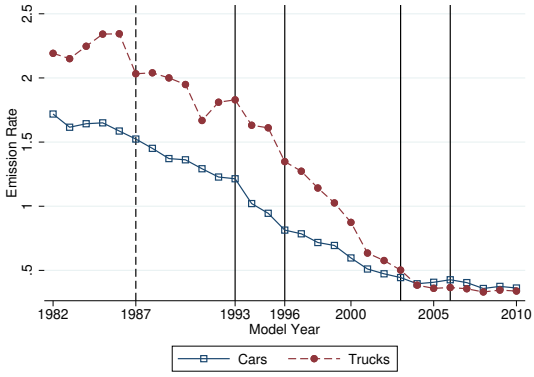
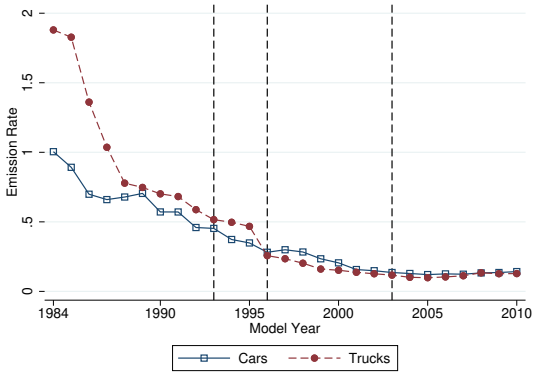
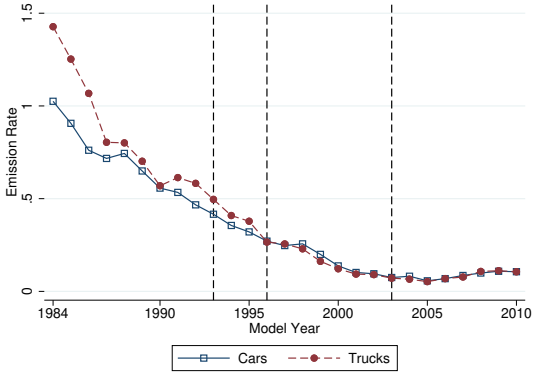


Figure A3: Exhaust Standards and Emission Rates, Cars Versus Trucks

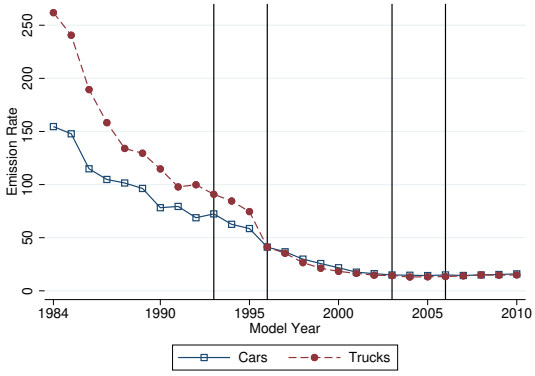
(g) Colorado remote sensing: CO



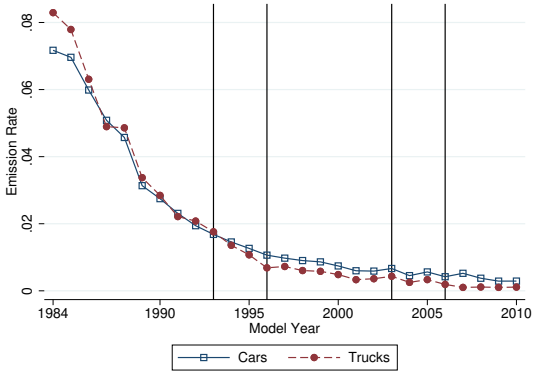
(h) Multi-state remote sensing: CO



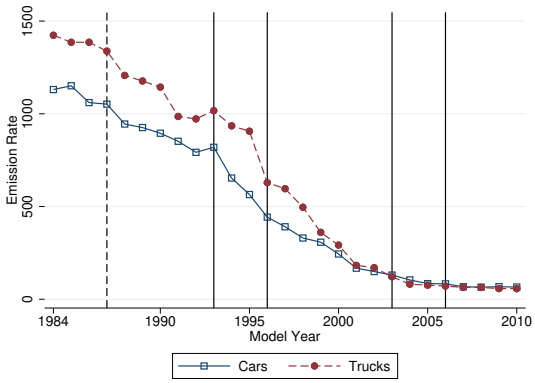
(i) Colorado remote sensing: HC



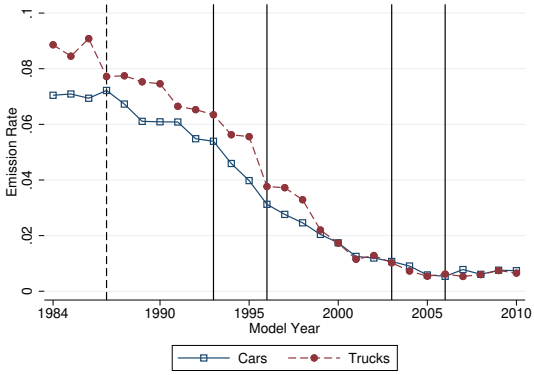
(j) Multi-state remote sensing: HC



(k) Colorado remote sensing: NO_x



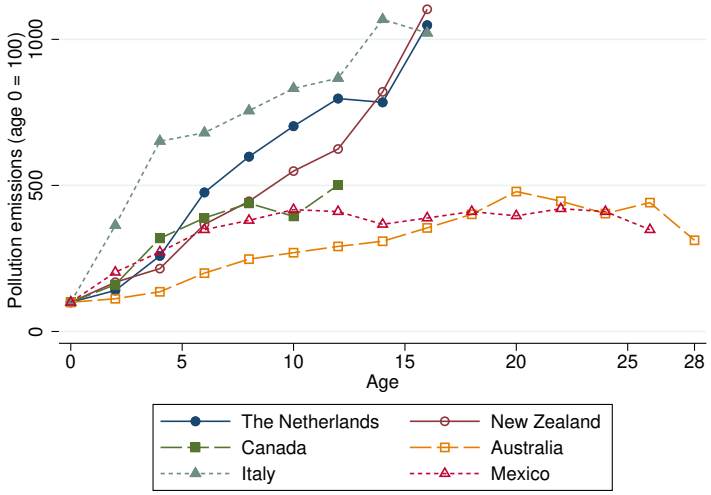
(l) Multi-state remote sensing: NO_x



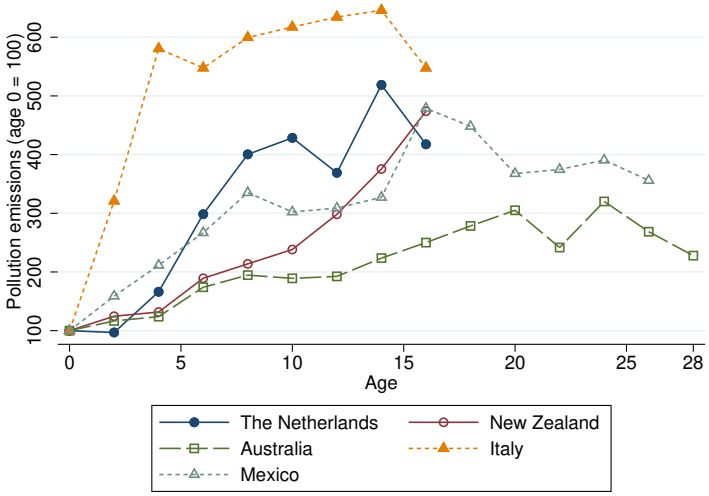
Notes: Each panel uses full sample, restricted to model years 1982-2010 (1984- for remote sensing). See text for explanation of mileage and age at which these standards apply, and for comparing different measures of HC over model years. Beginning in 1988 for NO_x and 1994 for other pollutants, graphs 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 lines show years when standards change for both cars and trucks.

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

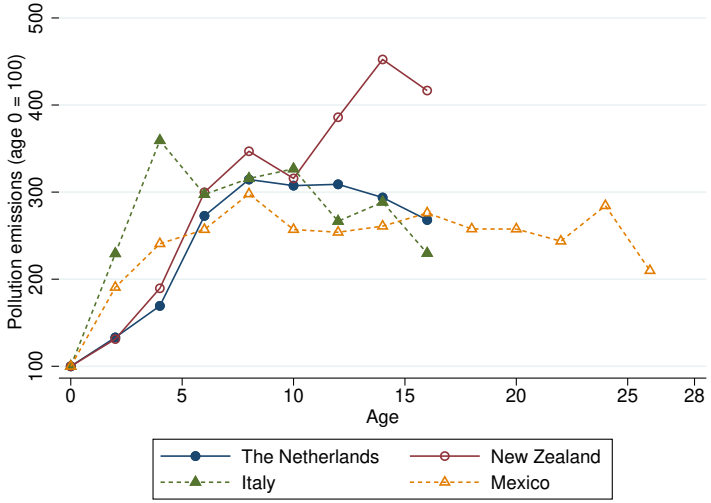
(A) By country: carbon monoxide (CO)



(B) By country: hydrocarbons (HC)



(C) By country: nitrogen oxide (NO)



(D) By country: carbon dioxide (CO_2)

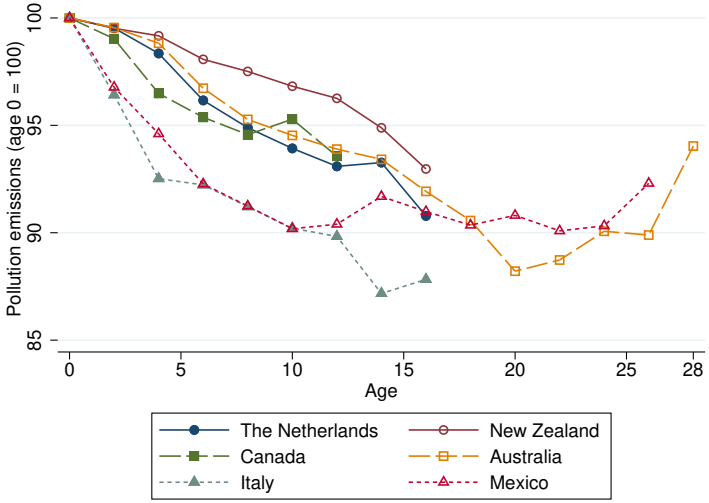
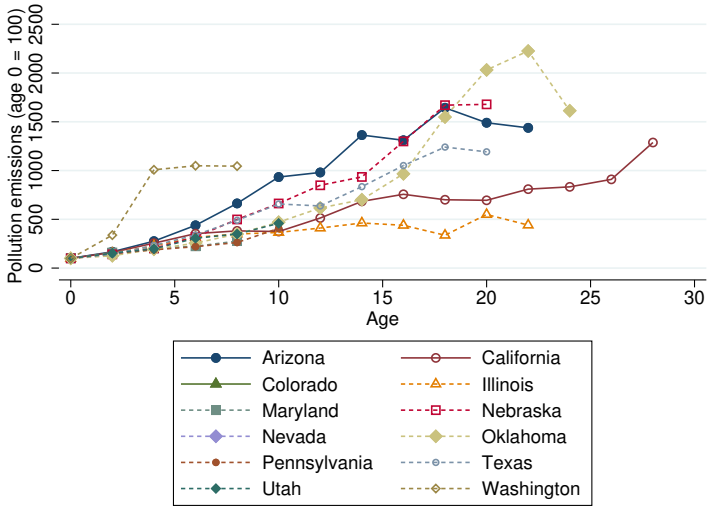
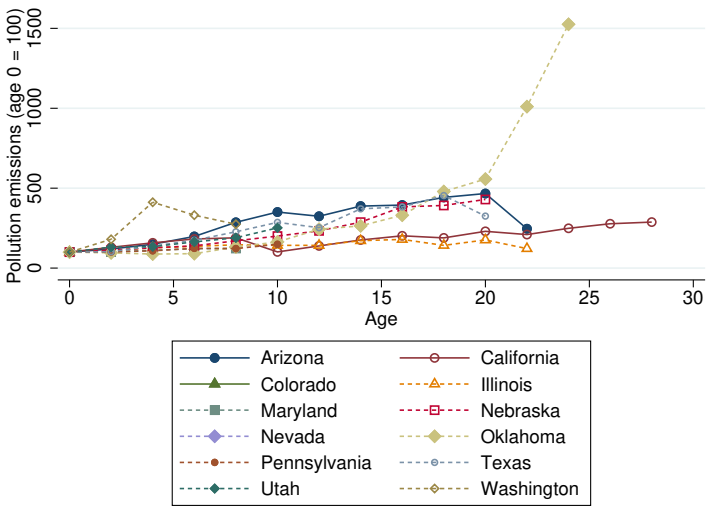


FIGURE A4: Emissions by Age, Separately by Country, State, and for Heavy-Duty Trucks (Continued)

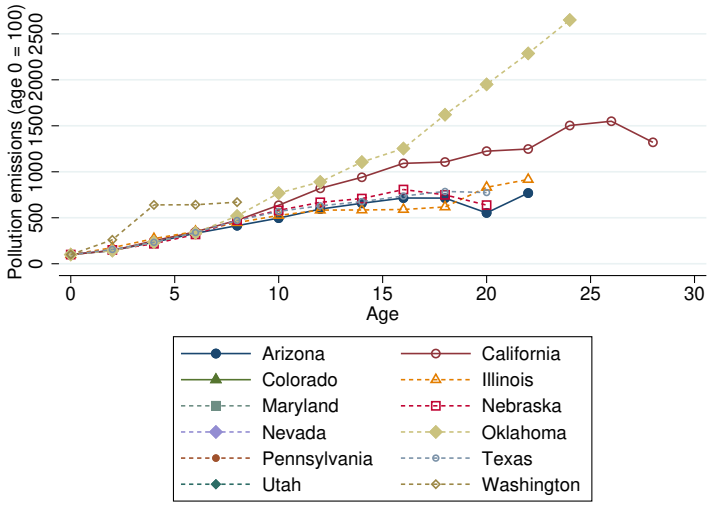
(E) By state: carbon monoxide (CO)



(F) By state: hydrocarbons (HC)



(G) By state: nitrogen oxide (NO)



(H) By state: carbon dioxide (CO_2)

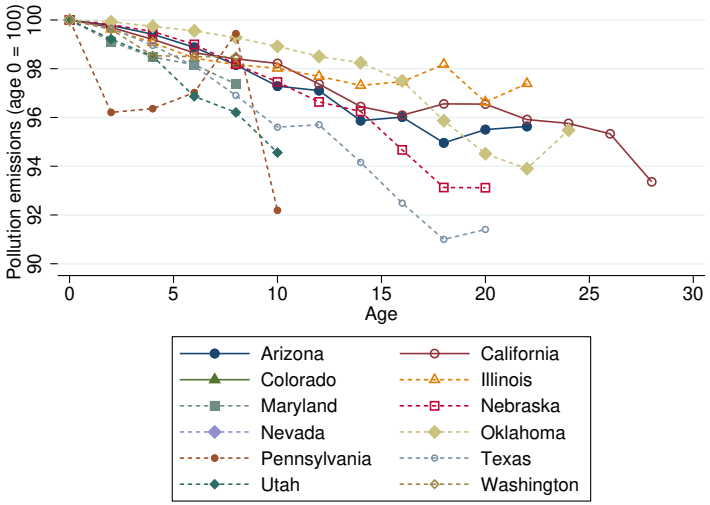
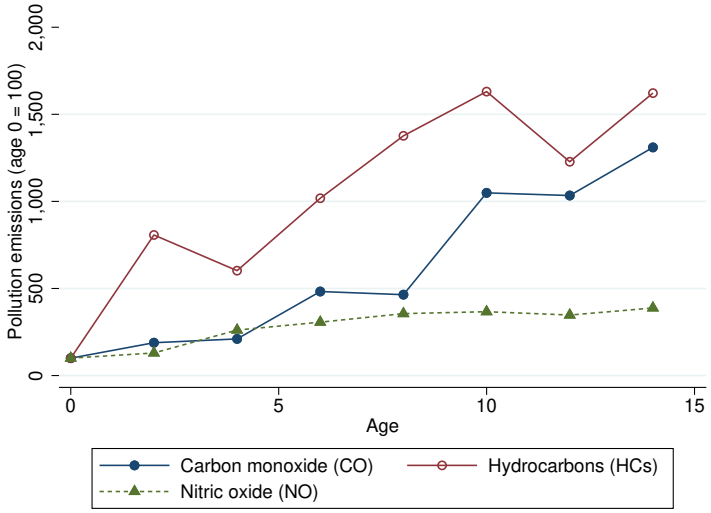


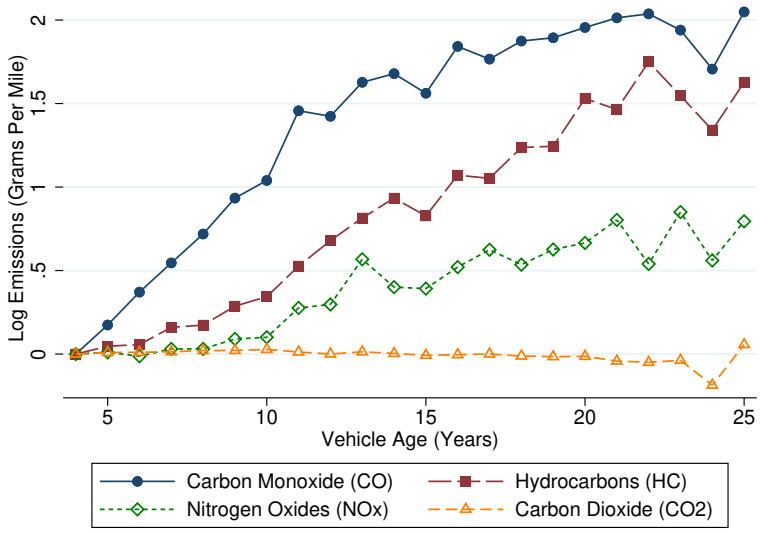
Figure A4: Emissions by Age, Separately by Country, State, and for Heavy-Duty Trucks (Continued)

(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.

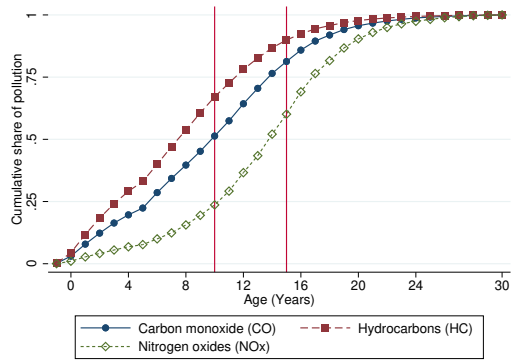
FIGURE A5: Air Pollution but Not CO_2 Increases with Vehicle Age



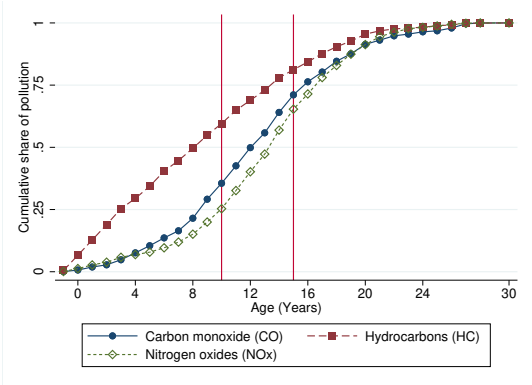
NOTES: Graph shows age fixed effects α_a from a regression including vehicle fixed effects μ_i and controls for odometer and odometer squared o: $E^u_{it} = \sum_j \alpha_j 1[age_{it} = j] + \gamma_1 o_{it} + \gamma_2 o^2_{it} + \mu_i + \epsilon_{it}$. Regression uses Colorado 240-second sample.

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

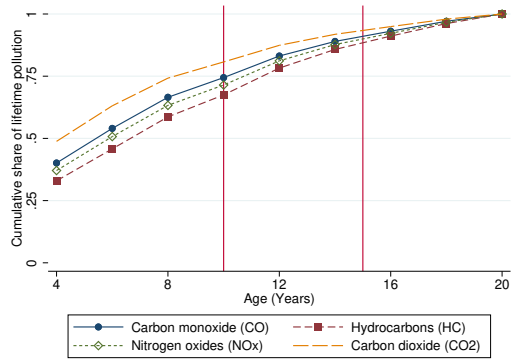
(a) 2014 fleet, Colorado remote sensing



(b) 2014 fleet, multi-state remote sensing

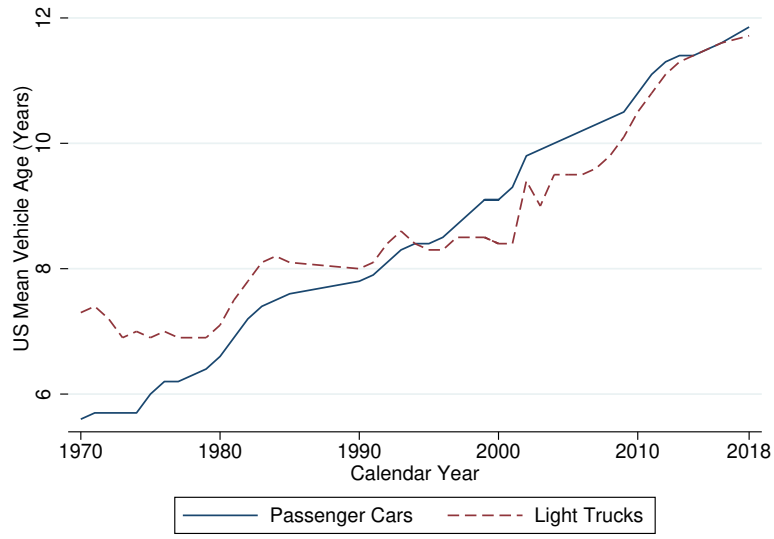


(c) 1993 cohort, Colorado inspections



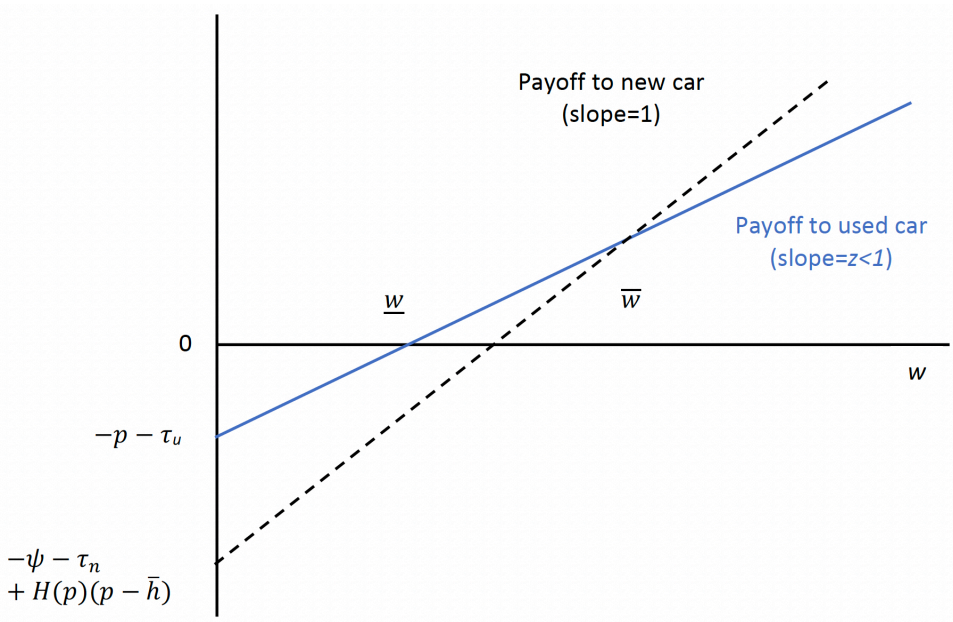
Notes: Each line shows the cumulative distribution for total pollution emissions from each age. Each pollutant is a separate line. Vertical lines at ages 10 and 15 show when exhaust standards stop applying. Pollution for an individual vehicle equals the emission rate measured in an individual test times 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).

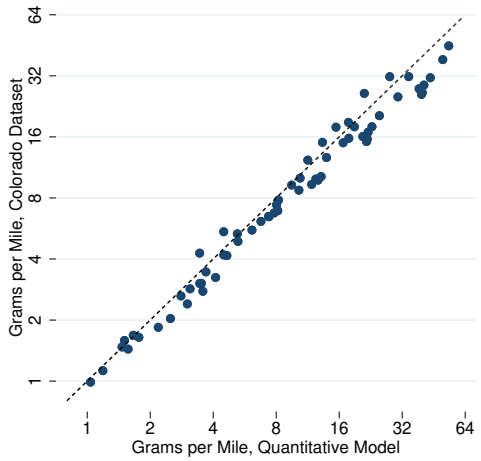
FIGURE A8: Schematic of Choice with an Outside Good



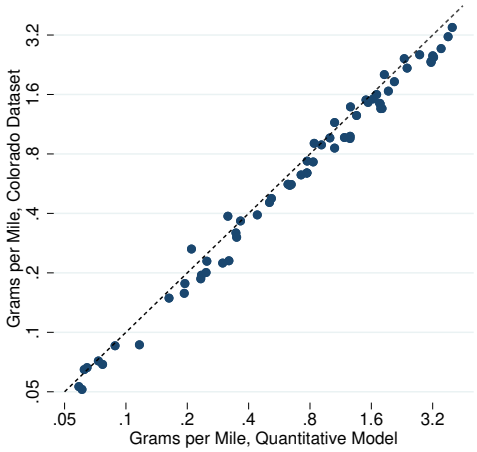
NOTES: Figure depicts the payoff to the three options of outside good, used vehicle, and new vehicle as a function of w. For $w < \underline{w}$, the outside good (value 0) will have the highest payoff. Between $\underline{w}$ and $\overline{w}$, the used vehicle has the highest value. If $w > \overline{w}$,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

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

(a) Carbon monoxide (CO)



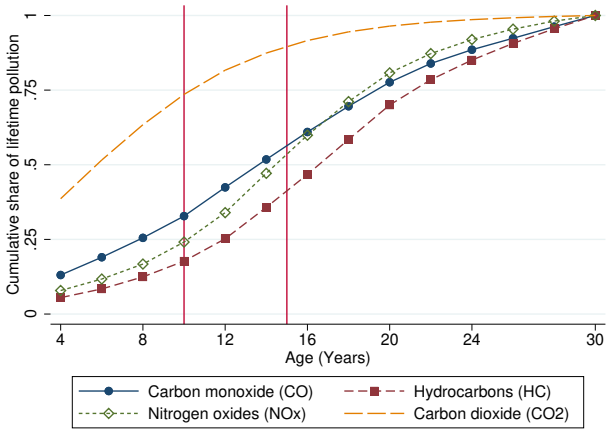
(b) Hydrocarbons (HC)



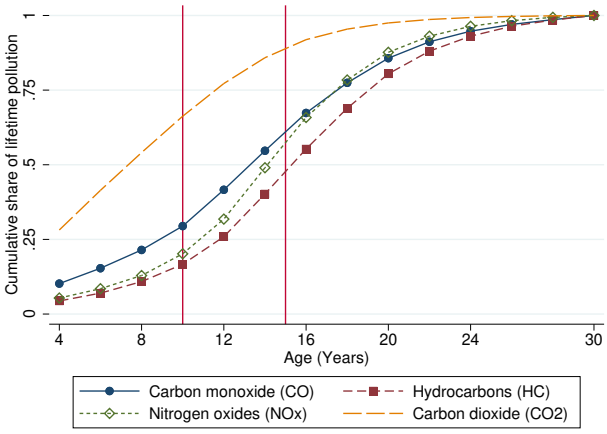
Notes: Figures compare the quantitative model calibration to the full sample from Colorado smog check data. Points represent mean emission rates in a given model year×age×vehicle class cell, averaged across all vehicles in the data. Axes have logarithmic scale.

Figure A10: Cumulative Emissions: Quantitative Model Versus Colorado Sample

(a) Quantitative Model



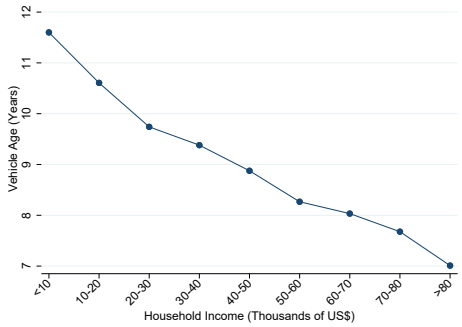
(b) Colorado Smog Check Data



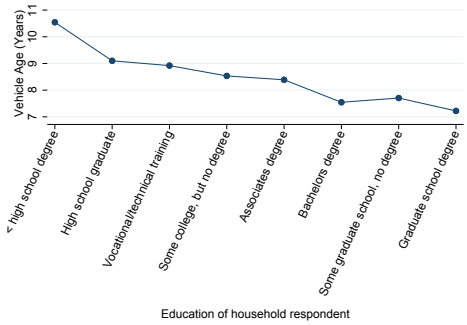
Notes: Figure compares cumulative emissions in the quantitative model (calibrated to the national vehicle age profile) to that in the Colorado smog check data.

Figure A11: Vehicle Age Across Demographic Groups

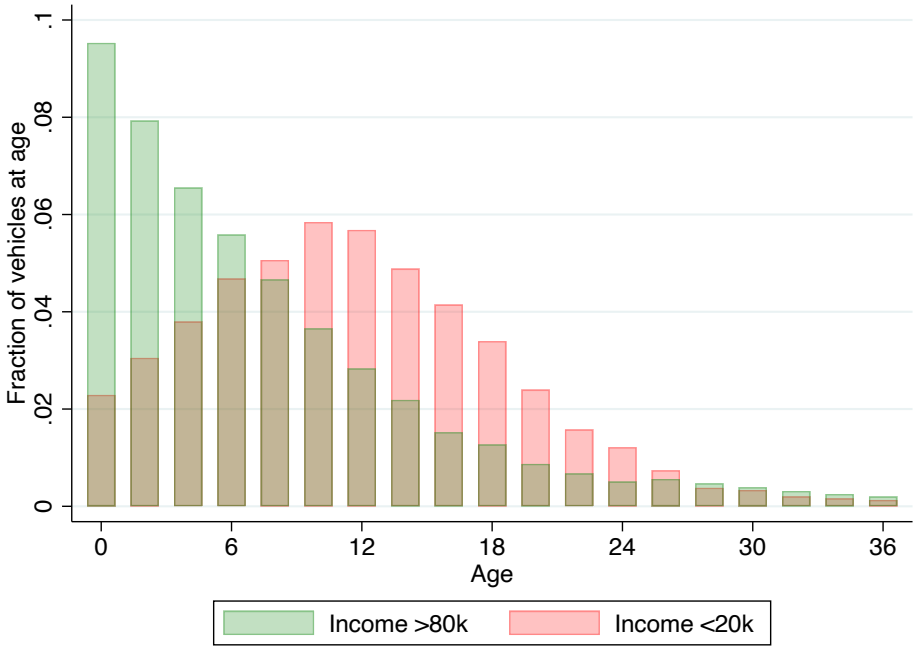
(a) Average Vehicle Age by Income



(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 × 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:	0.42***	0.38**	1.17***	1.40***
Within-engine changes	(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 IV 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 III
		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.