Discussion Paper Series – CRC TR 224

Discussion Paper No. 721
Project B 07

The Novel Vehicle Tax on Fine Particulate Matter Emissions

Nico Lukas Kasparetz[1]

December 2025

[1] Department of Economics, University of Mannheim. E-mail: nico.kasparetz@uni-mannheim.de

Support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

The Novel Vehicle Tax on Fine Particulate Matter Emissions*

Nico Lukas Kasparetz†

December 2, 2025

Abstract

Old diesel cars without modern emissions control technology substantially contribute to air pollution by emitting high amounts of fine particulate matter, which is known to be detrimental to human health. Periodic vehicle registration fees offer a potentially powerful lever to speed up the retirement of old and polluting vehicles, yet little empirical evidence exists on the matter. This paper analyzes how higher registration fees for old and polluting diesel vehicles in the Netherlands accelerate their outflow from the vehicle fleet. It leverages the staggered rollout of diesel particulate filters as factory-fitted equipment to create quasi-random variation in pollution levels across otherwise comparable diesel car models. By applying Synthetic Difference-in-Differences complemented with a hazard model, this paper establishes that the tax increase on old and polluting cars is effective at reducing their numbers, albeit at the cost of being a very regressive policy.

Keywords: Vehicle Retirement, Particulate Matter Emissions, Vehicle Registration Fee, Difference-in-Differences, Survival Analysis, Policy evaluation, The Netherlands, Vehicle Taxes, Externalities, Redistributive Effects, Environmental Taxes and Subsidies
JEL Classifications: H23, L62, Q52, R48

*I thank Ulrich J. Wagner and Mateus Souza for their support and guidance. I am grateful for comments and suggestions by Laura Grigolon, Ken Gillingham, Andreas Gerster, Johannes Gessner, Kevin Remmy, Waldemar Marz, Ayman Tibi, Adrià Mateu Romero, Alexander Abajian, as well as by seminar participants at the University of Mannheim, the CRC TR 224 Young Researcher Workshop in Bonn, the ENGAGE.EU Workshop for Early-stage Researchers in Toulouse, and the ECON SIG Seminar at Yale University.
Support by the German Research Foundation (DFG) through CRC TR 224 (Project B07) is gratefully acknowledged. I acknowledge funding from the Department of Economics at the University of Mannheim, and I acknowledge data access funded by CRC TR 224 to the Microdata Environment of Statistics Netherlands through project 8266 "Household decisions over the life-cycle: Finances, Education, and Care".
Results are based on calculations by Nico Lukas Kasparetz, University of Mannheim, in project number 8266 using non-public microdata from Statistics Netherlands. Under certain conditions, these microdata are accessible for statistical and scientific research. For further information: https://www.cbs.nl/en-gb/our-services/customised-services-microdata/microdata-conducting-your-own-research.
Furthermore, access to and use of German vehicle registration data was provided under contract no. 251001 with the Research Data Centre of the Federal Motor Transport Authority (Forschungsdatenzentrum im Kraftfahrt-Bundesamt, FDZ-KBA).
Declarations of interest: none.

†Department of Economics, University of Mannheim. E-mail: nico.kasparetz@uni-mannheim.de

1 Introduction

Ownership taxes are a potentially powerful policy instrument for influencing the duration of durable good holdings, yet little is known about their effectiveness. While more stringent regulations make new durable goods increasingly efficient, the efficiency of the overall stock improves only slowly when turnover rates are low. This is particularly true in the transportation sector where the vehicle stock evolves slowly. Although new vehicles have become very clean, overall vehicle pollutant emissions decrease only at a modest rate because many drivers hold onto their old and polluting vehicles. In Europe, old diesel vehicles without modern emissions control technology such as particulate filters are the primary contributors to traffic-related air pollution. This is a particularly huge concern in the Netherlands, which has one of the oldest active vehicle fleets in Western Europe (Tweede Kamer der Staten-Generaal, 2015). These vehicles are responsible for a large share in the 59,000 years of life lost in the Netherlands every year due to traffic-related air pollution (Fischer et al., 2015; Compendium voor de Leefomgeving, 2023; Soares et al., 2023).

In order to reduce air pollution levels, the Dutch government introduced a surcharge on the private ownership of diesel vehicles with high particulate matter (PM) emissions to incentivize their retirement and thereby decrease the number of highly polluting vehicles in the active fleet (Tweede Kamer der Staten-Generaal, 2015). This approach was new as policies in other countries had previously been limited to scrappage schemes and low emission zones. The Dutch government, however, ruled out both of these policies because, from their point of view, a scrappage scheme would have been too expensive to maintain, and low emission zones would be unjust due to their prohibitive nature (Tweede Kamer der Staten-Generaal, 2015). On January 1, 2020, one year later than in the initial schedule, the surcharge on old and polluting diesel vehicles came into effect.

This paper evaluates the effectiveness of the Dutch surcharge on polluting diesel vehicles in accelerating vehicle retirement. Beyond its aggregate impact, I analyze which types of households respond most strongly and how the effect differs across vehicle characteristics. This question is of great relevance for various durable goods markets beyond the economics of transportation. As of now, there is little empirical evidence on how changing the ownership costs of old durable goods affects their retirement. This paper fills this research gap by analyzing the Dutch surcharge on polluting vehicles with its explicit goal of "cleaning up" the vehicle fleet. It capitalizes on the fact that different manufacturers started the series production of diesel vehicles with particulate filters at different points in time because of supply bottlenecks in the filter industry. This natural experiment introduces variation in PM emissions and thus in surcharge eligibility across several dimensions of vehicle characteristics.

I combine two main data sources. Historical data on Dutch vehicles and vehicle ownership comes from Statistics Netherlands (CBS). CBS offers annual vehicle-level data on the entire fleet of registered vehicles in the Netherlands, making it possible to link every vehicle to its owner. I can further link individual-level socioeconomic data, such as age, income, education level, place of residence, and marital status, among others. Lastly, I complement the vehicle data from CBS with publicly available data on technical specifications of vehicles including make, model, and PM emissions from Dienst Wegverkeer (RDW), the Dutch motor vehicle agency. I combine the data from both sources to analyze how the introduction of the surcharge has affected the outflow of highly polluting diesel vehicles relative to otherwise comparable vehicles that pollute substantially less and are not subject to the policy.

Drawing on a decade of vehicle registration data, I show that the surcharge is inherently regressive. It applies only to diesel vehicles that were at least 12 years old when the surcharge came into effect. The owners of such vehicles are predominantly young individuals from lower-income households, and they predominantly live in regions with elevated reliance on individual motorized traffic. I then proceed to show that the surcharge significantly reduces the number of polluting vehicles by about 16 percent within the first three years of the policy, indicating that car owners indeed bring forward vehicle replacement decisions. This behavior is consistent with how car owners choose to replace an old durable good in a setting in which keeping the existing durable good entails an elevated cost in form of additional ownership fees (see Appendix C). I also show that the policy's regressivity intensifies over time because lower-income households are less able to advance vehicle retirement, and uncover further insightful heterogeneity in household responses to the soot tax.

This paper contributes to three strands of the literature: (i) the determinants of vehicle retirement, (ii) the effectiveness of registration fees in shaping consumer behavior, and (iii) the welfare effects of changes in the vehicle fleet. Within the first strand, several papers show that scrappage schemes are effective at speeding up vehicle retirement, improving average fuel economy, and reducing average pollution levels of the vehicle fleet (Lenski et al., 2010; Busse et al., 2012; Li et al., 2013; Grigolon et al., 2016; Laborda & Moral, 2019; Marin & Zoboli, 2020). However, their substantial costs render them unsuitable as a long-run policy instrument (Sandler, 2012). Low emission zones exclude the most polluting vehicles and thereby cause short-lived spikes in their scrap rates (Balaguer et al., 2023). Higher gasoline prices accelerate the scrappage of old vehicles (which are the most polluting) by reducing their used market value (Li et al., 2009; Jacobsen & van Benthem, 2015; Bento et al., 2018). I contribute to this literature by assessing how a novel policy that changes ownership costs affects vehicle retirement, for which there is little empirical evidence. Moreover, Jacobsen et al. (2023) derive a model of individual vehicle replacement decisions to predict that "environmental" ownership taxes based on tailpipe pollutant emissions improve welfare through their effect on vehicle scrappage rates. My framework empirically assesses this model prediction from Jacobsen et al. (2023) as well as its distributional impact in a concrete setting.

The second important branch of the literature analyzes the effect of vehicle taxes in particular. Many studies focus on how upfront taxes, which are levied at the time of purchase or first registration, impact the composition and nature of new car sales. These papers consistently show that higher upfront vehicle taxes on new cars with higher CO_2 emissions lead to reductions in the emission intensity of new vehicle registrations (Gallagher & Muehlegger, 2011; Kok, 2015; Ciccone, 2018; Bergantino et al., 2021; Østli et al., 2022). However, several papers emphasize that the impact, albeit statistically significant, tends to be economically modest (Rogan et al., 2011; Adamou et al., 2014; D'Haultfœuille et al., 2014; Kok, 2015; Klier & Linn, 2015; Alberini & Bareit, 2019). At the same time, tax cuts or exemptions for low-emission vehicles are comparatively inefficient while giving rise to substantial public expenditure (Kok, 2015; Sun et al., 2018). In contrast to the large literature on upfront taxes, the literature on periodic ownership taxes, commonly known as "registration fees", is sparse. As a notable exception, Alberini et al. (2018) analyze how annual registration fee surcharges for cars with high CO_2 emissions in Switzerland decrease their average lifetime and thereby reduce average CO_2 emissions of the vehicle fleet. My paper complements their work because the Dutch registration fee surcharge targets tailpipe PM emissions rather than CO_2, affecting a fundamentally different set of vehicles.

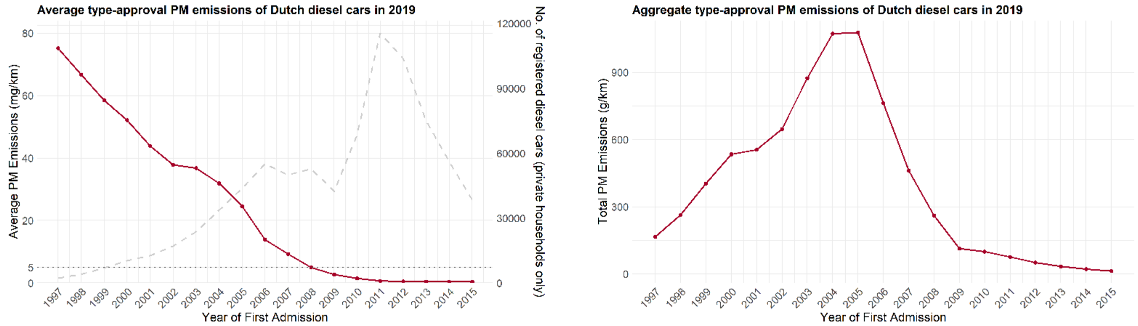
I also draw on quasi-random variation in treatment status due to the staggered rollout of particulate filters whereas Alberini et al. (2018) rely on cross-sectional and temporal variation in cantonal registration fees within Switzerland.

The third important branch of the literature investigates welfare. It is well-established that diesel cars without adequate emissions control technology, even when few in numbers, substantially contribute to air pollution and significantly aggravate health hazards (Degraeuwe et al., 2017; Jonson et al., 2017; Alexander & Schwandt, 2022). Removing the most polluting vehicles can significantly improve air quality and generate substantial welfare gains through associated health benefits. A large body of literature has shown this for low emission zones (Cyrys et al., 2014; Malina & Scheffler, 2015; Gehrsitz, 2017; Jiang et al., 2017; Pestel & Wozny, 2021; Margaryan, 2021). However, policies targeting highly polluting cars are generally regressive because it is usually lower-income households that own these cars (Nikodinoska & Schröder, 2016; Eliasson et al., 2018; Tovar Reaños & Sommerfeld, 2018; Jacobsen et al., 2023). I contribute to this literature by means of the extraordinary microdata at hand that establish individual links between vehicles and vehicle owners. This allows me to analyze household heterogeneity in terms of (i) who is initially affected by the soot tax, and (ii) who responds more strongly to the soot tax by bringing forward vehicle replacement.

The remainder of the paper is structured as follows. Section 2 provides the institutional background, showing how production decisions create the variation in surcharge eligibility more than a decade later that I exploit in the empirical analysis. Section 3 describes my data sources, provides some illustrative descriptive statistics about vehicle holdings in the Netherlands, and contains summary statistics for my samples. Section 4 elaborates on my methodology and how the setting at hand allows me to infer causal estimates of the effect of the registration fee surcharge on the vehicle stock and on individual vehicle retirement decisions. I present the results in Section 5, conduct both robustness checks and placebo tests to corroborate them, and introduce a toy model to rationalize my findings. Moreover, I carry out a back-of-the-envelope calculation to illustrate the implications of my results from Section 5 for total vehicle holdings and fleet-wide PM emissions in Section 6. Section 7 eventually concludes.

2 Institutional Background

Figure 1: Type-approval PM emissions of diesel cars by year of admission



(a) Average PM emissions in mg/km of diesel vehicles, by year of first admission

(b) Aggregate PM emissions in g/km of diesel vehicles, by year of first admission

Notes: The left graph shows the average type-approval PM emissions of diesel vehicles first admitted in the respective year in mg/km, as in the public RDW motor vehicle register (opendata.rdw.nl). The right graph multiplies these average PM emissions with the number of diesel vehicles from the respective cohorts that were registered on January 1, 2019. Numbers on registered diesel vehicles in 2019 come from the CBS Microdata Environment. Own calculations.

In 2015, Fischer et al. (2015) made were the first to establish a causal link between elevated levels of particulate matter with less than 10 micrometers in diameter ("PM" in the following) and various detrimental health outcomes for millions of Dutch citizens. A main source of PM emissions is the remaining fleet of old diesel vehicles without modern emissions control technology. The red line in Figure 1a plots (type-approval) PM emissions of diesel passenger cars by year of first admission. Newer diesel vehicles admitted after 2009 have very low average emissions of less than 0.5 milligrams per kilometer (mg/km). In contrast, the average diesel car from 2005 is more than 50 times as polluting in terms of (type-approval) PM emissions than its modern-day counterparts, and the average diesel car from 2000 is more than 100 times as polluting (see also Table A1). Figure 1b puts the emissions intensity of older vehicles into perspective by scaling per-vehicle PM emissions to the full diesel fleet using registration data from January 1, 2019 (the grey line in Figure 1a). Even though many of the older diesel vehicles produced before 2009 have already left the fleet by 2019, the remaining ones still account for the majority of overall PM emissions. Therefore, the continued presence of old and highly polluting diesel cars, even in small numbers, still imposes a substantial burden on air quality.

Vehicle taxes constitute a natural policy instrument to address air pollution from road traffic. The Netherlands, like most countries, levy two types of vehicle taxes: a one-time tax levied at the time of purchase of a new vehicle ("vehicle sales tax");

and a cyclical tax on vehicle ownership that the owner pays when renewing its registration to legally drive the vehicle on public roads. I will refer to this cyclical tax on vehicle ownership as a "registration fee" in line with the established literature (Jacobsen et al., 2023).

These two taxes serve very different purposes in the Netherlands. The Dutch vehicle sales tax (called BPM (*belasting van auto's en motorrijwielen*)) is levied when a new car is purchased or when a used car is imported to the country. Since the early 2010s, the sales tax liability of a vehicle depends on its carbon dioxide (CO_2) emissions and fuel type. The tax structure is designed to reflect each vehicle's environmental impact. First, the sales tax rises with vehicle CO_2 emissions to discourage the adoption of vehicles with high fuel consumption and high CO_2 emissions. Roughly speaking, the sales tax doubles when going from a moderately-sized engine with emissions of 120 g CO_2/km to a larger engine with 150 g CO_2/km. Second, the sales tax is about twice as high for diesel cars than for gasoline cars because diesel cars have higher pollutant emissions on average. Plug-in hybrid vehicles (PHEVs) face only a fraction of the tax liability of comparable gasoline vehicles because they have substantially lower emissions. Lastly, battery electric vehicles (EVs) are completely exempt as they have zero emissions.[1]

The registration fee MRB (*motorrijtuigenbelasting*), in contrast, is the registration fee, the periodic levy on vehicle ownership that must be paid in order to legally use the vehicle on public roads. In the following, I describe the Dutch registration fee system for passenger cars.[2] In the Netherlands, vehicle registration fees are due every three months. Before the introduction of the surcharge for polluting diesel vehicles in 2020, the registration fee had been based on two determinants only, the vehicle's (i) empty mass, and (ii) fuel type. Importantly, neither component is meant to address the environmental impact of the vehicle. The vehicle mass component accounts for the fact that heavier vehicles cause more wear and tear on road infrastructure so that their contribution to maintaining the road system is larger. It further serves as an incentive mechanism to adopt lighter and smaller cars, which are safer for pedestrians and cyclists and take up less public space. The Dutch government also charges higher registration fees on diesel than on gasoline passenger cars to counteract lower excise

[1] Tables A2, A3 and A4 explain the calculation in detail.
[2] There are separate calculations for delivery vans, company cars and campers, which are not the focus of this paper.

taxes on diesel fuel. In the Netherlands, like in most of Europe, diesel is subject to a lower excise tax than gasoline.[3] The lower excise tax on diesel serves as a subsidy for businesses because commercial vehicles such as trucks, delivery vans, and agricultural machinery almost universally run on diesel (Sterner, 2007; Ptak et al., 2024). Since the lower excise tax is not targeted towards private households, the registration fee offsets the lower cost of diesel at the gas station by levying extra charges for the private ownership of diesel cars.[4] Therefore, the introduction of the surcharge on old polluting diesel cars in 2020 marks a structural change in the registration fee system by introducing an environmentally motivated dimension.

Figure 2 shows annual registration fee payments for passenger cars. Solid lines refer to the 2019 payment schedule, and dashed lines to the 2024 schedule, which is higher due to annual inflation adjustments. One can see that vehicle registration fees account for a meaningful share of household expenses (Centraal Bureau voor de Statistiek, 2025). Medium-sized 1,500 kg gasoline cars face annual fees of about €870 in 2019 and €1,000 in 2024.[5] In contrast, equally heavy diesel cars face annual fees of about €1,660 in 2019 and €2,000 in 2024. These figures highlight the substantial cost of private diesel car ownership in the Netherlands, which levies the highest registration fees in the entire European Union. In 2024, a 1,500 kg diesel car is liable for an annual registration fee of €232 in Germany, €357 in Denmark, €536 in Austria, or €537 in Flanders, but about €2,000 in the Netherlands.

Since January 1, 2020, there is a third component in the registration fee schedule, the fine particulate matter surcharge (Dutch: *fijnstoftoeslag*), colloquially known as the "soot tax". For passenger cars, the soot tax applies to diesel vehicles with registered PM emissions[6] above 5 mg/km[7]. The soot tax was first proposed in June

---

[3] As of 2024, the excise tax on gasoline is €0.79 per liter, and the excise tax on diesel is €0.52 per liter (Rijksdienst voor Ondernemend Nederland, 2025)

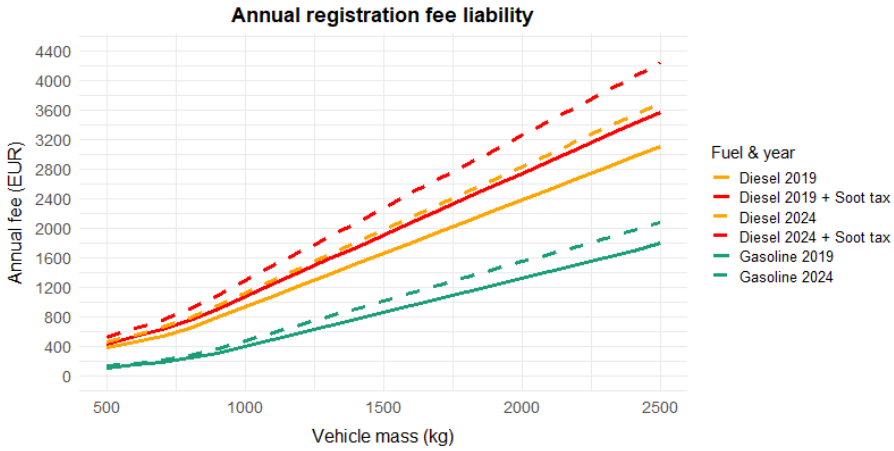
[4] The registration fee for commercial vehicles, in contrast, is independent of fuel type so that commercial diesel vehicles enjoy the excise tax discount without being penalized through higher registration fees.

[5] This calculation is on average because one part of total registration fee liability depends on a rate that every province sets independently and whose revenue flows to the provinces. This introduces variation in registration fee liability for identical vehicles between provinces by a few percent at most. However, this is not the focus of this analysis and can therefore be ignored for the remainder of the paper.

[6] RDW registers this emissions reading as "Emission Particles, light", which refers to particulate matter of less than 10 micrometers in diameter.

[7] The registered emissions refer to either type-approval emissions or measured vehicle emissions immediately after production. Vehicles without a PM emissions reading that were first registered before September 2009 are also subject to the soot tax, as are vehicles that were originally equipped

Figure 2: Registration fee schedules



Notes: The graph shows registration fees by vehicle mass, for diesel and gasoline passenger cars, in 2019 and 2024, averaged across provinces. The data show the quarterly (3-month) registration fee liability. A small portion of the registration fee goes to the provinces, and their charges and thus total registration fee liabilities differ by province. The displayed schedule is an average across provinces. The solid (dashed) lines refer to 2019 (2024) fee schedules. The green lines refer to gasoline cars, the orange lines to diesel cars, and the red lines to diesel cars subject to the soot tax.

2015 when the Dutch government laid forward a vehicle tax reform (the "Autobrief II") that included plans to introduce a surcharge for polluting diesel vehicles on top of the regular registration fee. The First Chamber of the Dutch parliament voted in favor of the Autobrief II in July 2016, planning to introduce the soot tax by January 1, 2019 (Eerste Kamer der Staten-Generaal, 2016). However, the rollout of the soot tax was later rescheduled to January 1, 2020, due to administrative delays. The soot tax aims to speed up the scrappage of old diesel vehicles lacking particulate filters because these vehicles have air pollutant emissions of as much as two orders of magnitude higher than modern diesel vehicles with filters (Tweede Kamer der Staten-Generaal, 2015).

The soot tax is a 15% surcharge on the vehicle-specific registration fee, increasing the relative tax burden equally across all affected vehicles (Tweede Kamer der Staten-Generaal, 2015). To maximize the salience of the soot tax, the government sent letters to all affected vehicle owners in early November 2019, less than two months before the soot tax was introduced, clearly stating that the new surcharge would be imposed as a penalty for owning a highly polluting diesel vehicle. Moreover, although formally a surcharge on the regular registration fee, the soot tax is billed separately

with a diesel particulate filter reducing emissions below 5 mg/km, but from which the filter has subsequently been removed. Retrofitting a diesel vehicle with a filter, however, does not exempt the vehicle from the soot tax.

to further enhance its salience and underscore its distinct purpose. Because of the high baseline fees, the soot tax constitutes a substantial extra burden to diesel car owners and, beyond that, is a considerable cost relative to the residual value of affected vehicles as it applies exclusively to relatively old diesel cars.

In practice, the eligibility threshold of (type-approval) PM emissions of 5 mg/km means that diesel cars are subject to the soot tax if they are not equipped with a particulate filter. These filters, installed directly upstream of the tailpipe, clean the exhaust and reduce PM emissions by up to two orders of magnitude. Particulate filters became a standard equipment in all diesel passenger cars with the introduction of the Euro 5 emission standard in September 2009. This development took place within a broader regulatory framework. Passenger car emissions have been regulated in the European Union since 1993, with standards becoming progressively more stringent every four to five years. Table A5 shows the PM emission limits of the respective Euro standards for diesel cars. Manufacturers were able to meet the diesel emission limits up to Euro 4 through exhaust gas recirculation and the use of catalytic converters, circumventing high-cost particulate filters. Accordingly, diesel vehicles of vintage[8] 2003 and earlier are all subject to the soot tax. These vehicles were manufactured under the considerably more lenient emission standards prior to Euro 5, at a time when particulate filters had not yet been introduced in passenger car production.

The introduction of the stringent Euro 5 PM limits effectively compelled manufacturers to adopt filter technology in diesel cars (Scoltock, 2014). Manufacturers therefore began adopting the technology as soon as Euro-5-compliant filters became available in 2004. However, supply constraints in the filter industry led to a gradual introduction of filters during the years leading up to Euro 5. In total, it took several years for particulate filters to become standard equipment on all new diesel cars. In 2005, only about 20% of new diesel cars were already equipped with particulate filters. This share gradually rose to roughly 50% in 2006, 63% in 2007, and 75% in 2008, before Euro 5 effectively mandated particulate filters starting in September 2009 (Figure A1). It is this staggered adoption of diesel particulate filters between 2005 and 2008, due to initial supply shortages and the fact that the expansion of

[8] In what follows, I equate a vehicle's year of production and a vehicle's year of first registration, and refer to both as its "vintage" in line with established academic literature (Barahona et al., 2020). Occasionally, some authors also refer to this a "model year".

production to serve the entire market took several years, that generates substantial variation in PM emission levels between diesel car models from the mid-2000s. I leverage this variation to estimate the causal effect of the soot tax on vehicle holdings because the lack of a factory-fitted particulate filter directly determines soot tax eligibility in practice.

3 Data

This paper draws on two main data sources. First, Statistics Netherlands (CBS) provides annual data on the entire Dutch vehicle stock as registered with RDW on January 1 each year from 2014 to 2023. I observe both the individual vehicle as well as the vehicle owner, and can follow the same vehicle over time as it changes owners, as well as the same owner over time as they replace vehicles. I focus on privately owned diesel passenger cars, which are the main target of the soot tax. CBS also provides various other data about the individual vehicle owners, such as their income, education, place of residence, and others, that I merge to the vehicle data using unique person identifiers. Following established practice in the literature, I treat all permanent deregistrations as vehicle retirements (Alberini et al., 2018), although the registration data do not reveal the ultimate fate of deregistered vehicles.

Whereas the information on vehicle owners is vast, the CBS data only contain little information on the vehicles themselves. They only provide a vehicle's year of first registration, fuel type, mass (empty weight), and number of cylinders. There is no data on make-model or on registered PM emissions to directly derive soot tax eligibility. I therefore complement the CBS data with publicly available data from RDW, which provide comprehensive information on the entire Dutch vehicle fleet registered on the day of access (in my case, May 4, 2024). The RDW data include more than 100 variables for each registered vehicle in the Netherlands and allow me to identify the make-model as well as registered PM emissions for a subset of vehicles in the CBS data. Based on the registered PM emissions, I can assign soot tax treatment status to all vehicles with an identifiable make-model.[9] Appendix B

[9] For example, I may observe 100 individual vehicles in the CBS data that have these characteristics: Year of first registration: 2007, Fuel type: Diesel, Vehicle mass: 1,400 kg, Number of cylinders: 4. I then check whether at least 95% of the vehicles in the RDW with these characteristics are of the same make-model, e.g. some version of the Volkswagen Golf with some value of type-approval PM emissions. If this is the case, I assign the make-model "Volkswagen Golf" as well a treatment status based on their registered emissions to all these vehicles in the CBS data.

explains this process in detail. In accordance with the literature, I call such an identified make-model a "car model" (Grigolon et al., 2016).

Because the RDW data is vehicle-specific, registered PM emissions occasionally differ between individual vehicles of the same car model. Some manufacturers conducted updates ("facelifts") to a car model, which sometimes included starting to install filters as a standard equipment. Other manufacturers sometimes offered the same car model in two otherwise identical versions, differing only in the presence or absence of a filter. I discard such car models from my analysis. Moreover, the soot tax also applies to vehicles that were originally equipped with a factory-installed filter but later had it removed, often to achieve marginal improvements in performance or fuel economy.[10] Although I cannot identify such vehicles in the CBS data, I consider post-production filter removal a minor concern for two reasons. First, it affects only a small share of vehicles. TNO (2018) reports a filter-removal rate of 1.2% in diesel cars and light-duty vehicles. Second, it introduces a small bias towards zero at worst, rendering my results a little conservative.

3.1 Sample Construction & Matching Quality

In my empirical analysis, I focus on a sample of diesel passenger cars produced between 2005 and 2008. Treatment variation in this sample arises from the staggered introduction of particulate filters prompted by the regulatory change in emission standards enacted more than a decade before the soot tax comes into effect.[11] Because I can only identify the make-model and PM emissions for a subsample of all vehicles, I first examine how this subsample compares to the entire vehicle fleet. Ideally, the identified subsample is representative of the entire fleet of Dutch vehicles. To evaluate representativeness, I use the 2019 vehicle data set, which is one year before the soot tax, and in which I identify make-model and PM emissions for about 30.66% of vehicles from the relevant 2005-2008 diesel vintages.

[10] Although removing the particulate filter is prohibited by law, owners of diesel cars produced before 2011 could notify the RDW that the filter had been removed. Once notified, the vehicle became subject to the soot tax, while the owner avoided the fine associated with the illegal removal itself.

[11] As a robustness check, I also draw on a sample of gasoline and diesel car models built between 1990 and 2004. The diesel car models in this sample are old enough to all be subject to the soot tax, yet not old enough to become exempt from all registration fees upon achieving classic-car status. Gasoline car models of the same 1990-2004 vintages comprise the natural control group in this case as these are never subject to the soot tax.

I use covariate distributions as a measure of representativeness and compare them between the identified and unidentified vehicles. Both groups are very much comparable in terms of owner characteristics. For instance, household disposable income and owner age feature almost the exact same distribution. Their means only differ by 0.5% for income and 2.3% for owner age (see Figures A2 and A3 in the Appendix). In contrast, the differences in terms of the actual vehicle characteristics are more pronounced. For instance, identified vehicles tend to be lighter than the vehicles for which I cannot identify make-model and PM emissions (Figure A4). Among identified vehicles, 55.9% weigh less than 1,025 kg, typical of small or light compact cars, compared with 37.1% of unidentified vehicles. In contrast, 57.6% of unidentified vehicles weight between 1,075 kg to 1,525 kg, the range of heavier compact cars, mid-size and executive cars, compared to 38.2% of identified vehicles. Therefore, my sample of identified vehicles is skewed towards lighter types of cars. The average identified vehicle weighs 1,069 kg, a statistically significant 8.1% difference compared to 1,156 kg for the average unidentified vehicle.

3.2 Descriptive statistics

The comprehensive individual-level vehicle data allow for some first descriptive insights into vehicle ownership in the Netherlands. I first use the CBS data to describe the Dutch vehicle fleet as a whole, and subsequently use the subset of identified vehicles to characterize the analytical sample.

3.2.1 The entire fleet

I first analyze the spatial distribution of diesel vehicles across Dutch municipalities. There is a clear divide in vehicle ownership between rural and urban areas. The data clearly reveal that higher population density is associated with lower diesel prevalence (see Figure A5). At the same time, there is also a north-south divide (see Figure A6). These differences in diesel ownership reflect the economical advantages of diesel engines over gasoline engines when people drive more frequently and when they regularly cover relatively long distances (Reif, 2014). Indeed, in the Netherlands, both (i) average everyday distances as well as (ii) the share of travel distances covered by car are on average larger in more sparsely populated areas (see Figures A7 and A8).

Next, I explore household heterogeneity in vehicle ownership. A common counterargument against tying registration fees to tailpipe pollutant emissions is the regressive nature of such a policy. For instance, Jacobsen et al. (2023) show how ownership of old cars is concentrated at lower-income households in the US. Figure A9 assesses this relation in the Netherlands by plotting vehicle mass against household disposable income for different age bins of diesel cars. The first eminent pattern is that the newest and cleanest diesel vehicles, produced since 2010 under Euro 5 (or 6), belong to the highest-income households, with a strikingly positive association between vehicle mass and income in this vehicle age group (Figure A9a). As we look at medium-aged, more polluting cars from the 2000s (Figure A9b), average household incomes are substantially lower. Old and highly polluting diesel cars produced in the 1990s (Figure A9c) belong the the lowest-income households of all households that own a car to begin with. Moreover, household incomes among these old diesel cars are low across all vehicle masses, highlighting how ownership of the most polluting diesel vehicles is concentrated on the lower end of the income distribution. Because the soot tax applies exclusively to diesel vehicles produced before 2009, these descriptive results highlight how the soot tax disproportionately affects lower-income households.

Furthermore, the data indicate that the soot tax predominantly affects younger car owners (see Figure A10). While owners of gasoline vehicles are on average in their early 50s and owner age varies little across vehicle masses, the average owner of a small or compact diesel car of roughly 1,000 kg is in her early 40s, nearly ten years younger than the owner of a gasoline car in the same weight category. Beyond that, average owner age increases sharply with vehicle mass for diesel cars, rising to 55 to 60 years for very heavy diesel cars beyond 2,000 kg. Overall, given that the soot tax primarily applies to lighter diesel cars (which, on average, received filters later) and to older diesel cars (which also tend to be lighter), this strongly suggests that the tax burden will fall mainly on younger car owners.

3.2.2 The samples

In my empirical analysis, I draw on diesel vehicles of 2005-2008 vintage with identified make-model for which I observe registered PM emissions and can therefore determine whether they are subject to the soot tax. I am able to identify a total of 342 car models in the data. The makes Volvo and Volkswagen alone constitute more than 40% of the car models in the sample. Their models often feature odd and unique vehicle masses (such as 1,063 kg for a 2005 Volkswagen Polo), making them uniquely identifiable. In contrast, the three German premium manufacturers BMW, Audi, and Mercedes-Benz constitute just below 20% of all the identified car models because their models often have the same round vehicle masses so that many of them remain unidentified.[12] I also include the previous year's values of the following household/owner covariates: household disposable income, vehicle owner age, a dummy for college education, marital status, a dummy for homeownership, and population density of place of residence.

I use this set of diesel vehicles to construct both a panel data set on the car model level as well as an individual-vehicle-level data set. The upper half of Table 1 contains summary statistics for the sample on the car model level. In this data set, I aggregate the individual-vehicle-level data up to the car model level. This results in a panel data set with 3,420 observations, in which I follow the 342 identified car models over ten years. 131 car models are in the treatment group since they become subject to the soot tax in 2020, while the other 211 car models make up the control group. The outcome of interest in this car model $\times$ year panel data set is the number of registered vehicles per car type. There are on average 97.27 registered vehicles per car model across the entire observation period, with 55 vehicles for the median car model. The average vehicle mass in the sample is 1,512 kg, about the mass of a typical 2000s mid-size diesel car. The average list price in the sample is €52,493 although 48 car models have missing values for this variable. Average fuel consumption comes in at 6.7 l/100 km, and CO_2 emissions average 177.5 g/km.

Household/owner covariates, averaged for every car model $\times$ year, show that household disposable income is somewhat right-skewed, with median disposable income at €60,189 and mean disposable income at €66,754. This is about €20,000 above the nationwide average during the same period, highlighting how vehicle ownership is generally concentrated among higher-income households. Owners are on average 48.8 years old, and the majority of owners are male, married, and own their home. About half have some form of college education.

[12] The CBS data only contain a vehicle's vintage, mass, fuel type, and number of cylinders. If two or more models share the same values for these four variables, then I cannot assign ("identify") a make-model to the individual vehicles with these characteristics.

Table 1: Summary statistics for 2005–2008 diesel car-model-level sample

Variable	Unit	No. of obs.	Mean	St.d.	Median
Treatment Group	Dummy	3,420	0.383	0.486	0
No. of vehicles	Count	3,420	97.27	113.4	55
Vintage	Year	3,420	2006.64	1.155	2007
Vehicle mass	kg	3,420	1,512	366.1	1,522
List price	€	2,940	52,493	27,131	48,612
Mileage	l/100 km	3,400	6.694	1.805	6.700
CO_2	g/km	3,400	177.5	48.18	177.7
Mean HH Disp. Inc.	€	3,420	66,941	29,733	60,278
Mean Owner Age	Years	3,420	48.82	6.128	49.20
Mean Pop. Density	Persons/km^2	3,420	1,225	280.3	1,195
Mean Male Owner	Dummy	3,420	0.798	0.112	0.822
Mean College Educ.	Dummy	3,420	0.502	0.187	0.500
Mean Married	Dummy	3,420	0.608	0.129	0.617
Mean Homeownership	Dummy	3,420	0.791	0.135	0.818

	Control group				Treatment group			
Variable	No. of obs.	Mean	St.d.	Median	No. of obs.	Mean	St.d.	Median
Treatment Group	2,110	0	0	0	1,310	1	0	1
No. of vehicles	2,110	94.18	114.7	51.00	1,310	102.3	111.2	63.00
Vintage	2,110	2007.13	0.943	2007	1,310	2005.84	0.995	2005
Vehicle mass	2,110	1,619	271.5	1,570	1,310	1,341	428.7	1,161
List price	2,020	61,008	24,725	60,722	920	33,797	22,398	25,172
Mileage	2,110	7.213	1.506	7.238	1,290	5.845	1.930	4.935
CO_2	2,110	190.9	40.17	192.2	1,290	155.7	52.08	132.1
Mean HH Disp. Inc.	2,110	72,136	29,464	66,169	1,310	58,572	28,226	51,493
Mean Owner Age	2,110	50.29	5.394	50.51	1,310	46.47	6.497	46.30
Mean Pop. Density	2,110	1,196	277.2	1,169	1,310	1,271	279.2	1,240
Mean Male Owner	2,110	0.840	0.083	0.850	1,310	0.731	0.121	0.750
Mean College Educ.	2,110	0.554	0.180	0.565	1,310	0.417	0.166	0.400
Mean Married	2,110	0.637	0.122	0.649	1,310	0.562	0.128	0.555
Mean Homeownership	2,110	0.841	0.102	0.862	1,310	0.710	0.143	0.715

Notes: Descriptive statistics on the final 2005-2008 diesel car model sample in which I observe the number of registered vehicles of 342 car models across ten years (2014-2023, January 1). 131 car models are treated, 211 are untreated. "Treatment Group" is one in all periods for car models with PM emissions above 5 mg/km, and zero otherwise. "No. of vehicles" counts the number of registered vehicles of an (identified) car model. "Vintage" is year of first admission in the data. Mean variables are average values of household/owner covariates across all individual vehicles of a car type.

The bottom half of Table 1 shows some underlying covariate imbalances between the treatment and control group. As a direct consequence of the staggered rollout of particulate filters, car models in the treatment group are on average older than their counterparts in the control group. Because manufacturers tended to install filters in heavier and more expensive models first, treated car models are on average lighter, have a lower list price, and feature lower fuel consumption as well as lower CO_2 emissions. In addition to that, the two groups of car models also differ in terms of their owner structure. Car models from the treatment group tend to belong to households with lower disposable incomes; their owners tend to be younger and live in more densely populated areas; and they are more likely to be female, without college education, unmarried, and renters. This highlights the necessity to control for underlying household/owner differences to account for different vehicle replacement cycles and potentially different responses to the soot tax.

Table 2: Summary statistics for 2005-2008 diesel individual-vehicle sample

Variable	Unit	No. of obs.	Mean	St.d.	Median
Treatment Group	Dummy	268,059	0.497	0.500	0
Event	Dummy	268,059	0.156	0.363	0
Vintage	Year	268,059	2006.24	1.085	2006
Vehicle mass	kg	268,059	1,490	366.6	1,476
Cylinder volume	cm^3	268,059	2,112	545.0	1,995
List price	€	221,186	48,206	25,631	43,085
Mileage	l/100 km	267,560	6.494	1.789	6.286
CO_2	g/km	268,059	172.6	47.70	164.6
HH No. Cars Owned	Count	268,059	1.795	1.166	2
HH Disp. Inc.	Years	268,059	64,204	113,306	53,668
Owner Age	Years	268,059	48.46	13.63	48
Population	Persons	268,059	120,633	180,713	50,290
Population Density	Persons/km^2	268,059	1,270	1,477	617
Male Owner	Dummy	268,059	0.792	0.406	1
Married	Dummy	268,059	0.630	0.483	1
Homeownership	Dummy	268,059	0.798	0.402	1

	Control group				Treatment group			
Variable	No. of obs.	Mean	St.d.	Median	No. of obs.	Mean	St.d.	Median
Treatment Group	134,798	0	0	0	133,261	1	0	1
Event	134,798	0.141	0.348	0	133,261	0.170	0.376	0
Vintage	134,798	2006.87	0.967	2007	133,261	2005.60	0.782	2005
Vehicle mass	134,798	1,625	275.5	1,570	133,261	1,354	395.7	1,262
Cylinder volume	134,798	2,360	509.1	2,400	133,261	1,862	459.1	1,896
List price	126,920	59,259	22,470	58,485	94,266	33,323	21,811	26,740
Mileage	134,750	7.215	1.545	7.3	132,810	5.762	1.721	5.2
CO_2	134,798	191.1	41.44	193	133,261	153.9	46.31	140
HH No. Cars Owned	134,798	1.849	1.140	2	133,261	1.741	1.189	2
HH Disp. Inc.	134,798	72,403	143,456	59,814	133,261	55,911	69,792	48,252
Owner Age	134,798	50.85	12.74	50	133,261	46.05	14.06	46
Population	134,798	112,460	170,370	48,506	133,261	128,901	190,249	54,474
Population Density	134,798	1,229	1,444	602	133,261	1,311	1,509	632
Male Owner	134,798	0.839	0.368	1	133,261	0.743	0.437	1
Married	134,798	0.669	0.470	1	133,261	0.590	0.492	1
Homeownership	134,798	0.870	0.336	1	133,261	0.725	0.446	1

Notes: Descriptive statistics on the final 2005-2008 individual-vehicle diesel sample. The sample is an unbalanced panel. I start with 58,541 vehicles in 2014, and some vehicles permanently disappear from the data set. “Treatment Group” is one in all periods for a vehicle with PM emissions above 5 mg/km, and zero otherwise. “Event” is one if the current period is the last one in which the car is present in the data (so that it is permanently deregistered before the next period). “Vintage” is year of first admission in the data.

I also create a sample on the individual vehicle level. This sample follows individual vehicles that are of one of the 342 identified car models. I restrict it to vehicles that are continuously observed in every period either until the end of the observation period or until they permanently disappear from the vehicle registry. This results in a panel data set with attrition that starts with 58,541 individual vehicles in 2014, some of which drop out in every subsequent year. It features 268,059 observations in total so that the average vehicle remains in the data set for 4.58 years.

The upper half of Table 2 shows summary statistics for the individual vehicle sample. The average vehicle weighs about 1.5 tonnes, has a list price of about €48,000, consumes 6.5 liters of diesel per 100 km, and emits 172.6 g CO_2/km. I also observe the absence or presence of other cars in the household. There are about 1.8 vehicles per household on average. The average population of the municipality in which the vehicle is registered is 120,633 while the median is at 50,290. The majority of car owners are male (79.2%), married (63%), and own a home (79.8%). Lastly, the lower half of Table 2 shows the prevalent covariate imbalance between the treatment and control group. On average, treatment group vehicles are older, lighter, less expensive, and consume less diesel fuel than control group vehicles. Their owners also have relatively less income, are younger, live in more urban areas, and are less likely to be male, married, or to own a home.

4 Methodology

I estimate the effect of the soot tax on the number of registered polluting diesel vehicles and on their propensity of being retired using a twofold strategy. First, I use the sample at the car model level to estimate the causal effect of the soot tax on the number of registered vehicles in a Generalized Difference-in-Difference (DiD) framework. This allows me to obtain an estimate on the "aggregate" effect, that is on how the soot tax affects the overall number of registered polluting vehicles. After all, the goal of the policy is to reduce the remaining stock of these vehicles. Second, I use the sample at the individual vehicle level to estimate a hazard model on how the soot tax affects an individual vehicle's risk of deregistration (which I consider equivalent to retirement in accordance with established literature (Alberini et al., 2018)). This allows me to granularly analyze heterogenous household responses and assess the distributional impact of the policy.

4.1 Difference-in-Differences

In the sample on the car model level, the outcome of interest Y is the log number of registered vehicles of car model i in period t. D is the dummy for the treatment group. The object of interest is the Average Treatment Effect on the Treated (ATT):

$$ATT = E\left[Y_{it}(1) - Y_{it}(0)|D_i = 1, \mathbf{Z}\right] \tag{1}$$

The ATT is the average difference between $Y_{it}(1)$, the outcome when there is treatment, and $Y_{it}(0)$, the outcome when there is no treatment, within the group of actually treated units ($D_i = 1$), in post-treatment period t, and conditional on covariates $\mathbf{Z}$ (Callaway & Sant'Anna, 2021). A convincing DiD approach to estimate the ATT as in Equation 1 relies on a parallel trends assumption such as

$$E\left[Y_{it}(0) - Y_{i,t-1}(0) \mid D_i = 1, \mathbf{Z}\right] = E\left[Y_{it}(0) - Y_{i,t-1}(0) \mid D_i = 0, \mathbf{Z}\right] \tag{2}$$

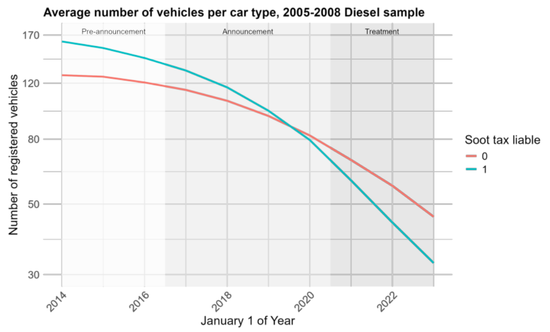
Here, think of $t-1$ as pre-treatment and t as post-treatment. Equation 2 says that, absent treatment, trends between the treated and untreated would have evolved in parallel. Under this assumption, Equation 1 can be written as

$$ATT = E\left[Y_{it}(1) - Y_{i,t-1}(0)|D_i = 1, \mathbf{Z}\right] - E\left[Y_{it}(0) - Y_{i,t-1}(0)|D_i = 0, \mathbf{Z}\right] \tag{3}$$

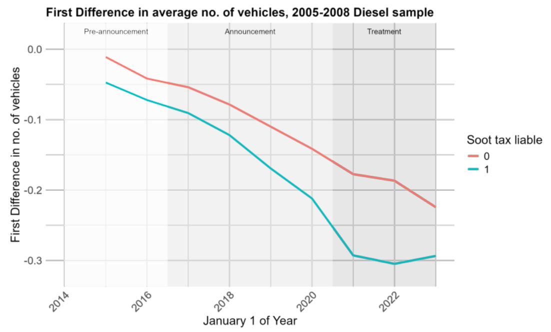
I can directly estimate Equation 3 using sample averages for the expectation terms because all its components are observed[13], and it corresponds to the ATT if parallel trends as in Equation 2 holds. A starting point to assess parallel trends is to look at unconditional trends, that is evaluating Equation 2 without conditioning on $\mathbf{Z}$. I do this in Figure 3a and clearly, the trends are not parallel. Vehicles from the treatment group already leave the fleet at a higher rate than their counterparts from the control group before treatment. Since conditional parallel trends is sufficient for identification, I examine trends conditional on $\mathbf{Z} = \begin{pmatrix} \mathbf{X_{it}} & \gamma_i & \mu_t \end{pmatrix}$. I do this by first regressing Y_{it} on $\mathbf{Z}$ using OLS and pre-treatment data only. Then, I use the residuals $\hat{\epsilon}_{it} = Y_{it} - \left(\hat{\beta}_0 + \hat{\mathbf{\Theta}}\mathbf{X_{it}} + \hat{\gamma}_i + \hat{\mu}_t\right)$ and calculate their mean value in the treatment ($\bar{\hat{\epsilon}}_{treated,t}$) and control group ($\bar{\hat{\epsilon}}_{control,t}$). The difference $\mathbf{\Delta_t} = \bar{\hat{\epsilon}}_{treated,t} - \bar{\hat{\epsilon}}_{control,t}$ should

[13] Both $Y_{i,t-1}(0)$ as well as $Y_{it}(1)$ are observed for the treatment group $D = 1$ because treatment only starts in period t. Likewise, because the control group $D = 0$ is never treated, both $Y_{it}(0)$ as well as $Y_{i,t-1}(0)$ are observed.

Figure 3: Number of registered vehicles in logs and log first differences



(a) Log number of registered vehicles, 2005–2008 diesel sample. Numbers from the CBS Microdata Environment.



(b) First difference in log number of registered vehicles, 2005–2008 diesel sample. Numbers from the CBS Microdata Environment.

be zero under conditional parallel trends. However, the resulting pattern of $\mathbf{\Delta_t}$ still strongly indicates a violation of parallel trends so that applying straightforward DiD would result in biased estimates of the ATT.

To account for the non-parallel trends, I apply the Synthetic Difference-in-Differences (SDiD) estimator (Arkhangelsky et al., 2021). SDiD combines standard DiD with the Synthetic Control Method (SCM) (Abadie et al., 2010), weighing control units to create a synthetic control group that runs in parallel with the treatment group during pre-treatment periods. Additionally, SDiD also upweighs (downweights) pre-treatment periods that are more (less) similar to post-treatment periods. Figure A11 shows an intuitive example of how SDiD combines DiD and SCM. The SDiD estimate of the ATT, $\hat{\beta}_1^{sdid}$, results from this optimization problem:

$$\left(\hat{\beta}_1^{sdid}, \hat{\beta}_0, \hat{\gamma}_i, \hat{\mu}_t, \hat{\mathbf{\Theta}}\right) = \underset{(\beta_1, \beta_0, \gamma, \mu, \mathbf{\Theta})}{\operatorname{argmin}} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \beta_0 - \gamma_i - \mu_t - \mathbf{\Theta X_{it}} - \beta_1 D_i Post_t)^2 \hat{\omega}_i^{sdid} \hat{\lambda}_t^{sdid} \right\} \tag{4}$$

SDiD assumes a standard Two Way Fixed Effects (TWFE) regression equation with covariates and minimizes the doubly-weighted sum of squared residuals, with unit weights ($\hat{\omega}_i^{sdid}$) and time weights ($\hat{\lambda}_t^{sdid}$), both of which are estimated separately beforehand. The vector $\hat{\omega}^{sdid}$ contains unit-specific weights $(\hat{\omega}_1, ..., \hat{\omega_{N_c}})$ for the N_c control group units plus a common intercept $\hat{\omega}_0$ so that, in each pre-treatment period $t = 1, ..., t_{pre}$, the weighted average of the control outcomes, $\hat{\omega}_0 + \sum_{i=1}^{N_c} \hat{\omega}_i Y_{it}$, runs as parallel as possible to the average outcome in the treatment group in a least-squares sense (Arkhangelsky et al., 2021). Additional regularization prevents the weight mass in $\hat{\omega}^{sdid}$ from collapsing onto a small subset of control units, ensuring stable inference (Arkhangelsky et al., 2021). In a similar fashion, the vector $\hat{\lambda}^{sdid}$ of time weights serves to minimize the squared gap between the weighted pre-treatment average outcome of control units and the average post-treatment outcome of the same units (Arkhangelsky et al., 2021).

I estimate standard errors using the jackknife algorithm. Jackknife standard errors combine computational feasibility with solid inference. They tend to be exact when treated units are similar to control units, and become conservative when differences between the treatment and control group are more pronounced (Arkhangelsky et al., 2021). They are calculated by running SDiD once, fixing the SDiD weights $\{\hat{\omega}_i^{sdid}, \hat{\lambda}_i^{sdid}\}_{i=1}^{N_c}$, and then estimating N different treatment effects by omitting one unit from the sample every time. The root of the sample variance of these N estimates is the jackknife standard error.

I also apply the more established SCM as a robustness check for the relatively new SDiD estimator. The estimator looks as follows:

$$(\hat{\beta}_1^{sc}, \hat{\beta}_0, \hat{\mu}_t, \hat{\mathbf{\Theta}}) = \underset{(\beta_1,\beta_0,\mu,\mathbf{\Theta})}{\arg\min} \left\{ \sum_{i=1}^{N} \sum_{t=1}^{T} \left(Y_{it} - \beta_0 - \mu_t - \mathbf{\Theta} X_{it} - \beta_1 D_i Post_t\right)^2 \hat{\omega}_i^{sc} \right\} \tag{5}$$

Equation 5 differs from Equation 4 in three ways. The unit fixed effects γ_i are omitted; the time weights $\hat{\lambda}_t^{sdid}$ are omitted; and the unit weights $\hat{\omega}_i^{sc}$ are estimated without an intercept term, which forces the average pre-treatment outcomes of the treated and synthetic control group to match exactly (Arkhangelsky et al., 2021). I calculate placebo standard errors as the root of the variation in point estimates of placebo treatment within the control group (Arkhangelsky et al., 2021).

To analyze treatment effect heterogeneity, I also estimate a TWFE model with group-specific linear time trends. The regression equation is:

$$Y_{it} = \beta_0 + \beta_1 D_i Post_t + \mathbf{\Theta X_{it}} + \alpha \left(D_i \cdot t\right) + \gamma_i + \mu_t + \epsilon_{it} \tag{6}$$

Angrist and Pischke (2009) champion this approach in the presence of non-parallel trends. Standard errors are clustered at the unit level to address autocorrelation in the errors (Grigolon et al., 2016). The extra term is $D_i \cdot t$, which provides a linear time trend for the treatment group throughout all periods. This is common practice in various empirical papers (Wolfers, 2006; Carpenter et al., 2017; Hansen et al., 2017; Hassell et al., 2020). Group-specific time trends can account for a violation of parallel trends in a dynamic TWFE setting and may enable causal identification under a "parallel trends-in-trends" (Egami & Yamauchi, 2023) assumption, which says that the divergence from parallel trends itself is linear (Strezhnev, 2024). Figure 3b shows the first differences in the outcome variable. Clearly, these run (approximately) in parallel during pre-treatment periods, which is something that group-specific linear time trends capture well (Strezhnev, 2024). Recent evidence such as by Steffens and Stuhler (2025) shows that including linear time trends substantially mitigates bias and yields consistent estimates in settings in which "differential trends are persistent at the treatment level" (Steffens & Stuhler, 2025).

The context of vehicle retirement decisions also provides a rationale for how linear time trends may capture pre-trends. The trends in the outcome variable of interest, the (log) number of vehicles per car model, are determined by how many of these old cars leave the fleet each period. As Jacobsen and van Benthem (2015) describe, owners of old vehicles are regularly confronted with a random repair cost shock. For illustrative purposes, assume that realized repair cost shocks are such that they induce scrappage of b% of vehicles in the fleet every year. Certainly, b will differ across subgroups of cars. First and foremost, older cars are more prone to higher repair cost shocks because their components are closer to the end of their lifecycle. Lower-quality vehicles may face higher repair cost shocks (relative to residual value) than higher-quality vehicles because their components are more likely to malfunction. Since treatment group car models are on average older and lighter, the latter of which proxies lower vehicle quality, than car models from the control group (Table 1), it is not surprising that the former leave the fleet at a higher rate than the latter already during pre-treatment periods. This in turn creates the constant gap in vehicle outflow rates between the two groups before treatment occurs (Figure 3b).

4.2 Hazard model

I complement the DiD analysis from the previous subsection with a hazard model estimated at the individual vehicle level. This provides insights into which households respond more or less strongly to the soot tax by examining individual vehicle retirements. I apply a version of the Cox proportional hazard model. This framework draws on a hazard function that describes the instantaneous risk that an event occurs at time t, conditional on the event not having occurred up to time t. The event of interest in this context is the permanent exit of a vehicle from the stock of registered vehicles. The hazard function $h_i(t)$ of unit i at time t in the Cox model looks as follows:

$$h_i(t) = h_0(t) \times \exp\left(X'_{it}\beta\right) \tag{7}$$

Vehicle i's hazard (risk of exit) at time t is the product of two components. The first component is the baseline hazard $h_0(t)$ that may vary over time but is common to all units. The second component is an exponentiated linear function of covariates. These covariates act multiplicatively on vehicle i's hazard, scaling the baseline hazard $h_0(t)$ by a factor of $\exp\left(X'_{it}\beta\right)$.

Accordingly, if two vehicles i and j only differ in their soot tax status, their post-treatment hazards differ by a factor of

$$\frac{h_i(t|X'_{it})}{h_j(t|X'_{jt})} = \frac{h_0(t) \times \exp\left(X'_{it}\beta\right)}{h_0(t) \times \exp\left(X'_{jt}\beta\right)} = \exp\left(X'_{it}\beta - X'_{jt}\beta\right) = \exp\left(\beta_{treat}\right) \tag{8}$$

where β_{treat} is the coefficient on the soot tax treatment indicator. In other words, if i is subject to the soot tax while j is not, the soot tax increases i's hazard by a factor of $\exp\left(\beta_{treat}\right)$. Importantly, the hazard itself is not a probability but the instantaneous rate at which the event of a vehicle's "death" occurs at time t, conditional on survival up to t. Because I have annual snapshots of the vehicle fleet, I apply the discrete version of the continuous proportional hazard model. Defining $\lambda_{it} = P\left(T_i = t \mid T_i \geq t\right)$, the probability of deregistration at t conditional on survival up to t, and the interval-specific cumulative hazard as $\Lambda_{it} = -\log\left(1 - \lambda_{it}\right)$, the regression equation of the discrete-time version of the Cox model, called complementary log-log (cloglog), is given by

$$\log\left(\Lambda_{it}\right) = \alpha_t + X'_{it}\beta \tag{9}$$

I draw on my sample of 58,541 privately owned diesel cars of vintage 2005 to 2008 as defined in Section 3.2.2. Because I observe vehicles every January 1 from 2014 to 2023, and the hazard event is "Does this vehicle disappear from the data until next year?", my sample contains years 2014 to 2022. The years 2020 to 2022 constitute the treatment period.

4.3 Identification

My identification strategy relies on variation in particulate filter installations across diesel car models produced between 2005 and 2008. The staggered rollout of filters in the production process determines whether a vehicle becomes subject to the soot tax, twelve to fifteen years after production. Causal identification requires that the timing in filter installations is uncorrelated to demand or unobserved factors that will co-determine later vehicle registrations. In particular, filter adoption should not have been strategically adjusted in response to contemporaneous demand factors or in anticipation of future registration fee reforms penalizing high emissions.

Manufacturers started installing filters as a standard equipment in anticipation of the Euro 5 emission standard, rather than in response to a broad shift in consumer demand. Likewise, variation in the timing of adopting filters as standard equipment across models arose from the initially limited availability of filters. Manufacturers primarily equipped larger and more expensive cars during the initial phase of the rollout of filters until late 2006. They did so to recover the relatively high fixed costs associated with the technology, especially while supply was still limited. Only from around 2007 onwards, once supply had significantly increased and market prices had declined, did it become economically viable to equip small and inexpensive cars with filters on a broader scale. Importantly, this implies that the timing of filter installations is not systematically linked to differences in demand. In fact, the most popular car models during those years were small and compact cars, and such cars were, on average, rather late to receive factory-fitted particulate filters.

One may wonder why manufacturers would prematurely install expensive filters years before it was required by regulation. On the supply side, car models are usually the product of relatively long development cycles during which major design decisions are taken long before pending regulation becomes binding (Bull, 2024). Locking in technological advancements early fit well into the development schedule of manufacturers. Moreover, using filters in diesel passenger cars was a new and unexplored technology at the time. Adopting the technology early allowed manufacturers to gain experience with real-world emissions performance and to avoid penalties if the filters proved insufficient. Lastly, car manufacturers sought to lock in supply chains with producers early to ensure sufficient supply of filters once Euro 5 became mandatory; and once secured, many began installing filters even before the regulation took effect.

On the demand side, recall that manufacturers in Europe produce cars for the entire European market. In many European countries, consumers valued particulate filters and had some additional willingness-to-pay for diesel vehicles with factory-fitted filters. For instance, low-emission zones (LEZs) were put in place in many cities starting in 2007, most prominently in Germany. Diesel cars with high PM emissions are completely barred from entering these zones (Margaryan, 2021). Consequently, purchasing a diesel car with the most up-to-date emissions control technology served as an insurance mechanism for German customers against potential urban driving restrictions. Conversely, this risk was virtually non-existent in the Netherlands where LEZs only restricted access for delivery vans and trucks, not for passenger cars (Regulations, 2025). While discussing the introduction of the soot tax, the government itself ruled out extending LEZs to cars as a measure to curb urban air pollution (Tweede Kamer der Staten-Generaal, 2015). In addition to that, the Dutch scheme of vehicle sales taxes was gradually revised during the mid- and late-2000s to incentivize the acquisition of vehicles with low CO_2 emissions, but not low tailpipe pollutant emissions (Kok, 2015). In fact, vehicles with small engines and thus low CO_2 emissions tended to have high PM emissions at that time because early filter adoption was concentrated in higher-class models with large engines.

Another central requirement for causal identification is the Stable Unit Treatment Value Assumption (Angrist et al., 1996). It requires that any effect of the soot tax is strictly confined to treated car models. This may not hold because of spillover effects or because of general equilibrium (GE) effects. Households that own an eligible diesel car might respond to the soot tax by replacing it with a similar yet ineligible car from the control group. GE effects on the second-hand market may further increase the value of ineligible diesel cars (Jacobsen & van Benthem, 2015). Both effects delay retirement of control group vehicles beyond what would occur without the soot tax.

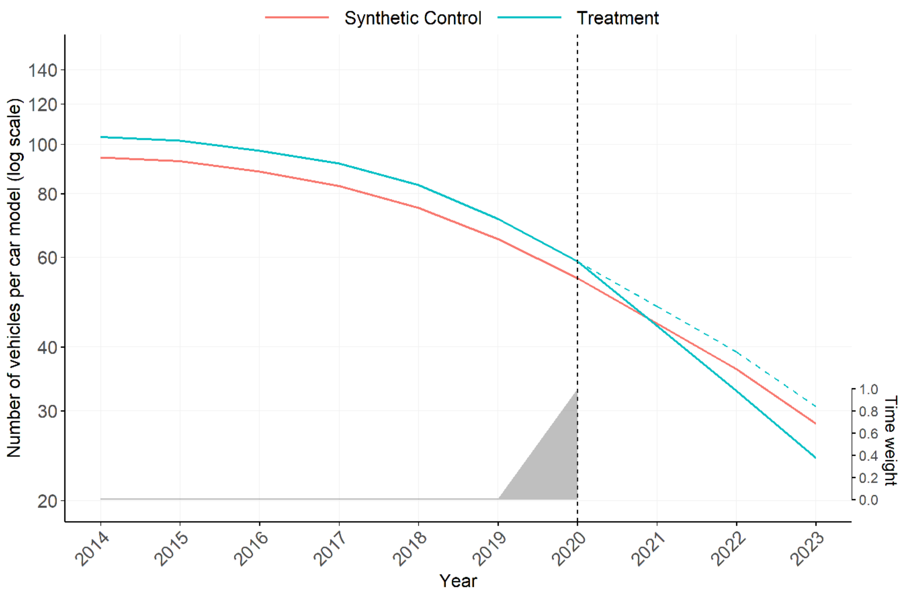
However, a closer inspection mutes the concern of a significant contamination of the control group. First, one can see in Figure 3a that the control group remains on a smooth trend after treatment starts in 2020. A contaminated control group (red line) would feature a visible kink similar to how the treatment group (blue line) features a notable kink in 2020. Second, old diesel vehicles are typically replaced with a vehicle that is several years younger. Replacing a car with one of similar or greater age is rare and occurs only in a small share of replacement events. Third, the data do not hint at the fact that owners of treated 2005-2008 diesel cars replaced these with untreated 2005-2008 diesel cars at a notable scale. If these control group cars had become more desirable replacement vehicles due to the soot tax, the share of these cars that undergo a change in ownership from one year to the next should peak after the introduction of the policy. However, the opposite happens. The share of control group diesel cars that change hands within one year are constant until 2020 and then fall rather than rise (Figure A12). Fourth, old diesel cars in the sample (those aged fifteen years or more) rarely change owners in general. Therefore, second-hand market volumes for such vehicles are very low. Their average ownership spell is about five years. This is much longer than in other countries, for example about three times as long as in the US (Jacobsen & van Benthem, 2015). GE effects in the second-hand market are therefore likely to be heavily constrained, due to the overall low trade volume in the second-hand market.

Alternatively, owners of treated diesel cars might replace their vehicle with the respective gasoline version. In Section 5.2.3, I conduct a placebo test with gasoline vehicles of similar age and do not find any evidence of such a spillover on gasoline vehicles similar to the treated diesel cars. Lastly, there is the possibility of self-selection into treatment and control group. Since the soot tax was announced in advance, vehicle owners may have selected themselves into treatment and control group. I assess such self-selection by predicting characteristics of vehicle owners using the treatment dummy. The results from this test are summarized in Appendix D. By comparing characteristics from 2016, before the soot tax was announced, with 2019, when it was initially scheduled to come into effect, I find little evidence for systematic self-selection during pre-treatment periods.

5 Results

This section contains the results from applying the empirical methods as described in Section 4 to the samples as described in Section 3. Section 5.1 starts with the results from applying DiD methods to the sample on the car model level in which I estimate the effect of the soot tax on the number of registered and polluting vehicles. Section 5.2 examines the robustness of these results. Then, I analyze individual replacement decisions by means of a hazard model in Section 5.3, using the sample on the individual vehicle level.

Figure 4: Synthetic Difference-in-Differences results, 2005-2008 diesel sample



Notes: The graph shows the results from the Synthetic Difference-in-Differences estimation, 2005-2008 diesel sample. The Synthetic Control Group is constructed to achieve parallel trends with the Treatment Group in the outcome variable during pre-treatment periods. The outcome is the natural logarithm of the variable "Number of registered vehicles" on the car model × year level.

5.1 DiD results, 2005-2008 diesel sample

I first assess how the soot tax affects the number of polluting vehicles in the vehicle stock. Figure 4 presents the outcome of the SDiD estimator (Equation 4) in the 2005-2008 diesel sample. The graph shows how the average number of registered vehicles per car model evolves over the observation period in the treatment (blue) and synthetic control (red) group. Because I run SDiD on the natural logarithm of this variable, I show the y-axis using a logarithmic scale.

On January 1, 2014, the treatment group contains, on average, 103 registered vehicles per car model, compared to 94 vehicles per car model in the control group. Initially, both series decline in parallel, reaching about 59 and 55 vehicles per model by January 1, 2020, which corresponds to a decrease of approximately 42%. Once the soot tax comes into effect in 2020, there is a clear kink in the treatment group, indicating that treated vehicles start leaving the fleet at a higher rate. By January 1, 2023, the number of vehicles in the treatment group is about 59% relative to January 1, 2020. In contrast, there is no such kink in 2020 for the synthetic control group. Instead, it appears to continue its smooth decline and experiences a further reduction of only about 48% over the subsequent three years.

The dashed blue line depicts the counterfactual trajectory for the treatment group, inferred from how the synthetic control evolves post-treatment. In the SDiD framework, the average divergence of the treatment group from its counterfactual corresponds to the estimated average treatment effect. Beyond that, the grey area on the bottom displays the SDiD time weights $\{\hat{\lambda}_t^{sdid}\}$ as in Equation 4. All weight is put on the last pre-treatment period so that estimated treatment effects are computed relative to January 1, 2020. In the case at hand, all weight rests on the last pre-treatment period because of the smooth parallel trends right until treatment starts. In settings with more volatile pre-treatment outcomes, SDiD generally puts positive weights on several pre-treatment periods.

One potential threat to identification is that the observed kink for the treatment group in 2020 might reflect a discrete, age-related increase in retirement rates rather than the policy effect itself. Treated car models are, on average, older than untreated car models (mostly 13 or 14 vs. mostly 11 or 12 years old). The observed kink could, in principle, be rationalized by a sudden, discontinuous jump in scrappage rates once vehicles reach an age of 13 to 14 years. However, three pieces of evidence speak against this interpretation. First, existing empirical evidence shows that scrappage rates increase smoothly and continuously with age, without any discontinuities around 13 to 14 years (Jacobsen & van Benthem, 2015; Bento et al., 2018; Laborda & Moral, 2019; Held et al., 2021; Greene & Leard, 2024). Second, no comparable kink appears for the control group roughly two years later when its car models reach a similar age. Third, I will show in Figure 6 in Section 5.2.1. that the kink also appears for even older diesel car models once the soot tax takes effect. Taken together, these observations strongly suggest that the SDiD estimates capture the causal impact of the soot tax rather than differences in age composition.

Column (2) in Table 3 displays the estimated treatment effect implied by Figure 4. Column (1) additionally shows the SDiD estimate without including covariates. The results are qualitatively robust to either choice. The point estimates indicate that the soot tax reduced the number of polluting vehicles by on average 16.5% (without covariates) or 16.4% (with covariates) over the course of the first three years of the policy. The estimated effect is highly significant. The jackknife standard error implies a 95 percent confidence interval ranging from 11 to 22 percent. Columns (3) and (4) display the results from using the SCM as in Equation 5. The point estimates indicate an average effect of -14.9% (without covariates) or -15.7% (with covariates) on the number of polluting vehicles within the first three years of the policy.[14] In total, both SDiD and SCM yield very similar quantitative results. The point estimates imply that about one in six to one in seven polluting diesel vehicles are additionally deregistered within three years because of the soot tax.

Table 3: Regression results for 2005–2008 diesel sample.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SDiD		SCM		FE regression			
Soot tax	-0.165***	-0.164***	-0.149***	-0.157***	-0.145***	-0.129***	-0.160***	-0.154***
	(0.027)	(0.027)	(0.045)	(0.037)	(0.022)	(0.025)	(0.020)	(0.023)
TG × Year					-0.021**	0.019*		
					(0.009)	(0.010)		
Unit × Year							✓	✓
Car type FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓		✓	
Vintage-year FE						✓		✓
Covariates		✓		✓	✓	✓	✓	✓
No of obs.	3,420	3,420	3,420	3,420	3,420	3,420	3,420	3,420
Adj. R^2					0.941	0.949	0.994	0.994
Within R^2					0.173	0.108	0.904	0.882

Notes: SDiD = Synthetic Difference-in-Differences; SCM = Synthetic Control Method; FE = Fixed Effects. The table contains point estimates for the effect of the soot tax on the log number of registered vehicles of treated car models in the sample of 2005-2008 diesel car models. Covariates contain a car model's average log household disposable income, log owner age, log population density of owner's home municipality, and dummies for being male, college-educated, married, and owning a home. Columns (1) and (2) are Synthetic Difference-in-Differences estimates (Equation 4), without and with covariates, respectively. Columns (3) and (4) are SCM estimates (Equation 5), without and with covariates, respectively. Columns (5) to (8) are fixed effect regression estimates (Equation 6). The specifications differ in whether they employ group- or unit-specific linear time trends, and whether they employ year fixed effects or vintage×year fixed effects. Standard errors are in parenthesis. SDiD standard errors are jackknife errors, SC standard errors are placebo errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The long version of this table is Table A6.

Columns (5) and (6) show results from running TWFE regressions with a group-specific linear time trend as an additional regressor, and columns (7) and (8) apply even more granular unit-specific linear time trends. The choice between unit- and group-specific linear time trends does not qualitatively change the estimates[15]. In the application at hand, point estimates remain statistically equal to the results in columns (1) to (4). Only the point estimate in column (6), combining group-specific trends with vintage × year fixed effects, is somewhat smaller at 12.9 percent.

Next, I investigate whether the effect grows over time. Vehicle replacement decisions are unlikely to be immediate due to market frictions and information asymmetries. Therefore, the average effect of 15 to 16 percent within the first three years likely accumulates from year to year. The fact that we see a change in slope rather than a pure level shift in Figure 4 already illustrates that the soot tax affects vehicle registrations gradually rather than instantaneously. Table 4 quantifies this insight.

[14] The underlying graph is Figure A13 in the Appendix.
[15] Steffens and Stuhler (2025) show this equivalence of group- and unit-specific linear time trends analytically.

Table 4: SDiD & SCM per period results for 2005-2008 diesel sample

	(1)	(2)	(3)	(4)	(5)	(6)
	SDiD		SCM		SDiD	
Soot tax	-0.165***		-0.149***		-0.168***	
	(0.027)		(0.045)		(0.030)	
TG × '20						-0.026*
						(0.014)
TG × '21		-0.088***		-0.070**		-0.128***
		(0.016)		(0.033)		(0.026)
TG × '22		-0.176***		-0.179***		-0.222***
		(0.029)		(0.048)		(0.039)
TG × '23		-0.233***		-0.197***		-0.284***
		(0.039)		(0.048)		(0.048)

Notes: The table shows the results from SDiD and SCM on the 2005-2008 diesel car model sample. "Soot tax" is the treatment indicator for all post-treatment years 2021-2023 whereas the other terms are per-period treatment terms. Columns (1) and (3) are the point estimates using only the treatment indicator as in Equations 4 and 5, respectively. Columns (2) and (4) contain the implied per-period estimates with standard errors as in Facure (2025). Columns (5) apply SDiD but define 2020 as a post-treatment year to incorporate the effect that the soot tax was initially scheduled to come into effect on January 1, 2019. The long version of this table is Table A7.

Column (2) shows that, according to the SDiD estimates, the soot tax has lowered the number of registered polluting vehicles by 8.8% within the first year of the policy. This effect doubles to 17.6% after two years, and increases further to 23.3% after three years. I estimate standard errors by (i) including only the respective post-treatment period in the data and (ii) then running the jackknife standard error algorithm in the synthdid R package from Arkhangelsky et al. (2021) as suggested by Facure (2025). All per-period estimates turn out statistically significant. I also run SCM (column (4)) as a robustness check. It corroborates both the accumulative nature of the soot tax as well as the per-period magnitudes.

The soot tax was initially scheduled to take effect on January 1, 2019, but its implementation was postponed to January 1, 2020, with this decision being made as late as the summer of 2018 due to administrative delays. Anecdotal evidence indicates that the majority of car owners became aware of the forthcoming soot tax only in November 2019 when the tax authorities sent official notifications to affected individuals (van Wingerden, 2019). Despite this, it remains plausible that some car owners anticipated the introduction of the soot tax at the originally scheduled launch date. To account for potential anticipatory responses, I re-define the treatment start date to 2019 and re-estimate the SDiD model accordingly. Column (5) of Table 4 shows that this altered definition of the treatment period leaves the estimated average effect virtually unchanged at -16.8%. As column (6) shows, I further find weak evidence that the soot tax already affected vehicle holdings in 2019. The point estimate indicates a 2.6% reduction in the number of polluting vehicles already in 2019. This estimated effect is statistically significant at the ten percent level. Further specifications with even earlier treatment start dates do not find any other statistically significant effects before 2019.

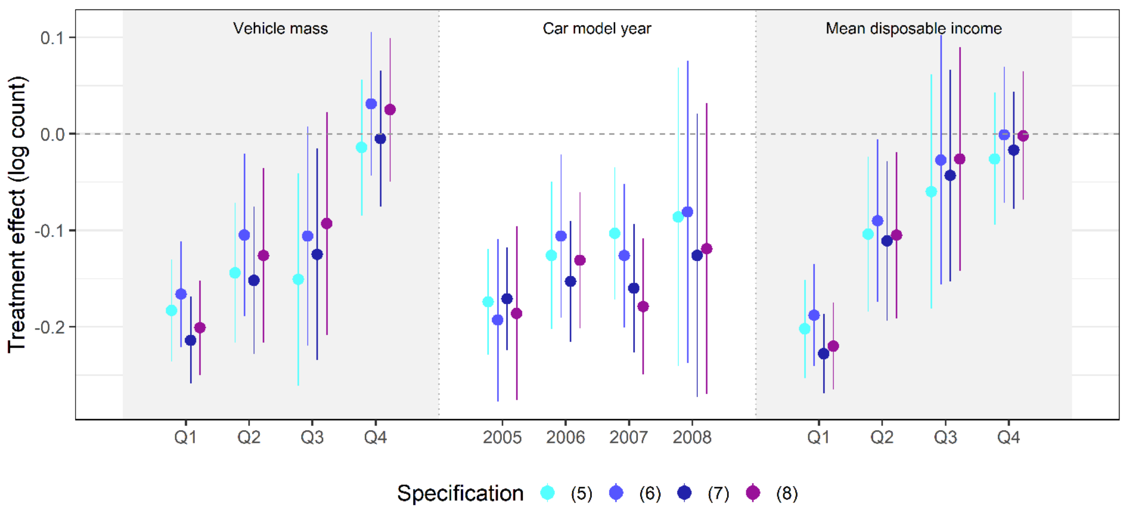
A further dimension of treatment effect heterogeneity pertains to variation in effect size across covariates. Because synthetic methods do not accommodate interaction terms, I estimate variations of the TWFE model in Equation 6, interacting the treatment indicator $D_i Post_t$ with covariates. The specifications differ in whether they include group- or unit-specific linear time trends and in whether they employ common time fixed effects μ_t or vintage-by-year fixed effects μ_{vt}. I assess treatment effect heterogeneity along several covariates. First, I differentiate treatment effects by vehicle mass because this is a close proxy for vehicle quality in the European car market. Light vehicles with low mass are generally inexpensive cars with few accessories; heavy vehicles are generally luxurious and expensive cars with high-quality accessories. Second, I differentiate by vehicle vintage. Whereas a vehicle's residual value decreases with vehicle age, the registration fee is independent of vehicle age. This means that the soot tax constitutes a larger increase in registration fee liability relative to residual value for older cars (= earlier vintages). The same repair cost shock should thus lead to more scrappage for older cars. Third, I differentiate by household disposable income. Because this is average income of vehicle owners within one car type, this measure is strongly correlated with vehicle mass (poorer households own lighter (more inexpensive) cars on average).

Figure 5 presents the estimated treatment effects of the soot tax across different covariate values. Specifications (5) to (8) relate to the TWFE specifications as in Table 3 with their respective choice of year or vintage × year fixed effects as well as group- or unit-specific linear trend terms. The left-hand panel of Figure 5 plots how the effect varies by vehicle mass. The estimates reveal that the policy's impact diminishes as car models become heavier. For the lightest quartile of car models, that is those with the lowest 25% of vehicle mass, the soot tax reduces the number of polluting vehicles by roughly 17 to 21%. This effect moderates to about 10 to 15% for models in the second and third quartiles, and becomes statistically indistinguishable from zero for the heaviest 25% of car models. Clearly, the soot tax most strongly accelerates the retirement of lighter car models, while its impact tapers off and approaches zero for heavier vehicles.

The center panel of Figure 5 illustrates how the impact of the soot tax also varies with vehicle vintage. The emerging pattern shows that the effect is larger for older car models (those of an earlier vintage). Across the four specifications, the soot tax reduces the number of polluting vehicles by about 18% among 2005 vintages, 13% among 2006 and 2007 vintages, and roughly 11% among 2008 vintages. However, the estimates for the 2008 vintages lack statistical significance due to the small number of treated car models in this late cohort just before the introduction of Euro 5. Overall, the results show that the soot tax accelerates vehicle retirement more strongly among older car models, which are expected to have lower residual values, consistent with the idea that the policy moved more of these vehicles across the margin from continued use to scrappage.

In the right-hand panel, I bin car models into four quartiles by the mean household disposable incomes of their owners. Clearly, the effect of the soot tax on vehicle deregistrations is the largest among those car models that tend to be owned by lower-income households. I estimate a 20% reduction in the number of polluting vehicles among car models in the bottom quartile, about 10% in the second quartile, and an effective zero for the third and fourth quartile. However, these results alone do not imply that lower-income households are inherently more responsive to the soot tax. Instead, it indicates that the effect is strongest for car models predominantly owned by lower-income households.

Figure 5: Fixed Effect regression, treatment effect heterogeneity, 2005-2008 diesel sample



Notes: The graph shows point estimates of heterogeneous treatment effects from the Fixed Effects regressions, 2005-2008 diesel sample. The analyzed heterogeneity is in terms of vehicle mass quartile (left), vintage (middle), and mean disposable income quartile (right).

I set up a model of vehicle replacement that explains these heterogeneous treatment effects in Appendix C. In this model, similar to Jacobsen and van Benthem (2015), the owner of an old car faces a random repair cost shock and must then decide whether to keep or replace the old car. The utility from keeping the old car depends on the car's residual value minus ownership costs, which is registration fee liability plus repair costs. When choosing to replace the old car, the owner obtains the car's scrap value and avoids paying both the registration fee and the repair cost.

I express the utility difference between keeping and replacing the car as a function of the car's ratio of residual value to registration fee, and derive that a lower ratio increases a car's replacement probability. This rationalizes why the soot tax, which in itself is an exogenous reduction in the ratio of residual value to registration fee, increases a vehicle's replacement probability and thereby decreases the number of old polluting vehicles. The model further rationalizes the finding that treatment group vehicles already leave the fleet at a higher rate than control group vehicles even before treatment starts (Figure 3b). One of the structural differences between treatment and control diesel car models is their vehicle mass. I show that in the Dutch registration fee schedule, a lower vehicle mass is associated with higher registration fee liability relative to vehicle residual value. I approximate vehicle residual values of the relevant 2005-2008 diesel car models, compute model-specific registration fee liability, and find that the annual registration fee liability of light car models below 1,200 kg is on average about 25 to 30 percent of their residual value, whereas it falls to only about 15 percent for heavy car models above 1,500 kg (Figure A14).

I then use the model to rationalize why lighter car models respond more strongly to the soot tax than heavier car models (see the left-hand panel of Figure 5) by means of a vehicle quality measure. Vehicle mass and quality are strongly positively correlated in the European car market so that comparing light and heavy vehicles essentially amounts to comparing lower- and higher-quality vehicles. Vehicle quality affects the decision to keep or replace an old vehicle through three channels. Higher quality (i) reduces the risk of a high repair cost draw and (ii) is associated with a higher residual-value-to-fee ratio, both of which make keeping the car more likely, but (iii) also raises scrap value, which makes replacing more likely. I use the finding that higher vehicle quality is associated with lower pre-soot-tax replacement rates to show that, overall, the effect of the soot tax on vehicle replacement is larger when the vehicle is of lower quality. Intuitively, lower-quality vehicles are closer to the replacement margin. Even modest increases in ownership costs, such as the soot tax, are enough to trigger scrappage. In contrast, higher-quality vehicles on average have higher residual value and face lower repair cost shocks relative to residual value, which implies that they tend to remain well below that margin. Accordingly, the same shock to ownership costs induces a larger change in replacement probability for low-quality (light) vehicles than for high-quality (heavy) ones.

5.2 Robustness checks & Placebo tests

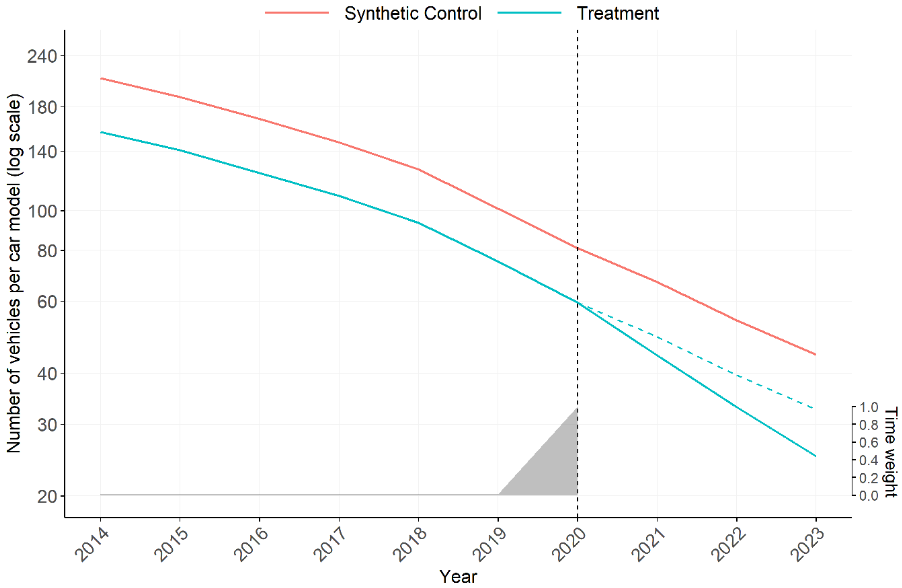
In this section, I examine the robustness of my empirical approach and my results from Section 5.1. First, I replicate the main results from the DiD car-model-level analysis using a sample of older diesel car models, comparing their numbers to equally old but untreated gasoline car models. In addition to that, I conduct two placebo tests in which I assign a fictitious ("placebo") treatment to samples consisting entirely of untreated units. Under the identifying assumptions of no treatment spillovers and no concurrent confounding shocks, estimated effects in these placebo samples are zero in expectation. Insignificant placebo estimates therefore support the credibility of my identification strategy.

5.2.1 Robustness check: 1990-2004 sample

In Section 5.1, I compare diesel car models of vintages 2005–2008 and find that the soot tax reduces the number of polluting vehicles by about 16% on average across specifications. The identifying variation comes from the gradual introduction of particulate filters during their production process. To assess the robustness of these findings, also with regard to the covariate imbalance in the 2005–2008 diesel sample, I draw on a secondary sample of passenger car models with vintages 1990–2004.

Diesel car models produced until 2004 universally lack factory-fitted filters and are therefore all subject to the soot tax. Gasoline car models of the same vintages serve as a natural control group because the soot tax is never levied on gasoline vehicles. Treatment assignment in this sample is determined solely by fuel type. Accordingly, the identifying variation may be contaminated by unobserved factors that influence scrappage and correlate with fuel choice, potentially yielding biased estimates. Nonetheless, the sample allows me to assess the impact of covariate

Figure 6: Synthetic Difference-in-Differences results, 1990-2004 sample



Notes: The graph shows the results from the Synthetic Difference-in-Differences estimation, 1990-2004 sample. The Synthetic Control Group is constructed to achieve parallel trends with the Treatment Group in the natural logarithm of the variable "Number of registered vehicles" on the car model × year level.

imbalance because it runs in the opposite direction compared to the 2005–2008 diesel sample (see Table A8). Therefore, obtaining qualitatively similar results using this sample reinforces the credibility of the baseline findings and suggests that the empirical approach adequately accounts for remaining covariate imbalances.

The results from running SDiD in this sample of 1990-2004 can be seen in Figure 6. The treatment and synthetic control group evolve in parallel during pre-treatment periods. The number of vehicles decreases by about 62% between January 1, 2014 and 2020 in both groups. Once the soot tax comes into effect, there is a notable kink in the treatment group (just like the kink in Figure 4 in Section 5.1) whereas the control group continues on its smooth downward trend.[16] The trends clearly begin to diverge once the soot tax takes effect and continue to diverge thereafter. Another 58% of diesel vehicles are deregistered during the first three years of the policy, compared to 45% of gasoline vehicles over the same period.

Table 5 shows that the implied point estimates are of similar magnitude to those in Section 5.1. If anything, they are somewhat larger, ranging mostly from 16.4 to 19.6%, with one notably higher estimate of 23.8% obtained from the SCM with

[16] Since the y-axis is in logs, a straight line corresponds to a constant percent reduction from period to period.

Table 5: Regression results for 1990–2004 sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SDiD		SCM		FE regression			
Soot tax	-0.196***	-0.196***	-0.192***	-0.238***	-0.193***	-0.176***	-0.185***	-0.164***
	(0.024)	(0.024)	(0.029)	(0.029)	(0.016)	(0.016)	(0.016)	(0.015)
TG × Year					-0.065***	-0.087***		
					(0.007)	(0.007)		
Unit × Year							✓	✓
Car type FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓		✓	
Vintage-year FE						✓		✓
Covariates		✓		✓	✓	✓	✓	✓
No of obs.	11,590	11,590	11,590	11,590	11,590	11,590	11,590	11,590
Adj. R^2					0.957	0.962	0.995	0.996
Within R^2					0.297	0.332	0.926	0.933

Notes: SDiD = Synthetic Difference-in-Differences; SCM = Synthetic Control Method; FE = Fixed Effects. The table contains point estimates for the effect of the soot tax on the log number of registered vehicles of treated car models in the sample of 1990-2004 car models. Covariates contain a car model's average log household disposable income, log owner age, log population density of owner's home municipality, and dummies for being male, college-educated, married, and owning a home. Columns (1) and (2) are Synthetic Difference-in-Differences estimates (Equation 4), without and with covariates, respectively. Columns (3) and (4) are SCM estimates (Equation 5), without and with covariates, respectively. Columns (5) to (8) are fixed effect regression estimates (Equation 6). The specifications differ in whether they employ group- or unit-specific linear time trends, and whether they employ year fixed effects or vintage×year fixed effects. Standard errors are in parenthesis. SDiD standard errors are jackknife errors, SC standard errors are placebo errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The long version of this table is Table A9.

covariates. Intuitively, it makes sense that treatment variation based on fuel type rather than quasi-random filter adoption gives rise to slightly biased estimates. On average, diesel vehicles accumulate roughly twice the annual mileage of gasoline vehicles, which structurally lowers their residual value even at the same vehicle age. A lower residual value, in turn, amplifies the relative burden of the soot tax and yields an estimate that is biased away from zero (see Appendix C). Nonetheless, the point estimates remain similar in magnitude to those from Section 5.1.

Furthermore, when assessing treatment effect heterogeneity, I again find that the effect of the soot tax is larger for lighter (lower-quality) car models (see Figure A15 in the Appendix). In contrast to Section 5.1., I find that the effect of the soot tax in the 1990-2004 sample tends to decrease with vehicle age. It is largest at more than 20 percent for Euro 3 cars (vintages 2000-2004), smaller but still significant at 8 to 14 percent for Euro 2 cars (vintages 1996-1999), yet insignificant for pre-Euro 2 cars from 1995 and before. Overall, this pattern aligns with prior evidence that scrappage rates increase with vehicle age up to a peak at around 20 years, after which they begin to decline (Greene & Leard, 2024).

As my model in Appendix C shows, lower baseline replacement probabilities translate into a smaller effect of the soot tax. The finding that the effect starts to become smaller for vehicle ages beyond 20 years and approaches zero for vehicle ages beyond 25 years is presumably connected to endowment effects and loss aversion (a large positive η_i in the model in Appendix C). Moreover, residual values of very old, well-maintained vehicles may rise again as such cars become increasingly scarce. This plausible mechanism shrinks the baseline replacement probability sufficiently small to render the effect of the soot tax insignificant.

5.2.2 Placebo test: Diesel car models in Germany

The first placebo test simulates a placebo soot tax in the German vehicle fleet. Germany is a suitable country to conduct a placebo test for several reasons. Most importantly, it did not introduce a registration fee surcharge for polluting diesel vehicles in 2020. Beyond that, the Dutch and German car markets follow the same European legal baselines and vehicle standards. Diesel penetration in the two countries has historically followed similar trends. Both countries feature high incomes, well-developed road networks, widespread public transit alternatives, and comparable car ownership rates. Lastly, the German registration fee schedule is particularly well-suited for a placebo test for two additional reasons. First, German registration fees are much lower than those in the Netherlands, diminishing their overall importance in vehicle retirement decisions. Second, German registration fees do not feature surcharges for diesel cars that fail to meet Euro 5 standards. With regard to PM emissions, registration fees are only slightly higher for diesel cars that fail to comply with Euro 3 standards. Permitted PM emissions under Euro 3 are ten times higher than under Euro 5, and crucially, the relevant 2005–2008 diesel vintages were produced under Euro 4 standards so that none of these vehicles are subject to elevated fees.

The German Federal Motor Transport Authority (KBA) provides me with a representative 40% subsample of the registered German vehicle fleet on every January 1 from 2015 to 2023 (Kraftfahrt-Bundesamt (KBA), 2023). The KBA data only feature little information about the car holders. It includes month and year of birth, gender, a dummy for being self-employed, state of residence, and a variable indicating region type. I approximate additional sociodemographic characteristics using municipal- or county-level data on educational attainment, purchasing power, and marital status. Conversely, the KBA data contain much more granular vehicle covariates. This allows me to identify more diesel car models. Restricting the data to privately owned passenger cars, I ultimately obtain a panel data set of 1,342 car models over 9 years. Table A10 contains the relevant summary statistics.

I run SDiD on this sample and obtain Figure 7. The blue line corresponds to the placebo treatment group of diesel models with PM emissions above the soot tax eligibility threshold of 5 mg/km. The red line is the synthetic control group. SDiD achieves parallel trends before placebo treatment starts in 2020. There are on average 148 vehicles per car model in the placebo treatment group and 187 in the synthetic control group in 2015. These numbers decline in parallel by about 44% until 2020, a reduction strikingly similar to that observed in the Netherlands. The crucial contrast is that, in Germany, the two groups continue to evolve in parallel in the subsequent years. The SDiD results imply an insignificant ATT of -1.0% (see Table A11). Any divergence between the two groups following the placebo treatment remains statistically insignificant, even after three years. This statistically insignificant effect in the German registration data corroborates that it is indeed the Dutch soot tax that has driven the accelerated outflow of old polluting diesel vehicles from the Dutch fleet since 2020. I further run both the SCM as well as TWFE regressions on the sample. All resulting point estimates oscillate around zero and turn out statistically insignificant at conventional significance levels (see Table A11).

At the same time, the German data feature the same pattern of non-parallel trends. Placebo-treated diesel car models in Germany have higher outflow rates than control-group models already prior to 2020. When running TWFE with linear time trends, the point estimates on the linear time trend of the placebo treatment group, "TG $\times$ Year", are negative and significant, and vintage $\times$ year fixed effects only partially absorb the non-parallel trends.[17] This underscores that the non-parallel trends arise from some fundamental differences between treated and untreated car models, not from some peculiarity of the Dutch vehicle fleet. Creating a synthetic control group accounts for these imbalances and thereby allows for clean causal identification.[18]

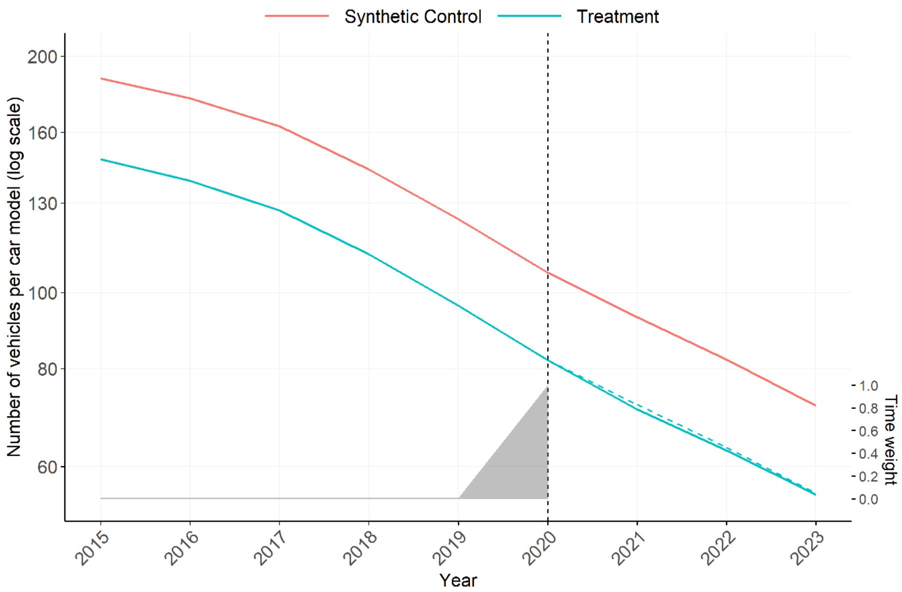
5.2.3 Gasoline car models in the Netherlands

The second placebo test draws on a sample of Dutch gasoline vehicles, which are never subject to the soot tax. In principle, two ways in which car owners could respond to the soot tax is that they either switch to a non-treated diesel car or to a gasoline car similar to their treated diesel car. The placebo test among gasoline

[17] See columns (5) and (6) of Table A11.

[18] Adding linear time trends to TWFE regressions acts as an imperfect parametric correction that removes the first-order (linear) bias due to non-parallel trends.

Figure 7: Synthetic Difference-in-Differences results, sample of German 2005-2008 diesel car models



Notes: The graph shows the results from the Synthetic Difference-in-Differences estimation in the sample of German 2005-2008 diesel car models. I assign a placebo treatment to all identified car models with registered PM emissions above 5 mg/km. The Synthetic Control Group is constructed to achieve parallel trends with the Placebo Treatment Group in the outcome variable up until 2020. The outcome is the natural logarithm of the variable "Number of registered vehicles" on the car model × year level.

car models therefore serves as a test for spillover effects. Car owners might perceive the soot tax as a signal to shift away from diesel technology altogether. In this case, some affected car owners might replace their diesel car with its gasoline counterpart, extending the lifetime of these gasoline vehicles and increasing their post-treatment numbers in the vehicle fleet. Correspondingly, an estimated zero treatment effect among gasoline car models indicates the absence of such spillover effects.

To run this placebo test, I draw on gasoline car models of vintages 2005-2008. I restrict the sample to those gasoline models that have a diesel counterpart. I assign placebo treatment status to gasoline car models based on their similarity to the diesel car models in the actual treatment and control groups in the 2005-2008 diesel sample, assuming larger similarity equals more appealing replacement. I measure similarity in terms of both vehicle characteristics and owner characteristics. Specifically, I construct a similarity score based on the car model covariates vehicle mass, number of vehicle cylinders, vintage, as well as on the owner covariates household disposable income, owner age, population density of place of residence, and dummies for having college education, being male, being married, and homeownership.

I compute standardized pre-treatment covariate means at the car model level and estimate two multivariate kernel densities: one for the treatment-group covariates (tax-eligible diesel models) and one for the control-group covariates (tax-ineligible diesel models). Then, for each gasoline car model, I evaluate its covariate profile under both kernel density functions to obtain a continuous measure of its similarity to the treated and control diesel models. The placebo treatment propensity score for gasoline car model i with covariate profile x_i is then given by

$$p_i = \frac{f_{treat}(x_i)}{f_{treat}(x_i) + f_{control}(x_i)}$$

$f_{treat}(x_i)$ and $f_{control}(x_i)$ denote the kernel density estimates of the covariate distributions for the treated and control diesel car models, respectively, evaluated at x_i. Intuitively, $f_g(x_i)$ will be larger the more similar gasoline car model i is to car models in group g. In particular, $f_{treat}(x_i)$ will be large relative to $f_{control}(x_i)$ when gasoline car model i more closely resembles the treated diesel car models, yielding a propensity score p_i close to one. Conversely, when a gasoline car model more closely resembles the control-group diesel models, its propensity score will be near zero.

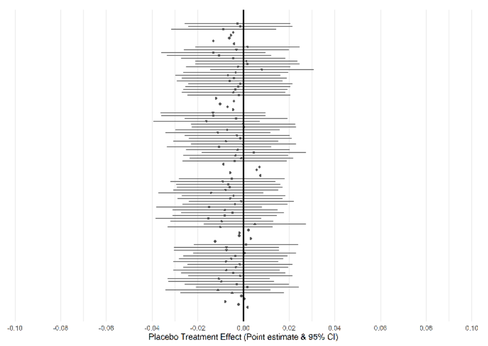
I use the propensity score as the probability of being assigned placebo treatment. Specifically, for each gasoline car model i, I draw a random treatment status according to

$$Z_i \sim \text{Bernoulli}(p_i)$$

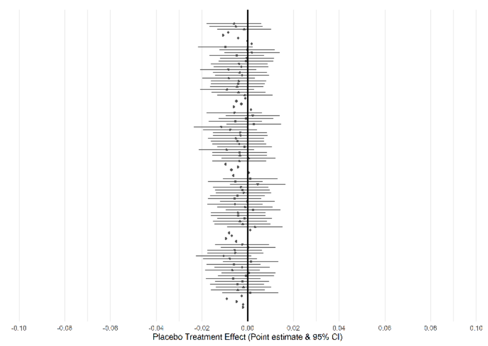
I repeat this random draw 100 times to generate 100 distinct placebo splits. This provides me with 100 placebo treatment samples of 2005-2008 gasoline car models. I use these samples to replicate the analysis as in Section 5.1. The results from the placebo tests are displayed in Figure 8. I consider four different specifications: (1) TWFE without group-specific linear time trends, the same specification as column (6) of Table 3 (top-left); (2) TWFE with group-specific linear time trends, the same specification as column (8) of Table 3 (top-right); (3) SDiD (bottom-left); and (4) the SCM (bottom-right). The point estimates from the fixed effects regressions are statistically insignificant across all 100 placebo splits.

The point estimates from standard TWFE (Figure 8a) fluctuate between -0.016 and 0.008, with a mean of -0.005 and average standard errors of 0.012. When adding group-specific linear time trends (Figure 8b), the confidence intervals become smaller, but point estimates remain statistically zero at the five percent level across all 100 placebo splits. Figure 8c contains the point estimates and confidence intervals from the SDiD estimates. The average placebo treatment effect is estimated at -0.008 with a mean standard error of 0.006. Four out of the 100 point estimates are significant at the five percent level, which is consistent with the number of false positives expected by chance under a five percent significance threshold. Lastly, the SC point estimates in Figure 8d fluctuate more heavily and tend to be larger in absolute value, but larger standard errors render all point estimates insignificant. Overall, the placebo test in the 2005-2008 gasoline sample consistently yields zero estimated treatment effects across all empirical methods, supporting the notion (i) that the soot tax did not spill over to similar gasoline car models, and (ii) that the estimated effect from Section 5.1 in fact arises because of the soot tax.

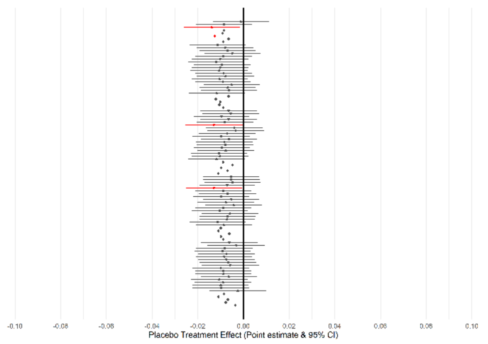
Figure 8: Placebo test results using the 2005-2008 gasoline sample



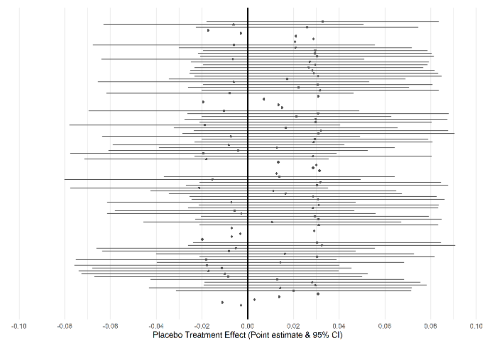
(a) Fixed effects estimation, without linear time trends



(b) Fixed effects estimation, with linear time trends



(c) SDiD estimation



(d) SC estimation

Notes: The graphs show estimated treatment effects in the 2005-2008 gasoline sample after assigning 100 placebo treatments. Placebo status is assigned based on a propensity score measuring car model similarity to treated and untreated diesel car models from the 2005-2008 diesel sample. The solid black vertical line indicates zero. Black confidence bands indicate statistical insignificance at five percent; red confidence bands indicate statistical significance.

5.3 Hazard model results

The previous results have shown that the soot tax significantly reduces the number of polluting vehicles. This decline is predominantly driven by light-weight cars, older vintages, and car models that are, on average, more prevalent among lower-income households. Importantly, the last finding does not mean that lower-income households are inherently more tax-sensitive. It may instead reflect differences in the kinds of cars they hold. To examine heterogeneous responses among car owners, which directly relate to distributional effects, I draw on a vehicle-level hazard model.

In my baseline specification, I run OLS on the cloglog model as in Equation 9. I control for vehicle mass, cylinder volume, and CO_2 emissions (all in log), as well as for the household/owner covariates household disposable income, population and population density of their hometown, owner age, the number of vehicles that the household owns, and dummies for being married, being male, and owning a home. I further include time and brand fixed effects.

Column (1) of Table 6 presents the results of this baseline specification. The reported coefficients are presented as hazard ratios. The estimated coefficient for the soot tax is 0.166, which corresponds to an increase in the hazard ratio of $\exp(0.166) = 1.181$ (the first row of column (1)). In other words, the soot tax increases a vehicle's hazard, that is its risk of deregistration at any point in the post-treatment period, by 18.1%. This effect is highly significant: the 95 percent confidence interval stretches from 1.134 to 1.227.

The point estimate on the treatment group indicator (last row with point estimates) is also significantly larger than one, indicating that treatment group vehicles generally face a higher baseline risk of deregistration. This is exactly in line with the non-parallel trends observed when following the number of registered vehicles at the car-model level in Section 5.1. Treatment group vehicles already have higher outflow rates (hence higher "realized" per-period hazards) even before treatment starts.

The results further reveal that other covariates also significantly affect a vehicle's hazard (see Table A12). A vehicle's risk of deregistration rises when it is not the only vehicle that the household owns. Households with the financial capability to own several cars plausibly have the means to also replace them more regularly so that the presence of other cars in the household raises an old car's hazard in general.

Table 6: Hazard model, cloglog results using 2005–2008 diesel vehicles

Dep. var.: Log(hazard rate)	(1)	(2)	(3)
Soot tax	1.181		1.208
	[1.134, 1.227]		[1.089, 1.327]
Soot tax, 2020		1.224	
		[1.162, 1.286]	
Soot tax, 2021		1.176	
		[1.095, 1.257]	
Soot tax, 2022		1.092	
		[1.001, 1.183]	
Soot tax × High income			1.078
			[1.005, 1.151]
Soot tax × High pop dens			1.091
			[1.016, 1.167]
Soot tax × Home			0.936
			[0.871, 1.001]
Soot tax × High pop			1.002
			[0.933, 1.071]
Soot tax × High age			0.963
			[0.902, 1.025]
Soot tax × Married			1.026
			[0.959, 1.093]
Soot tax × Male			0.964
			[0.895, 1.034]
Soot tax × 2 cars			0.949
			[0.882, 1.017]
Soot tax × 3 cars			0.899
			[0.802, 0.997]
Soot tax × 4 cars			0.986
			[0.810, 1.161]
Soot tax × 5+ cars			0.980
			[0.761, 1.199]
Treatment group	1.083	1.083	1.086
	[1.054, 1.111]	[1.055, 1.111]	[1.055, 1.117]
Household controls	✓	✓	✓
Vehicle controls	✓	✓	✓
Time FE	✓	✓	✓
Brand FE	✓	✓	✓
No. of obs.	268,059	268,059	268,059

Notes: FE = Fixed Effects. The table contains point estimates for the effect of the soot tax on a vehicle's hazard of permanent deregistration within the next year (from this year's to next year's January 1). The point estimates are the exponentiated estimated coefficients. A value of one corresponds to a zero effect. Interpretation is in percent, e.g. 1.181 means "The hazard increases by 18.1%". All columns refer to Equation 9. Household controls are household disposable income, population and population density of home municipality, owner age, as well as dummies for homeownership, married, male, as well as dummies for how many cars the household owns overall, one (omitted), two, three, four, or five and more cars. Vehicle controls are vehicle mass, cylinder volume, and CO_2 emissions. Column (1) contains only a treatment indicator. Column (2) contains separate indicators for every post-treatment period. Column (3) contains the treatment indicator in simple form as well as interacted with binary variables for above-median ("high") 95% confidence intervals are in square brackets. I report confidence intervals instead of standard errors because H_0 posits a coefficient value of one, which complicates interpreting significance by eyeballing coefficient values and standard errors. Confidence interval provide a more intuitive understanding of statistical significance in this setting. The long version of this table is Table A12.

Alongside this, the point estimates on the year dummies consistently grow from year to year, highlighting how the risk of deregistration continuously rises as a vehicle ages (see Figure A16). A higher vehicle mass also increases the vehicle's hazard whereas a larger engine (measured by cylinder volume) slightly reduces it.

Moreover, a vehicle's hazard varies significantly with owner characteristics. The risk of vehicle retirement is higher when the vehicle is registered in a municipality with higher population, and when the owner is married or male. At the same time, the vehicle's baseline hazard is lower when the household has a higher disposable income, when the owner is older, and when the owner resides in their own home (instead of being a renter). Full regression results are reported in Table A12 in the Appendix.

Next, I use the flexibility of the cloglog model to assess treatment heterogeneity on the vehicle level. I first assess heterogeneous treatment effects across time. Column (2) of Table 6 shows that the effect of the soot tax is strongest in the first post-treatment period, but remains positive in subsequent periods, gradually tapering off over time. The soot tax increases the vehicle's hazard by 22.4% in the first year, with the effect declining to 17.6% in the second year and 9.2% in the third.

Second, I analyze if and how the effect of the soot tax differs by household and owner characteristics. We already know that, upon introduction, most of the affected diesel vehicles were owned by lower-income households. The hazard model now allows me to examine how responses to the soot tax vary across affected households. These heterogeneous responses are central to assessing the tax's distributional effects, including the extent of its regressivity. I investigate such treatment heterogeneity by interacting the treatment indicator with dummy variables for marital status, gender, homeownership, above-median household disposable income, above-median owner age, above-median population, above-median population density, and with dummies for the number of vehicles in the household.

The estimates for these interaction terms are in column (3) of Table 6. Several coefficients are statistically significant. Households with above-median income respond more strongly than their below-median counterparts. The effect of the soot tax on a treated vehicle's hazard is larger by 7.8 percentage points if the household has above-median instead of below-median disposable income. This alludes to the argument that higher income eases financial constraints and thereby allows households to bring forward vehicle replacement decisions more flexibly.

Likewise, the effect of the soot tax on a treated vehicle's hazard is larger by 9.1 percentage points in municipalities with above-median population density compared to those with below-median population density. Since driving a diesel is less economical in more densely populated areas, the soot tax more strongly incentivizes earlier vehicle retirement in such areas. Homeownership substantially reduces the effect of the soot tax by 14.3 percentage points, potentially because homeownership proxies wealth, and wealthy people may be more insensitive to vehicle operating costs.

The results provide further evidence that the effect of the soot tax is smaller among households that own additional cars besides the affected diesel vehicle. The interaction term for households with three cars is significant at the five percent level, and the coefficient for those with two cars is marginally significant, with a p-value just above 0.1. One possible explanation is that multi-vehicle households consider their cars' total registration fee liability rather than individual, car-specific liability when making replacement decisions. When only one out of several cars becomes subject to the soot tax, this potentially dampens these households' responsiveness.

These results have direct implications for the distributional effects of the soot tax. When first introduced, the burden of the soot tax falls predominantly on lower-income households because they own most of the affected vehicles. In addition to that, as column (3) of Table 6 shows, lower-income households are significantly less responsive to the soot tax than their higher-income counterparts.[19] This means that the former tend to be exposed to the soot tax for a longer time on average, which in turn tends to aggravate the policy's regressivity. In other words, the soot tax appears regressive in two ways. It predominantly affects lower-income households upon introduction, and these households are more likely to continue paying the soot tax in the following years given that they are less likely to retire the affected polluting vehicle.

6 Reductions in PM emissions and vehicle lifetime

To gauge the broader implications of the results from Section 5, I conduct two back-of-the-envelope calculations linking the previous estimates to changes in total vehicle holdings, aggregate PM emissions, and vehicle lifetimes. To quantify the

[19] Strictly speaking, the hazard model results show that a vehicle's hazard is lower when its household has below-median disposable income.

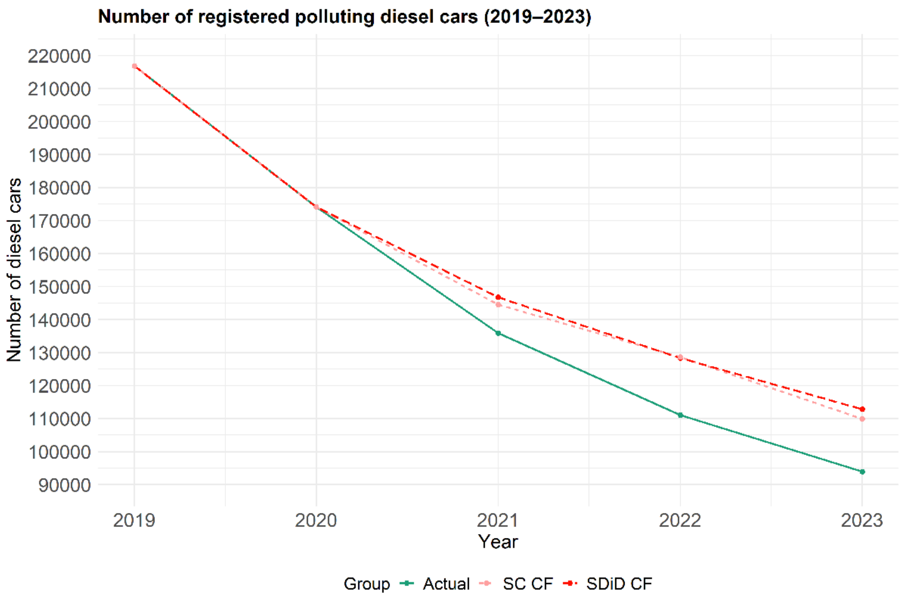
reduction in fleet-wide PM emissions resulting from the accelerated scrappage of polluting diesel vehicles, I assume that the ratio of polluting to non-polluting diesel vehicles in my identified 2005-2008 diesel sample is representative for the entire fleet of 2005-2008 diesel vehicles. I utilize the per-period estimates from columns (2) and (4) in Table 4 for the back-of-the-envelope calculation. For simplification, I assume that diesel cars built before 1990 are universally exempt from registration fees due to classical car status. Vintage-specific PM emissions of diesel cars are calculated using the public vehicle registration data from RDW. Data on annual vehicle miles traveled by vehicle age is taken from Centraal Bureau voor de Statistiek (2023b).

The green line in Figure 9a shows how the number of polluting diesel vehicles has actually developed between 2019 and 2023, whereas the dashed red lines show how it would have developed if the soot tax had never been implemented, according to the per-period SDiD (dark red) and SCM (light red) estimates from Table 4. "Polluting" is defined according to the soot tax eligibility threshold of tailpipe PM emissions exceeding the Euro 5 limit of 5 mg/km.

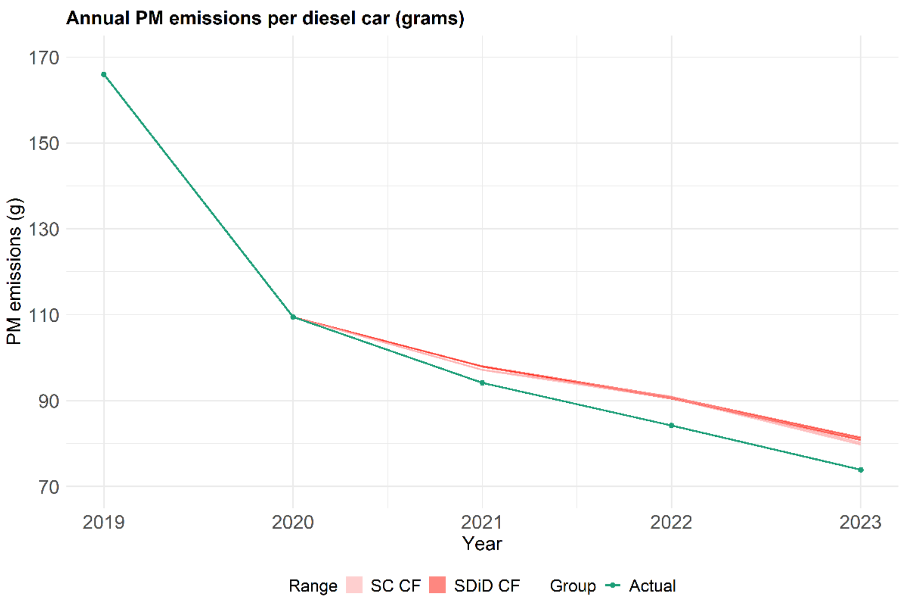
Figure 9a shows that there were about 217,000 such polluting diesel passenger cars registered in the Netherlands on January 1, 2019. This number has dropped by almost 57% within four years to about 94,000 vehicles on January 1, 2023. The results from Table 4 imply that this outflow of polluting vehicles would have been significantly slower without the soot tax. If the policy had never been introduced, there would have been about 110,000 to 113,000 polluting diesel cars in the registered vehicle fleet instead of 94,000 by January 1, 2023. This amounts to 17–20% more polluting vehicles remaining in the fleet than are actually observed. In other words, the soot tax has lead to the early retirement of 16,000 to 19,000 polluting diesel cars three years into the policy.

In the counterfactual without the soot tax, this higher number of polluting diesel cars would then substantially increase overall PM emissions. Figure 9b shows the development of annual tailpipe PM emissions of the average Dutch diesel car. Considering the entire set of diesel passenger cars in 2019, the average car emitted a total of 165.98 grams of PM over the course of that year. Thanks to the continuing retirement of old and polluting diesel vehicles (and a reduction in vehicle miles travelled since 2020 due to Covid), annual PM emissions per diesel car have gone down by more than 55% to an average of 73.89 grams until 2023.

Figure 9: Back-of-the-envelope calculation of the effect of the soot tax on the number of polluting diesel cars and PM emissions



(a) Number of polluting (> 5 mg PM/km) diesel vehicles



(b) Annual PM emissions in g/km of the average Dutch diesel car

Notes: Graph (a) shows the number of polluting diesel vehicles in the Netherlands every January 1 from 2019 to 2023. "Polluting" is defined as being eligible for the soot tax = PM emissions higher than 5 mg/km. Graph (b) shows the amount of PM in g/km that the average Dutch diesel car emits within one year, based on type-approval PM emissions and annual mileage. The green line corresponds to the actual data in which the soot tax came into effect on January 1, 2020. The dark (light) red line/area corresponds to how both outcomes would have evolved in absence of the soot tax according to the SDiD (SC) estimates from Table 4. Without the soot tax, there would have been more polluting diesel cars and higher annual PM emissions past 2020. Graph (b) shows a range of counterfactual PM emissions. In graph (b), I first calculate every diesel car's total annual PM emissions as the product of its registered PM emissions per kilometer times vehicle kilometers travelled according to CBS ((Centraal Bureau voor de Statistiek, 2023a)), and then divide this by the total number of registered diesel cars to obtain "Annual PM emission per diesel car (grams)", what I display in the graph. Table A13 contains implied per-year numbers.

Some of this reduction only occurred because of the soot tax. I calculate an upper and lower bound for counterfactual PM emissions by considering two extreme scenarios of vehicle replacement. The upper bound corresponds to a maximum abatement scenario, assuming that when vehicle owners respond, they replace their polluting diesel vehicle with a zero PM emission vehicle. Therefore, the upper-bound estimate is calculated by simply inflating the number of polluting diesel vehicles in post-treatment years.

The lower-bound estimate corresponds to a minimum abatement scenario, assuming that when vehicle owners respond, they replace their polluting diesel vehicle with one that has PM emissions of exactly 5 mg/km, the highest permitted emission level without being eligible for the soot tax. I subsequently subtract the emissions of a hypothetical quantity of such vehicles from the vehicle fleet on a one-to-one-basis when adding back the polluting vehicles that would have never left the fleet in absence of the soot tax. This provides me with a lower-bound estimate.

Figure 9b shows two trajectories for the implied emissions according to the SDiD and SCM estimates, spanning the range from the upper- to lower-bound estimates. Because vehicles that are not liable for the soot tax are so much cleaner in terms of tailpipe PM emissions than the old diesel cars that they replace, the resulting trajectories turn out to be very narrow. Without the soot tax, the average diesel car in 2023 would have emitted between 79.7 and 81.4 grams of particulate matter annually, compared to 73.9 grams under the policy, because more polluting vehicles would have remained on the road. These numbers translate into a 7.9 to 10.2% increase in overall tailpipe PM emissions from diesel passenger cars by 2023 if the soot tax had never been introduced.

For my second back-of-the-envelope calculation, I utilize the hazard model estimates from Section 5.3 to calculate the implied reduction in vehicle lifetime, similar to Alberini et al. (2018). I use the cloglog model to predict the log cumulative hazard η_k for year k to calculate the hazard of retirement during year k as $h_k = 1 - \exp(-\exp(\eta_k))$. The discrete-time survivor function $S_k = \prod_{j=1}^{k} (1 - h_j)$ then provides the probability of vehicle survival up to the start of year $k+1$ (because the observations are always from January 1). The average estimated vehicle lifetime in the three-year window from January 2020 to January 2023 then equals $S_{2020} + S_{2021} + S_{2022}$, the sum of the survival probabilities over the treatment years.

Table 7: Implied reduction in lifetime

	Scenario	Avg estimated lifetime in three-year window, in years		
		Without soot tax	With soot tax	% change due to soot tax
(1)	Avg. car, VW, 1 car	1.741	1.565	-10.11%
(2)	Avg. car, VW, 2 cars	1.603	1.471	-8.24%
(3)	Avg. car, VW, 1 car, p20 income	1.690	1.549	-8.34%
(4)	Avg. car, VW, 1 car, p80 income	1.762	1.544	-12.37%
(5)	Avg. car, VW, 1 car, rural	1.824	1.696	-7.02%
(6)	Avg. car, VW, 1 car, urban	1.733	1.513	-12.69%

Notes: The table shows the average lifetime that are implied by estimates from the hazard model with interaction terms (column (3) in Table 6). "Avg. car" refers to average values of numeric covariates for treatment group vehicles as on Jan 1, 2019. The table considers only treatment group diesel vehicles to show how the soot tax affects their lifetime by comparing average estimated lifetime without and with the soot tax. In rows (1) and (2), all numeric covariates are at their average values, whereas I manually set the factor variables "brand" and "number of cars in the household". In rows (3) and (4), I further set household disposable income to its 20th and 80th percentile by assuming a lognormal distribution and calculating percentiles based on the known mean, median, and standard deviation. In rows (5) and (6), I set population and population density to 15,000 and 200 persons/km^2 in row (5) and to 150,000 and 2,000 persons/km^2 in row (6).

Calculated average vehicle lifetimes are shown in Table 7. The underlying specification is the cloglog model with covariate interaction terms as in column (3) of Table 6. I use mean covariate values of treatment group vehicles as a benchmark. In addition, I use "Volkswagen" for the factor variable "Vehicle brand" for two reasons. First, it is the most prevalent brand among diesel car models from the mid-2000s, and second, the estimated coefficient on the factor "Volkswagen" is statistically zero, highlighting that this is an "average" brand in terms of remaining vehicle lifetime.

I look into different scenarios to uncover heterogeneous effects of the soot tax on vehicle lifetime in Table 7. In row (1), I consider a Volkswagen car with average treatment group covariate values in a single-car household. Absent the soot tax, the hazard model results imply an average remaining lifetime of about 1.74 years within the three-year window from January 2020 to 2023 for this car. Its average remaining lifetime drops by about 10% with the soot tax in place. If the household owns one other car, as in row (2), two things change. First, the vehicle's average remaining lifetime is generally lower because multi-vehicle households replace their vehicles more frequently. Second, the effect of the soot tax on vehicle lifetime is smaller. The soot tax reduces the polluting vehicle's lifetime by only 8.24%.

In rows (3) and (4), I further differentiate by household disposable income, comparing the 20$^{\text{th}}$ and 80$^{\text{th}}$ percentile of the disposable income distribution within the treatment group.[20] Remaining lifetime of the polluting diesel car is higher at a high-income than a low-income household absent the soot tax, implying that

[20] I consider household disposable income in the set of eligible vehicles, assume a lognormal distribution of incomes, use the mean, median, and standard deviation of observed incomes in this set, and then calculate the 20th and 80th percentile based on that.

low-income households replace their old vehicles more quickly. However, high-income households are more responsive to the soot tax so that vehicle lifetimes virtually equalize with the soot tax in place. The reduction in remaining vehicle lifetime is larger for the high-income household at 12.37%, compared with 8.34% for the low-income household. Presumably, high-income households have greater flexibility to adjust their replacement behavior in response to fiscal incentives.

In rows (5) and (6), I differentiate by place of residence. I compare a vehicle that is registered in a rural area to one in an urban area.[21] The baseline remaining lifetime for a polluting diesel vehicle in a rural area is already higher at about 1.82 years, almost 0.09 years more than for the otherwise identical urban vehicle. With the soot tax in place, this gap widens and doubles in size to 0.18 years. Remaining vehicle lifetime of the urban vehicle drops by 12.69%, compared to 7.02% for the same vehicle in a rural area. Because lower-income households also tend to live in more rural regions, these findings imply that the policy's regressivity tends to aggravate over time. A lower-income household in a rural area is not only the most likely to be affected by the soot tax at the moment of its introduction, but also the least likely to bring forward vehicle retirement in the following years.

Lastly, the different effect sizes for rural and urban households also matter with regard to the policy's potential health benefits. Harmful PM concentrations are usually an urban phenomenon. The effect of the soot tax is particularly pronounced in urban areas. This pattern indicates that the policy incentivizes the retirement of polluting vehicles precisely where PM concentrations tend to be high and the health benefits from reducing PM pollution are greatest.

7 Conclusion

As our understanding of the harmfulness of local air pollutants such as fine particulate matter has vastly improved in recent years, policymakers in many countries have adopted strict standards for pollutant emissions from new vehicles. However, the existent fleet of old diesel vehicles without modern emissions control technology is untouched by such regulation, and continues to contribute substantially to air pollution. To incentivize the accelerated removal of these vehicles from the roads,

[21] I define "rural" as a municipality of 15,000 inhabitants & 200 persons/km^2, and "urban" as a municipality of 150,000 inhabitants & 2,000 persons/km^2.

the Netherlands introduced an explicit surcharge in its registration fee system that applies exclusively to highly polluting diesel vehicles without particulate filters.

Using a decade of microdata from Dutch vehicle registrations, I quantify the impact of this surcharge on the size of the polluting vehicle fleet and examine how vehicle retirement responses vary across households. Eligibility for the surcharge is determined by whether a diesel passenger car has a factory-installed particulate filter in place. I leverage the staggered rollout of factory-installed filters due to upstream supply constraints as quasi-random treatment variation to identify the causal effect of the registration fee surcharge on vehicle holdings. In doing so, this paper is the first to empirically examine how tying registration fees to pollutant emissions affects the presence of older, highly polluting vehicles in the fleet.

I apply state-of-the-art methods from the DiD literature to find that the surcharge has reduced the number of polluting vehicles by about 16 percent within the first three years, with the effect accumulating over time. I further set up a simple model of vehicle replacement to rationalize the finding that the effect is larger for older and lower-quality vehicles. A back-of-the-envelope calculation shows that, within the first three years of the policy, the soot tax caused between 16,000 to 19,000 premature deregistrations of polluting diesel vehicles. This in turn reduced aggregate tailpipe PM emissions from the stock of diesel cars by about 8 to 10%.

In addition to that, I estimate a hazard model to unveil the importance of household characteristics in vehicle retirement decisions. Such heterogeneity has important implications for the policy's distributional effects. I find that the policy is highly regressive. Firstly, the initial tax burden upon introduction predominantly falls on lower-income households, and secondly, the burden tends to grow over time for these households because these tend to be less responsive to the policy. In general, the policy causes the largest reductions in remaining vehicle lifetime among higher-income and among single-vehicle households; those that would have held onto their polluting vehicles the longest absent the policy. Moreover, the policy generates larger reductions in remaining vehicle lifetime for urban households than for rural households, despite urban households already replacing their polluting vehicles sooner. Nonetheless, the large effect of the soot tax on vehicle retirement in urban areas presumably entails relatively large health benefits because urban areas tend to be those that are most affected by elevated levels of PM pollution.

On a higher level, these findings demonstrate the potential of targeted ownership taxes to accelerate the transition toward a cleaner stock of durable goods, while also highlighting the importance of accounting for household heterogeneity in policy design. Since ownership taxes on old and inefficient durable goods are presumably regressive in many contexts, policymakers should carefully consider complementary measures that mitigate distributional impacts while preserving environmental effectiveness. Future work could assess such complementary policies or explore longer-term consequences, for instance for technology adoption and market structure. Another path for future research with regard to vehicle registration fees concerns alternative fee schedules that better account for the substantial heterogeneity in PM emissions among older diesel vehicles. The present policy applies a flat surcharge regardless of whether a car is marginally or massively more polluting. This suggests meaningful scope for more differentiated taxation and associated welfare gains.

References

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Adamou, A., Clerides, S., & Zachariadis, T. (2014). Welfare implications of car feebates: A simulation analysis. *The Economic Journal*, *124*(578), F420–F443. https://doi.org/10.1111/ecoj.12094

Alberini, A., & Bareit, M. (2019). The effect of registration taxes on new car sales and emissions: Evidence from switzerland. *Resource and Energy Economics*, *56*, 96–112. https://doi.org/10.1016/j.reseneeco.2017.03.005

Alberini, A., Bareit, M., Filippini, M., & Martinez-Cruz, A. L. (2018). The impact of emissions-based taxes on the retirement of used and inefficient vehicles: The case of switzerland. *Journal of Environmental Economics and Management*, *88*, 234–258. https://doi.org/10.1016/j.jeem.2017.12.004

Alexander, D., & Schwandt, H. (2022). The impact of car pollution on infant and child health: Evidence from emissions cheating. *The Review of Economic Studies*, *89*(6), 2872–2910. https://doi.org/10.1093/restud/rdac007

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*, *91*(434), 444. https://doi.org/10.2307/2291629

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton Univ. Press.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Balaguer, J., Pernías, J. C., & Ripollés, J. (2023). Is vehicle scrapping affected by low-emission zones? the case of madrid. *Transportation Research Part A: Policy and Practice*, *172*, 103668. https://doi.org/10.1016/j.tra.2023.103668

Barahona, N., Gallego, F. A., & Montero, J.-P. (2020). Vintage-specific driving restrictions. *Review of Economic Studies*, *87*, 1646–1682. https://doi.org/10.1093/restud/rdz031

Belastingdienst. (2025). Bpm tariff passenger car. https://www.belastingdienst.nl/wps/wcm/connect/bldcontenten/belastingdienst/individuals/cars/bpm/calculate_and_pay_bpm/bpm_tariff/bpm-tariff-passenger-car

Bento, A., Roth, K., & Zuo, Y. (2018). Vehicle lifetime and scrappage behavior: Trends in the u.s. used car market. *The Energy Journal*, *39*(1), 159–184. https://doi.org/10.5547/01956574.39.1.aben

Bergantino, A. S., Intini, M., & Percoco, M. (2021). New car taxation and its unintended environmental consequences. *Transportation Research Part A: Policy and Practice*, *148*, 36–48. https://doi.org/10.1016/j.tra.2021.03.026

Bull, N. (2024). Agility in car design: How shorter product lifecycles are disrupting the automotive industry. https://escalent.co/blog/agility-in-car-design-how-shorter-product-lifecycles-are-disrupting-the-automotive-industry/

Busse, M. R., Knittel, C. R., Silva-Risso, J., & Zettelmeyer, F. (2012). Did “cash for clunkers” deliver? the consumer effects of the car allowance rebate system. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=d26c5b51a61b0c8fb9d63ad0111547c53e4141cc

Callaway, B., & Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, *225*, 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Carpenter, C. S., McClellan, C. B., & Rees, D. I. (2017). Economic conditions, illicit drug use, and substance use disorders in the united states. *Journal of health economics*, *52*, 63–73. https://doi.org/10.1016/j.jhealeco.2016.12.009

Centraal Bureau voor de Statistiek. (2023a). Documentatie motorvoertuigenpark met kenmerken van voertuigen op kentekenniveau (rdwnpacttab). https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/microdatabestanden/rdwnpacttab-motorvoertuigenpark

Centraal Bureau voor de Statistiek. (2023b). Verkeersprestaties personenauto’s, brandstof uitgebreid, leeftijd. https://www.cbs.nl/nl-nl/cijfers/detail/85405NED

Centraal Bureau voor de Statistiek. (2025). Welfare of households; key figures. https://www.cbs.nl/en-gb/figures/detail/83739eng

Ciccone, A. (2018). Environmental effects of a vehicle tax reform: Empirical evidence from norway. *Transport Policy*, *69*, 141–157. https://doi.org/10.1016/j.tranpol.2018.05.002

Compendium voor de Leefomgeving. (2023). Emissies naar lucht door verkeer en vervoer, 2021. https://www.clo.nl/indicatoren/nl012937-emissies-naar-lucht-door-verkeer-en-vervoer-2021

Cyrys, J., Peters, A., Soentgen, J., & Wichmann, H. E. (2014). Low emission zones reduce pm10 mass concentrations and diesel soot in german cities. *Journal of the Air & Waste Management Association (1995)*, *64*(4), 481–487. https://doi.org/10.1080/10962247.2013.868380

Degraeuwe, B., Thunis, P., Clappier, A., Weiss, M., Lefebvre, W., Janssen, S., & Vranckx, S. (2017). Impact of passenger car nox emissions on urban no2 pollution – scenario analysis for 8 european cities. *Atmospheric Environment*, *171*, 330–337. https://doi.org/10.1016/j.atmosenv.2017.10.040

D'Haultfœuille, X., Givord, P., & Boutin, X. (2014). The environmental effect of green taxation: The case of the french bonus/malus. *The Economic Journal*, *124*(578), F444–F480. https://doi.org/10.1111/ecoj.12089

Eerste Kamer der Staten-Generaal. (2016). Stemming wet uitwerking autobrief ii. https://www.eerstekamer.nl/verslagdeel/20160705/wet_uitwerking_autobrief_ii

Egami, N., & Yamauchi, S. (2023). Using multiple pretreatment periods to improve difference-in-differences and staggered adoption designs. *Political Analysis*, *31*(2), 195–212. https://doi.org/10.1017/pan.2022.8

Eliasson, J., Pyddoke, R., & Swärdh, J.-E. (2018). Distributional effects of taxes on car fuel, use, ownership and purchases. *Economics of Transportation*, *15*, 1–15. https://doi.org/10.1016/j.ecotra.2018.03.001

Facure, M. (2025). Synthetic difference-in-differences [In *Python Causality Handbook*. Retrieved August 2025].

Fischer, P. H., Marra, M., Ameling, C. B., Hoek, G., Beelen, R., de Hoogh, K., Breugelmans, O., Kruize, H., Janssen, N. A. H., & Houthuijs, D. (2015). Air pollution and mortality in seven million adults: The dutch environmental longitudinal study (duels). *Environmental health perspectives*, *123*(7), 697–704. https://doi.org/10.1289/ehp.1408254

Gallagher, K. S., & Muehlegger, E. (2011). Giving green to get green? incentives and consumer adoption of hybrid vehicle technology. *Journal of Environmental Economics and Management*, *61*(1), 1–15. https://doi.org/10.1016/j.jeem.2010.05.004

Gehrsitz, M. (2017). The effect of low emission zones on air pollution and infant health. *Journal of Environmental Economics and Management*, *83*, 121–144. https://doi.org/10.1016/j.jeem.2017.02.003

Greene, D. L., & Leard, B. (2024). Trends in scrappage and survival of u.s. light-duty vehicles. *Transportation Research Part A*, *180*, 103982. https://doi.org/10.1016/j.tra.2024.103982

Grigolon, L., Leheyda, N., & Verboven, F. (2016). Scrapping subsidies during the financial crisis — evidence from europe. *International Journal of Industrial Organization*, *44*, 41–59. https://doi.org/10.1016/j.ijindorg.2015.10.004

Hansen, B., Sabia, J. J., & Rees, D. I. (2017). Have cigarette taxes lost their bite? new estimates of the relationship between cigarette taxes and youth smoking. *American Journal of Health Economics*, *3*(1), 60–75. https://doi.org/10.1162/AJHE_a_00067

Hassell, H. J. G., Holbein, J. B., & Baldwin, M. (2020). Mobilize for our lives? school shootings and democratic accountability in u.s. elections. *American Political Science Review*, *114*(4), 1375–1385. https://doi.org/10.1017/S0003055420000635

Held, M., Rosat, N., Georges, G., Pengg, H., & Boulouchos, K. (2021). Lifespans of passenger cars in europe: Empirical modelling of fleet turnover dynamics. *European Transport Research Review*, *13*(1). https://doi.org/10.1186/s12544-020-00464-0

Jacobsen, M. R., Sallee, J. M., Shapiro, J. S., & van Benthem, A. A. (2023). Regulating untaxable externalities: Are vehicle air pollution standards effective and efficient? *The Quarterly Journal of Economics*, *138*(3), 1907–1976. https://doi.org/10.1093/qje/qjad016

Jacobsen, M. R., & van Benthem, A. A. (2015). Vehicle scrappage and gasoline policy. *American Economic Review*, *105*(3), 1312–1338. https://doi.org/10.1257/aer.20130935

Jiang, W., Boltze, M., Groer, S., & Scheuvens, D. (2017). Impacts of low emission zones in germany on air pollution levels. *Transportation Research Procedia*, *25*, 3370–3382. https://doi.org/10.1016/j.trpro.2017.05.217

Jonson, J. E., Borken-Kleefeld, J., Simpson, D., Nyíri, A., Posch, M., & Heyes, C. (2017). Impact of excess no x emissions from diesel cars on air quality, public health and eutrophication in europe. *Environmental Research Letters*, *12*(9), 094017. https://doi.org/10.1088/1748-9326/aa8850

Klier, T., & Linn, J. (2015). Using taxes to reduce carbon dioxide emissions rates of new passenger vehicles: Evidence from france, germany, and sweden. *American Economic Journal: Economic Policy*, *7*(1), 212–242. https://doi.org/10.1257/pol.20120256

Kok, R. (2015). Six years of co2-based tax incentives for new passenger cars in the netherlands: Impacts on purchasing behavior trends and co2 effectiveness. *Transportation Research Part A: Policy and Practice*, *77*, 137–153. https://doi.org/10.1016/j.tra.2015.04.009

Kraftfahrt-Bundesamt (KBA). (2023). Vehicle registration statistics (zfzrbs), versions 2015.1–2023.1 [Data releases used: 10.25525/kba-fdzoff.zfzrbs.2015.1; 10.25525/kba-fdzoff.zfzrbs.2016.1; 10.25525/kba-fdzoff.zfzrbs.2017.1; 10.25525/kba-fdzoff.zfzrbs.2018.1; 10.25525/kba-fdzoff.zfzrbs.2019.1; 10.25525/kba-fdzoff.zfzrbs.2020.1; 10.25525/kba-fdzoff.zfzrbs.2020.2; 10.25525/kba-fdzoff.zfzrbs.2021.1; 10.25525/kba-fdzoff.zfzrbs.2022.1; 10.25525/kba-fdzoff.zfzrbs.2023.1]. https://www.kba.de/DE/Statistik/Fahrzeuge/fahrzeuge_node.html

Laborda, J., & Moral, M. J. (2019). Scrappage by age: Cash for clunkers matters! *Transportation Research Part A: Policy and Practice*, *124*, 488–504. https://doi.org/10.1016/j.tra.2019.04.014

Lenski, S. M., Keoleian, G. A., & Bolon, K. M. (2010). The impact of 'cash for clunkers' on greenhouse gas emissions: A life cycle perspective. *Environmental Research Letters*, *5*(4), 044003. https://doi.org/10.1088/1748-9326/5/4/044003

Li, S., Linn, J., & Spiller, E. (2013). Evaluating "cash-for-clunkers": Program effects on auto sales and the environment. *Journal of Environmental Economics and Management*, *65*(2), 175–193. https://doi.org/10.1016/j.jeem.2012.07.004

Li, S., Timmins, C., & von Haefen, R. H. (2009). How do gasoline prices affect fleet fuel economy? *American Economic Journal: Economic Policy*, *1*(2), 113–137. https://doi.org/10.1257/pol.1.2.113

Malina, C., & Scheffler, F. (2015). The impact of low emission zones on particulate matter concentration and public health. *Transportation Research Part A: Policy and Practice*, *77*, 372–385. https://doi.org/10.1016/j.tra.2015.04.029

Margaryan, S. (2021). Low emission zones and population health. *Journal of Public Economics*, *76*, 102402. https://doi.org/10.1016/j.jhealeco.2020.102402

Marin, G., & Zoboli, R. (2020). Effectiveness of car scrappage schemes: Counterfactual-based evidence on the italian experience. *Economics of Transportation*, *21*, 100150. https://doi.org/10.1016/j.ecotra.2019.100150

Nikodinoska, D., & Schröder, C. (2016). On the emissions–inequality and emissions–welfare trade-offs in energy taxation: Evidence on the german car fuels tax. *Resource and Energy Economics*, *44*, 206–233. https://doi.org/10.1016/j.reseneeco.2016.03.001

Østli, V., Fridstrøm, L., Kristensen, N. B., & Lindberg, G. (2022). Comparing the scandinavian automobile taxation systems and their co 2 mitigation effects. *International Journal of Sustainable Transportation*, *16*(10), 910–927. https://doi.org/10.1080/15568318.2021.1949763

Pestel, N., & Wozny, F. (2021). Health effects of low emission zones: Evidence from german hospitals. *Journal of Environmental Economics and Management*, *109*, 102512. https://doi.org/10.1016/j.jeem.2021.102512

Ptak, M., Neneman, J., & Roszkowska, S. (2024). The impact of petrol and diesel oil taxes in eu member states on co2 emissions from passenger cars. *Scientific Reports*, 14–52. https://doi.org/10.1038/s41598-023-50456-y

Regulations, U. A. (2025). Schemes per country - netherlands. https://urbanaccessregulations.eu/countries-mainmenu-147/netherlands-mainmenu-88

Reif, K. (2014). *Dieselmotor-management im überblick*. Springer. https://doi.org/10.1007/978-3-658-06555-3

Rijksdienst voor Ondernemend Nederland. (2025). Verlaging brandstofaccijns verlengd tot eind 2025. https://ondernemersplein.overheid.nl/verlaging-brandstofaccijns-verlengd-tot-eind-2025/

Rogan, F., Dennehy, E., Daly, H., Howley, M., & Ó Gallachóir, B. P. (2011). Impacts of an emission based private car taxation policy – first year ex-post analysis. *Transportation Research Part A: Policy and Practice*, *45*(7), 583–597. https://doi.org/10.1016/j.tra.2011.03.007

Sandler, R. (2012). Clunkers or junkers? adverse selection in a vehicle retirement program. *American Economic Journal: Economic Policy*, *4*(4), 253–281. https://doi.org/10.1257/pol.4.4.253

Scoltock, J. (2014). Diesel particulate filter: Psa peugeot citroën was the first to bring in particulate filters to help make diesels cleaner. *Automotive Engineer*, *1*(4).

Soares, J., Plass, D., Kienzler, S., González Ortiz, A., Gsella, A., & Horálek, J. (2023). Health risk assessment of air pollution: Assessing the environmental burden of disease in europe in 2021 [Eionet Report – ETC HE 2023/7, revised version, June 2024]. https://www.eionet.europa.eu/etcs/all-etc-reports

Steffens, C., & Stuhler, J. (2025). How to control for (pre-)trends? https://github.com/csteffenseco/Website-Files/blob/main/Steffens_Stuhler_Pretrends.pdf

Sterner, T. (2007). Fuel taxes: An important instrument for climate policy. *Energy Policy*, (35), 3194–3202. https://doi.org/10.1016/j.enpol.2006.10.025

Strezhnev, A. (2024). *Group-specific linear trends and the triple-differences in time design*. https://doi.org/10.31235/osf.io/dg5ps

Sun, F., Yang, R., & Yuan, D. (2018). Green stimulus tax incentives in china's automobile market. https://acsweb.ucsd.edu/~f4sun/pdfs/Car_Sun.pdf

Tovar Reaños, M. A., & Sommerfeld, K. (2018). Fuel for inequality: Distributional effects of environmental reforms on private transport. *Resource and Energy Economics*, *51*, 28–43. https://doi.org/10.1016/j.reseneeco.2017.10.007

Tweede Kamer der Staten-Generaal. (2015). Maatregelen op het gebied van autobelastingen. https://zoek.officielebekendmakingen.nl/kst-32800-27.html

van Wingerden, J. (2019). Per 2020 hogere wegenbelasting voor oude diesels. *Autoweek*. https://www.autoweek.nl/autonieuws/artikel/per-2020-hogere-wegenbelasting-voor-oude-diesels/

Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? a reconciliation and new results. *American Economic Review*, *96*(5), 1802–1820. https://doi.org/10.1257/aer.96.5.1802

Appendix

A Tables and Figures

Table A1: Average type-approval PM emissions of registered diesel passenger cars in the Netherlands as of May 2024.

Year of admission	Avg. mg PM / km	Year of admission	Avg. mg PM / km
1997	75.17	2005	24.59
1998	66.76	2006	13.85
1999	58.45	2007	9.24
2000	52.12	2008	4.94
2001	43.86	2009	2.67
2002	37.85	2010	1.45
2003	36.78	2011	0.66
2004	31.88	2012 & later	< 0.5

Notes: The numbers in the table correspond to the average registered PM emissions of all diesel passenger cars with non-missing values and with the same year of admission. Based on own calculations using data from opendata.rdw.nl on May 4, 2024.

Table A2: BPM for gasoline cars

CO_2 min	CO_2 max	Formula	BPM at CO_2 min	BPM at CO_2 max
0	80	$2 \cdot CO_2 + 440$	N.A.	600
80	104	$76 \cdot (CO_2 - 80) + 600$	600	2,424
104	145	$167 \cdot (CO_2 - 104) + 2,424$	2,424	9,271
145	161	$274 \cdot (CO_2 - 145) + 9,271$	9,271	13,655
161	∞	$549 \cdot (CO_2 - 161) + 13,655$	13,655	

Table A3: BPM for diesel cars

CO_2 min	CO_2 max	Formula	BPM at CO_2 min	BPM at CO_2 max
0	71	$2 \cdot CO_2 + 440$	N.A.	582
71	80	$(2 \cdot CO_2 + 440) + ((CO_2 - 71) \cdot 106.07)$	582	1,554.63
80	104	$(76 \cdot (CO_2 - 80) + 600) + ((CO_2 - 71) \cdot 106.07)$	1,554.63	5,924.31
104	145	$(167 \cdot (CO_2 - 104) + 2,424) + ((CO_2 - 71) \cdot 106.07)$	5,764.31	17,120.18
145	161	$(274 \cdot (CO_2 - 145) + 9,271) + ((CO_2 - 71) \cdot 106.07)$	17,120.18	23,201.30
161	∞	$(549 \cdot (CO_2 - 161) + 13,655) + ((CO_2 - 71) \cdot 106.07)$	23,201.30	

Table A4: BPM for PHEV cars

CO_2 min	CO_2 max	Formula	BPM at CO_2 min	BPM at CO_2 max
0	34	$28 \cdot CO_2$	N.A.	952
34	60	$100 \cdot (CO_2 - 34) + 952$	952	3,552
60	∞	$239 \cdot (CO_2 - 60) + 3,552$	3,552	

Source: Belastingdienst (2025), own depiction and calculation (Tables 2, 3, 4).
Notes: The tables show the BPM (vehicle sales tax) amounts based on the CO_2 emissions of the new passenger car.

Table A5: EU Emission Standards 1-5 for diesel passenger cars, mg/km

Tier	Date (Type approval)	Date (1^{st} registration)	CO	NO_x	$HC+NO_x$	PM
Euro 1	Jul 1992	Jan 1993	2,720	-	970	140
Euro 2	Jan 1996	Jan 1997	1,000	-	700	80
Euro 3	Jan 2000	Jan 2001	660	500	560	50
Euro 4	Jan 2005	Jan 2006	500	250	300	25
Euro 5	Sep 2009	Jan 2011	500	180	230	5
Euro 6	Sep 2014	Sep 2015	500	80	170	4.5

Notes: The table shows the pipetail emission limits of air pollutants for diesel passenger cars in the European Union, known as "Emission Standards". The standards set limits for vehicle pollutant emissions. The regulated pollutants for diesel cars are carbon monoxide (CO), nitrogen oxides (NO_x), hydrocarbons (HC), and particulates with a diameter of no more than 10 micrometers (PM). The type approval date refers to the date from which onwards new car models only receive type approval if they abide by the respective standard. The date of first registration refers to the deadline to apply for registration of a new passenger car at the local department of motor vehicles.

Table A6: Regression results for 2005-2008 diesel sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SDiD		SCM		FE regression			
Soot tax	-0.165***	-0.164***	-0.149***	-0.157***	-0.145***	-0.129***	-0.160***	-0.154***
	(0.027)	(0.027)	(0.045)	(0.037)	(0.022)	(0.025)	(0.020)	(0.023)
TG × Year					-0.021**	0.019*		
					(0.009)	(0.010)		
ln(inc)					-0.125**	-0.125**	-0.025	-0.031*
					(0.058)	(0.055)	(0.017)	(0.017)
ln(owner age)					-1.631***	-1.572***	-0.249**	-0.268***
					(0.258)	(0.226)	(0.099)	(0.100)
ln(pop.dens.)					-0.124**	-0.160***	-0.019	-0.022
					(0.054)	(0.048)	(0.022)	(0.023)
Male					0.438**	0.354*	0.008	0.014
					(0.219)	(0.204)	(0.087)	(0.087)
College					0.203*	0.188**	-0.042	-0.037
					(0.106)	(0.088)	(0.031)	(0.031)
Married					0.118	-0.057	-0.007	-0.014
					(0.124)	(0.114)	(0.045)	(0.045)
Home					0.474***	0.532***	0.087	0.100*
					(0.180)	(0.168)	(0.059)	(0.060)
Unit × Year							✓	✓
Car type FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓		✓	
Vintage-year FE						✓		✓
SDID Covariates		✓		✓				
No of obs.	3,420	3,420	3,420	3,420	3,420	3,420	3,420	3,420
Adj. R^2					0.941	0.949	0.994	0.994
Within R^2					0.173	0.108	0.904	0.882

Notes: SDiD = Synthetic Difference-in-Differences; SCM = Synthetic Control Method; FE = Fixed Effects. The table contains point estimates for the effect of the soot tax on the log number of registered vehicles of treated car models in the sample of 2005-2008 diesel car models. Covariates are a car model's average log household disposable income ("ln(inc)"), log owner age ("ln(owner age)"), log population density of owner's home municipality ("ln(pop.dens.)"), and dummies for being male, college-educated, married, and owning a home. Columns (1) and (2) are SDiD estimates (Equation 4), without and with covariates, respectively. Columns (3) and (4) are SCM estimates (Equation 5), without and with covariates, respectively. Columns (5) to (8) are fixed effect regression estimates (Equation 6). The specifications differ in whether they employ group- or unit-specific linear time trends, and whether they employ year fixed effects or vintage×year fixed effects. Standard errors are in parenthesis. SDiD standard errors are jackknife errors, SCM standard errors are placebo errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table A7: Regression results for 2005-2008 diesel sample, per-period effects

	(1)	(2)	(3)	(4)	(5)	(6)
	SDiD		SCM		SDiD	
Soot tax	-0.165***		-0.149***		-0.168***	
	(0.027)		(0.045)		(0.030)	
TG × '14		0.016		0.000		0.015
TG × '15		0.016		0.000		0.014
TG × '16		0.017		0.000		0.014
TG × '17		0.024		0.000		0.019
TG × '18		0.027		0.000		0.018
TG × '19		0.015		0.000		0.000
TG × '20		0.000		0.000		-0.026*
						(0.014)
TG × '21		-0.088***		-0.070**		-0.128***
		(0.016)		(0.033)		(0.026)
TG × '22		-0.176***		-0.179***		-0.222***
		(0.029)		(0.048)		(0.039)
TG × '23		-0.233***		-0.197***		-0.284***
		(0.039)		(0.048)		(0.048)

Notes: SDiD = Synthetic Difference-in-Differences; SCM = Synthetic Control Method. The table contains per-period point estimates for the effect of the soot tax on the log number of registered vehicles of treated car models in the sample of 2005-2008 diesel car models. Specifications are as in Table 4. Standard errors are in parenthesis. SDiD standard errors are jackknife errors, SCM standard errors are placebo errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Estimates in columns (2) and (4) are relative to 2020, estimates in column (6) are relative to 2019. There is no established method to calculate standard errors for pre-treatment periods in SDiD and SCM.

Table A8: Summary statistics for the 1990-2004 car-model-level sample

Variable	Unit	No. of obs.	Mean	St.d.	Median
Treatment Group	Dummy	11,590	0.230	0.421	0
No. of vehicles	Count	11,590	263.2	561.3	85
Vintage	Year	11,590	1999.94	3.589	2001
Vehicle mass	kg	11,590	1,323	354.9	1,252
List price	€	1,980	47,956	30,818	38,634
Mileage	l/100 km	8,710	8.027	2.328	7.900
CO_2	g/km	9,190	197.0	53.81	191.6
Mean HH Disp. Inc.	€	11,590	54,000	18,416	50,501
Mean Owner Age	Years	11,590	48.61	5.828	48.54
Mean Pop. Density	Persons/km^2	11,590	1,400	282.0	1,405
Mean Male Owner	Dummy	11,590	0.729	0.148	0.768
Mean College Educ.	Dummy	11,590	0.361	0.169	0.328
Mean Married	Dummy	11,590	0.534	0.116	0.521
Mean Homeownership	Dummy	11,590	0.709	0.133	0.704

	Control group				Treatment group			
Variable	No. of obs.	Mean	St.d.	Median	No. of obs.	Mean	St.d.	Median
Treatment Group	8,920	0	0	0	2,670	1	0	1
No. of vehicles	8,920	301.4	626.9	90	2,670	135.6	183.0	72.5
Vintage	8,920	1999.56	3.639	2000	2,670	2001.19	3.101	2002
Vehicle mass	8,920	1,270	312.2	1,218	2,670	1,502	423.3	1,502
List price	1,350	46,067	32,294	34,940	630	52,002	26,969	49,490
Mileage	6,460	8.615	2.188	8.100	2,250	6.338	1.844	5.700
CO_2	6,810	205.8	52.50	194.0	2,380	172.1	49.52	162.0
Mean HH Disp. Inc.	8,920	53,582	16,097	50,027	2,670	53,398	24,580	51,996
Mean Owner Age	8,920	48.37	5.653	48.29	2,670	49.37	6.319	49.55
Mean Pop. Density	8,920	1,433	259.8	1,438	2,670	1,291	322.8	1,249
Mean Male Owner	8,920	0.714	0.157	0.755	2,670	0.780	0.096	0.798
Mean College Educ.	8,920	0.342	0.156	0.313	2,670	0.425	0.194	0.391
Mean Married	8,920	0.522	0.113	0.508	2,670	0.572	0.117	0.573
Mean Homeownership	8,920	0.703	0.130	0.696	2,670	0.731	0.138	0.733

Notes: Descriptive statistics on the final 1990-2004 car model sample in which I observe the number of registered vehicles of 1,159 car models across ten years (2014-2023, January 1). 267 are treated diesel car models, 892 are untreated gasoline car models. "Treatment Group" is one in all periods for diesel car models, and zero otherwise. "No. of vehicles" counts the number of registered vehicles of an (identified) car model. "Vintage" is year of first admission in the data. Mean variables are average values of household/owner covariates across all individual vehicles of a car type.

Table A9: Regression results for 1990–2004 sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SDiD		SCM		FE regression			
Soot tax	-0.196***	-0.196***	-0.192***	-0.238***	-0.193***	-0.176***	-0.185***	-0.164***
	(0.024)	(0.024)	(0.029)	(0.029)	(0.016)	(0.016)	(0.016)	(0.015)
TG × Year					-0.065***	-0.087***		
					(0.007)	(0.007)		
ln(inc)					-0.270***	-0.172***	-0.035**	-0.036**
					(0.061)	(0.054)	(0.015)	(0.015)
ln(owner age)					-2.693***	-2.309***	-0.600***	-0.438***
					(0.178)	(0.174)	(0.065)	(0.062)
ln(pop.dens.)					0.278***	0.194***	0.061***	0.035**
					(0.047)	(0.048)	(0.019)	(0.017)
Male					-0.572***	-0.447***	-0.016	0.012
					(0.126)	(0.124)	(0.051)	(0.047)
College					0.072	0.243***	-0.078***	-0.076***
					(0.082)	(0.078)	(0.029)	(0.026)
Married					0.304***	0.272***	0.125***	0.080**
					(0.091)	(0.090)	(0.035)	(0.032)
Home					-0.057	0.430***	-0.083*	-0.101**
					(0.112)	(0.108)	(0.043)	(0.040)
Unit × Year							✓	✓
Car type FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓		✓	
Vintage-year FE						✓		✓
Covariates	✓		✓		✓	✓	✓	✓
No. of obs.	11,590	11,590	11,590	11,590	11,590	11,590	11,590	11,590
Adj. R^2					0.957	0.962	0.995	0.996
Within R^2					0.297	0.332	0.926	0.933

Notes: SDiD = Synthetic Difference-in-Differences; SCM = Synthetic Control Method; FE = Fixed Effects. The table contains point estimates for the effect of the soot tax on the log number of registered vehicles of treated car models in the sample of 1990-2004 car models. Covariates are a car model's average log household disposable income ("ln(inc)"), log owner age ("ln(owner age)"), log population density of owner's home municipality ("ln(pop.dens.)"), and dummies for being male, college-educated, married, and owning a home. Columns (1) and (2) are SDiD estimates (Equation 4), without and with covariates, respectively. Columns (3) and (4) are SCM estimates (Equation 5), without and with covariates, respectively. Columns (5) to (8) are fixed effect regression estimates (Equation 6). The specifications differ in whether they employ group- or unit-specific linear time trends, and whether they employ year fixed effects or vintage×year fixed effects. Standard errors are in parenthesis. SDiD standard errors are jackknife errors, SCM standard errors are placebo errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table A10: Summary statistics for the German 2005–2008 diesel car-model-level sample

Variable	Unit	No. of obs.	Mean	St.d.	Median
Placebo Treatment Group	Dummy	12,078	0.423	0.494	0
No. of vehicles	Count	12,078	222.1	356.7	103
Vintage	Year	12,078	2006.62	1.140	2007
Vehicle mass	kg	12,078	1,601	295.8	1,599
Displacement	cc	12,078	2,112	562.0	1,995
Emission Class	Factor	12,078	3.894	0.426	4
Mean Purchasing Power	€	12,078	25,308	2,279	24,459
Mean Owner Age	Years	12,078	50.83	4.367	50.50
Mean Pop. Density	Persons/km^2	12,078	543.4	128.3	532.3
Mean Male Owner	Dummy	12,078	0.718	0.110	0.737
Mean Abitur	Dummy	12,078	0.317	0.018	0.318
Mean Non-Single HH	Dummy	12,078	0.606	0.012	0.608
Mean Homeownership	Dummy	12,078	0.713	0.250	0.800

	Control group				Placebo Treatment group			
Variable	No. of obs.	Mean	St.d.	Median	No. of obs.	Mean	St.d.	Median
Placebo Treatment Group	6,975	0	0	0	5,103	1	0	1
No. of vehicles	6,975	250.1	395.4	114	5,103	183.8	291.4	90
Vintage	6,975	2007.07	0.972	2007	5,103	2006.01	1.072	2006
Vehicle mass	6,975	1,668	218.6	1,670	5,103	1,509	356.5	1,436
Displacement	6,975	2,295	560.5	1,998	5,103	1,861	457.3	1,896
Emission Class	6,975	4.075	0.263	4	5,103	3.647	0.478	4
Mean Purchasing Power	6,975	25,365	2,271	24,527	5,103	25,231	2,289	24,345
Mean Owner Age	6,975	50.66	4.457	50.28	5,103	51.05	4.231	50.77
Mean Pop. Density	6,975	551.9	126.6	543.4	5,103	531.7	129.6	518.3
Mean Male Owner	6,975	0.748	0.099	0.765	5,103	0.677	0.111	0.689
Mean Abitur	6,975	0.314	0.017	0.315	5,103	0.320	0.019	0.321
Mean Non-Single HH	6,975	0.605	0.012	0.607	5,103	0.607	0.013	0.609
Mean Homeownership	6,975	0.712	0.249	0.798	5,103	0.715	0.251	0.802

Notes: Descriptive statistics on the German 2005-2008 diesel car model sample in which I observe the number of registered vehicles of 1,342 car models across nine years (2015-2023, January 1). 567 car models are placebo treated, 775 are untreated. "Placebo treatment Group" is one in all periods for car models with PM emissions above 5 mg/km, and zero otherwise. "No. of vehicles" counts the number of registered vehicles of an (identified) car model. "Vintage" is year of first admission in the data. Mean variables are average values of household/owner covariates across all individual vehicles of a car type.

Table A11: Regression results for a Placebo soot tax in Germany using German 2005–2008 diesel sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	SDiD		SCM		FE regression			
Soot tax	-0.010	-0.010	0.009	-0.026	0.005	-0.005	0.005	-0.005
	(0.010)	(0.010)	(0.051)	(0.031)	(0.008)	(0.009)	(0.008)	(0.009)
TG × Year					-0.028***	-0.006*		
					(0.003)	(0.003)		
ln(pp)					-2.480***	-2.552***	0.155	0.161
					(0.321)	(0.314)	(0.258)	(0.256)
ln(owner age)					-0.346***	-0.365***	-0.158***	-0.158***
					(0.083)	(0.078)	(0.056)	(0.056)
ln(pop.dens.)					0.012	-0.007	0.039**	0.041**
					(0.025)	(0.024)	(0.020)	(0.020)
Male					-0.135**	-0.101*	-0.027	-0.028
					(0.142)	(0.053)	(0.041)	(0.041)
High school					-0.941***	-0.720**	-0.439*	-0.454*
					(0.332)	(0.314)	(0.264)	(0.264)
2+ person HH					1.558**	1.158*	0.894	0.917
					(0.712)	(0.696)	(0.577)	(0.574)
Home					0.379	0.431	0.321	0.329
					(0.311)	(0.299)	(0.253)	(0.253)
Unit × Year							✓	✓
Car type FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓		✓	
Vintage-year FE						✓		✓
Covariates		✓		✓	✓	✓	✓	✓
No of obs.	12,078	12,078	12,078	12,078	12,078	12,078	12,078	12,078
Adj. R^2					0.977	0.980	0.992	0.992
Within R^2					0.075	0.034	0.711	0.667

Notes: SDiD = Synthetic Difference-in-Differences; SCM = Synthetic Control Method; FE = Fixed Effects. The table contains point estimates for the effect of a placebo soot tax on the log number of registered vehicles of treated car models in the sample of 2005-2008 diesel car models in Germany. Covariates contain a car model's average household purchasing power, log owner age, log population density of owner's home region, and dummies for being male, having a high school degree, living in a multiple-person-household, and owning a home. "Owner age" and "Male" are directly in the KBA data. The other variables are on the state × region-type level, that is the average value in the state × region-type level in which the owner is registered to live. Concretely, the KBA data may tell us that the owner of a certain vehicle lives in a peripheral rural area in the state of Bavaria. All covariates except "Owner age" and "male" are then measured as the population-weighted average of Bavarian municipalities defined as "peripheral rural". These covariates are: log purchasing power per household ("ln(pp)"), log population density ("ln(pop.dens.)"), share of persons that hold a high school diploma, by gender ("High school"), share of persons living in a household with at least two people ("2+ person HH"), and share of persons living in a family home, either single-family home or semi-detached homes ("Home").
Columns (1) and (2) are SDiD estimates (Equation 4), without and with covariates, respectively. Columns (3) and (4) are SCM estimates (Equation 5), without and with covariates, respectively. Columns (5) to (8) are fixed effect regression estimates (Equation 6). The specifications differ in whether they employ group- or unit-specific linear time trends, and whether they employ year fixed effects or vintage×year fixed effects. Standard errors are in parenthesis. SDiD standard errors are jackknife errors, SCM standard errors are placebo errors. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table A12: Hazard model, cloglog results using 2005-2008 diesel vehicles

Dep. var.: Log(hazard rate)	(1)	(2)	(3)
Soot tax	1.181		1.208
	[1.134, 1.227]		[1.089, 1.327]
Soot tax, 2020		1.224	
		[1.149, 1.299]	
Soot tax, 2021		1.176	
		[1.083, 1.269]	
Soot tax, 2022		1.092	
		[0,994, 1.189]	
Soot tax × High income			1.078
			[1.005, 1.151]
Soot tax × High pop dens			1.091
			[1.016, 1.167]
Soot tax × Home			0.936
			[0.871, 1.001]
Soot tax × High pop			1.002
			[0.933, 1.071]
Soot tax × High age			0.963
			[0.902, 1.025]
Soot tax × Married			1.026
			[0.959, 1.093]
Soot tax × Male			0.964
			[0.895, 1.034]
Soot tax × 2 cars			0.949
			[0.882, 1.017]
Soot tax × 3 cars			0.899
			[0.802, 0.997]
Soot tax × 4 cars			0.986
			[0.810, 1.161]
Soot tax × 5+ cars			0.980
			[0.761, 1.199]
Treatment group	1.083	1.083	1.086
	[1.052, 1.113]	[1.052, 1.114]	[1.055, 1.117]
Household owns 2 cars	1.155	1.155	1.160
	[1.129, 1.180]	[1.129, 1.180]	[1.133, 1.187]
Household owns 3 cars	1.227	1.227	1.241
	[1.184, 1.270]	[1.184, 1.271]	[1.195, 1.287]
Household owns 4 cars	1.269	1.269	1.268
	[1.193, 1.344]	[1.193, 1.344]	[1.188, 1.349]
Household owns 5+ cars	1.231	1.232	1.232
	[1.134, 1.328]	[1.135, 1.328]	[1.130, 1.335]
Vehicle mass (in 100 kg)	1.043	1.043	1.043
	[1.033, 1.052]	[1.033, 1.052]	[1.034, 1.053]
Cylinder volume (in 100 cc)	0.994	0.994	0.994
	[0.989, 0.999]	[0.989, 0.999]	[0.989, 0.999]
CO_2 (in g/km)	0.999	0.999	0.999
	[0.999, 1.000]	[0.999, 1.000]	[0.999, 1.000]
log(income)	0.932	0.931	0.928
	[0.913, 0.950]	[0.913, 0.950]	[0.908, 0.947]
log(pop)	1.037	1.037	1.037
	[1.023, 1.051]	[1.023, 1.051]	[1.023, 1.052]
log(pop dens)	1.012	1.012	1.008
	[1.000, 1.024]	[1.000, 1.024]	[0.995, 1.020]
Owner age	0.986	0.986	0.986
	[0.985, 0.987]	[0.985, 0.987]	[0.985, 0.987]
Married	1.052	1.052	1.048
	[1.029, 1.075]	[1.029, 1.075]	[1.024, 1.072]
Male owner	1.142	1.142	1.147
	[1.114, 1.170]	[1.114, 1.170]	[1.117, 1.178]
Homeownership	0.849	0.849	0.857
	[0.827, 0.871]	[0.827, 0.871]	[0.834, 0.881]
Time FE	✓	✓	✓
Brand FE	✓	✓	✓
No. of obs.	268,059	268,059	268,059

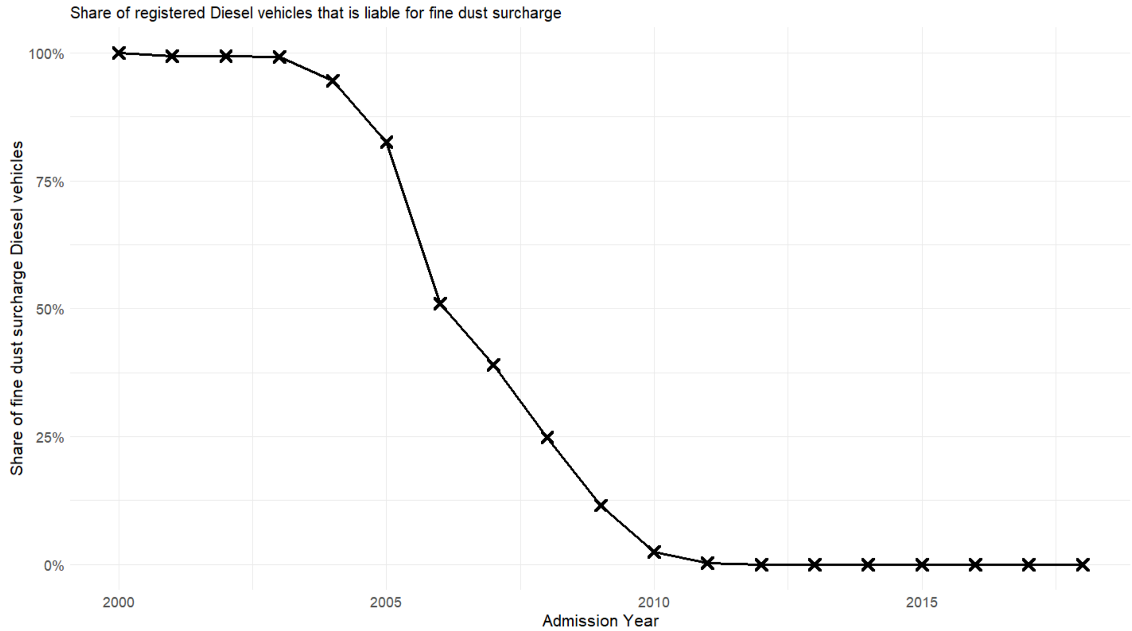
Notes: FE = Fixed Effects. The table contains point estimates for the effect of the soot tax on a vehicle's hazard of permanent deregistration within the next year (from this year's to next year's January 1). The point estimates are the exponentiated estimated coefficients. A value of one corresponds to a zero effect. Interpretation is in percent, e.g. 1.181 means "The hazard increases by 18.1%". All columns refer to Equation 9. Household controls are household disposable income, population and population density of home municipality, owner age, as well as dummies for homeownership, married, male, as well as dummies for how many cars the household owns overall, one (omitted), two, three, four, or five and more cars. Vehicle controls are vehicle mass, cylinder volume, and CO_2 emissions. Column (1) contains only a treatment indicator. Column (2) contains separate indicators for every post-treatment period. Column (3) contains the treatment indicator in simple form as well as interacted with binary variables for above-median ("high") 95% confidence intervals are in square brackets.

Table A13: Back-of-the-envelope calculation using SDiD/SCM estimates

Year	Actual	Scenario	SDiD CF	SCM CF
Number of polluting diesel passenger cars				
2019	216,876		216,876	216,876
2020	174,119		174,119	174,119
2021	135,882		146,726	144,508
2022	111,024		128,330	128,625
2023	93,948		112,835	109,917
Annual PM emissions per diesel passenger car in grams				
2019	165.98		165.98	165.98
2020	109.47		109.47	109.47
2021	94.13	U	97.99	97.21
		L	97.92	97.14
2022	84.21	U	90.85	90.96
		L	90.45	90.55
2023	73.89	U	81.41	80.27
		L	80.79	79.72

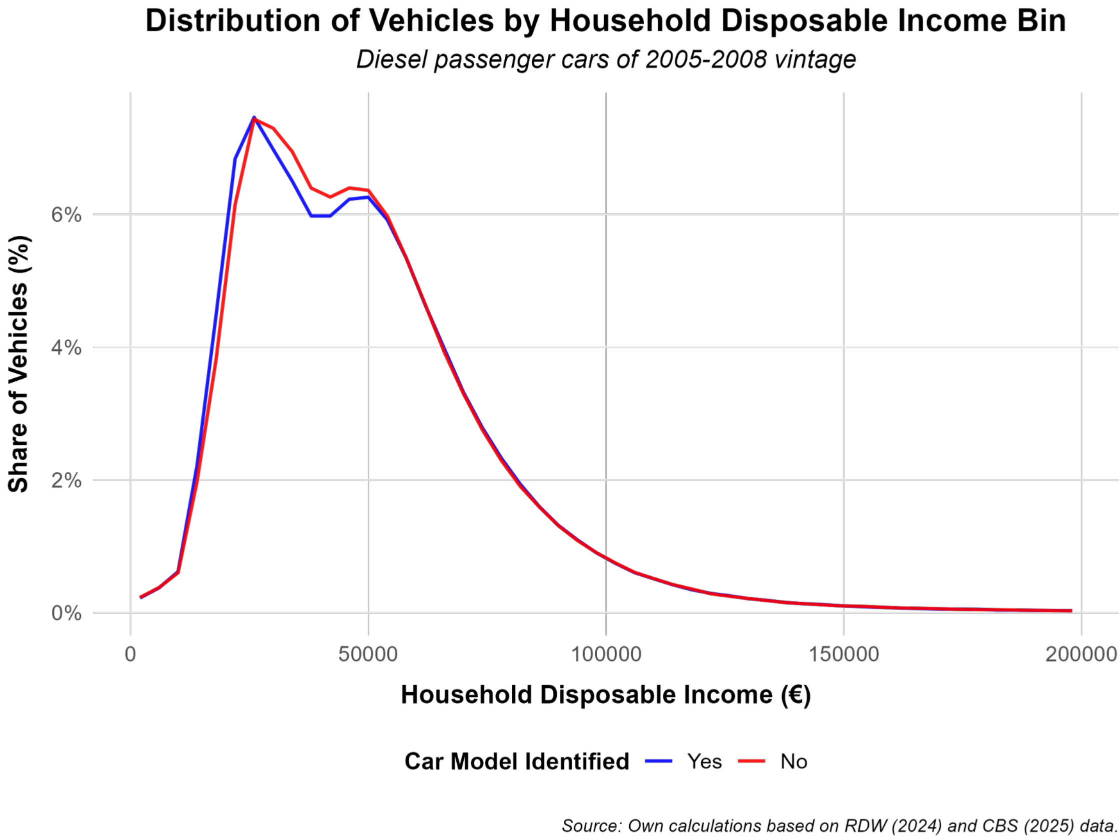
Notes: The counterfactuals (CF) are based on the SDiD and SCM estimates in columns (1) (SDiD) and (2) (SCM) in Table **??**. I assume that diesel cars from before 1990 are not affected by the soot tax due to classical car status, which makes them exempt from registration fees and thus from the soot tax. I assume two extremes regarding the replacement behavior of vehicle owners to calculate annual PM emissions. The upper-bound ("U") estimate assumes that when people replace their eligible diesel car because of the soot tax, they never replace with another diesel car, maximizing the causal reduction in PM emissions (assuming zero PM emissions of whatever vehicle they replace their diesel car with). The lower-bound ("L") estimate assumes that people always replace their eligible diesel car with another diesel car that is as polluting as possible without being eligible (assuming 5 mg/km of PM emissions). This minimizes the causal reduction in PM emissions because the gain in lower PM emissions is made as small as possible, conditional on people replacing to "escape" the soot tax.

Figure A1: Share of diesel vehicles eligible for the soot tax, by year of admission



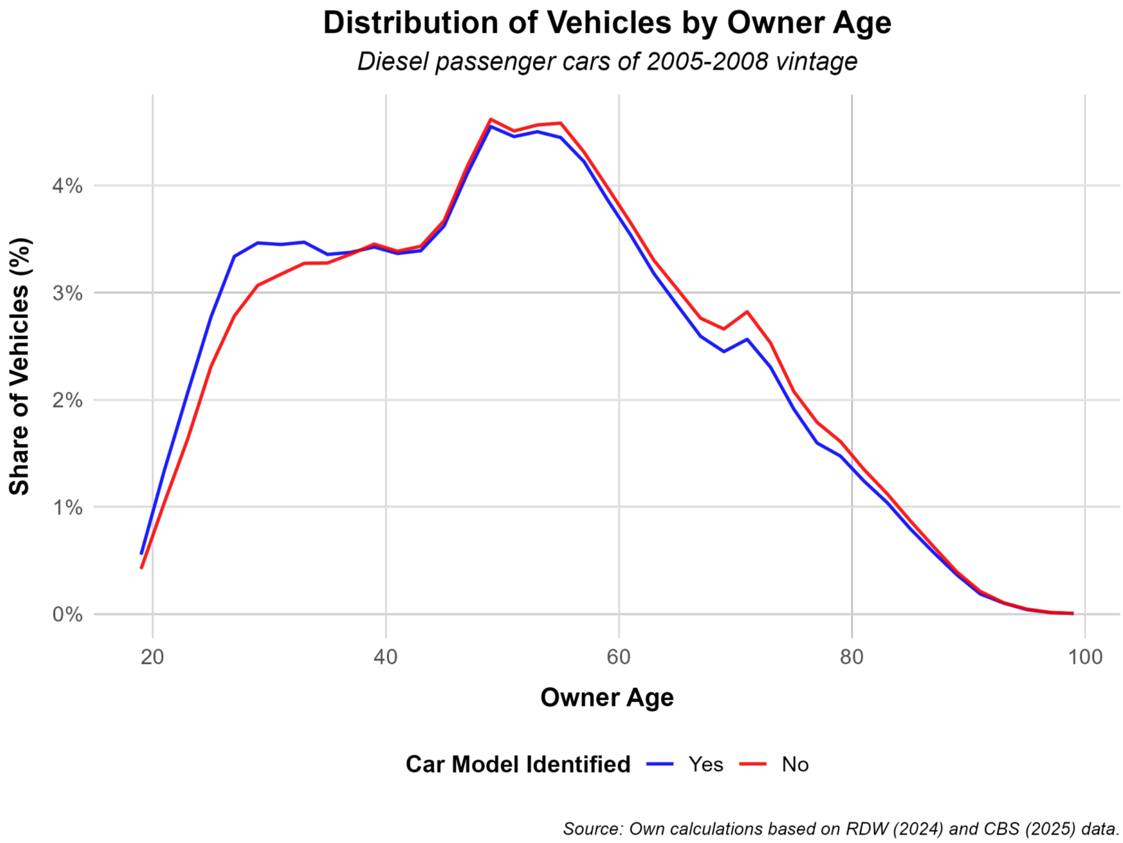
Notes: The figure shows the share of diesel vehicles in each vintage that are eligible for the soot tax. The figure is based on emission data from opendata.rdw.nl on May 4, 2024. I only exclude cars with a removed particulate filter to focus on cars that are eligible because they were produced without a particulate filter.

Figure A2: Distribution of household disposable income by whether a vehicle is of an identified car model



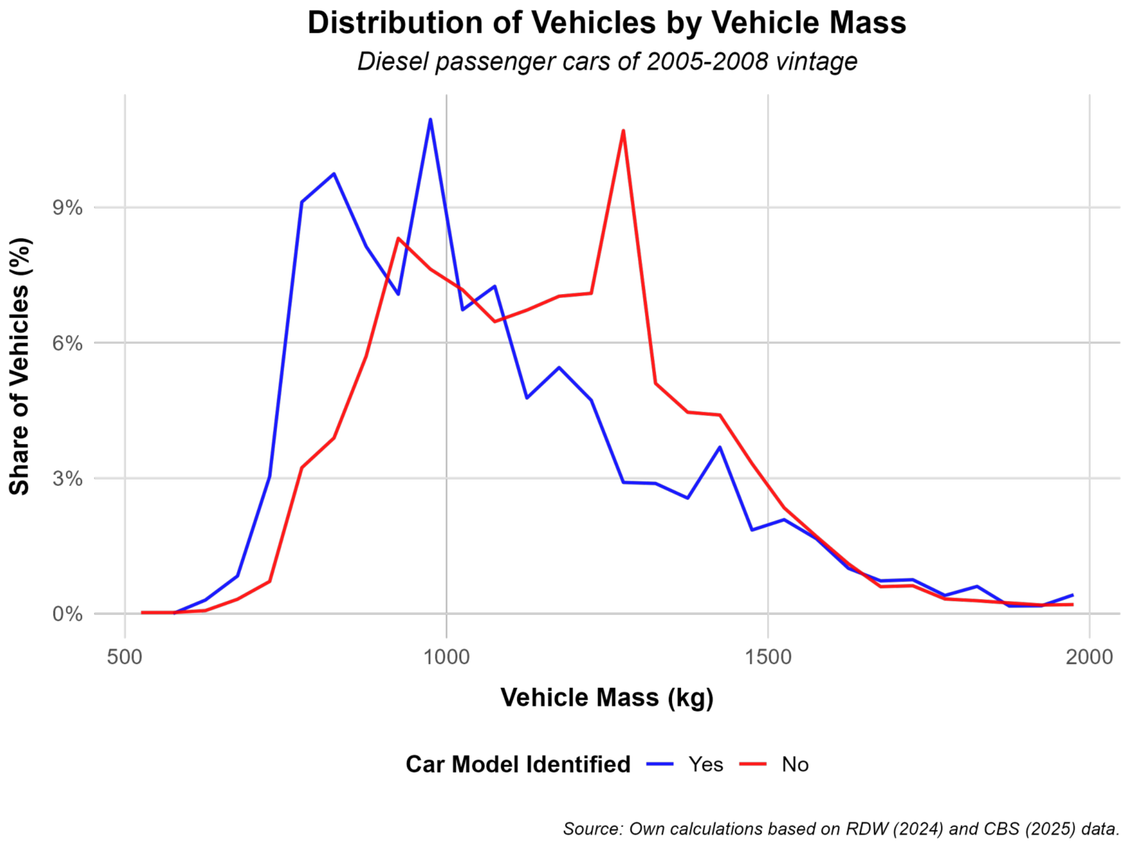
Notes: The figure shows the density of household disposable incomes for registered diesel passenger cars with year of first admission between 2005 and 2008, as of January 1, 2019. Diesel cars are grouped by whether I can assign them an identified make-model. The graph is based on data from the CBS Microdata Environment.

Figure A3: Distribution of owner age by whether a vehicle is of an identified car model



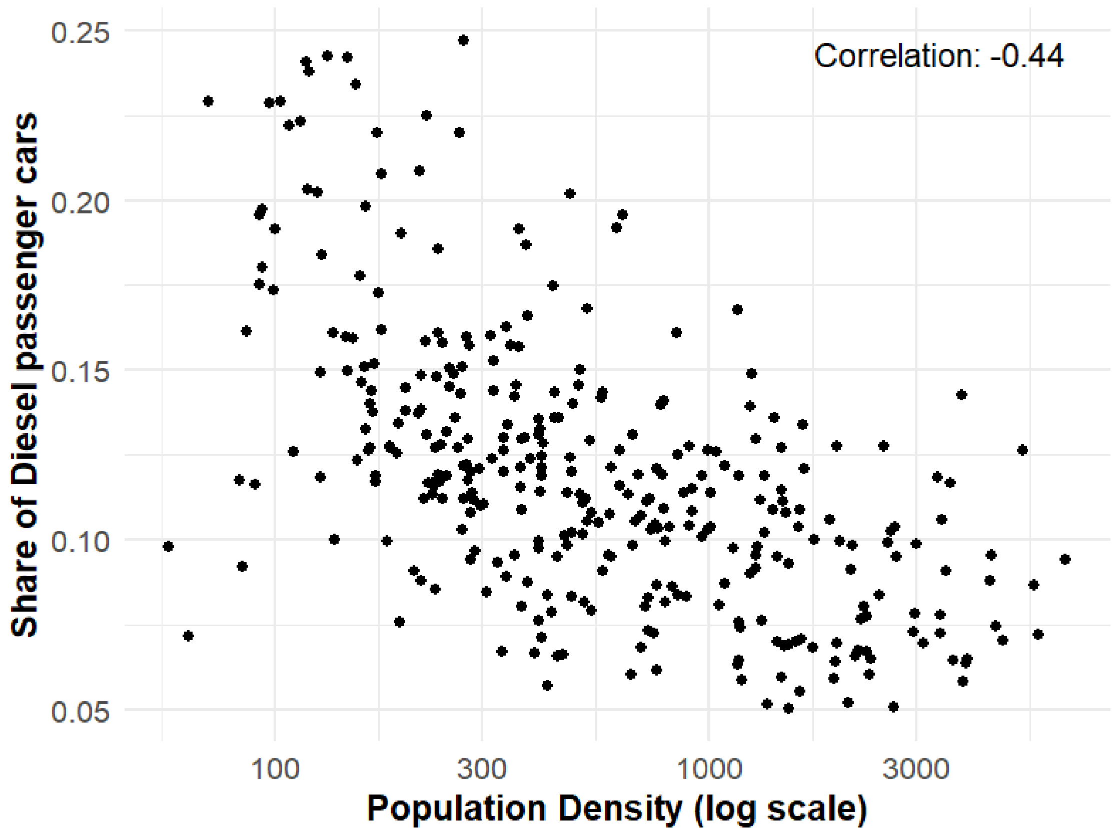
Notes: The figure shows the density of ages of car owners for registered diesel passenger cars with year of first admission between 2005 and 2008, as of January 1, 2019. Diesel cars are grouped by whether I can assign them an identified make-model. The graph is based on data from the CBS Microdata Environment.

Figure A4: Distribution of vehicle mass by whether a vehicle is of an identified car model



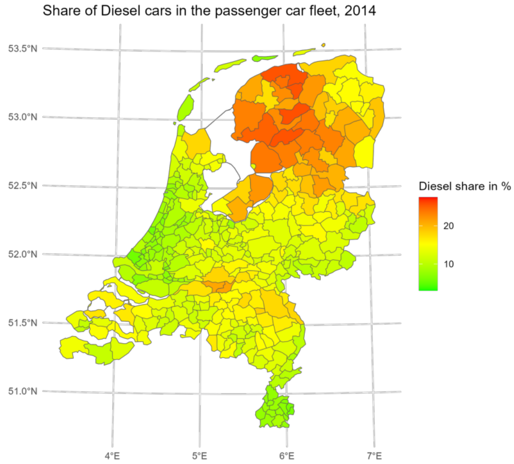
Notes: The figure shows the density of vehicle empty masses in kilograms for registered diesel passenger cars with year of first admission between 2005 and 2008, as of January 1, 2019. Diesel cars are grouped by whether I can assign them an identified make-model. The graph is based on data from the CBS Microdata Environment.

Figure A5: Share of diesel passenger cars vs. population density, per municipality, 2019

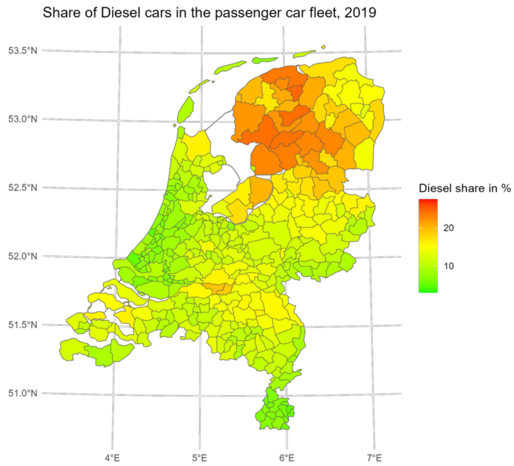


Notes: The scatterplot is based on the vehicle registration data from January 1, 2019. One dot is one municipality in the Netherlands. The plot shows the relationship between population density (x-axis, log scale) and the share of diesel cars in the registered passenger car fleet (y-axis). The graph is based on data from the CBS Microdata Environment.

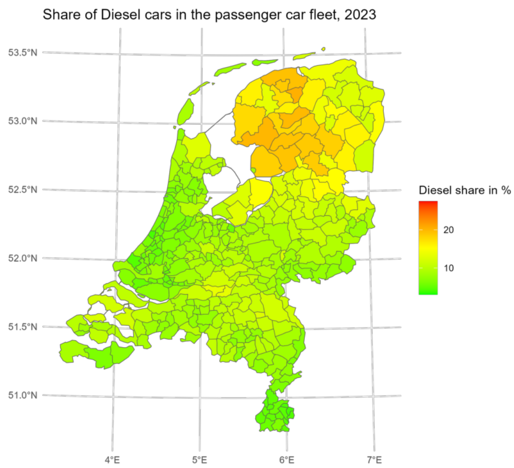
Figure A6: Share of diesel passenger cars by municipality, 2014, 2019, 2023



(a) Jan 1, 2014



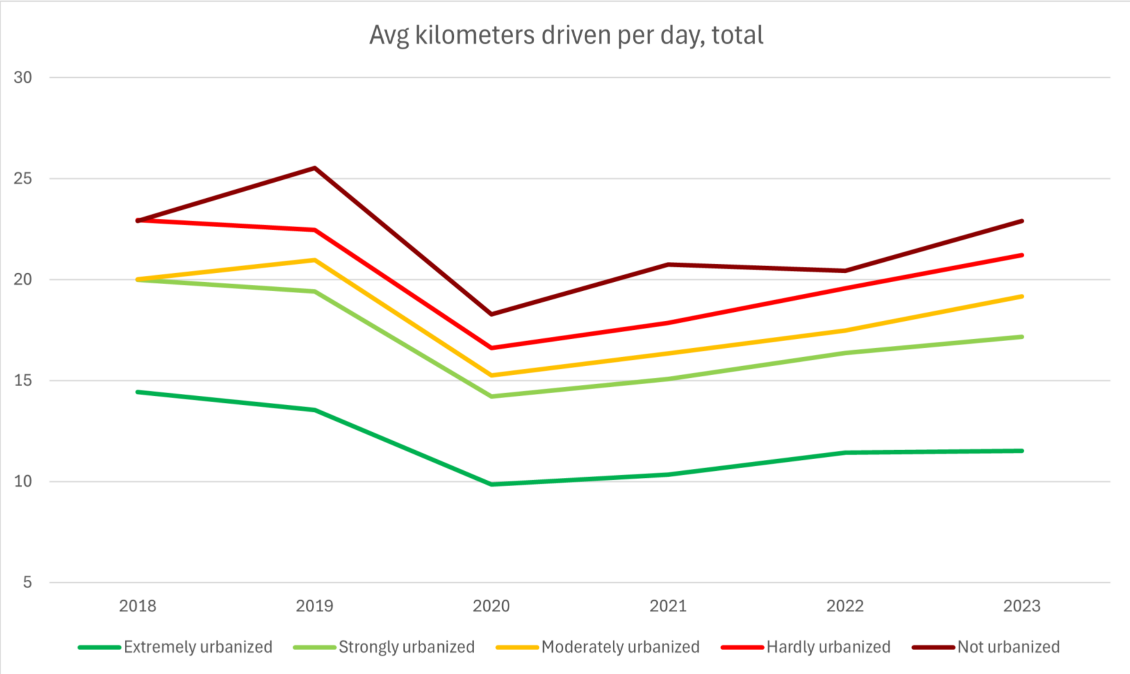
(b) Jan 1, 2019



(c) Jan 1, 2023

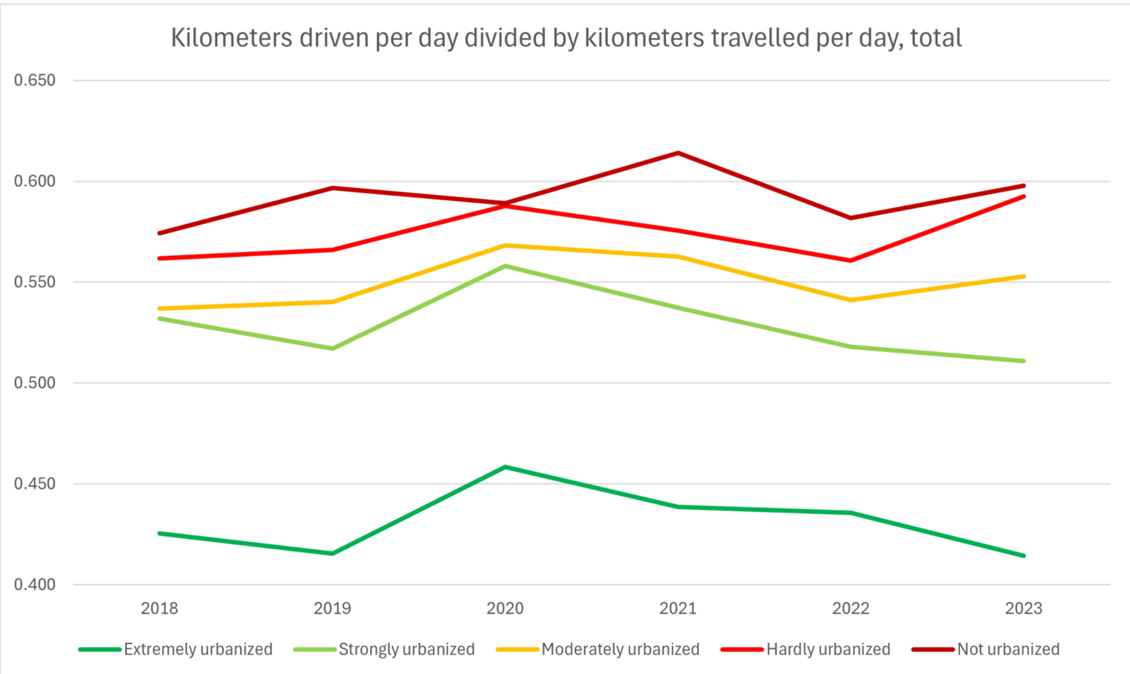
Notes: The maps show the prevalence of diesel passenger cars in the entire passenger car fleet per municipality on Jan 1, 2014 (top), 2019 (middle), and 2023 (bottom). The maps are based on data from the CBS Microdata Environment.

Figure A7: Average kilometers driven per day, all trips (total)



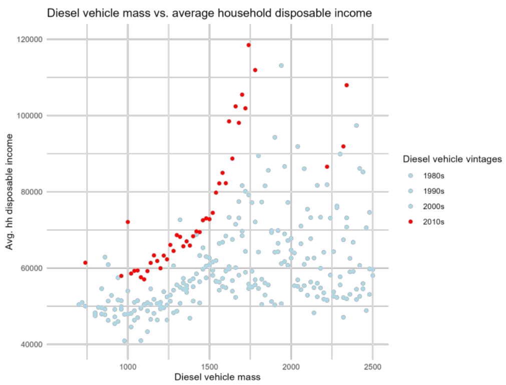
Source: https://opendata.cbs.nl/CBS/en/dataset/84710ENG, own depiction.
Notes: The figure shows kilometers driven per person per day, for the average person in the Netherlands aged 12 and above, including all trips, in every year from 2018 to 2023.

Figure A8: Share of kilometers travelled per day that are done driving, all trips

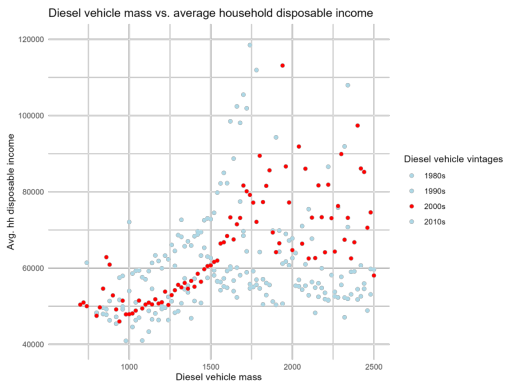


Source: https://opendata.cbs.nl/CBS/en/dataset/84710ENG, own depiction.
Notes: The figure shows the share of daily kilometres travelled that are covered by car, for all individuals aged 12 and above in the Netherlands, across all trips, from 2018 to 2023.

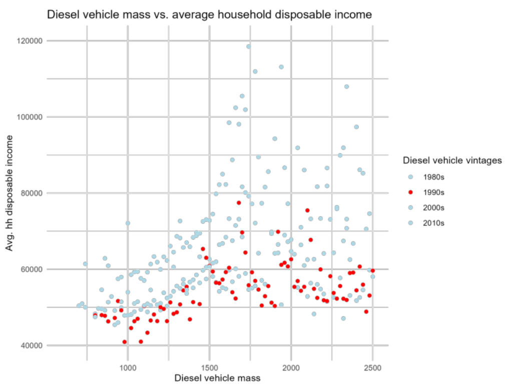
Figure A9: Relationship between vehicle mass and household disposable income, by vintages, 2019



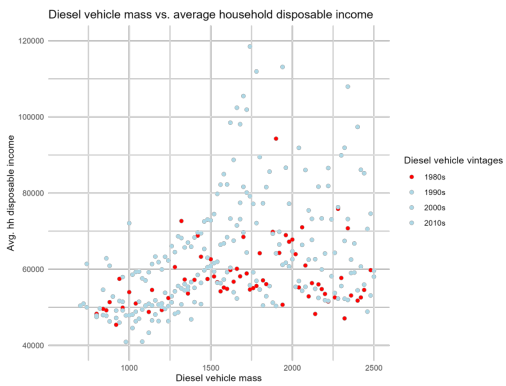
(a) Diesel vintages 2010-2019



(b) Diesel vintages 2000-2009



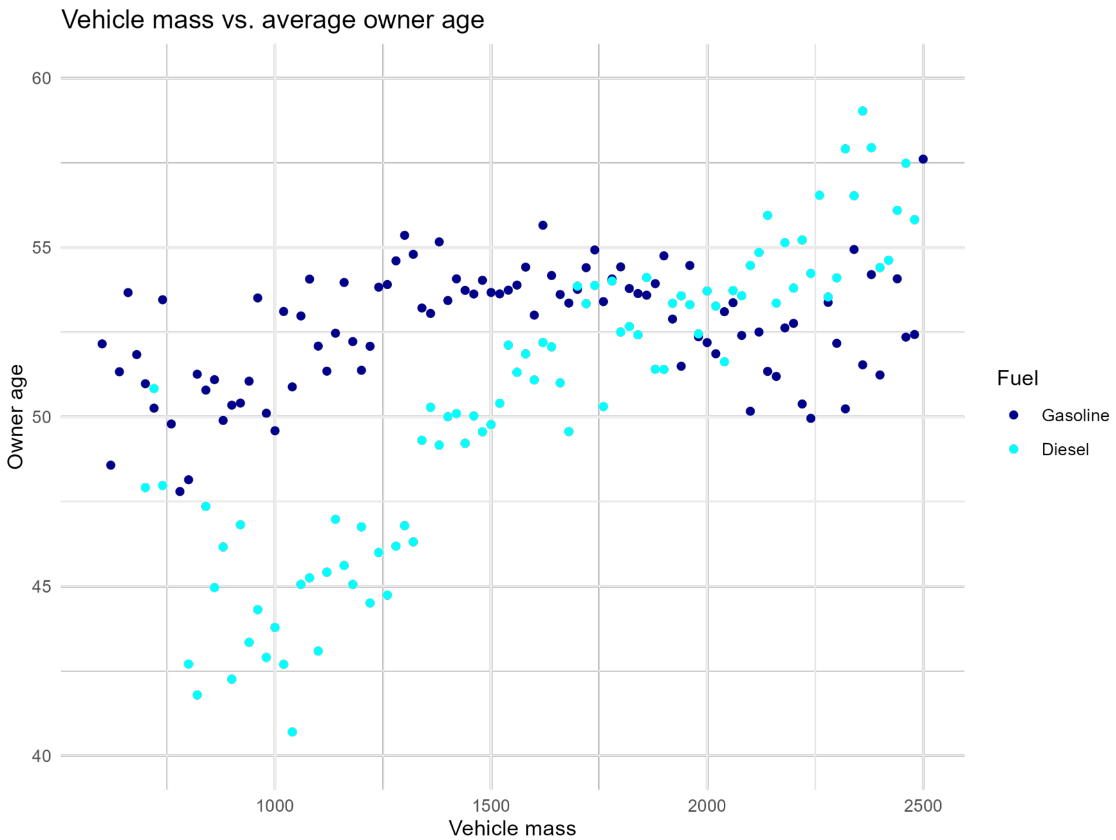
(c) Diesel vintages 1990-1999



(d) Diesel vintages 1980-1989

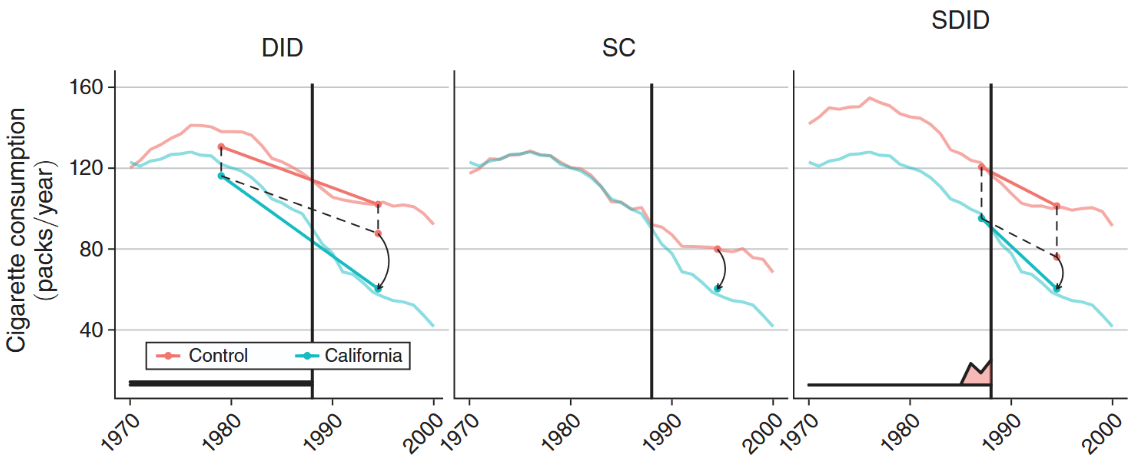
Notes: The graphs show the relationship between vehicle mass and household disposable income on January 1, 2019, separately by vehicle vintage. Disposable income is for the calendar year 2018. I aggregate all passenger cars, separately by fuel type and by vintage (binned by decade), in bins of twenty kilograms and then calculate the average household disposable income of all car owners in every fuel × vehicle mass bin × vintage bin. The graphs are based on data from the CBS Microdata Environment.

Figure A10: Relationship between vehicle mass and owner age, 2019



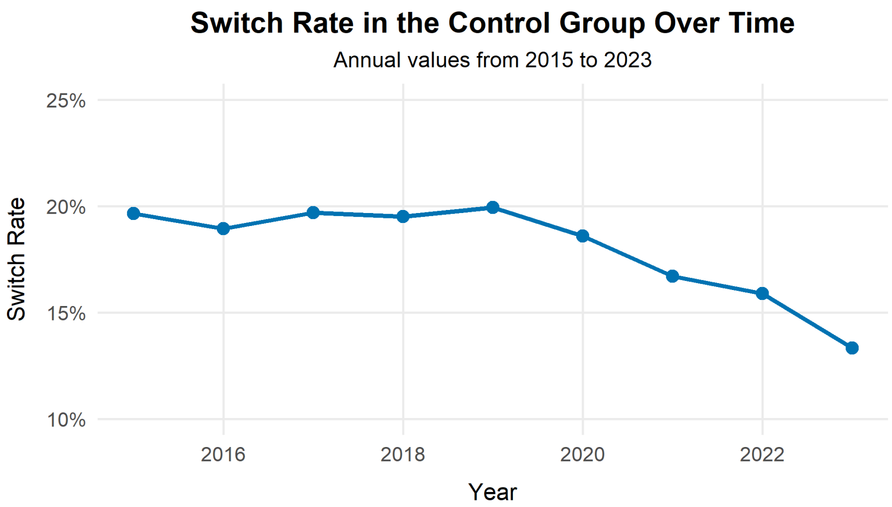
Notes: The graph shows the relationship between vehicle mass and car owner age on January 1, 2019. Disposable income is for the calendar year 2018. I aggregate all passenger cars, separately by fuel type, in bins of twenty kilograms and then calculate the average age of all car owners in every fuel × vehicle mass bin. The graph is based on data from the CBS Microdata Environment.

Figure A11: Synthetic Difference-in-Differences



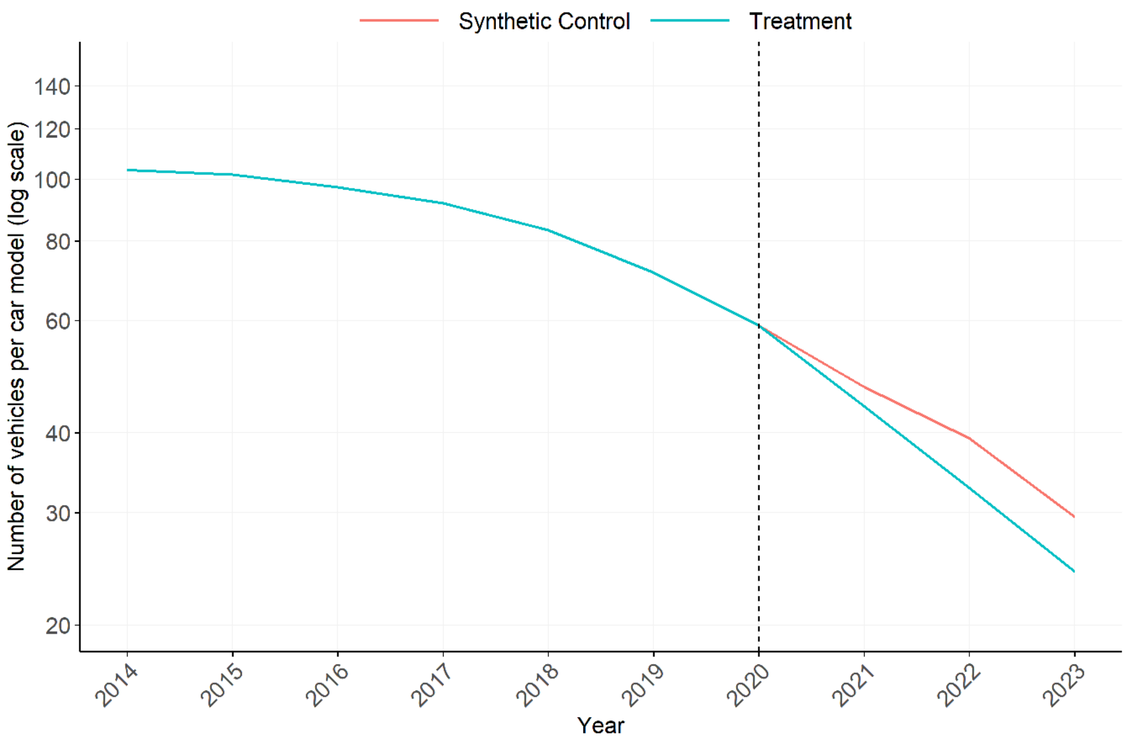
Source: Arkhangelsky et al. (2021).
Notes: The figure shows how Synthetic Difference-in-Differences combines regular Difference-in-Differences and the Synthetic Control Method.

Figure A12: Switch rates in the control group over time



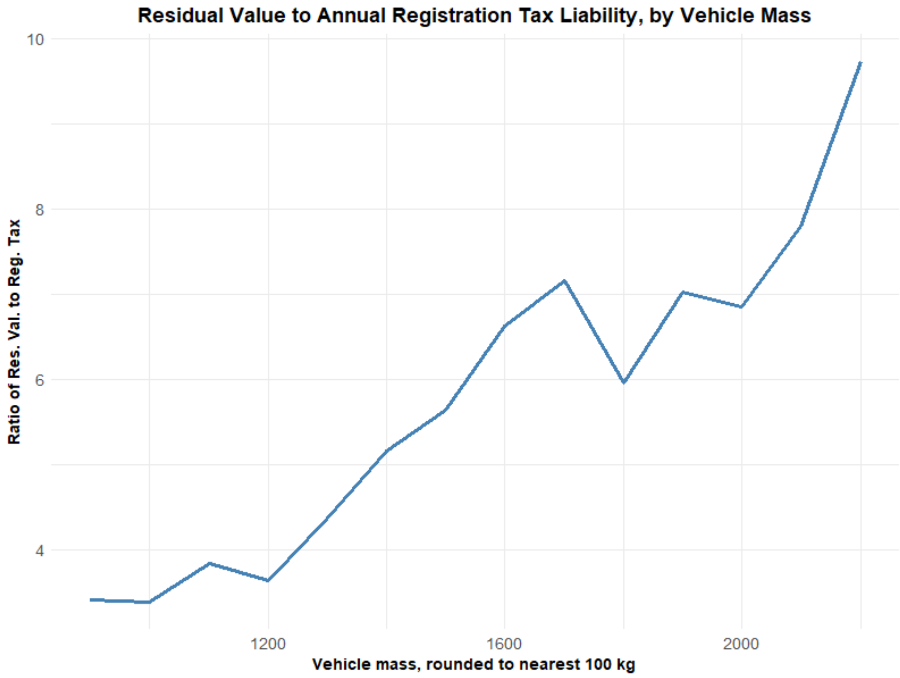
Notes: The graph shows switch rates for the control group in the 2005–2008 diesel sample (non-eligible diesel vehicles from the 2005–2008 vintages). The switch rate is calculated as the share of diesel vehicles that have changed ownership between January 1 of last year and January 1 of this year. Since the vehicle registration data is always collected on January 1, the switch rate for 2015, for instance, is the share of diesel vehicles that switch owner sometime between January 1, 2014 and January 1, 2015. The graph is based on data from the CBS Microdata Environment.

Figure A13: Synthetic Control Method results, 2005-2008 diesel sample.



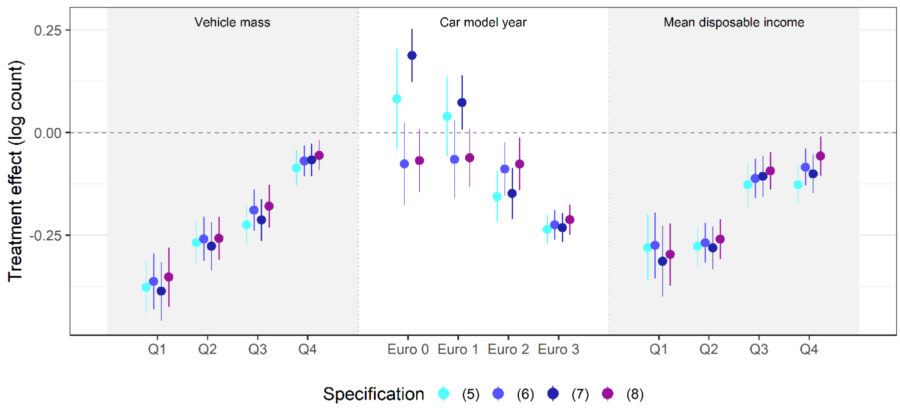
Notes: The graph shows the results from the SCM, 2005- 2008 diesel sample. The Synthetic Control Group is a weighted average of control group units that achieves an equal outcome with the Treatment Group during pre-treatment periods. The outcome is the natural logarithm of the variable “Number of registered vehicles” on the car model × year level. The graph is based on data from the CBS Microdata Environment.

Figure A14: Average ratio of residual car value to annual registration fee liability, 2005-2008 diesel sample, by vehicle mass



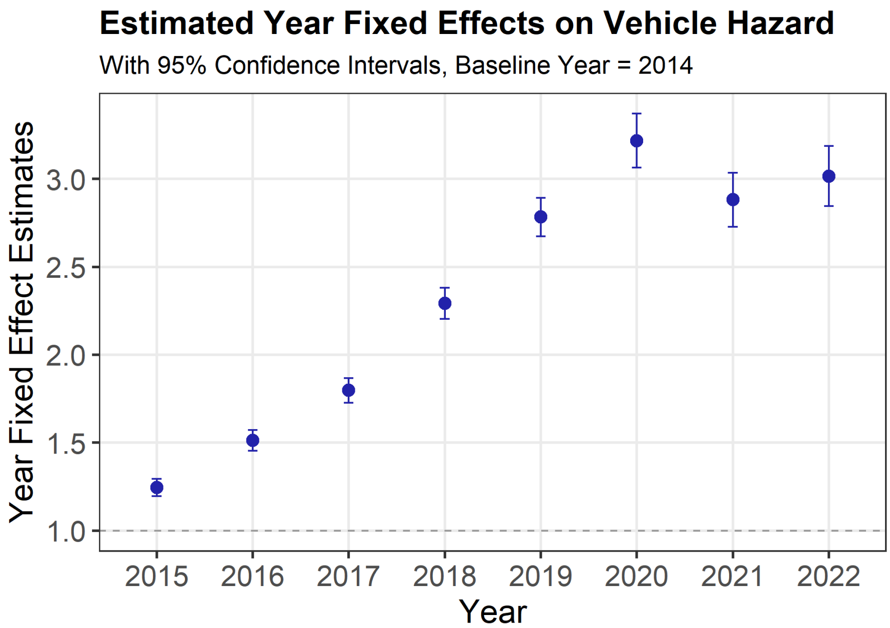
Notes: The graph shows the relationship between a car's ratio of residual value to registration fee liability and vehicle mass, rounded to the nearest 100 kilograms. I estimate each vehicle's residual value by applying a depreciation function to its original list price, applying differentiated depreciation factors depending on brand, fuel type, and vehicle age, with premium brands retaining more value, diesels less, and all vehicles depreciating exponentially over time. I calculate registration fees (averaged across provinces) as these depend only on vehicle mass and fuel type. I do not consider the soot tax because the graph shall depict the ratio of residual value to registration fee liability before the soot tax came into effect. The graph is based on data from the CBS Microdata Environment.

Figure A15: Fixed Effect regression, treatment effect heterogeneity, 1990-2004 diesel sample



Notes: The graph shows point estimates of heterogeneous treatment effects from the Fixed Effects regressions, 1990-2004 sample. The analyzed heterogeneity is in terms of vehicle mass quartile (left), vintage bins by which Euro emission standard was in place during the vehicle's year of first registration (middle), and mean disposable income quartile (right).

Figure A16: Estimated Year Fixed Effects on Vehicle Hazard



Notes: The graph shows Year Fixed Effects estimates from the Hazard Model, Section 5.3., Specification (1), Table 6. The estimates are exponentiated coefficients. A coefficient of one equals a zero effect because the covariates enter multiplicatively so that an exponentiated coefficient of one means that the hazard rate "changes" by a factor of one (hence remains unchanged). 95% confidence intervals are calculated using exponentiated standard errors.

.

B Connecting the data sets

Table A14: Exemplary snippet of CBS data

	Year	Vehicle ID	Fuel Type	Mass	No. Cylinders	First Admission	Owner ID
(1)	2019	10594820	Gasoline	1,205	4	2016	f8431235
(2)	2019	86499389	Diesel	1,050	4	2007	r1343121
(3)	2019	47483920	Gasoline	890	3	2013	t6548952
(4)	2019	51658952	Gasoline	1,300	4	1999	b3215912
(5)	2019	12354849	Diesel	1,560	6	2010	y1115626
(6)	2019	23598566	Electric	1,725	NA	2018	w8411112
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Notes: This table is based exclusively on publicly available information from the RDWNPACTTAB handbook on the CBS website, which describes the corresponding vehicle registration microdata files accessible in the confidential CBS Microdata Environment. It does not contain any information from the Microdata Environment. In particular, the "Vehicle ID" and "Owner ID" columns are randomly generated numbers/letter-number-combinations.

In this section, I explain how I connect my two data sources (RDW and CBS) to identify treated and untreated diesel car models. I use the RDWNPACTTAB data sets from CBS. These data cover the entire universe of registered passenger cars in the Netherlands every year on January 1 from 2014 to 2023. This data set contains information on individual vehicles plus an identifier for their owners. Table A14 provides an exemplary snippet of how the CBS data sets are structured, assuming data from January 1, 2019 for illustration. One row in Table A14 corresponds to one vehicle, identified by a dimensionless ID variable "Vehicle ID". For every vehicle, CBS provides, among other variables, the vehicle's fuel type, vehicle mass in kilograms, number of cylinders, and the year of first admission. CBS also provides information on who owns the vehicle by means of an individual's identifier variable "Owner ID". Based on "Owner ID", I can match covariates such as household disposable income, gender, education level, marital status, place of residence, etc.

In terms of vehicle characteristics, the CBS data **only** contain information on vehicles' fuel type, mass, number of cylinders, and the year of first admission. In the following, I call these the "common variables" because these variables also show up in the data from RDW that I subsequently utilize. The CBS data do not contain information on make, model, and most importantly no information on registered PM emissions. In order to figure out vehicles' make, model, and PM emissions, I first group vehicles based on the four common variables, and then ask when this grouping uniquely identifies a car model. Consider, for example, the 1,050 kg, 4-cylinder diesel from 2007 in row (2) in Table A14. Let us assume I observe 150 individual vehicles with these characteristics, as shown in Table A15.

Table A15: Exemplary snippet of CBS data

	Year	Vehicle ID	Fuel Type	Mass	No. Cylinders	First Admission	Owner ID
(1)	2019	86499389	Diesel	1,050	4	2007	r1343121
(2)	2019	51416322	Diesel	1,050	4	2007	h4165468
(3)	2019	20321599	Diesel	1,050	4	2007	n1654681
(4)	2019	21215463	Diesel	1,050	4	2007	w1356423
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
(149)	2019	65216096	Diesel	1,050	4	2007	e8976589
(150)	2019	84326918	Diesel	1,050	4	2007	a4685981

Notes: This table is based exclusively on publicly available information from the RDWNPACTTAB handbook on the CBS website, which describes the corresponding vehicle registration microdata files accessible in the confidential CBS Microdata Environment. It does not contain any information from the Microdata Environment. In particular, the "Vehicle ID" and "Owner ID" columns are randomly generated numbers/letter-number-combinations.

The natural question is whether all these vehicles are of the same make-model and have the same level of PM emissions. If so, I can assign them a make-model and a soot tax treatment status. The RDW data set on registered vehicles in the Netherlands helps me with this endeavor. It contains the common variables (fuel type, mass, number of cylinders, and year of first admission) **plus** more than seventy other variables, such as the actual license plate number, list price, fuel consumption, number of doors, and importantly make, model, and PM emissions. The RDW data are updated daily to always describe the current state of the Dutch vehicle fleet. I downloaded all available data on RDW's Open Data website (opendata.rdw.nl) on May 4, 2024.

Table A16: Exemplary snippet of RDW data

	Day	License Plate	Fuel Type	Mass	No. Cylinders	First Admission	Brand	Model	PM emissions	...
(1)	24-05-04	VES42A	Gasoline	1,150	4	2017	AUDI	A1	NA	...
(2)	24-05-04	BPE26W	Gasoline	920	4	2004	FORD	FIESTA	NA	...
(3)	24-05-04	VXA31A	Diesel	1,565	6	2013	BMW	5	0.0015	...
(4)	24-05-04	GPS75R	Electric	1,700	4	2023	TESLA	MODEL 3	NA	...
(5)	24-05-04	KWJ32N	Diesel	1,370	4	2004	PEUGEOT	307	0.0210	...
(6)	24-05-04	REE96F	Gasoline	1,450	4	2015	SEAT	IBIZA	NA	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table A16 provides a sense of how the RDW data looks like. I use this data in the following way. Consider again the 150 vehicles from Table A15 with the characteristics 1,050 kg, four cylinders, running on diesel, first admitted in 2007. The natural question is "Are all these cars of the same make-model, and can I attribute a certain level of PM emissions to them (which then determines treatment status)?". I do this by filtering the RDW data for all cars that have the same values of the common variables. If I am lucky, I will get something like this:

Table A17: Exemplary snippet of RDW data

	Day	License Plate	Fuel Type	Mass	No. Cylinders	First Admission	Brand	Model	PM emissions	...
(1)	24-05-04	BYI42G	Diesel	1,050	4	2007	VW	POLO	0.0240	...
(2)	24-05-04	MSN34Q	Diesel	1,050	4	2007	VW	POLO	0.0240	...
(3)	24-05-04	PAL44D	Diesel	1,050	4	2007	VW	POLO	0.0240	...
(4)	24-05-04	ANC23T	Diesel	1,050	4	2007	VW	POLO	0.0240	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
(N-1)	24-05-04	LAQ74F	Diesel	1,050	4	2007	VW	POLO	0.0240	...
(N)	24-05-04	SLS39N	Diesel	1,050	4	2007	VW	POLO	0.0240	...

This is the best case. All N cars in the RDW data with these values of the common variables are (i) the same make-model ("Volkswagen Polo" in this example) and (ii) have a PM emissions reading above or below the soot tax threshold of 0.005 g/km (above in this example). In this case, I would then assign to all 150 vehicles in the CBS data (Table A15) the make-model "VW POLO" and put them in the treatment group because the inferred PM emissions are above the soot tax threshold. When this assignment of a car model and a treatment status works, I call these vehicles and this car model "identified", because I can identify vehicles' make-model and treatment status in the RDW data based on the common variables.

Several things can go wrong that do not allow me to "identify" a car model and/or its soot tax eligibility. First, a combination of values of the common variables may correspond to more than one make-model. This sometimes happens, in particular with the German premium brands Audi, BMW and Mercedes-Benz. Their models often have the same vehicle mass and number of cylinders, which makes it impossible to distinguish them solely based on the common variables. I disregard these cases because it would render fixed effect terms in Equation 4 or Equation 6 meaningless.

Second, a combination of values of the common variables may not uniquely identify treatment status. Even within the same car model, some vehicles may have PM emissions above the soot tax threshold while others may be below the threshold. For example, some manufacturers have sold the same model both with and without particulate filters for some time, sometimes at the buyer's discretion. In such cases, I cannot assign treatment status based on the common variables.

I allow for an error margin of five percent in my identification. That is, I consider a car model and treatment status "identified" if at least 95% of vehicles with the same values of the common variables (i) are of the same car model in the RDW data, and (ii) have PM emissions either above or below the soot tax threshold. For example, if I observed 100 vehicles in the RDW data with the same values of the common variables, 97 of which are Audi A3 with PM emissions below the threshold and 3 are Skoda Octavia with PM emissions above the threshold, I would consider this within my error margin and still assign an identified car model "Audi A3" and an identified treatment status of "untreated".

At the end of this process, I have an unbalanced panel data set with identified vehicles spanning the years 2014 to 2023. In my main analysis, I restrict this data set to diesel passenger cars with year of first admission between 2005 and 2008. These are the vintages of diesel vehicles for which I have quasi-random variation in filter status, and thus treatment status, due to the gradual rollout of filters in the production process that constitutes my natural experiment. This data set on the individual vehicle level is used in the hazard model analysis in Section 4.2 (once I restrict the sample to vehicles that I continuously observe every year, either until they disappear from the data permanently or until the last observation period).

This data set looks like the exemplary snippet in Table A18. For example, there is diesel car 86499389, an identified Volkswagen Polo from 2007. I observe this car every year until 2020 after which it disappears from the data. My interpretation of this is that the event of "deregistration" (or equivalently "retirement") occurs sometime during the year 2020 because the data is always from January 1. In other words, I observe the car on January 1, 2020, but not anymore on January 1, 2021, so that deregistration must have occurred sometime during the year 2020. I also learn that the car switched hands sometime during 2016 from owner g2629269 to owner r1343121, and that its type-approval PM emissions of 0.024 g/km exceed the soot tax threshold of 0.005 g/km.

In contrast, the car 51658952, a 2008 Peugeot 308, has PM emissions below the soot tax threshold, has never switched hands (always registered in the name of owner p7546245), and is not deregistered during the observation period. It is still registered at the last time of observation on January 1, 2023.

Table A18: Exemplary snippet of individual vehicle level data set

	Year	Vehicle ID	Fuel Type	Mass	Cylinders	First Admission	Owner ID	Brand	Model	PM
(1)	2014	86499389	Diesel	1,050	4	2007	g2629269	VW	POLO	0.0240
(2)	2015	86499389	Diesel	1,050	4	2007	g2629269	VW	POLO	0.0240
(3)	2016	86499389	Diesel	1,050	4	2007	g2629269	VW	POLO	0.0240
(4)	2017	86499389	Diesel	1,050	4	2007	r1343121	VW	POLO	0.0240
(5)	2018	86499389	Diesel	1,050	4	2007	r1343121	VW	POLO	0.0240
(6)	2019	86499389	Diesel	1,050	4	2007	r1343121	VW	POLO	0.0240
(7)	2020	86499389	Diesel	1,050	4	2007	r1343121	VW	POLO	0.0240
(8)	2014	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(9)	2015	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(10)	2016	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(11)	2017	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(12)	2018	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(13)	2019	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(14)	2020	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(15)	2021	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(16)	2022	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(17)	2023	51658952	Diesel	1,300	4	2008	p7546245	PEUGEOT	308	0.0020
(18)	2014	32592631	Diesel	1,775	6	2006	s4284631	BMW	5 SERIES	0.0003
(19)	2015	32592631	Diesel	1,775	6	2006	s4284631	BMW	5 SERIES	0.0003
(20)	2016	32592631	Diesel	1,775	6	2006	s4284631	BMW	5 SERIES	0.0003
(21)	2017	32592631	Diesel	1,775	6	2006	s4284631	BMW	5 SERIES	0.0003
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮			

Notes: This table is based exclusively on publicly available information from the RDWNPACTTAB handbook on the CBS website, which describes the corresponding vehicle-registration microdata files accessible in the confidential CBS Microdata environment. It does not contain any information from the Microdata Environment. In particular, the "Vehicle ID" and "Owner ID" columns are randomly generated numbers/letter-number-combinations.

Now, to obtain the data on the car model level that I use in my Difference-in-Differences analysis, I aggregate the individual vehicle level data as in Table A18 on the car model level, creating a variable "Number of registered vehicles" that counts how many individual vehicles of each identified car model are registered in each period. Having aggregated the data on the car model level, I get a sample as in Table A19. The data set becomes a balanced panel data set on the car model × year level. I follow every identified car model over the ten years 2014 to 2023, counting the number of registered vehicles ("No. of vehicles") in every year. Because these are old cars from the mid-2000s, their numbers tend to go down from year to year.

The natural logarithm of "No. of vehicles" is the outcome on which all the results in Section 5.1. are eventually based on, providing me with estimates for the percent effect of the soot tax on total vehicle holdings.

Table A19: Exemplary snippet of individual vehicle level data set

	Year	Fuel Type	Mass	Cylinders	First Admission	Brand	Model	PM	No. of vehicles
(1)	2014	Diesel	1,050	4	2007	VW	POLO	0.0240	169
(2)	2015	Diesel	1,050	4	2007	VW	POLO	0.0240	162
(3)	2016	Diesel	1,050	4	2007	VW	POLO	0.0240	156
(4)	2017	Diesel	1,050	4	2007	VW	POLO	0.0240	150
(5)	2018	Diesel	1,050	4	2007	VW	POLO	0.0240	144
(6)	2019	Diesel	1,050	4	2007	VW	POLO	0.0240	137
(7)	2020	Diesel	1,050	4	2007	VW	POLO	0.0240	131
(8)	2021	Diesel	1,050	4	2007	VW	POLO	0.0240	117
(9)	2022	Diesel	1,050	4	2007	VW	POLO	0.0240	102
(10)	2023	Diesel	1,050	4	2007	VW	POLO	0.0240	85
(11)	2014	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	242
(12)	2015	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	233
(13)	2016	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	224
(14)	2017	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	212
(15)	2018	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	199
(16)	2019	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	186
(17)	2020	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	175
(18)	2021	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	161
(19)	2022	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	148
(20)	2023	Diesel	1,300	4	2008	PEUGEOT	308	0.0020	136
(21)	2014	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	208
(22)	2015	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	200
(23)	2016	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	193
(24)	2017	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	183
(25)	2018	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	175
(26)	2019	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	167
(27)	2020	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	161
(28)	2021	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	151
(29)	2022	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	139
(30)	2023	Diesel	1,775	6	2006	BMW	5 SERIES	0.0003	208
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮		

C A exemplary model of vehicle replacement

Consider a simple random utility formulation that rationalizes vehicle replacement decisions. Let $T_i = T(q_i)$ be today's residual value of old car i that positively depends on vehicle quality q_i. Let $t_i = t(q_i)$ be the registration fee that is due when its owner decides to keep old car i. In the spirit of Jacobsen and van Benthem (2015), car owners face a recurring random repair cost shock. The random repair cost shock in my model, z_i, is defined as a share of the car's residual value so that z_i lies in the unit interval, $z_i \in (0, 1]$. Realized repair costs c_i are then given by $c_i = T_i \cdot z_i$. I assume that the distribution of the random repair cost shock z_i depends on vehicle quality q_i, hence $z_i \sim F_z(\cdot \mid q_i)$. I further assume that $F_z(\cdot \mid q_i)$ is decreasing in q in the first-order stochastic dominance (FOSD) sense:

$$q' < q'' \;\Rightarrow\; F_z(z|q') \leq F_z(z|q'')\;\forall z \;\wedge\; \exists z:\; F_z(z|q') < F_z(z|q'')$$

This property says that draws of z are stochastically larger for lower values of vehicle quality q because, by the nature of low quality, such vehicle parts tend to break down more frequently, requiring higher repair costs (relative to residual value).

A simple random-utility formulation for the utility of **keeping** the old car i is

$$\begin{aligned}
U_i^K &= a + \gamma \log(T_i - c_i) - \gamma \log(t_i) + \eta_i + \epsilon_i^K \\
&= a + \gamma \log(T_i (1 - z_i)) - \gamma \log(t_i) + \eta_i + \epsilon_i^K \\
&= a + \gamma \log(T_i) + \gamma \log(1 - z_i) - \gamma \log(t_i) + \eta_i + \epsilon_i^K \\
&\approx a + \gamma \log(T_i) - \gamma z_i - \gamma \log(t_i) + \eta_i + \epsilon_i^K \\
&= a + \gamma (\log(T_i) - \log(t_i) - z_i) + \eta_i + \epsilon_i^K
\end{aligned}$$

The utility of keeping the car depends positively on the car's residual value net of repair costs, $T_i - c_i$, and negatively on the registration fee t_i. For simplification, I assume that both factors share the same coefficient γ, and I use $\log(1 - x) \approx -x$ for small x in the fourth row. The logarithmic structure of the utility captures that residual values generally develop as a percentage of the initial new car value. An x% lower residual value should have the same effect on utility regardless of the car's absolute residual value. Logs also naturally capture diminishing sensitivity because a loss of, for instance, €1,000 in residual value should have a larger effect on utility the lower the baseline residual value is. Logs further reduce skewness to right-hand side outliers and turn multiplicative measurement errors into additive terms. This matters in particular with respect to measurement error in vehicle residual value, which is likely to be multiplicative rather than additive. Lastly, the unobserved term η_i captures both vehicle- and owner-specific factors.

The utility of **replacing** old car i is given by

$$U_i^R = b_i + \epsilon_i^R$$

The utility of replacing is expected scrap value $b_i = b(q_i)$ plus some idiosyncratic shock to the scrap value ϵ_i^R. The expected value depends positively on vehicle quality b_i; higher-quality vehicles yield higher scrap value in expectation. It follows that the difference between the utility of keeping and replacing is then given by

$$\Delta U_i = U_i^K - U_i^R = (a - b_i) + \gamma\left(\log\ T_i - \log\ t_i - z_i\right) + \eta_i + \left(\epsilon_i^K - \epsilon_i^R\right)$$

I define $r_i \equiv \frac{T_i}{t_i}$ as the ratio of car i's residual value to its registration fee liability. I collect terms in $\alpha_i \equiv a + \eta_i - b_i$ and $\epsilon_i \equiv \epsilon_i^K - \epsilon_i^R$. We get the following compact formulation for the utility difference:

$$\Delta U_i \ = \ \alpha_i \ + \ \gamma\left(\log\ r_i \ - z_i\right) \ + \ \epsilon_i$$

The straightforward decision rule is that the owner keeps old car i if $\Delta U_i > 0$, replaces it if $\Delta U_i < 0$, and is indifferent if $\Delta U_i = 0$. Therefore, given a realized repair cost shock z_i, the probability that the owner replaces is given by

$$\begin{aligned}
\mathrm{P}\left(\text{replace} \mid z_i, q_i\right) &= \Pr\left(\Delta U_i < 0 \mid z_i, q_i\right) \\
&= \Pr\left(\alpha_i \ + \ \gamma\left(\log\ r_i \ - z_i\right) \ + \ \epsilon_i \ < \ 0\right) \\
&= \Pr\left(\epsilon_i \ < \ \gamma z_i - \alpha_i - \gamma \log r_i\right) \\
&= \Lambda(\underbrace{\gamma z_i - \alpha_i - \gamma \log r_i}_{x_i(z_i,q_i)}) \\
P_i(z_i, q_i) &= \Lambda\left(x_i(z_i, q_i)\right)
\end{aligned}$$

The last row follows from my assumption that the error ϵ_i is a logistic error so that the CDF of P_i is of the form $\Lambda(x) = \frac{1}{1+e^{-x}}$. The replacement probability

· increases with the repair cost shock z_i

(more expensive repairs $\Rightarrow$ replacement more likely);

· increases with scrap value b_i

(higher scrap value b_i $\Rightarrow$ lower α_i $\Rightarrow$ replacement more likely);

· decreases with ratio r_i

(higher residual value relative to registration fee $\Rightarrow$ replacement less likely).

Because I assume $z \mid q \ \sim \ F_z\left(\cdot \mid q\right)$, I can express the replacement probability unconditioned on z as

$$P_i(q_i) \ = \ \mathbb{E}_{z|q_i}\left[\Lambda\left(x_i(z, q_i)\right)\right]$$

Now, I can show how the replacement probability $P_i(q_i)$ increases when the (log) ratio of residual value to registration fee liability shrinks. That is, both a lower residual value and a higher registration fee liability make replacement more likely:

$$
\begin{aligned}
\frac{\partial P_i\left(q_i\right)}{\partial \log r_i} &= \frac{\partial}{\partial \log r_i} \mathbb{E}_{z|q_i}\left[\Lambda\left(x_i(z, q_i)\right)\right] \\
&= \mathbb{E}_{z|q_i}\left[\Lambda'\left(x_i(z, q_i)\right) \cdot \frac{\partial x_i(z, q_i)}{\partial \log r_i}\right] \\
&= \mathbb{E}_{z|q_i}[\Lambda\left(x_i(z, q_i)\right)\left(1-\Lambda\left(x_i(z, q_i)\right)\right) \cdot (-\gamma)] \\
&= -\gamma \cdot \mathbb{E}_{z|q_i}[\Lambda\left(x_i(z, q_i)\right)\left(1-\Lambda\left(x_i(z, q_i)\right)\right)] \\
&\equiv -\gamma \cdot M(q_i) \\
&< 0
\end{aligned}
$$

In the second-to-last line, I define $M(q_i) \equiv \mathbb{E}_{z|q_i}\left[\Lambda\left(x_i(z, q_i)\right)\left(1-\Lambda\left(x_i(z, q_i)\right)\right)\right]$. Now, we can use this derivative to elicit the effect of the soot tax on a vehicle's replacement probability. The soot tax τ exogenously increases the registration fee from t to $(1+\tau)\, t$ and thereby changes the residual-value-to-fee ratio from $r = \frac{T}{t}$ to $r' = \frac{T}{(1+\tau)t}$. This is equivalent, in terms of $\log r$, to $\Delta \log r = \log r' - \log r = -\log(1+\tau)$. Accordingly, I use a first-order approximation for the change in the individual replacement probability P_i, namely $\Delta P_i = P_i\left(r'\right) - P_i\left(r\right)$, when $\log r_i$ changes by $\Delta \log r_i = -\log\left(1+\tau\right)$:

$$
\begin{aligned}
\Delta P_i\left(q_i\right) &\approx \frac{\partial P_i\left(q_i\right)}{\partial \log r_i}|_{r_i=r} \cdot \Delta \log r_i \\
&= -\gamma \cdot M\left(z, q_i\right) \cdot \left(-\log(1+\tau)\right) \\
&= \gamma \cdot \log\left(1+\tau\right) \cdot M\left(q_i\right)
\end{aligned}
$$

Note that $\Delta P_i < 0$ when $\tau < 0$. The individual replacement probability P_i goes down, which means that keeping the car becomes more appealing, when the registration fee decreases. In parallel, when $\tau > 0$ so that $\Delta P_i > 0$, the higher registration fee increases the utility of replacing so that car i's replacement probability P_i goes up.

Now, we can look into the empirical finding that lighter cars respond more strongly to the soot tax than heavier cars. Besides vehicle mass, a light car l differs from a heavy car h mainly in its quality. In Europe, the average light car l has a substantially lower vehicle quality than the average heavy car h, $q_l < q_h$. Because $T_i = T(q_i)$ and $\frac{\partial T_i}{\partial q_i} > 0$, $q_h > q_l \;\Rightarrow\; T_h > T_l$. The heavy and higher-quality car h has a higher residual value than the light and lower-quality car l (at the same vehicle age).

The Dutch registration fee schedule penalizes vehicle mass so that $t_h > t_l$. The registration fee for heavy car h is higher than for light car l. The data (Figure A14) indicate that $\frac{T_h}{T_l} > \frac{t_h}{t_l} \Leftrightarrow \frac{T_h}{t_h} > \frac{T_l}{t_l} \Leftrightarrow r_h > r_l$. Car h's residual-value-to-fee ratio r_h is higher than car l's ratio r_l because residual values diverge more than registration fees.

We can now see how quality differences between the light car l and the heavy car h explain both their differing baseline replacement probabilities as well as their differential responses to the soot tax. First, because $P_i(q_i) = \mathbb{E}_{z|q_i}\left[\Lambda\left(x_i\left(z, q_i\right)\right)\right] = \mathbb{E}_{z|q_i}\left[\Lambda\left(\gamma z - \alpha_i - \gamma \log r_i\right)\right]$, the effect of vehicle quality on the baseline replacement probability is threefold because (i) the distribution of z, (ii) α_i, and (iii) r_i all depend on q_i. Differentiating with respect to vehicle quality q and defining $h(z, q) = \Lambda\left(x(z, q)\right)$ for brevity yields

$$
\begin{aligned}
\frac{\partial P_i}{\partial q} &= \frac{\partial}{\partial q}\mathbb{E}_{z|q}\left[h(z,q)\right] \\
&= \frac{\partial}{\partial q}\int h(z,q)\; f(z|q)dz \\
&= \int \frac{\partial}{\partial q}\left[h(z,q)\; f(z|q)\right]dz \\
&= \int \frac{\partial h(z,q)}{\partial q}\; f(z|q)dz \;+\; \int \frac{\partial f(z|q)}{\partial q}\; h(z,q)dz \\
&= \int \frac{\partial h(z,q)}{\partial q}\; f(z|q)dz \;+\; \int h(z,q)\; \frac{\partial \log f(z|q)}{\partial q}\; f(z|q)\; dz \\
&= \mathbb{E}_{z|q}\left[\partial_q h(z,q)\right] + \mathbb{E}_{z|q}\left[h(z,q)s(z|q)\right]
\end{aligned}
$$

I apply the Leibniz rule in the second line and use the definition of the score $s(z|q) = \partial_q \log f(z|q)$ of a conditional distribution in the last line.

The first term in the last line is the expectation of $\partial_q h(z, q)$ conditional on q. This is

$$
\begin{aligned}
\partial_q\; h(z,q) &= \partial_q\; \Lambda(x(z,q)) \\
&= \Lambda'(x)\;\cdot\;\partial_q\; x(z,q) \\
&= \Lambda'(x)\;\cdot\;\left(-\alpha'(q) - \gamma\frac{r'(q)}{r(q)}\right)
\end{aligned}
$$

$\Lambda'(x) \geq 0$ holds true by the nature of a CDF. We further have $\alpha'(q) < 0$ because $\alpha(q) = a - b(q)$ and $b'(q) > 0$ (higher quality raises scrap value), and we have $r'(q) > 0$ because of the empirical observation displayed in Figure A14 that higher vehicle quality is associated with a residual-value-to-fee ratio r. A priori, the sign of $\mathbb{E}_{z|q}\left[\partial_q h(z,q)\right]$ is unclear.

The second term is $\mathbb{E}_{z|q}\left[h(z,q)s(z|q)\right] = \mathbb{E}_{z|q}\left[\Lambda\left(x(z,q)\right)s(z|q)\right]$. It captures how the entire distribution of repair cost shocks z moves with q. A marginal increase in q shifts z to the left in the sense of FOSD. In expectation, the product $\Lambda\left(x(z,q)\right)\cdot s(z|q)$ is negative under a leftward (FOSD) shift of z for larger q. Accordingly, we have

$$\frac{\partial P_i}{\partial q} = \underbrace{\mathbb{E}_{z|q}\left[\partial_q h(z,q)\right]}_{\gtreqless 0} + \underbrace{\mathbb{E}_{z|q}\left[h(z,q)s(z|q)\right]}_{<0}$$

A priori, the effect of vehicle quality on the replacement probability is unclear. Higher vehicle quality reduces the expected repair cost shock and increases the residual-value-to-fee ratio, both of which make vehicle replacement less likely, but raises scrap value and thereby makes vehicle replacement more likely. Empirically, I find consistent evidence for $\frac{\partial P_i}{\partial q} < 0$. For instance, Figure 3b clearly shows that treatment group car models, which are on average lighter and of lower quality, have higher baseline replacement probabilities than the heavier and higher-quality car models from the control group even before the soot tax was introduced.

To rationalize this, recall that the sign of $\mathbb{E}_{z|q}\left[\partial_q h(z,q)\right]$ depends on the relative size of the "scrap value channel" and the "residual value channel" (how vehicle quality affects both scrappage as well as residual value, respectively). It is plausible that the residual value channel dominates the scrap value channel. High-quality cars are often of a reputable brand, carrying value beyond their pure vehicle parts. This makes the residual value of the heavy, high-quality car h larger than that of the light, low-quality car l because of the higher value vehicle parts **and** because of its brand. The scrap value, in contrast, is only determined by the value of the actual vehicle parts regardless of brand. Once it is disassembled and scrapped, the brand premium ceases to exist and the vehicle's scrap value will only depend on the raw materials.

Accordingly, the difference in scrap value between cars l and h is smaller than the difference in residual value while the car is still in one piece. Hence, $|r'(q)| > |\alpha'(q)|$, which in turn makes it relatively likely that $\mathbb{E}_{z|q}\left[\partial_q h(z,q)\right] < 0$. This in combination with $\mathbb{E}_{z|q}\left[h(z,q)s(z|q)\right] < 0$ (the higher expected repair cost shock for car l relative to car h) rationalizes the empirical finding that $\frac{\partial P_i}{\partial q} < 0$, from which it follows that $P_l(q_l) > P_h(q_h)$.

Now, I can finally show that soot tax has a larger effect on light cars l than on heavy cars h. The soot tax τ exogenously decreases r, keeping everything else constant. As shown before, the approximate change in the replacement probability due to the soot tax τ is

$$\Delta P_i(q_i) \approx \gamma \cdot \log(1+\tau) \cdot M(z, q_i)$$

How does this change with q? It holds true for the derivative of ΔP_i to q that

$$\begin{aligned}
\frac{\partial \Delta P_i}{\partial q} &\propto \frac{\partial M(z,q)}{\partial q} \\
&\propto \frac{\partial}{\partial q} \mathbb{E}_{z|q} \left[\Lambda\left(x(z,q)\right)\left(1-\Lambda\left(x(z,q)\right)\right)\right] \\
&\propto \mathbb{E}_{z|q}\left[1-\Lambda(x(z,q))\right]\left[\frac{\partial}{\partial q}\mathbb{E}_{z|q}\left[\Lambda(x(z,q))\right]\right] + \mathbb{E}_{z|q}\left[\Lambda(x(z,q))\right]\left[\frac{\partial}{\partial q}\mathbb{E}_{z|q}\left[1-\Lambda(x(z,q))\right]\right] \\
&\propto \left[\frac{\partial}{\partial q}\mathbb{E}_{z|q}\left[\Lambda(x(z,q))\right]\right] \cdot \left[\mathbb{E}_{z|q}\left[1-\Lambda(x(z,q))\right] - \mathbb{E}_{z|q}\left[\Lambda(x(z,q))\right]\right]
\end{aligned}$$

We know from before that $\frac{\partial P_i}{\partial q} = \frac{\partial}{\partial q}\mathbb{E}_{z|q}\left[h(z,q)\right] = \frac{\partial}{\partial q}\mathbb{E}_{z|q}\left[\Lambda(x(z,q))\right] < 0$. In addition to that, actual outflow rates of old diesel cars are in the range of 10 to 30% per year. I interpret outflow rates as realizations of replacement probability to narrow the relevant range of values of the replacement probability $P_i(q_i) = \mathbb{E}_{z|q_i}\left[\Lambda\left(x_i\left(z,q_i\right)\right)\right]$ to an interval $P_i \in (b_{min},\ b_{max})$ bounded by some $b_{min} > 0$ and $b_{max} < 0.5$. This restriction provides us with the sign of $\frac{\partial \Delta P_i}{\partial q}$:

$$\frac{\partial \Delta P_i}{\partial q} \propto \underbrace{\left[\frac{\partial}{\partial q}\mathbb{E}_{z|q}\left[\Lambda(x(z,q))\right]\right]}_{<\,0} \cdot \left[\underbrace{\mathbb{E}_{z|q}\left[1-\Lambda(x(z,q))\right]}_{>\,0.5} - \underbrace{\mathbb{E}_{z|q}\left[\Lambda(x(z,q))\right]}_{<\,0.5}\right] < \ 0$$

Hence, the soot tax induces a larger change in the replacement probability when vehicle quality is lower. This is exactly what I find in my heterogeneity analysis in Section 5.1. Figure 5 shows that the effect of the soot tax is larger for lighter car models, which are of lower quality on average. Similarly, when differentiating by average household disposable income of vehicle owner, the same patterns arises because owners of light (heavy) car models have low (high) disposable incomes on average.

D Endogenous sorting

One threat to identification is endogenous sorting into treatment and control group after the soot tax was announced but before it was actually implemented. This could happen if, for instance, the announcement of the soot tax alone changes the residual value of eligible and ineligible cars in the second-hand market. Owners of polluting diesel vehicles may replace them with similarly aged but clean diesel vehicles to avoid the soot tax. This could inflate the lifetime of control group vehicles. In the absence of the soot tax, previous owners might have scrapped these vehicles. With the tax in place, however, they might choose to sell them instead, since owners of polluting diesels may now have a higher willingness to pay for control-group cars. At the same time, although presumably less likely, the soot tax might also induce certain individuals to acquire a polluting diesel vehicle on the second-hand market and bear the soot tax if these individuals have strong enough preferences for certain diesel models.

One way to test for such endogenous sorting into treatment and control group is to predict characteristics of vehicle owners using the treatment dummy. I do so with a DiD regression on the vehicle level. The soot tax was announced in July 2016, so registration data from January 1, 2016 should accurately represent the vehicle stock before any sorting could have occurred. The soot tax was initially planned for January 1, 2019, but the government announced its postponement only a few months beforehand, so any anticipatory sorting should have occurred before that date. I therefore assess sorting by comparing the vehicle stock on January 1, 2016 to that on January 1, 2019.

I draw on the 342 car types that I use to derive my main results in Section 5.1. I restrict the sample to vehicles observed in every year from 2016 to 2019. This yields 26,964 vehicles belonging to 342 car types from the 2005–2008 diesel sample. I run the following regression equation, using 2016 as the pre- and 2019 as the post-sorting period:

$$char_{it} = \gamma_0 + \gamma_1 D_i + \gamma_2 PostSorting_t + \gamma_3 D_i PostSorting_t + u_{it} \tag{10}$$

My outcome is household/vehicle owner characteristic *char* of vehicle i in year t. I regress this in a straightforward DiD design on D_i, a dummy indicating whether vehicle i is in the treatment group (will be subject to the soot tax), a dummy indicating post-sorting year 2019, and their interaction. The coefficient of interest, γ_3, should be statistically equal to zero in the absence of sorting.

Table A20 summarizes the results, column by column, for all outcomes of interest. The first column looks at the outcome log owner age. We observe a few things. First, the significant and negative coefficient on the treatment group indicator D shows what we have already seen in Table 2: owners of treated diesel vehicles are on average younger than those of control diesel vehicle. Second, the insignificant coefficient on $PostSorting$ indicates that these diesel vehicles tend to be sold to younger owners. If all owners had hold onto their vehicles, the coefficient would be significant and positive since all owners would have become three years older. Third, and most importantly, the coefficient on the interaction term $D \times PostSorting$ is insignificant. The owner age structure between treatment and control group vehicles has not significantly changed between 2016 and 2019, so there is no evidence for endogenous sorting.

Column (2) considers the outcome log household disposable income. The coefficient on D highlights that owners of treatment group vehicles generally have lower incomes than those of control group vehicles. We look at nominal incomes and these increase over time, as indicated by the positive and significant coefficient on $PostSorting$. Most importantly, the coefficient on the interaction term $D \times PostSorting$ is insignificant, so we do not find evidence for endogenous sorting with regard to income.

In contrast to the first two columns, the coefficients on $D \times PostSorting$ turn out statistically significant in columns (3) to (6). By 2019, owners of treatment group vehicles are more likely to be married, be male, live in more densely populated areas, and are less likely to be homeowners. At the same time, these results can be plausibly rationalized with other factors than sorting based on future soot tax treatment.

Column (3) indicates that owners of treated diesel vehicles are more unlikely to be married, but that this gap significantly shrinks over time. The group of treated diesel car owners, however, is inherently more likely to get married over the three-year window. They are on average younger than owners of control group cars and more likely to be of an age when people generally marry (for the first time). The same phenomenon may also explain the significant coefficient on the interaction term in column (4). Insurers in the Netherlands often offer discounts called "family car arrangements" when multiple vehicles in the same household are insured under the same policyholder. Because owners of treatment group vehicles are more likely to get married during the observation period, some of these vehicles may have changed ownership on paper to benefit from these insurance discounts, which could in turn help explain the shift in the share of male owners.

Column (5) considers the outcome homeownership. The estimated coefficient of $D \times PostSorting$ takes on the opposite sign of what sorting would imply. Homeownership rates are lower in the treatment group already in 2016. If owners of treatment group diesel cars switch to control group vehicles, this would reduce differences in homeownership rates and generate a positive coefficient for TG $\times$ Post, the opposite of what we find.

It remains the outcome population density in column (6). In 2016, there are no differences in population density between treatment and control group diesel cars, however, by 2019, the former tend to relocate to more densely populated areas. This pattern is consistent with sorting based on fuel expenses. Diesel prices in the Netherlands rose sharply by almost 30% between January 2016 and January 2018, before declining again by about 10% until January 2019. Higher diesel prices may have heightened consumers' sensitivity to fuel efficiency, encouraging households in denser areas to opt for smaller and more fuel-efficient diesel models. Since treatment group diesel cars tend to be lighter and more fuel-efficient than their control group counterparts, sorting based on changing fuel prices, not the soot tax, rationalizes well the positive and significant coefficient on $D \times PostSorting$ in column (6).

In general, household disposable income and owner age should have the largest impact on vehicle replacement decisions. Income is a direct determinant of the frequency and price sensitivity of households' vehicle turnover. Owner age presumably matters because of lifecycle preferences. Whereas young people tend to acquire used cars and plausibly delay vehicle replacement because of a higher risk tolerance, middle-aged people may have grown more risk-averse, therefore preferring newer cars and more frequent replacements. Given that the estimated coefficients on the interaction term in columns (1) and (2) are insignificant, and given that the significant estimates in columns (3) to (6) are rationalizable with other circumstances, I view the results as providing little support for the notion of endogenous sorting.

Table A20: Results for endogenous sorting

	(1) ln(age)	(2) ln(inc)	(3) married	(4) male	(5) homeowner	(6) ln(pop dens)
Intercept	3.893***	10.99***	0.677***	0.835***	0.889***	6.459***
	(0.002)	(0.005)	(0.004)	(0.003)	(0.003)	(0.009)
D	-0.107***	-0.229***	-0.097***	-0.116***	-0.124***	0.005
	(0.004)	(0.007)	(0.006)	(0.005)	(0.005)	(0.014)
$PostSorting$	0.002	0.029***	-0.041***	-0.006	-0.039***	0.004
	(0.003)	(0.006)	(0.005)	(0.005)	(0.004)	(0.013)
$D \times PostSorting$	0.003	0.004	0.002**	0.013*	-0.047***	0.086***
	(0.005)	(0.010)	(0.008)	(0.007)	(0.007)	(0.020)
No of obs.	53,928	53,928	53,928	53,928	53,928	53,928
Adj. R^2	0.031	0.036	0.009	0.017	0.040	0.001

E Administrative Data

This paper utilizes administrative microdata provided by Statistics Netherlands (CBS), which covers the entire population of Dutch residents. The data comprise a range of individual- and household-level registers accessible through a secure remote access environment.[22] All datasets can be linked at the individual level via encrypted personal identifiers. My analysis draws on several key sources. First, I utilize annual data on vehicle ownership in the Netherlands. This record allows me to identify which individual person owns which vehicle at what point in time.

The vehicle data is matched to household income data of the household the individuals are a part of, individuals' gender and age, education level, marital status, and place of residence. All these mergers occur based on the variable RINPERSOON, which is an encrypted identifier for an individual across all used datasets in the CBS microdata environment. The following section provides further detail on the primary datasets used in the analysis. Documentation for each dataset is accessible via the embedded hyperlinks in their names. Table A21 displays all the data sources.

RDWNPACTTAB: This dataset contains information on every individual vehicle of the entire fleet of registered vehicles in the Netherlands, on January 1 of the respective year. The data comes directly from RDW, the Dutch motor vehicle agency responsible for vehicle registration and MVT collection. This data registers a vehicle's owner, the vehicle's fuel type, year of first registration, number of cylinders, empty mass, color, and vehicle type.

[22] See the CBS website for a detailed catalogue of all available administrative datasets (in Dutch): https://www.cbs.nl/nl-nl/onze-diensten/maatwerk-en-microdata/microdata-zelf-onderzoek-doen/catalogus-microdata

INHATAB: This data set contains information on a household's income, including taxable, gross, and disposable income, as well as on household composition (one-person household, two-person household, multi-person household, with or without children that are minors or adults) and a household head id to identify which individuals live in the same household.

GBAPERSOONTAB: This dataset contains basic demographic information about individuals, including their gender and their year of birth.

HOOGSTEOPLTAB: This dataset contains information on the highest finalized education of an individual.

GABHUISHOUDENSBUS: This dataset contains information on an individual's marital status by indicating the household type, partitioned among several dimensions, one of which is marital status.

GBAADRESOBJECTBUS: This dataset contains information on individual addresses. It contains identifiers for the actual building someone lives in, the quarter, the municipality, and the region.

Table A21: Administrative datasets

Key Variables Included	Name of Datasets	Years Used
Vehicle empty mass	RDWNPACTTAB	2013-2023
Vehicle first year of registration		
Vehicle number of cylinders		
Vehicle fuel type		
Vehicle encrypted plate number		
Vehicle type		
Vehicle owner type		
Vehicle number of cylinders		
Household disposable income	INHATAB	2013-2023
Household composition		
Household head id		
Vehicle owner's gender	GBAPERSOONTAB	2013-2023
Vehicle owner's year of birth		
Vehicle owner's education level	HOOGSTEOPLTAB	2013-2023
Vehicle owner's marital status	GBAHUISHOUDENSBUS	2013-2023
Vehicle owner's place of residence	GBAADRESOBJECTBUS	2013-2023