# Portfolio Complementarities and Electric Vehicle Adoption* 

Bjørn Gjerde Johansen ${ }^{\dagger} \quad$ Anders Munk-Nielsen ${ }^{\ddagger}$

October 2022


#### Abstract

We formulate and estimate a model of car portfolio choice and driving, which allows two cars to be substitutes or complements, both in ownership and driving. We estimate the model using Norwegian register data, which features rich policy variation. We find significant portfolio synergies between an EV and a combustion vehicle (CV), and for $16 \%$ of households the synergies are so strong that the two become strict Hicksian complements. This implies that EVs tend to come as additions to existing CVs rather than as replacements. Failure to account for portfolio synergies leads to an overly optimistic assessment of EV adoption incentives.


Keywords: Car choice, electric vehicles, driving, emissions, bundle choice, complementarity, fuel tax.

[^0]
## 1 Introduction

Transportation is one of the key sectors in reaching the climate goals set by most developed countries today. Towards this end, many countries have enacted EV adoption policies of some form or another, with the EU effectively banning sales of internal combustion engine vehicles (CVs) by 2035. The motivation behind this is the desire to shift driving towards the lower pollution of EVs. However, the success of such a policy rests on the assumption that an EV is a substitute for a CV. This paper argues that there is considerable portfolio synergy between an EV and a CV. This means that to a certain extent, EVs do not replace existing CVs but instead come as additions.

If car types are not simply substitutes, environmental policy needs to take it into account. This counts doubly so if environmental policies do not directly target their key objective - reducing emissions - but instead rely on an intermediate step, such as stimulating EV adoption. We show that failing to account for portfolio synergy leads to an overly optimistic estimate of the abatement cost of EV adoption policies.

To account for this synergy, we develop and estimate a discrete-continuous model of car choice where portfolio complementarities can arise explicitly from driving. We find that bundle synergies are particularly strong between an EV and a CV. This means that households respond to EV incentives by adding an EV to their car portfolio, rather than replacing their CV. Crucially, we do not find evidence that the driving in the newly added EV significantly reduces driving in the existing CV, resulting in a near doubling of the total driving of the household. Thus, EV incentives fail to reduce fuel consumption but instead increase total driving and through that other externalities like congestion, accidents, and local pollution.

We find evidence of different forms of car portfolio synergies. One form that is common to all car types is driving specialization. Two-car households prefer to assign more driving to the car that is younger and has a more powerful engine. Intuitively, diversification of their car portfolio gives households access to specialized cars for different trips: a large car for family trips and a smaller car for single-person shopping.

However, this effect is not sufficient to explain the synergy between an EV and a CV, which is stronger than what can be accounted for by driving. Intuitively, one source of the complementarity may be the widely cited phenomenon of "range anxiety": that households are overly focused on infrequent but very long trips, exceeding the EV's maximum range. For such households, also having access to a traditional vehicle for those trips would, perhaps, alleviate the range concern. Even so, the net environmental effect is still ambiguous as it depends on driving by one- vs.
two-car households. If total driving does not increase much (i.e. strong satiation in driving demand) and most driving is shifted strongly towards the EV, then a policy that pushes a household with a single CV to choose an (EV, CV)-portfolio can still result in lower emissions.

Our model allows total driving by two car households to depend freely on the number of cars. Empirically, we estimate an economically small degree of satiation. So while we estimate strong preferences for shifting, it tends to shift driving towards the youngest and heaviest car and to a lesser extent towards the car that is cheapest to drive. The result is a negative bottomline environmental effect of pushing consumers from CV to (EV, CV)-portfolios. The validity of our driving model is both consistent with raw descriptives as well as external survey evidence for Norway. ${ }^{1}$

We are the first to point out the interplay between EVs and CVs in two-car portfolios for both the extensive (ownership) and intensive (usage) margins. Furthermore, the discussion above highlights why it is crucial to model both margins simultaneously to assess the cost effectiveness of EV adoption incentives.

To complement our structural model, we also present evidence based on raw descriptives and using different notions of complementarity. First off, Figure 1 shows new CV and EV sales by whether the household has one or two cars. Clearly, EVs are vastly overrepresented in two-car portfolios. Next, we consider the complementarity criteria proposed by Manzini et al. (2019), which do not rely on price-variation and are model-free. We evaluate these criteria for different partitions of Norwegian households and confirm that complementarity does not hold for all households but for some subsamples. We show that comparing choice frequencies among one- and two-car portfolios, the strongest overrepresentation in the data occurs for the same portfolio identified as having the strongest portfolio synergy by our model: a small EV and a large CV.

In our structural model, households make a discrete car portfolio choice and a continuous driving choice for each car they own. Households can choose between 20 car types, which allows for 231 different combinations of car portfolios. The model extends the discrete bundle complementarity model of Gentzkow (2007) with a continuous driving choice in the spirit of Dubin and McFadden (1984). Our methodological contribution lies in tying the bundle synergy parameters to driving in the two cars in a way that allows flexible preferences for driving by two-car households. This way, portfolio synergies are allowed to explicitly depend on driving.

To estimate the model, we use detailed register data for the full Norwegian

[^1]Figure 1: Car Ownership Over Time


Note: The right panel shows that in 2017, close to 4.5 percent of all households owned an electric vehicle - about 75 percent of these were two-car households. The same year, a bit more than 60 percent of the households were car-owning households without an EV in their portfolio - about 40 percent of these were two-car households. The figure excludes households that do not own a car (around 30 percent).
population covering 2005-2017. We observe household demographics, including data on each spouse's home and work locations, as well as car ownership status for all Norwegians, and the driving at vehicle safety checks when the vehicle is aged 4, 6, 8, etc. Norway is of special interest for EV adoption policy questions since it has achieved the highest penetration in the world.

Identification of the model relies on a combination of parametric assumptions and exogenous variation in three key monetary variables: ${ }^{2}$ fuel prices, road tolls, and registration taxes. First, fuel prices in Norway are largely determined by world oil market prices, which are unrelated to local Norwegian policy initiatives. ${ }^{3}$

Second, road tolls in Norway are substantial and EVs are exempt. We show that toll exposure is an extremely powerful predictor of EV adoption, nearly doubling the purchase probability. This is the same difference in ownership propensity as that between the richest and poorest decile of households. There is rich variation in toll exposure both cross-sectionally but also over time as more cities have enacted tolls to combat congestion and local pollution. We embed this directly into our model.

Third, registration taxes make up $25-50 \%$ of the purchase price and there were several major reforms of the formula during our period. The tax is attribute-based and went from targeting weight to emissions. We show that while car attributes

[^2]account for $96 \%$ of the variation in the pre-tax car price, they only account for $70 \%$ of the variation in the tax component. Thus, the majority of our new car price (MSRP) variation was induced by policy reforms. ${ }^{4}$

Our estimated model is able to fit key moments of the data and produces elasticities that are consistent with the literature. One key take-away is that the small EV and a large new diesel car are Hicksian complements to $16 \%$ of households. However, complements only arise for subsets of the population and all cars are strict substitutes in aggregate. Nevertheless, there are still strong synergies between an EV and a CV and we will demonstrate that this synergy is crucial for our key policy results.

Our main analysis is a counterfactual assessment of the cost of $\mathrm{CO}_{2}$-reductions from three policies: a higher fuel tax, a lower EV tax, or a higher CV tax. We find that fuel taxes are more cost-effective than registration taxes, at just 897 NOK per tonne of $\mathrm{CO}_{2}$. For the CV tax, the cost per tonne is 20 times higher. This is not surprising as the fuel tax targets the environmental externality directly and allows consumers to respond on both the extensive and intensive margins.

For purchase taxes, we find that lower taxes on EVs costs twice as much per tonne as raising taxes on CVs. This cost contains consumer welfare, tax revenue and other driving externalities but ignores the distributional concerns. EVs in Norway are exempt from VAT and we find that such a policy is strongly regressive, whereas the fuel tax is progressive.

Finally, we demonstrate that portfolio synergies explain why the EV adoption incentives are so much more costly than CV taxes. To show this, we estimate a model without portfolio effects and show that when the data is interpreted through such a lens, the two policies have equal cost per tonne of $\mathrm{CO}_{2}$. The intuition is that without range anxiety, a single EV is a good substitute for a single CV. That way, the EV adoption policy succeeds in moving driving from CVs to EVs. In our full model, however, the primary effect is that households go from a single CV to the mixed (EV, CV)-bundle, resulting in a dramatic increase in driving.

### 1.1 Existing Literature and Our Contributions

Our contributions are both methodological and substantive. Methodologically, we extend the discrete-continuous choice models with bundle choices. Discrete-continuous models date back to Dubin and McFadden (1984) and have been used successfully

[^3]to analyze the car market specifically, although for one-car models only. ${ }^{5}$ Discrete choice models with bundles have been used going back to Manski and Sherman (1980), but the framework proposed by Gentzkow (2007) has proven powerful for analyzing the notion of complementarity, sparking a growing literature. ${ }^{6}$ Our contribution is to provide a parametric model for how the bundle parameter depends on driving in a flexible way that also permits an tractable solution to the two-car driving problem. Perhaps the closest model methodologically is Thomassen et al. (2017) where the discrete choice is supermarket location and the consumption quantities are continuous. ${ }^{7}$

Another seemingly innocuous methodological departure from previous work on complementarity in discrete choice is that we allow consumers to buy two of the same car, something that frequently happens in our empirical setting. We refer to this as a specialized bundle, and it has been assumed away in all prior work where discrete demand has been taken as binary (buy or don't). We prove formally that incorrectly assuming binary demand will result in a bias towards finding Hicksian complements. And in fact, when our data is viewed through the binary lens, EVs and CVs appear to be strict complements at the market level.

We are not the first to consider portfolio effects with regard to cars purchases (Wakamori, 2015; Archsmith et al., 2020) or driving (De Borger et al., 2016). However, we are the first to do so jointly and the first to point out the synergy between an EV and a CV as different from those between two CVs. Furthermore, we use recent model-free criteria for establishing bundle complementarities by Manzini et al. (2019) which rely solely on portfolio market shares. It is well-known that econometric analysis of portfolio choice is complicated if one only observes market share data (Iaria and Wang, 2021) so we also evaluate criteria on subsamples of households. This illuminates the role of heterogeneity in shaping the portfolio synergies we document.

Substantively, we contribute to a large literature on environmental car policy evaluation. ${ }^{8}$ Our key contribution is that we highlight a novel feature of the demand for EVs: a demand synergy between an EV and a CV. Our estimates are consistent with the popular notion of "range anxiety."

This makes us related to recent empirical work on EV adoption. One strand emphasizes the dynamics of the charging network and how charging subsidies interact

[^4]with more direct EV policies (Springel, 2021; Li, 2019). ${ }^{9}$ A second strand of EV literature focuses the difference between the average and marginal $\mathrm{CO}_{2}$ emissions of electricity production (see e.g. Archsmith et al., 2015a; Holland et al., 2016). ${ }^{10} \mathrm{~A}$ third strand has sought to identify what an EV "replaces". Xing et al. (2021); Johansen (2020) show that EV purchases tend to divert households away from hybrids and CVs that on average are more environmentally friendly.

In terms of our policy conclusions, we find fuel taxes to be superior to car taxes working through the acquisition. This is consistent with prior work comparing it to attribute-based taxes (Grigolon et al., 2018) or CAFE standards (Jacobsen, 2013). ${ }^{11}$ This fact is not surprising, given that fuel taxes is the only true Pigouvian policy instrument; i.e., directly proportional to the environmental externality (Ito and Sallee, 2018).

## 2 Data and Institutional Setting

### 2.1 Car Policies and Reforms

This section presents the institutional setting regarding car taxation and policies in Norway. We will emphasize the cross-sectional and time-series variation generated by the policies and how they have been updated and reformed during our sample period. Apart from providing an understanding of the institutional setting, this variation is crucial for sources of exogenous variation in prices.

Norway taxes car ownership and driving heavily - around 60 billion kroner (almost $2 \%$ of GDP) are paid in car related taxes each year (Fridstrøm, 2019). Taxes fall into four categories: Purchase taxes, fuel taxes, annual ownership taxes, and tolls. In 2017, these four taxes summed up 15, 20, 10, and 10 billion NOK in revenue respectively, so none are negligible. EVs are exempt from purchase taxes and tolls and pay much lower annual taxes. In the following, we will briefly describe each of these policies in that order, and lastly cover EV-specific policies.

The purchase tax is paid for new car sales but not for resale of used cars. It consists of a flat $25 \%$ VAT and a registration tax component that depends on car

[^5]attributes. The registration tax varies between $0 \%$ and $100 \%$ of the producer price. ${ }^{12}$ There is substantial time-series variation in the formula, which has shifted gradually from a focus on engine power to $\mathrm{CO}_{2}$ emissions so that by 2017, the most efficient gasoline and diesel cars pay virtually zero registration fee. EVs are exempt from both VAT and registration taxes and have been throughout our sample period. Additional details are in Appendix A.4.1.

Fuel taxes on diesel and gasoline are comprised of a fixed and proportional term. Electricity is taxed in a similar fashion and there are no special provisions for EV use compared to any other private use. There have been no notable policy changes so all time-series variation comes from input prices, which are largely due to world market conditions in oil and electricity. Details are in Appendix A.4.2.

The annual tax is paid by all car owners, regardless of whether the car is new or used. The tax primarily differs between EV and non-EV, averaging 455 and 3,000 NOK respectively in 2017. Within these fuel segments and across time, there is virtually no variation in the annual tax except for minor updates that follow the CPI.

Toll payments constitute a growing share of car related tax payments, having increased from 3 billion NOK in 2005 to 10 billion in 2017. This increase is largely due to a gradual increase in the use of toll rings ("cordons") around major urban areas. A second type of toll payments is for bridges or tunnels, which are non-negligible in Norway due to the geography. In total, the share of households encountering a toll on their way to work went from $15 \%$ to $37 \%$ over our sample period. Apart from increased prevalence of toll rings, the rates per trip have also gone up. A key feature of the toll policy is the EV exemption. Figure B. 5 shows maps of toll exposure and EV adoption. Apart from the time-series variation, toll exposure has considerable cross-sectional variation depending on home and work location, which we observe from linked employer-employee data.

The final set of car policies specifically targeting EV adoption. These include reduced ferry rates, access to restricted bus lanes, subsidies to charging stations (starting in 2009), and access to free and restricted parking spots. Fridstrøm (2019) argues that apart from the purchase and toll tax exemptions, the access to bus lanes has been highly impactful. We will proxy for bus lane and parking access with a dummy for working in the city in our model. However, we do not include charging station network in our empirical model. ${ }^{13}$

[^6]
### 2.2 Data Sources and Constructed Variables

Our dataset consists of register data on all households in Norway over the years 20052017. We have four primary data sources, relating to demographics, car ownership, driving, and prices. We will describe each source in turn before turning to three constructed variables: work distance, tolls, and aggregated car types.

First, the annual demographic registers include information such as age, family composition, income, residential and work location. Second, car ownership spells are recorded in the central motor vehicle register, with a start and end date as well as a person identifier. ${ }^{14}$ The central motor vehicle register also contains technical information about the cars, such as make, model, fuel type, engine effect, etc. Third, driving data comes from mandatory safety inspections, occurring when the car is four years and biannually thereafter. At this inspection, the odometer is recorded from which we can compute the average daily driving in km . There is no requirement for a safety test when a car changes owners, so the ownership and driving periods are asynchronous.

Finally, we utilize several auxiliary datasets, including monthly data on fuel prices/taxes and quarterly data on electricity prices/taxes from Statistics Norway; annual merchant suggested retail prices (MSRPs) of new cars; annual rates of carrelated taxes from the Norwegian tax authorities; expected maintenance costs by car age based on data from US mechanics; and local external costs of driving from the Norwegian Institute of Transport Economics. For used-car prices, we employ data from the Norwegian tax authorities which is used to compute registration taxes on imported used cars. See Appendix A for more information.

We now turn to the construction of two key variables: the work distance and toll payments associated with the commute. We observe the residential and work locations of households at the level of "basic statistical units", henceforth neighborhoods. There are about 14,000 neighborhoods in Norway, with an average population of less than 200 households. We calculate the shortest paths (in travel time) between all centroids using the Norwegian road network. ${ }^{15}$ We set the commute distance to zero for unemployed and we drop approximately $10 \%$ of households where the location of the firm is missing (details for these are in Appendix A.1.3) Work distance is averaged between spouses in couples, and it is strictly positive for about $60 \%$ of

[^7]households. The second constructed variable is toll exposure. Using the shortest path from the home to work neighborhood centroids, we add up the tolls for each gate along the route. This is done separately for each year, so toll rates and exposure vary both cross-sectionally and over time.

Finally, we will describe the discretization of cars into a smaller set of types that will form the choiceset in our structural model. Since the model is computationally demanding, we partition car models into 20 types based on the values of car characteristics: we partition by propulsion system (diesel/gasoline/electric), weight (small/large) ${ }^{16}$ and four car age categories. ${ }^{17}$ For each car type in each year, we compute the population-weighted average characteristics of all cars falling in that category. Thus, characteristics will change from year to year. Not all EV types are available in all years. The first small EV became available in 2011, and the large EV in 2013 when Tesla Model S arrived. Furthermore, since there are no policies targeting hybrids specifically, we simply include those in their corresponding gasoline or diesel segment. ${ }^{18}$ Additional details about variables and the discretization into types as well as characteristics for the 2017 choiceset are presented in Appendix A. 2 .

### 2.3 Final Estimation Samples

Our model will make predictions on two margins, so we will construct two datasets: one for car ownership, and one for driving driving.

In the car ownership dataset the unit of observation is a household-year. In each year, we take the stock of cars owned by the household ultimo as the household's ownership decision. Since our model can handle 0 , 1 , or 2 cars, we drop cars beyond the first two, keeping only the youngest cars. We also omit cars older than 25 years, and cars registered for other purposes than personal use such as taxis. In terms of households, we drop observations with missing values, which happens most frequently for work distance ( $10 \%$ ). Our final sample contains $89 \%$ of the raw households (approximately 2 million annually). See Appendix A.1.1 for more details. A summary of car ownership and demographic variables for 2017 can be seen in

[^8]Table 1.
In the driving dataset, the unit of observation is a household in given period. Since our model will predict a different level of driving for a one- and a two-car household, we partition driving periods each time one of the following occurs: the car changes owner, the owning household changes its portfolio (e.g. buying a second car), or the driving period ends. For each sub-period, we compute the weighted average of the household yearly demographics using as weights the fraction of the driving period that covers a given year. Similarly, we compute the fuel price as a weighted average of the monthly (country-wide) averages over the driving period. Our model for driving is not computationally constrained and we are not forced to discretize car types so for the driving data, we use the actual micro-level car attributes to preserve as much variation in the data as possible.

To construct our final driving period sample, we first drop periods beginning before 1 Jan 2005 or where the car has missing attributes (mostly for very old cars). This admits 15.9 m driving sub-periods. From these, we drop periods violating one of the following criteria: driving is between 0 and $200 \mathrm{~km} /$ day, safety inspections are not unnaturally early or late, no more than $70 \%$ of the total driving period is dropped for other reasons (e.g. owner unobserved for part of the period). This leaves us with 13.4 m periods in the final sample. See Appendix A. 3 for details.

### 2.4 Descriptives

We now present descriptive evidence on car ownership and driving in Norway. To begin with, Figure 1 shows the overall change in new car sales. EV sales began in 2011 with the "small EV, new" (the modal car being a Nissan Leaf), followed in 2013 by the first large EV (the Tesla Model S). By 2017, EV sales made up $30 \%$ of all new car sales, and $4.5 \%$ of Norwegian households owned at least one EV.

Figure 1 also highlights that households with EVs are more likely to have two than one car. Thus, it is relevant to ask whether our demographics are strong predictors of the number of cars owned by a household. It turns out that income and work distance are extremely powerful predictors. Broadly, the effects of the two are similar: higher income or work distance is associated with higher car ownership and driving. For instance, comparing the poorest to the richest decile of households, the share without a car goes from over $80 \%$ to under $20 \%$, and average driving in one-car households from 25 to $40 \mathrm{~km} /$ day (Figures B. 2 and B.3). Our structural model will be able to fit these stark differences across household segments (Figure $3)$.

In terms of fuel types, higher income or work distance is associated with a shift

Table 1: Summary Statistics, Ownership Dataset, 2005 and 2017

| Portfolio type | 2005 |  | 2017 |  |
| :---: | :---: | :---: | :---: | :---: |
|  | Mean | Std dev | Mean | Std dev |
| Share no-car households | 0.304 | 0.460 | 0.304 | 0.460 |
| Share (CV) households | 0.446 | 0.497 | 0.403 | 0.491 |
| Share (CV,CV) households | 0.250 | 0.433 | 0.248 | 0.432 |
| Share (EV) households | 0 | - | 0.011 | 0.104 |
| Share (EV,EV) households | 0 | - | 0.002 | 0.045 |
| Share (EV,CV) households | 0 | - | 0.032 | 0.176 |
| Demographics | Mean | Std dev | Mean | Std dev |
| Household disposable income (100,000 NOK) | 4.822 | 6.312 | 5.106 | 6.862 |
| Age (household average) | 51.786 | 18.190 | 52.423 | 18.551 |
| Couple (dummy) | 0.556 | 0.497 | 0.546 | 0.498 |
| City (dummy for living in a major city) | 0.250 | 0.433 | 0.245 | 0.430 |
| Work distance (km, household average) | 8.492 | 16.520 | 8.336 | 16.340 |
| Work distance (excluding zeros) | 15.134 | 19.643 | 15.601 | 19.657 |
| Toll (NOK, one-way, sum of household members) | 2.085 | 9.038 | 8.441 | 21.780 |
| Toll (excluding zeros) | 24.713 | 20.278 | 42.629 | 30.631 |
| Observations (households) | 1,849,058 |  | 2,169,769 |  |

Notes: Summary statistics are for the "Ownership Dataset," where the unit of observation is a household-year. Only the 2005 and 2017 cross-sections are shown in this table. All monetary variables are measured in 2015 NOK. Disposable income is the sum of labor and capital income minus taxes plus transfers. Disposable income and tolls are summed over spouses, while age and work distance are averaged across spouses.
from gasoline towards diesel and EV. In terms of driving, diesel cars are driven approximately $10 \mathrm{~km} /$ day more than gasoline cars, even conditional on income or work distance decile. Unconditionally, the distribution of EV driving looks indistinguishable from gasoline car driving (Figure B.4), but conditionally they are somewhere in between (Figure B.3).

By far the strongest predictor of EV adoption, however, is toll exposure. Figure 2 shows for each decile of work distance the fraction of households that own (at least) one EV; panel A for one-car households and panel B for two-car households. Firstly, we see that EV ownership is more than three times as high for two-car households. But more importantly, toll exposure more than doubles the share that owns an EV. Furthermore, while EV adoption is virtually independent of work distance for households with no toll, those with toll exposure have nearly twice as high EV shares at higher work distances compared to lower.

Of course, income and work distance are correlated. Thus, we have estimated a set of linear regressions aiming to control for these and other variables simulta-

Table 2: Summary Statistics, Driving Dataset, all years

| Car data | Diesel |  | Gasoline |  | Electric |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Mean | Std dev | Mean | Std dev | Mean | Std dev |
| Driving ( km/day) | 41.532 | 19.325 | 31.303 | 17.801 | 37.115 | 16.801 |
| Length of driving period (years) | 2.271 | 0.669 | 2.128 | 0.478 | 3.395 | 0.750 |
| Length of sub-period (years) | 1.653 | 0.816 | 1.658 | 0.700 | 2.281 | 1.119 |
| Car age (years) | 6.862 | 3.252 | 9.296 | 3.826 | 2.436 | 1.313 |
| Car weight (tonnes) | 1.531 | 0.245 | 1.215 | 0.190 | 0.995 | 0.190 |
| Engine effect ( $100 \mathrm{~kW} \mathrm{)}$ | 0.903 | 0.225 | 0.803 | 0.227 | 0.833 | 0.530 |
| Demographics | Mean | Std dev | Mean | Std dev | Mean | Std dev |
| Household owns second car (dummy) | 0.629 | 0.483 | 0.522 | 0.500 | 0.811 | 0.391 |
| Household disposable income (100,000 NOK) | 7.116 | 9.270 | 6.103 | 8.845 | 9.779 | 13.339 |
| Age (household average) | 49.789 | 13.418 | 52.455 | 15.315 | 45.399 | 10.569 |
| Couple (dummy) | 0.820 | 0.370 | 0.721 | 0.435 | 0.893 | 0.291 |
| City (dummy for living in a major city) | 0.155 | 0.359 | 0.210 | 0.405 | 0.302 | 0.454 |
| Work distance (km, household average) | 14.659 | 20.858 | 11.512 | 18.554 | 18.353 | 18.604 |
| Work distance (excluding zeros) | 18.403 | 21.858 | 16.078 | 20.183 | 19.727 | 18.572 |
| Toll (NOK, one-way, sum of household members) | 9.843 | 22.161 | 7.765 | 18.572 | 30.758 | 36.656 |
| Toll (excluding zeros) | 33.433 | 29.652 | 28.789 | 25.953 | 44.319 | 36.538 |
| Driving periods | 2,553,075 |  | 3,522,844 |  | 12,308 |  |
| Driving sub-periods | 6,057,246 |  | 7,345,844 |  | 33,844 |  |
| Mid-year of driving period (average) | 2013 |  | 2011 |  | 2015 |  |

Notes: Summary statistics are for the "Driving Dataset," where the unit of observation is a sub-period, defined as a subset of the driving period where the car's owner and its car portfolio is unchanged. All years are pooled, and monetary variables are measured in 2015 NOK. Disposable income is the sum of labor and capital income minus taxes plus transfers. Disposable income and tolls are summed over spouses, while age and work distance are averaged across spouses.
neously (Table B.4). These reveal that urban residency and toll exposure tend to move households away from diesel cars and towards EVs. Thus, the prototypical diesel owner is rich and lives outside the city but commutes quite far, whereas the prototypical EV owner is also rich but more likely to live in the city and most likely has a toll along the commute to work.

Naturally, one may be concerned about endogeneity of toll exposure due to residential sorting. To explore this, we estimate a linear probability model where the outcome is a dummy for owning an EV, the key explanatory variable is the toll (in NOK/trip). We estimate a pooled regression in addition to one that includes fixed effects for neighborhoods, where identifying variation only comes from comparisons of neighbors that have different work locations. The results are in Table B.5. The estimated effect of toll exposure is $0.12 \%$-points without neighborhood fixed effects, and $0.10 \%$-points with effects. While statistically significantly different, the two effects are economically very close. This is particularly striking given the primary variation in tolls is across cities with and without tolls and over time. We take this

Figure 2: EV ownership by toll exposure, cars owned and work distance deciles, 2017

A: One-car households


B: Two-car households

-— Exposed to tolls on work route - No tolls on work route

Notes: 2017 households with positive work distances are divided in work distance deciles. The share of households owning an EV is then calculated separately for one-car and twocar households depending on whether at least one household member is required to pass a toll gate on her way to/from work.
as evidence that residential sorting is adequately controlled for by our rich set of demographic controls.

### 2.5 Portfolio Aspects

We now turn to descriptive evidence focusing explicitly on two-car aspects. We start with driving before turning to ownership.

The distribution of driving by cars in one- and two-car portfolios looks remarkably similar (Figure B.4), except that those in two-car portfolios are driven more. This holds even after controlling for demographics and car characteristics in a linear regression. ${ }^{19}$

We find more important effects pertaining to portfolio shifting: that is, a given car is driven differently depending on what other car the household has access to. To show this, we have run a regression of driving on demographics and car characteristics, where we crucially have included the difference in characteristics between the car in question and the second car in the portfolio. The estimated coefficients reveal that households tend to shift driving towards the car that is youngest, heaviest, and has the most horsepower. This precise mechanism will be built into our structural model in Chapter 4.

We now turn to the ownership margin. We have along the way seen some hints

[^9]that the type choices of two-car households may be different than those of one-car households, and we will now provide more direct descriptives. First, we have already seen in Figure 1 that EVs tend to belong to two-car portfolios. More precisely, if we select an EV at random from the 2017 dataset, then the probability that it belongs to a two-car portfolio is $75.6 \%$.

Next, if we partition the cars into just two types, EV or CV, then we can examine the market shares for all six possible portfolios, which we denote $\left(s_{0}, s_{\mathrm{EV}}\right.$, $s_{\mathrm{CV}}, s_{\mathrm{EV}, \mathrm{EV}}, s_{\mathrm{CV}, \mathrm{CV}}, s_{\mathrm{EV}, \mathrm{CV})}$. Table 1 presents these for 2017 . We see that out of all households, the (EV, CV) bundle is owned by $3.19 \%$ of households, while the singleton EV portfolio is only owned by $1.10 \%$. The symmetric bundle, (EV, EV), is very marginal at only $0.20 \%$, and we will not pay much attention to it throughout the rest of the paper. Conversely, the symmetric (CV, CV) bundle is the most popular two-car option, at $24.8 \%$. The full table of all portfolio market shares in Appendix Table B.1. This reveals that the most popular (EV, CV) combination is the Small used EV and the Large diesel 5-11 years, owned by $0.53 \%$ of households.

## 3 Complementarity

The key insight in this paper is that there are strong portfolio synergies between an EV and a CV, so strong that the two are strict Hicksian complements for some households. This section will provide simple descriptive evidence to back this up without relying on our full structural model. To do this, we present three types of evidence of increasing complexity, all based on portfolio market shares alone.

Let us start by defining some notation. We let $d \in \mathcal{D}$ denote a discrete portfolio choice, and $j, k \in \mathcal{J}$ denote individual cars. Then $d=0$ is no car, $d=j$ is a one-car portfolio, wheres $d=(j, k)$ denotes a bundle of two cars. We will let $s_{d}$ denote observed market shares for $d \in \mathcal{D}$, and let s denote a full vector of market shares. To simplify the exposition in this section, we will aggregate cars into just two, $\mathcal{J}=\{\mathrm{EV}, \mathrm{CV}\}$, and we will be showing market shares for the full population as well as within various subsamples of households in Table 3.

Our first piece of evidence asks whether an EV is chosen more by two-car households than (similar) one-car households:

$$
\text { C. } 1 \text { The Dependence Criterion: } \frac{s_{\mathrm{EV}, \mathrm{CV}}}{s_{\mathrm{EV}, \mathrm{CV}}+s_{\mathrm{EV}, \mathrm{EV}}+s_{\mathrm{CV}, \mathrm{CV}}}>2 \frac{s_{\mathrm{EV}}}{s_{\mathrm{CV}}+s_{\mathrm{EV}}} \frac{s_{\mathrm{CV}}}{s_{\mathrm{CV}}+s_{\mathrm{EV}}}
$$

This criterion is particularly relevant given that many previous structural one-car models of the car market implicitly make the assumption that the choices of two-car households are independent. Table 3 shows that this criterion is satisfied for every
single subset of the population. That means that the mixed bundle is chosen more often by 2-car households than what two randomly selected 1-car households would do. Of course, as Gentzkow (2007) discussed, part of this reason may be due to correlated preferences, but we note that the criterion is satisfied even within the richest decile of households or among the households exposed to road tolls.

Table 3: Market Shares and Complementarity Criteria

|  | $N$ | Market shares |  |  |  |  |  | Compl. Criteria |  |  | $\hat{\Gamma}_{\text {EV,CV }}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $s_{0}$ | $s_{\text {EV }}$ | $s_{\text {CV }}$ | $s_{\text {EV,EV }}$ | $s_{\text {CV,CV }}$ | $s_{\mathrm{EV}, \mathrm{CV}}$ | C. 1 | C. 2 | C. 3 |  |
| Income decile 1 | 216980 | 83.25 | 0.22 | 14.38 | 0.01 | 2.05 | 0.09 | $\checkmark$ | $\checkmark$ | $\checkmark$ | 0.87 |
| Income decile 5 | 216977 | 24.28 | 1.15 | 57.33 | 0.04 | 16.46 | 0.73 | - | $\checkmark$ | - | -1.31 |
| Income decile 10 | 216976 | 7.06 | 2.64 | 32.34 | 1.04 | 44.35 | 12.57 | - | $\checkmark$ | - | 0.04 |
| Urban $=0$ | 1638780 | 25.51 | 0.83 | 41.45 | 0.18 | 28.93 | 3.10 | $\checkmark$ | $\checkmark$ | - | 0.83 |
| Urban $=1$ | 530989 | 45.47 | 1.94 | 36.84 | 0.27 | 12.02 | 3.46 | $\checkmark$ | $\checkmark$ | - | 0.79 |
| Toll $=0$ | 1740109 | 33.84 | 0.67 | 40.25 | 0.10 | 23.38 | 1.77 | $\checkmark$ | $\checkmark$ | - | 0.79 |
| Toll $=1$ | 429660 | 16.44 | 2.84 | 40.61 | 0.64 | 30.52 | 8.94 | - | $\checkmark$ | - | 0.24 |
| $\mathrm{Wd}=0$ | 1010380 | 46.81 | 0.43 | 39.43 | 0.04 | 12.64 | 0.65 | $\checkmark$ | $\checkmark$ | - | 0.58 |
| $\mathrm{Wd}=1$ | 1159389 | 16.09 | 1.69 | 41.09 | 0.35 | 35.38 | 5.41 | - | $\checkmark$ | - | 0.23 |
| Couple $=0$ | 984615 | 53.66 | 0.76 | 39.92 | 0.02 | 5.33 | 0.32 | - | $\checkmark$ | - | -0.56 |
| Couple $=1$ | 1185154 | 11.07 | 1.39 | 40.65 | 0.36 | 40.96 | 5.57 | - | $\checkmark$ | - | 0.09 |
| Full sample | 2169769 | 30.40 | 1.10 | 40.32 | 0.20 | 24.79 | 3.19 | $\checkmark$ | $\checkmark$ | - | 0.78 |

Notes: A " $\checkmark$ " indicates that the corresponding complementarity criterion is satisfied for the corresponding market share vector. The coefficient, $\hat{\Gamma}_{\mathrm{EV}, \mathrm{CV}}$, is one of the five estimates, ( $\left.\hat{U}_{\mathrm{EV}}, \hat{U}_{\mathrm{CV}}, \hat{\Gamma}_{\mathrm{EV}, \mathrm{EV}}, \hat{\Gamma}_{\mathrm{CV}, \mathrm{CV}}, \hat{\Gamma}_{\mathrm{EV}, \mathrm{CV}}\right)$, for the model (1), exactly identified by a single market share vector.

The second piece of evidence is based on the Correlation Criterion of Manzini et al. (2019). This asks whether the following conditional probability is greater than the unconditional: $\operatorname{Pr}(\mathrm{EV} \mid \mathrm{CV})>\operatorname{Pr}(\mathrm{EV})$.

$$
\text { C. } 2 \text { The Correlation Criterion: } \frac{s_{\mathrm{EV}, \mathrm{CV}}}{s_{\mathrm{EV}, \mathrm{CV}+}+s_{\mathrm{CV}}+s_{\mathrm{CV}, \mathrm{CV}}}>s_{\mathrm{EV}, \mathrm{CV}}+s_{\mathrm{EV}}+s_{\mathrm{EV}, \mathrm{EV}}
$$

Table 3 shows that this criterion is satisfied for the full sample and nearly half of the subsamples of households. Clearly, C. 2 is much stricter than C.1. This is because conditional on owning a CV, a randomly selected household is most likely to own just a single CV, followed by two CVs, both of which are in the denominator.

Our third criterion is the Hicksian criterion. This gives the cleanest criterion for complements, but it necessitates a model. Following Gentzkow (2007), we set up the following random utility model (RUM), which will later be the backbone for our
full structural model.

$$
\begin{align*}
V_{i 0} & =U_{0}+\varepsilon_{i 0} \\
V_{i j} & =U_{j}+\varepsilon_{i j}  \tag{1}\\
V_{i j k} & =U_{j}+U_{k}+\Gamma_{j k}+\varepsilon_{i j}
\end{align*}
$$

where $\left(\varepsilon_{i d}\right)_{d \in \mathcal{D}}$ are IID Extreme Value Type 1. Our full structural model, presented in the next section, will simply be an extension where all coefficients are heterogeneous and depend on optimally chosen driving. In this model, the expected quantity of each good is

$$
\begin{equation*}
Q_{j} \equiv \operatorname{Pr}(j)+\sum_{k \in \mathcal{J}} \operatorname{Pr}(j, k)+\operatorname{Pr}(j, j) \tag{2}
\end{equation*}
$$

Thus, the third criterion is simply
C. 3 Hicksian Criterion: $\frac{\partial}{\partial U_{\mathrm{CV}}} Q_{\mathrm{EV}}>0$.

With the location normalization that $U_{0}=0$, the remaining parameters, $\left(U_{0}, U_{\mathrm{EV}}\right.$, $\left.U_{\mathrm{CV}}, \Gamma_{\mathrm{EV}, \mathrm{EV}}, \Gamma_{\mathrm{CV}, \mathrm{CV}}, \Gamma_{\mathrm{EV}, \mathrm{CV}}\right)$, are exactly identified from a single vector of market shares, $\mathbf{s}=\left(s_{d}\right)_{d \in \mathcal{D}}$, and can be computed in closed form as $\hat{U}_{d}:=\log s_{d}-\log s_{0}$ for $d \neq 0$, and $\hat{\Gamma}_{j k}=\hat{U}_{j k}-\hat{U}_{j}-\hat{U}_{k}$.

We see from Table 3 that C. 3 is the strictest criterion, being only satisfied by a single subsample, namely the poorest decile of households. It is not surprising that this criterion is the strictest: in Appendix C.1, we prove that C. 2 always implies C.3, but that the reverse only holds when demand is binary (buy or don't), which happens when $\Gamma_{j j}=-\infty$ for all $j \in \mathcal{J}$ (implying that $s_{j j}=0$ ). This was the assumption in Gentzkow (2007) where the two goods were subscriptions to online and printed news papers respectively. In that context, and in many others, there is no meaningful quantitative aspect to single good demand. In our setting, however, a non-trivial number of households buy cars of the same type, which we refer to as a specialized bundle, $d=(j, j)$. With binary demand, we furthermore prove that two goods $j$ and $k$ are complements if and only if $\Gamma_{j k}>0$. With non-binary demand $\left(\Gamma_{j j}, \Gamma_{k k}>-\infty\right)$, we need strictly higher values of $\Gamma_{j k}$ to satisfy C.3. Intuitively, we can interpret $\Gamma_{j k}$ as the demand synergy and $\Gamma_{j j}$ as the single-good (negative) satiation in demand. The less satiation there is, the stronger synergy is required to satisfy C.3.

Table 3 shows that there is positive synergy ( $\Gamma_{\mathrm{EV}, \mathrm{CV}}>0$ ) between an EV and a CV for all samples of household with the exception of two: the middle-income and single households. It makes sense that the two cars would not be complements single households since they are so unlikely to buy two cars in general and since all cars would be strict substitutes if the two-car bundles were unavailable (i.e. in a one-car model).

Before finishing, it is worth commenting on the fact that the only group of households that satisfies all complementarity criteria is the poorest decile of households. This may seem counterintuitive given that very few households in this group own two cars (2.1\%). But on the other hand, most of them own no car (83\%) which means that in spite of the apparent EV-CV synergies, the household group is unlikely to play a large role for the aggregate effects of policies. Nevertheless, we shall see that our structural model also identifies complementarities frequently among the poorest households.

## 4 Model

In this section, we present our model of household choice of car and driving. 4.1 presents the model briefly, while 4.2 discusses the parameterization and intuition.

### 4.1 A Discrete-Continuous Model

We now extend the random utility model from before so that mean utilities and portfolio effects are functions of the choice of driving.

$$
\begin{aligned}
V_{i 0} & =U_{i 0}+\varepsilon_{i 0} \\
V_{i j} & =U_{i j}\left(x_{i j}\right)+\varepsilon_{i j} \\
V_{i j k} & =U_{i j}\left(x_{i j}\right)+U_{i k}\left(x_{i k}\right)+\Gamma_{i j k}\left(x_{i j}, x_{i k}\right)+\varepsilon_{i j k}
\end{aligned}
$$

where ( $x_{i j}, x_{i k}$ ) denotes the chosen driving (in km ) in car $j$ and $k$ respectively, and $\left(\varepsilon_{i d}\right)_{d \in \mathcal{D}}$ are IID Extreme Value Type 1. The outside option is the only unchanged utility, which we specify as

$$
U_{i 0}=\omega_{i}
$$

capturing that households are heterogenous in their access to and relative preference for alternative modes of transport. We do so in a two-period framework following Gillingham (2012). The (ex ante) utility from choosing car $j$ is

$$
\begin{equation*}
U_{i j}(x)=u_{i j 1}+\beta \mathbb{E}\left[u_{i j 2}(x)\right], \tag{3}
\end{equation*}
$$

where $\beta$ is the discount rate, $x$ is driving in km , and $u_{i j 1}, u_{i j 2}$ denote the flow utilities for period 1 and 2 to be described in the following. Expectations are shown to highlight the distinction between fuel prices at the time of purchase and driving. We will assume that households have myopic expectations, i.e. that discrete choices are made conditional on observed fuel prices in the year of purchase. ${ }^{20}$ Utility will

[^10]be quasi-linear in outside expenditures, with fixed marginal utility of money equal to $\gamma_{i}$.

Our specification heavily relies on the assumption that households are forwardlooking and internalize all monetary aspects of car ownership. The literature has found some evidence for (Busse et al., 2013; Grigolon et al., 2018) and some against this assumption (Gillingham et al., 2021).

In the first period, the car is purchased, yielding utility

$$
u_{i j 1}=-\gamma_{i} P_{i j}
$$

where $P_{i j}$ is the purchase price. In the second period, the car is enjoyed, driven $x$ km , fuel expenditures and other flow costs $\mathrm{FC}_{i j}$ (annual taxes) are incurred, before the car is resold at the end of the period at the depreciated price $\tilde{P}_{i d}<P_{i d}$. Thus,

$$
u_{i d 2}=\gamma_{i}\left(\tilde{P}_{i j}-\mathrm{FC}_{i j}-p_{i j}^{\mathrm{km}} x\right)+v_{i j}(x)+\xi_{i j}
$$

where $v_{i j}(x)$ is the utility from driving, and $\xi_{i j}$ is the part of car ownership utility unrelated to driving (a linear index of car characteristics), $p_{i j}^{\mathrm{km}}$ is the cost of driving one km . Note that the price of driving, $p_{i d}^{\mathrm{km}}$, and annual ownership costs, $\mathrm{FC}_{i d}$, can vary across households due to the presence of road tolls, and with $d$ because EVs are exempt from these, and because annual car taxes vary with the car type.

Finally, the utility of driving is specified as a quadratic,

$$
v_{i d}(x)=\alpha_{1 i j} x+\alpha_{2} x^{2},
$$

which will admit a computationally simple form for optimal driving. ${ }^{21}$ Allowing $\alpha_{1 i j}$ to depend on demographics and car attributes permits households with higher work distances to gain higher utility from driving, and cars with larger engines to be more enjoyable to drive. ${ }^{22}$

Optimal driving for one-car portfolios is found from the first-order condition: ${ }^{23}$

$$
\begin{equation*}
x_{i j}^{*}=\frac{1}{-2 \alpha_{2}}\left(\alpha_{1 i j}-\gamma_{i} p_{i j}^{\mathrm{km}}\right) . \tag{4}
\end{equation*}
$$

With two-car portfolios, optimal driving in the cars become linked due to $\Gamma_{i j k}\left(x_{i j}, x_{i k}\right)$ :

[^11]Optimal driving, $\mathbf{x}_{i j k}^{*} \equiv\left(x_{i j}^{*}, x_{i k}^{*}\right)$ should then solve the system

$$
\begin{aligned}
& x_{i j}^{*}=\frac{1}{-2 \alpha_{2}}\left[\alpha_{1 i j}-\gamma_{i} p_{i j}^{\mathrm{km}}+\left.\frac{\partial \Gamma_{i j k}\left(x_{i j}, x_{i k}^{*}\right)}{\partial x_{i j}}\right|_{x_{i j}=x_{i j}^{*}}\right] \\
& x_{i k}^{*}=\frac{1}{-2 \alpha_{2}}\left[\alpha_{1 i k}-\gamma_{i} p_{i k}^{\mathrm{km}}+\left.\frac{\partial \Gamma_{i j k}\left(x_{i j}^{*}, x_{i k}\right)}{\partial x_{i k}}\right|_{x_{i k}=x_{i k}^{*}}\right] .
\end{aligned}
$$

Taken jointly, optimal portfolio driving is $\left(\mathbf{x}_{i d}^{*}\right)_{d \in \mathcal{D}}$. When plugged back into the ex ante random utility, (3), we obtain the "indirect utility" which households base their car choice on: $d_{i}=\arg \max _{d \in \mathcal{D}} U_{i d}\left(\mathbf{x}_{i d}^{*}\right)+\varepsilon_{i d}$.

### 4.2 Parameterization

We now describe and discuss how we parameterize the random coefficients of our model, $\left(\omega_{i}, \gamma_{i}, \alpha_{1 i j}, \xi_{i j}, \Gamma_{i j k}(\cdot, \cdot)\right)$. These coefficients depend on household characteristics, denoted $\mathbf{z}_{i}$, and car characteristics, denoted $\mathbf{q}_{j}$, both of which we describe in Section 5.2.

$$
\begin{equation*}
\omega_{i}=\boldsymbol{\omega}^{\prime} \mathbf{z}_{i}, \quad \gamma_{i}=\gamma^{\prime} \mathbf{z}_{i}, \quad \alpha_{1 i j}=\boldsymbol{\alpha}_{z}^{\prime} \mathbf{z}_{i}+\boldsymbol{\alpha}_{q}^{\prime} \mathbf{q}_{j}+\varphi_{1 i j}, \quad \xi_{i j}=\boldsymbol{\xi}_{q}^{\prime} \mathbf{q}_{j}+\varphi_{2 i j} \tag{5}
\end{equation*}
$$

The coefficients $\left(\varphi_{1 i j}, \varphi_{2 i j}\right)$ handle EV-specific aspects related to driving and ownership respectively, and we cover those in 5.2.1.

Note that since $\mathbf{z}_{i}$ shifts the utility of having no-car through $\omega_{i}$, it is not included in $\xi_{i j}$. In discrete choice models without bundle options, this choice makes no difference. With bundles, however, the demographic shifter would be added twice for two-car choices, but only once for one-car choices. Therefore, we prefer to have it enter in the outside option utility.

The parameterization of the portfolio effect, $\Gamma$, is one of the key contributions of our paper. We set
$\Gamma_{i j k}\left(x_{i j}, x_{i k}\right)=\Gamma_{1}+\mathbf{1}\{j \neq k\} \Gamma_{2 j k}+\Gamma_{3}\left(x_{i j}+x_{i k}\right)+\Gamma_{4}^{\prime}\left(\mathbf{q}_{j}-\mathbf{q}_{k}\right)\left(x_{i j}-x_{i k}\right)+\Gamma_{5} x_{i j} x_{i k}$, and below we describe each of these five terms in turn before discussing alternative options. First, $\Gamma_{1}$ captures the general satiation in utility for the second car compared to the first car, in a way that is unrelated to driving. Thus, $\Gamma_{1}$ controls the overall proportion of two- to one-car households. $\Gamma_{2}$ will instead permit demand synergies between distinct car types in a way that does not relate to driving. $\Gamma_{3}$ allows satiation in driving in the sense that the marginal utility of driving can be different in two-car portfolios. $\Gamma_{4}$ captures preferences for specialization of driving in the two cars, causing households to shift driving towards the car with the highest characteristic $q_{j m}$ if $\Gamma_{4 m}>0$. Finally, $\Gamma_{5}$ allows for a non-zero cross-price derivative of driving, so that driving can be substitutes (if $\Gamma_{5}<0$ ) or complements ( $\Gamma_{5}>0$ ).

To see more clearly how the different portfolio parameters work, let us first consider the optimal driving with a $d=(j, k)$ bundle, supposing that $\Gamma_{5}=0$ :

$$
\begin{equation*}
\bar{x}_{i j}=\frac{1}{-2 \alpha_{2}}\left[\alpha_{1 i j}-\gamma_{i} p_{i j}^{\mathrm{km}}+\Gamma_{3}+\Gamma_{4}^{\prime}\left(\mathbf{q}_{j}-\mathbf{q}_{k}\right)\right] \tag{6}
\end{equation*}
$$

and vice versa for $\bar{x}_{i k}$. Here, it is clear that $\Gamma_{3}$ simply affects the mean driving of $j$ regardless of what $k$ is, whereas the $\boldsymbol{\Gamma}_{4}$-term shifts driving, i.e. it reduces the driving of one by the same amount that it increases the driving in the other car. However, this shift is independent of the prices of driving in the two cars, $\left(p_{i j}^{\mathrm{km}}, p_{i k}^{\mathrm{km}}\right)$. Permitting $\Gamma_{5} \neq 0$ then yields

$$
\begin{equation*}
x_{i j}^{*}=\frac{1}{1-\left(\frac{\Gamma_{5}}{2 \alpha_{2}}\right)^{2}}\left(\bar{x}_{i j}-\frac{\Gamma_{5}}{2 \alpha_{2}} \bar{x}_{i k}\right) . \tag{7}
\end{equation*}
$$

From this, we can directly see that $\Gamma_{5}$ alone will determine the sign of any cross-price derivative, and thus in particular whether driving in car $j$ and $k$ are substitutes or complements.

## 5 Empirical strategy

In this section, we first explain the econometric methodology for taking our model to the data and estimating parameters and conducting inference. Next, we discuss identification, both intuitively regarding what variation in the data pins down each parameter, but also specifically concerning exogenous variation in prices and other monetary variables. Finally, we discuss some practical choices, particularly in regard to how EVs are treated in the implementation of the model.

### 5.1 Likelihood Function

Our model features two decisions: the discrete car choice and the continuous driving choice. The full information likelihood will thus have a contribution for each term: a logit part for the discrete choice, and a Gaussian for the continuous choice.

For the discrete car choice, we assume that agents solve

$$
\max _{d \in \mathcal{D}} U_{i d}\left(x_{i d}^{*}\right)+\varepsilon_{i d},
$$

where $\varepsilon_{i d}$ are distributed IID Extreme Value Type I, and where $\mathcal{D}$ denotes the set of all singleton cars, $\mathcal{J}$, and all unordered pairs, $(j, k)$, of which there are $\frac{1}{2} J(J-1)$, as well as the outside option of having no car. For choices involving two cars, the notation subsumes the driving in both cars, i.e. $x_{i d}^{*}=\left(x_{i j}^{*}, x_{i k}^{*}\right)$ for bundle choices $d=(j, k)$. The agent observes $\varepsilon_{i d}$, but the econometrician does not, leading to the
usual logit choice probabilities

$$
\operatorname{Pr}_{i}(d)=\frac{\exp \left[U_{i d}\left(x_{i d}^{*}\right)\right]}{\sum_{d^{\prime} \in \mathcal{D}} \exp \left[U_{i d^{\prime}}\left(x_{i d^{\prime}}^{*}\right)\right]} .
$$

Regarding driving, we assume that we observe the optimal driving with additive Gaussian measurement error, i.e.

$$
x_{i}^{\mathrm{data}}=x_{i d_{i}}^{*}+\eta_{i}, \quad \eta_{i} \sim \operatorname{IID} \mathcal{N}\left(0, \sigma_{x}^{2}\right),
$$

where $d_{i}$ denotes the chosen car alternative for household $i$, barring the fact that no driving is observed when $d_{i}=0$.

We use full information maximum likelihood and thus

$$
\begin{equation*}
\mathcal{L}(\theta)=\sum_{i=1}^{N}\left[\log \operatorname{Pr}_{i}\left(d_{i}\right)+\sum_{j \in d_{i}} \log \frac{1}{\sigma_{\eta}} \phi\left(\frac{x_{i j}^{*}-x_{i j}^{\mathrm{data}}}{\sigma_{\eta}}\right)\right], \tag{8}
\end{equation*}
$$

where in an abuse of notation, we use $d_{i}$ as a set of one or two elements, and there is no driving contribution when $d_{i}=0$.

The exposition above has been simplified for expositional clarity, with details relegated to Appendix C.2. The exposition ignores two aspects: first, the asynchronous nature of our driving data. That is, a household can purchase an second car midway through the first car's driving period, or the car can change owners between inspections. The driving sub-likelihood that we take to the data captures this precisely, by performing predictions at the sub-period level (where ownership and portfolio is constant), and aggregating to the driving period level, weighted by the fraction of the driving period covered by each sub-period.

The second simplification is that we work with unequal sample sizes. For computational reasons, we are forced to reduce the number of discrete choice observations to a random $0.2 \%$ subsample ( $N=52,739$ ). However, we are much less constrained in terms of the driving sub-likelihood, and because the driving and discrete choice components of (8) are additively separable, we choose to use a larger sample of $S=2,588,591$ driving periods. With these sample restrictions, numerical optimization wrt. 47 parameters takes several days.

### 5.2 Implementation

This section will describe how we implement our model using the Norwegian data. Specifically, we will describe our choices of demographics, $\mathbf{z}_{i}$, car characteristics, $\mathbf{q}_{j}$, specification for fixed costs of car ownership, $\mathrm{FC}_{i d}$, and the two EV-specific effects in (5), $\varphi_{1 i j}, \varphi_{2 i j}$. Apart from this, the annual discount rate, $\beta$, is the only parameter we do not estimate but fix at 0.95. ${ }^{24}$

[^12]First, the vector of household observables is

$$
\mathbf{z}_{i}=(\log (\text { net income }), W D, \text { age, } \mathbf{1}\{\text { live in major city }\}, \mathbf{1}\{\text { couple }\}, \text { constant })^{\prime},
$$

where net income is summed over spouses, while work distance (WD) and age are averaged. Note that we do not include employment as a demographic. Employment will mainly affect household decision making through (1) increased purchase power and (2) increased driving demand due to commuting. These channels are already taken into account through income and work distance.

The vector of car attributes is
$\mathbf{q}_{j}=\left(\text { car age, car age }{ }^{2}, \text { weight, engine effect, } \mathbf{1}\{E V\} \times \text { engine effect, } \mathbf{1}\{E V\}, \mathbf{1}\{\text { diesel }\}\right)^{\prime}$. We experimented with additional engine characteristics such as displacement, but found it to be nearly collinear with weight and engine effect.

### 5.2.1 Special Treatment of EVs

In the vector of car characteristics, $\mathbf{q}_{j}$, we allow engine effect to have a different effect for EVs compared to CV. There are two reasons for this: First, electric engines are based on different technology - one difference is that they tend to achieve a higher acceleration than internal combustion engines with similar effects. Second, it enables us to more flexibly fit the difference in preferences for small and large/luxury EVs. ${ }^{25}$

Other than the separate engine effect for EVs, we treat EVs differently in three regards: the EV's range can affect driving and ownership utility ( $\varphi_{1 i j}$ and $\varphi_{2 i j}$ ), and then we include a term to allow for demand synergies between an EV and a CV. First, $\varphi_{1 i j}$ should allow for range to affect the utility of driving. Specifically, having a long work distance (WD) and a low range should make the EV impractical for daily commutes. To capture this, we set $\varphi_{1 i j} \equiv \varphi_{1} \mathbf{1}\left\{\mathrm{WD}_{i}>0\right\} \mathbf{1}\{j$ is EV $\} \frac{\text { range }_{j}}{\mathrm{WD}_{i}}{ }^{26}{ }^{2}$ $\varphi_{1 i j}<1$ would mean that the car must recharge at least once on the way to work. Relating the variable to WD has a further effect in that otherwise, range and engine effect are very highly correlated, so this way we avoid near multicolinearity.

The second variable that captures EV aspects, $\varphi_{2 i j}$, relates to local incentives. We want the model to capture the non-price policies affecting EV ownership. These

[^13]include the benefits of free parking and charging in cities, as well as access to public transport lanes to avoid congested roads in rush hours. Since these benefits are best captured by an agent working in an urban area, we include the following dummy variable: $\varphi_{2 i j}=\mathbf{1}\{j$ is EV$\} \times \mathbf{1}\{i$ works in major city $\} .{ }^{27}$

The third and final variable that captures EV aspects, $\Gamma_{2 i j k}$, should capture demand synergies between an EV and a CV relating to the concept of range anxiety. We view this as a household's reluctance to purchase an EV because it may make very long trips, e.g. for holidays, unbearable due to the large number of required stops to wait while recharging. The fact that such trips are rare means that they might not be visible in average driving over two to four years, but might still be (overly) salient to consumers. A simple way to do this would be to set $\Gamma_{2 i d}$ equal to a dummy for an (EV, CV) portfolio. To maintain the interpretation as range anxiety, we instead set it to

$$
\Gamma_{2 i d}=\Gamma_{2}[\mathbf{1}\{d=(\mathrm{EV})\}+\mathbf{1}\{d=(\mathrm{EV}, \mathrm{EV})\}],
$$

so that $\Gamma_{2}$ will be a penalty imposed for choosing an EV without also having a CV. This way, the same range anxiety that keeps one-car households from choosing an EV will keep two-car households from choosing a double-EV portfolio. Note that there is a dummy for EV in $\mathbf{z}_{j}$.

### 5.2.2 Price Variables

Finally, we will describe the cost elements in the model. Ownership cost has two components. First, the purchase price of the car portfolio, $P_{i d}$, net of the resale price, $\tilde{P}_{i d}$. There is no cross-sectional variation in car prices, so the $i$ subscript only signifies that households face different prices at different points in time. Second, other fixed costs associated with the car portfolio, $\mathrm{FC}_{i d}$. These are specified as

$$
\begin{equation*}
\mathrm{FC}_{i d}=\theta_{\text {toll }} \mathrm{toll}_{i d}+\text { ownership tax }{ }_{i d}+{\text { maintenance } \operatorname{costs}_{d},} \tag{9}
\end{equation*}
$$

where ownership taxes vary over time and between EVs and CVs, and expected maintenance costs only vary with the age of the car, ${ }^{28}$ and toll ${ }_{i d}$ denotes road toll payments. We assume toll $i_{i d}$ to be zero if the household owns an EV, and otherwise equal to the monetary cost implied by the household's toll exposure assuming 220 days of commuting by car each year. As we do not know the average fraction of days toll-exposed households commute by car, $\theta_{\text {toll }}$, we estimate this parameter from

[^14]the data. That is, if $\theta_{\text {toll }}$ is one, it implies that households cross the toll road two times per day, 220 days per year, whereas a value of zero implies that they never cross it. We have thoroughly explored whether households treat tolls as a variable or a fixed cost of driving and even estimated both versions of our model, ultimately concluding that the fixed cost version yields a superior fit along all dimensions and picking that for our preferred specification. ${ }^{29}$ The implication is that tolls affect only the extensive and not the intensive margin. Finally, note that we do not take into account road tolls paid on non-work trips.

Driving cost per kilometer is defined as $p_{i j}^{k m}=$ fuel price ${ }_{i j} /$ fuel efficiency $_{i j}$, where fuel is either gasoline, diesel, or electricity. The purchase and driving decisions occur at different points in time which we account for: for the purchase, we take the average fuel prices from the previous year, assuming that households have static expectations. For the driving, we compute the realized average price over the course of the days covered by the specific driving period.

### 5.3 Identification

We now turn to the question of identification, both in a statistical sense and in an intuitive sense. We pursue a full information maximum likelihood method strategy, we leverage all aspects of behavior to identify the parameters. However, similarly to the Heckman (1979) selection model, it is possible to think of the model as composed of two separate parts: a linear regression for driving and a logit model of car choice. Discussing identification in these two avenues is thus more familiar and simpler. Finally in Subsection 5.3.3, we discuss the sources of exogenous variation in price variables in our data.

### 5.3.1 Driving $\left(\gamma_{i}, \alpha_{1 i j}\right)$

From data on driving of one-car households alone, the parameters $\left(\alpha_{1 i j}, \gamma_{i}, \sigma_{\eta}\right)$ are identified. If we fix the value of $\alpha_{2}$, then (4) is just a linear equation, so we can estimate parameters by linear regression. We simply regress driving on household and car characteristics, $\left(\mathbf{z}_{i}, \mathbf{q}_{j}\right)$, and an interaction with the price of driving, $p_{i j}^{\mathrm{km}} \times$ $\mathbf{z}_{i}$. The intuition for identification is thus straightforward: e.g. $\gamma_{i}$ is identified by

[^15]observing how different households change their driving in response to fuel cost shocks.

Adding data on driving by two-car households, we can furthermore identify $\left(\Gamma_{3}, \Gamma_{4}, \Gamma_{5}\right)$. First, $\Gamma_{3}$ is identified from the difference in driving of the same car whether it is the only car of the household or belongs to a portfolio, other things equal. As discussed in Section 2.5, cars in two-car portfolios are on average driven slightly more than in one-car portfolios. Conversely, we saw strong effects of portfolio shifting, whereby the household redistributes driving towards the car that is youngest and has the most powerful engine in its portfolio. This effect is captured by $\boldsymbol{\Gamma}_{4}$. Intuitively, it can be identified from the change in driving of the first car of a household that swaps the second car.

The final coefficient, $\Gamma_{5}$, is what allows for a non-zero cross-price elasticity of driving between the two cars, thus allowing driving to be either complements or substitutes in the two cars. Thus, it can be identified by comparing the driving of a household with the same portfolio as fuel prices change over time.

Note that $\left(\alpha_{1 i j}, \gamma_{i}, \Gamma_{3}, \boldsymbol{\Gamma}_{4}\right)$ can be estimated by linear regression because (6) is linear. The results hereof are presented in Appendix Table B.3. Once we add $\Gamma_{5}$, the driving equation, (7), becomes non-linear, so estimation must be conducted by Maximum Likelihood.

### 5.3.2 Discrete choice $(\omega, \xi, \Gamma)$

The parameters $\left(\omega_{i}, \xi_{i j}, \alpha_{2}, \Gamma_{1}, \Gamma_{2}\right)$ are identified solely from discrete choices. We will discuss these first before turning to the benefit of joint estimation by full information likelihood.

First, $\omega_{i}$ shifts the utility of the outside option. Therefore, it directly fits the share of no-car ownership for households with demographics $\mathbf{z}_{i}$. Conversely, $\xi_{j}$ shifts the utility of car $j$ all else equals. In a purely one-car model without two-car options, this directly controls the market shares of each car.

The coefficient $\alpha_{2}$ controls the curvature of driving utility but is not identified from driving data alone. It only becomes identified once discrete choice data is added, and then it controls the importance of driving utility relative to other sources of utility. ${ }^{30}$

Finally, two of the portfolio parameters are only identified from discrete choice data, $\Gamma_{1}$ and $\Gamma_{2}$. The first, $\Gamma_{1}$, captures satiation in the number of cars owned.

[^16]Conditional on the remaining parameters of the model, it directly affects the market share of the symmetric bundle $d=(j, j)$ relative to the corresponding one-car choice, $d=j$. Conversely, the coefficient $\Gamma_{2}$ controls the market shares of (EV), (EV,EV). ${ }^{31}$

### 5.3.3 Exogenous Price Variation

In this section, we discuss identification of the marginal utility of money, $\gamma_{i}$. Partly, identification comes from behavioral restrictions such as rationality, fungibility of money, and our assumption of static expectations for fuel prices. The other part is the variation in monetary variables in our data. Below, we go through the sources of variation for the following three key variables: fuel prices, road tolls, and consumer car prices.

First, fuel prices varied considerably over the sample period (see Figure A.4), between 12 and 15 NOK per liter of gasoline, and between 0.9 and 1.2 NOK per kWh . Since no significant reforms of fuel taxes occurred, the time-series variation is largely due to the world market oil prices.

Second, road tolls are an important local monetary policy in Norway, making up $\frac{1}{6}$ th of total car related tax revenue. Road tolls as well as the EV exemption from tolls provide identifying variation in the cost of an EV relative to a CV along two dimensions: time-series variation in toll rates and toll locations, and cross-sectional variation in toll exposure on work trips depending on home and work locations.

There is substantial variation in toll levels, both between and within different geographical areas - for instance, two neighbors might have similar work distances but commute in different directions. In 2017, around $20 \%$ of households (and $40 \%$ of households with positive work distances) were exposed to tolls on their way to work. The toll amount is substantial: the top ventile ( $5 \%$ of the data) faced 30,000 NOK and the second-highest ventile 15,000 NOK annually. ${ }^{32}$ Given that a small new gasoline car costs 46,166 NOK annually in depreciation and taxes (Table A.2), this is a considerable amount. As we discussed in Section 2.4, endogeneity due to residential sorting does not appear to play a big role when we explore variation within and across households in neighborhoods.

Our third and equally important monetary variable is the new car price. The registration tax in Norway typically makes up between 25 and 50 percent of the final price, with EVs being fully exempt. As discussed in Section 2.1, the tax composition faced dramatic reforms in our sample period, shifting from an emphasis on engine power towards $\mathrm{CO}_{2}$ emissions.

[^17]We argue that these reforms of the registration tax scheme were sufficiently large as to be the dominant source of price variation in our data. In order to investigate this, we run a regression of price on exogenous car attributes, $\mathbf{q}_{j}$. We run two version of this regression, one for each component of the final consumer price: one for the total registration tax payment of the car, and one for the price net of the tax (the producer price). Results are in Appendix Table A. 4 and the key takeaway is that the $R^{2}$ is 0.96 for the producer price but only 0.70 for the tax component. This means that after accounting for changes in car characteristics over time, there is only $4 \%$ residual variation in the firm's price; conversely, there is $30 \%$ residual variation in the tax payment. ${ }^{33}$ We conclude that while price endogeneity may of course still determine the remaining $4 \%$ of producer price variation, the vast majority of the variation in prices is policy-induced.

We emphasize that the reason why there is so little residual variation in producer prices and so much in taxes is due to our coarse aggregation of cars into just 20 types. The approach works because registration taxes vary greatly across these types. Conversely, if product quality is independently distributed across the characteristics space (as assumed by most state-of-the-art empirical IO methods), then averaging across products within a segment is likely to take away this idiosyncratic variation. Note that this also means that our approach cannot be used to study anything that is firm-specific (e.g. market power), since we average over many firms within car types.

Finally, we will discuss some of the limitations of our approach. First, we do not observe transaction prices either for new or used cars, but are instead forced to rely on suggested retail prices and fixed depreciation schedules. Second, we ignore some niches of the car market like company cars and leasing due to lack of data. Third, our data on maintenance cost is very limited and only varies with the car's age, again due to data limitations.

### 5.4 Methodology for Counterfactual Simulations

This section provides the details underlying our counterfactual simulations. We will use these both to compute the marginal effects on various outcomes in Section 6, and for our policy analysis in Section 7.

All our counterfactual simulations are based on the population in 2017. All

[^18]reported outcomes, $a_{i d}$, are averaged over the households in the sample and expected with respect to the discrete choice probabilities:
\[

$$
\begin{equation*}
a^{s}=N^{-1} \sum_{i=1}^{N} \sum_{d \in \mathcal{D}_{i}} \operatorname{Pr}_{i}^{s}(d) a_{i d}^{s}, \tag{10}
\end{equation*}
$$

\]

where $s=0,1$ denotes the scenario: baseline and counterfactual, respectively. Examples of such outcomes, $a_{i d}$, include the number of cars, km , or $\mathrm{CO}_{2}$ emitted.

Furthermore, we take into account externalities due to driving. We take monetary constant marginal externalities per kilometer driven from Rødseth et al. (2019), as described in Appendix A.4.3. All outcomes are computed at the annual level.

Car taxes consists of the registration tax and VAT (paid for new cars only), fuel taxes, tolls on the work route and the annual ownership tax. We assume $100 \%$ passthrough of both fuel and registration taxes, which is empirically plausible. ${ }^{34}$ Furthermore, we assume that used-car prices always follow the fixed depreciation schedule of $12.5 \%$ annually. This implies a full "proportional" passthrough in the sense that when the price of a brand new car falls by $20 \%$, all used cars of that type also fall by $20 \% .{ }^{35}$ Robustness checks to this assumption is included in Appendix ??.

The consumer welfare measure is the usual "logsum" (Small and Rosen, 1981)

$$
\begin{equation*}
\mathrm{CS}_{i}^{s}=\frac{1}{\gamma_{i}} \log \sum_{d \in \mathcal{D}} \exp \left[U_{i d}^{s}\left(x_{i d}^{*}\right)\right] . \tag{11}
\end{equation*}
$$

Our social welfare measure is then

$$
\text { social welfare }=\text { tax revenue }{ }_{i}+\mathrm{CS}_{i}-\text { externalities }_{i} .
$$

Since utility is quasi-linear, there are no income effects and so our model is neutral to redistribution. Instead, we will simply report whether the policies we simulate are progressive or regressive, leaving it to the policy maker to judge the relative weight that should be attached to distributional concerns.

## 6 Results

We will now present our parameter estimates and their marginal effects on outcomes of interest, and discuss the model fit and implied elasticities.

[^19]
### 6.1 Parameter Estimates

Table 4 displays the parameter estimates from our preferred specification. All parameters have the expected sign and are statistically significant as indicated by the $t$-values. For instance, richer households have a lower marginal utility of money, and households with longer work distances have a higher utility of driving. Households living in urban areas have a higher utility of the outside option of not owning a car, likely due to better access to alternative modes of transportation. And households with a high work distance have higher $\alpha_{1 i j}$ but lower $\gamma_{i}$, consistent with higher driving needs for commuting and that commuting needs are less price sensitive than other trip types on average.

The portfolio estimates also have the expected signs: $\Gamma_{1}$ is negative, so there is diminishing utility from the second compared to the first car. Estimates also indicate strong range anxiety in ownership utility $\left(\Gamma_{2}<0\right)$ and to a smaller extent related to driving ( $\varphi_{1 i j}<0$ ), but the presence of local EV incentives are important to consumers ( $\varphi_{2 i j}=1.58$ ).

The estimate of $\Gamma_{5}$ is negative, -. 001, implying that in two-car portfolios, driving in the two cars are substitutes. This confirms earlier findings by De Borger et al. (2016), although our estimates imply a low cross-price elasticity of 0.006 . Moreover, the estimate (combined with $\Gamma_{3}$ ) implies that most cars are driven slightly less in two than one-car portfolios. For example, the "large new EV" is driven 51 km daily by the average household when alone, 45 km each when owning two copies of it (see Appendix Table D.3). The estimated vector $\boldsymbol{\Gamma}_{4}$ reveal strong portfolio shifting, with households preferring to allocate more driving to the car that is youngest and has the most powerful engine. This shifting implies that households have preferences for diversification in car characteristics.

For use in our counterfactual analysis later, we have also estimated an alternative specification where we set all 11 synergy coefficients in $\Gamma_{i j}$ to zero. Those estimates are presented in Appendix Table D.4. Unsurprisingly, a likelihood ratio test of the null that $\Gamma_{i j}=0$ yields a value of 961.96 , which is clearly rejected.

### 6.2 Model Fit

We evaluate the model's fit out of sample on a $10 \%$ random sample not including our estimation sample. ${ }^{36}$ We start with the discrete ownership choice, and then show the fit of driving.

Figure 3 shows that the model provides a reasonable fit of the number of cars

[^20]Table 4: Parameter Estimates

| Demographics: | Driving ( $\alpha_{z}$ ) |  | Outside option $\left(\omega_{z}\right)$ |  | Utility of money ( $\gamma_{z}$ ) |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\ln$ (household income) | -0.0162 | (18.3) | -0.372 | (-7.58) | -0.537 | (-10.8) |
| work distance (kms) | 6.86e-4 | (47.3) | $9.54 \mathrm{e}-4$ | (12.2) | 0.00966 | (20.6) |
| age (avg) | -0.00110 | (-14.9) | -0.0100 | (-2.61) | -0.0195 | (-25.2) |
| 1 \{city\} | -0.0359 | (-12.0) | 0.186 | (10.8) | -0.380 | (-4.3) |
| 1 \{couple\} | 0.0198 | (7.93) | 0.252 | (13.0) | -0.324 | (-6.95) |
| cons | 0.00938 | (8.87) | 8.22 | (12.9) | 9.61 | (14.0) |
| Car attributes: | Driving ( $\alpha_{q}$ ) |  | Car ownership $\left(\xi_{q}\right)$ |  | Portfolio shifting ( $\Gamma_{4}$ ) |  |
| car age | 1.91e-4 | (-15.7) | 0.411 | (49.1) | $3.44 \mathrm{e}-4$ | (11.9) |
| car age ${ }^{2}$ | -3.30e-4 | (-12.7) | -0.00860 | (-11.8) | -4.95e-05 | (-8.05) |
| engine effect ( $100 \mathrm{~kW} \mathrm{)}$ | 0.0412 | (7.44) | 1.70 | (7.76) | 0.0131 | (6.19) |
| engine effect $\times 1\{\mathrm{EV}\}$ | -0.0255 | (-3.94) | -2.21 | (-5.99) | -0.00479 | (-9.79) |
| total weight (tonnes) | 0.0561 | (15.0) | -4.24 | (-39.7) | -0.00157 | (-7.20) |
| 1 \{diesel\} | 0.0540 | (19.1) | -1.99 | (-15.5) | -0.00364 | (-12.0) |
| 1 \{EV\} | 0.0599 | (20.4) | -2.36 | (-12.7) | -0.0312 | (-11.3) |
| Other variables |  |  | Portfolio effects |  |  |  |
| Range by work distance ( $\varphi_{1 i j}$ ) | $3.61 \mathrm{e}-4$ | (22.4) | Ownership | satiation ( $\Gamma_{1}$ ) | -6.08 | (-19.0) |
| Local EV incentives ( $\varphi_{2 i j}$ ) | 1.58 | (11.3) | Range an | ety $\left(\Gamma_{2}\right)$ | -0.757 | (-5.38) |
| Driving squared ( $\alpha_{2}$ ) | -0.00330 | (-34.0) | Driving s | iation ( $\Gamma_{3}$ ) | 0.0325 | (15.6) |
| SD of error term, driving ( $\sigma_{x}$ ) | 29.505 | (533.1) | Driving s | stitution ( $\Gamma_{5}$ ) | -0.00102 | (-19.3) |
| Realised toll payment share ( $\theta_{\text {toll }}$ ) | 0.445 | (5.22) |  |  |  |  |

Notes: Parentheses show the t-statistics corresponding to each parameter value using the standard errors from the Sandwich formula in Equation 14. The results are based on a random sub-sample of $N=52,739$ discrete choice observations. The driving dataset is a $30 \%$ random sample of all driving periods, yielding $S=2,588,591$ odometer readings with an average of 2.2 sub-periods (due to ownership or portfolio changes).
owned across income and work distance distributions, with two-car ownership increasing in both income and work distance. The model explains this by the fact that richer households have lower marginal utility of money, combined with a stronger dislike of the outside option.

In Appendix ??, we present additional fit graphs, including car attributes by income and work distance (Figure D.2); car ownership over time (Figure D.3); and car ownership by the $J=20$ car types (Figure D.4). To summarize, the model is able to capture the cross-sectional fit of ownership remarkably well, while performing slightly worse when it comes to capturing changes over time. This is not surprising, as no time trends or time specific coefficients are included - changes over time are solely attributed to changes in demographics, car attributes or the choice set.

Figure 4 shows the data and model fit for selected driving-related outcomes. The top panels show histograms of observed and predicted driving for one-car (left) and two-car (right) households respectively. The dispersion of model predictions reveals that a remarkable amount of heterogeneity is captured by the model through covariates (similarly to a high $R^{2}$ for the driving equation). This means that the model can accurately capture which households have high or low driving needs, which is

Figure 3: Number of cars by income and work distance

crucial for capturing the selection into car types based on anticipated driving.
The bottom panels highlight portfolio aspects of driving. For instance, the left panel shows that if one of the two cars is new and the other is ten years old, then the new car is driven 45 km per day, and the old car is driven just under 35 km per day, on average. The right panel shows a corresponding figure but with the difference in the weight of the two cars, showing that households tend to allocate more driving into heavier cars. In Appendix Figure D.1, we demonstrate the fit of driving for one-car portfolios with respect to car age, car weight, income, and work distance.

### 6.3 Elasticities

The parameter effects themselves can be interpreted to some extent, but calculating marginal elasticities is more directly comparable to existing literature. To compute these elasticities, we first compute expected outcomes according to (10), and then re-compute after increasing the exogenous forcing variable in question by $1 \%$. Table 5 presents the elasticities of our model evaluated on the 2017 cross section.

The first two columns show the elasticity with respect to income and work distance. Mostly, the two variables have the same qualitative effect, but the income elasticity is numerically larger. For instance, raising income by $1 \%$ results in $0.42 \%$ more cars, while increasing work distance by $1 \%$ only results in $0.03 \%$ more cars. A key difference between income and work distance, though, is that when a household becomes richer, it shifts towards the diesel segment, whereas a longer work distance implies a greater shift towards an EV.

The next three columns increase fuel prices: first for CV alone, then EV alone, and then both. The last three columns do the same, but for the new car prices instead. We assume that the price increases were due to increased taxes under full passthrough to new car prices. We assume that used-car prices always obey the

Figure 4: Driving


Notes: The top panels show the distribution of driving in the data and the distribution of predicted driving. In the bottom two panels, the x -axis is a difference in characteristics between the two cars for two-car households: age and weight (in tonnes) respectively. Driving is measured in kilometers per day per car.
fixed depreciation rate of $87.5 \%$ annually described in Section A.
Turning to driving, we see that the overall price elasticity of driving is -0.14 . This is slightly below some estimates such as the -0.3 found by Gillingham and Munk-Nielsen (2019) for Denmark, but above some estimates for the US. We also note that in aggregate, EVs and CVs are substitutes: the cross-price elasticity of the number of EVs with respect to the CV purchase price is 0.35 , and for the fuel price it is 0.21 .

The results in Table 5 indicate aggregated own-price elasticities of -0.10 and -0.31 for the gasoline and diesel cars segments, and -0.97 for EVs. This is seemingly quite far from typical estimates of the own-price elasticity of car demand but is due to the aggregation to fuel segment. Instead, Appendix Table D. 1 shows a full matrix of cross-price elasticities for the sales of individual cars. For new cars, we see ownprice elasticities more in line with the literature, ranging from -1.7 (small gasoline) to -3.4 (large diesel) for traditional cars, and -1.5 for the small EV and -3.8 for the large EV. Similarly, the cars that lose most demand to the outside option are the cheapest cars (e.g. 0.093 for the "large diesel, 5-11 years"), and lowest for the "large new EV" (0.0017). This is due to the demographic composition of the marginal households
in the market. We have also disaggregated the cross-price elasticities by income (results available upon request) and the poorer half of households have an average own-price elasticity that is nearly twice that of the richest half of households.

Finally, we turn to the question of portfolio synergies and complementarity. First, we note that no two cars are Hicksian complements to all households, and none are on average. However, most are to some subset of households. Most notably, the small EV and the large, new diesel car are Hicksian complements to $16 \%$ of households. (Appendix Table D.2). Conversely, only about $1 \%$ of households have sufficiently strong synergies between typical (CV, CV)-pairs for them to be complements. ${ }^{37}$

However, there can still be strong portfolio synergy between two cars without it being strong enough to imply strict complements. Table 6 visualizes a measure of portfolio synergy given by the average supermodularity of utility:

$$
\begin{equation*}
\Delta_{i j k} \equiv\left[U_{i j k}\left(x_{i j \mid 2}^{*}, x_{i k \mid 2}^{*}\right)-U_{i j}\left(x_{i j \mid 1}^{*}\right)\right]-\left[U_{i k}\left(x_{i k \mid 1}^{*}\right)-U_{i 0}\right], \tag{12}
\end{equation*}
$$

where $x_{i j \mid 2}^{*}, x_{i j \mid 1}^{*}$ denotes the optimal driving in car $j$ when it is in the $(j, k)$ portfolio and in the singleton portfolio respectively. There are two reasons why we prefer to show $\Delta_{i j k}$ rather than $\Gamma_{i j k}$. Firstly, $\Delta_{i j k}$ takes into account that a car $j$ is driven differently when combined with $k$ compared to alone, and this is precisely the deep form of synergy our model is meant to capture. And secondly, we have parameterized range anxiety as a penalty to the single-EV portfolio rather than a bonus to the (EV, CV) combination. Note that if there were no driving, homogenous parameters, $U_{i 0}=0$, and no range anxiety, then $\Delta_{i j k}=\Gamma_{i j k}$. ${ }^{38}$

Table 6 shows that portfolio synergy is strongest precisely between the small EVs and the larger CVs. Within the CV segment, synergy is strongest between the most diverse cars: the small, oldest, cheapest gas car and the large, luxurious, new diesel car.

## 7 Counterfactuals

In this section, we investigate the relative cost-effectiveness of three environmental policies in use in Norway through counterfactual simulations: (1) the VAT exemption for EVs currently in place, (2) increasing fuel taxes for diesel and gasoline, and (3) increasing purchase taxes for CVs. Each of these policies are motivated by

[^21]environmental goals of reducing $\mathrm{CO}_{2}$ emissions, and each contribute toward this end. In order to make the policies comparable in magnitude, we set the rates in (2) and (3) so that they yield the same $\mathrm{CO}_{2}$ reduction as (1), which we find to be 7.3 kg per household in 2017. To achieve this, we raise the fuel tax proportionally until emissions have also fallen by 7.3 kg , at which point consumer fuel prices have increased by $2.06 \%$. We do the same for the CV VAT, and the final increase amounts to $1.62 \%$ higher car prices. These magnitudes line up well with the elasticity of $\mathrm{CO}_{2}$ emissions wrt. the fuel and purchase prices in Table 5.

The results are shown in Table 7. The bottom line is that the most cost-effective policy tool is the CV fuel tax, followed by the CV purchase tax, and finally the EV exemption from VAT. The abatement costs per ton of $\mathrm{CO}_{2}$ are 897, 13,481, and 26,740 NOK per tonne respectively.

It is not surprising that the fuel tax is able to achieve a given reduction in $\mathrm{CO}_{2}$ with lower distortions given that it allows consumers to respond on both the intensive (driving) and extensive (type choice) margins. Conversely, a uniform increase in the VAT is mainly a tax on the most expensive cars, so the primary way this affects emissions is by pushing households out of car ownership entirely. This is clear from the fact that the reduction in total driving is $35 \%$ greater with purchase taxes compared to fuel taxes. That is, the purchase tax achieves emission reductions by reducing car ownership altogether whereas the fuel tax shifts driving towards more efficient cars, whereby the same reduction can be achieved with a smaller reduction in driving. ${ }^{39}$

The EV exemption from VAT is similar to the CV purchase tax in that it targets the vehicle purchase, rather than use. However, ex ante one might have had higher hopes for this policy given that it is incentivizing a vehicle type with zero emissions (given Norway's high reliance on renewables). Nevertheless, we see that the EV incentive resulted in a drastic increase in overall car ownership of 3,769 cars. The biggest change in portfolio choice is (CV, EV) which increases by 8,060 households. Conversely, one-car EV-ownership only increases by 2,773 . This is the first indication that the EV-CV complementarity plays a role but we will focus more on this later.

An additional disadvantage of the EV exemption from VAT is that it results in an increase in total driving. Even though this driving is in zero emission vehicles (hence the reduction in $\mathrm{CO}_{2}$ ), driving also entails non-environmental externalities such as accidents, congestion, and noise.

Before moving on, a brief note on the level of the abatement cost in relation to

[^22]Figure 5: Welfare components by income decile

A: Increased fuel taxes


B: VAT exemption for EVs


Notes: Panels de-compose the welfare effect for deciles of the income distribution. The top line is the change in consumer surplus. The second line is the change in annual tax payments. The third line is the uniform lump-sum transfer required for the Government budget constraint to hold, i.e. the negative of the average change in tax revenue. The fourth line is the change in local exernalities. The final line is net welfare, i.e. the change in consumer surplus net of transfers and local externalities. All outcomes are measured in 2015-kroners per household per year along the y-axis.
the literature. The costs we find are generally higher than what has been reported in the literature reviewed by Gillingham and Stock (2018): 160-410 NOK/tonne for fuel taxes (Knittel and Sandler, 2018) and 3000-5700 NOK/tonne for direct EV subsidies (Archsmith et al., 2015b). One reason for our higher estimated costs is perhaps the much higher overall taxes on fuel and registrations compared to the US, since one would expect marginal deadweight losses to be increasing in the level of taxes.

### 7.1 Distributional effects

Since the model is quasi-linear in income, it is neutral to redistribution. Nevertheless, the counterfactuals are not budget neutral so we analyze the distributional consequences under a uniform lump-sum net transfer to all households.

Figure 5 shows the average welfare, tax payments, and externalities across the income distribution, by deciles. We see that for the fuel tax, consumer welfare net of taxes is positive for the poorest five deciles, and increasingly negative for richer households, clearly illustrating the progressivity. Conversely, the EV exemption from VAT is strongly regressive with only the richest decile of households enjoying a net benefit; while consumer surplus increases for the remaining households, it is more than offset by the fall in tax revenue. So in conclusion: not only is the fuel tax
more cost effective than the EV incentives, it also has more desirable distributional consequences. In practice, the regressivity of EV incentives may be obscured by the fact that car sales vary so much from year to year, making it difficult to notice the fall in tax revenue.

### 7.2 The Role of Complementarities

We now assess what role complementarities play in driving the policy conclusions above. One might think of simulating counterfactuals where we set $\Gamma_{i j k}:=0$ but keep all other parameters fixed. That is not optimal, however, because that immediately implies a dramatic increase in two-car ownership. This is because $\Gamma_{i j k}$ captures two things, the overall level of two-car ownership and the degree of portfolio (anti-) synergy between two specific cars. Thus, we choose instead to re-estimate all parameters except $\left(\Gamma_{1}, \ldots, \Gamma_{5}\right)$, which are fixed at zeros. We then run the same counterfactuals as in Table 7. This exercise answers the following question: What does optimal policy look like when the data is interpreted through the lens of a model without portfolio effects? The restricted estimates can be found in Appendix Table D. 4 while the counterfactuals are in Appendix Table D.5.

Table 8 shows the key excerpts from the comparison. Most importantly, ignoring portfolio effects does not alter the fact that fuel taxes are the most cost-effective tool for reducing $\mathrm{CO}_{2}$ emissions. However, it changes the relative balance of EV incentives vs. CV taxes: In a restricted model without portfolio effects, the policy maker would be roughly indifferent between the two ( 16,856 vs. $16,520 \mathrm{NOK} /$ tonne $)$, while EV incentives are twice as expensive as CV taxes in the full model ( 26,552 vs. 13,481).

Looking at the shifts in ownership portfolios, we see that the single-EV portfolio is too responsive in the model without portfolio effects. Both models agree that the number of cars increased by approximately 3,700 , and they agree that there will be just over 8,000 additional households with the combined (EV,CV)-bundle. However, the model without portfolio effects predicts nearly twice as many households becoming single-EV owners (4,562 vs. 2,773). In other words, the single-EV portfolio is a much closer substitute to the other portfolios in a model without portfolio effects. The net change in the number of cars is roughly the same, but without portfolio effects, 9,563 CV disappear (and get replaced by EV), compared to just 7,299 in the full model. The difference in replacement rates is almost fully explained by the difference in the number of new households becoming single-EV owners (4,562 vs. 2,773 ). With a greater replacement of CV by EV, households are shifted more into zero-emission driving, which implies a greater reduction in $\mathrm{CO}_{2}$.

Another way of seeing the same pattern is to count the replacement ratio: in both models, there are more EVs and fewer CVs in response to the VAT exemption for EVs. However, in the full model, there are 1.52 new EVs per discarded CV, while the restricted model has 1.38 EV per CV. ${ }^{40}$ A ratio of 1.0 would imply a one-for-one replacement of CVs by EVs, which the restricted model gets closer to.

One of the main drivers behind the willingness to switch to single-EV ownership in the model is the range anxiety parameter, $\Gamma_{2}$. This parameter is a utility penalty for owning an EV but no matching CV in the portfolio: i.e. a penalty to the (EV) and (EV,EV) portfolios. This captures the descriptive pattern we saw that EVs are overrepresented in two- relative to one-car portfolios: both in the full sample and narrowly in income decile groups.

It is important to note that a priori, it was possible that households could have purchased the (EV,CV) portfolio but chosen to drive both cars half as much. However, our estimates show that once a household owns two cars, it tends to use both quite a lot, the result is increased local externalities and a smaller reduction in $\mathrm{CO}_{2}$ emissions. ${ }^{41}$ This finding is consistent with Kverndokk et al. (2020) who survey CV owners that have purchased an EV: only $10.8 \%$ state that they end up driving the CV less.

## 8 Conclusion

In this paper, we have examined the effects of EV adoption policies in Norway, and compared them to alternative environmentally motivated car policies. To do so, we have developed a joint model for car ownership and driving where households can own and drive zero, one, or two cars. For two car-portfolios, the model explicitly allows cars to be substitutes or complements on both the extensive (ownership) and intensive (driving) margins.

No cars are complements for all households. However, in some groups of households, we estimate strong synergies between EVs and CVs. In particular, we find that the small EV and the large, new diesel car are complements to 16 percent of households in our sample (in the sense of a negative cross-price elasticity).

Portfolio synergies arise from two channels: first, households are able to specialize their driving by choosing cars suited for different trip types. Our structural model

[^23]explicitly accounts for this, and with the example from above, the household would shift more driving towards the large diesel car when it is complemented with the small EV compared to when it is owned alone.

The second source of synergy is through the ownership utility. Households are reluctant to choose car portfolios with EVs if the portfolio does not include a CV as well. We interpret this as broadly consistent with the popular notion of "range anxiety," whereby consumers are overly focused on the limited range of an EV even though it will not affect their daily driving. This explains why we tend to see EVs over-represented in two-car portfolios relative to one-car portfolios.

These forms of synergy are important for EV adoption policies, because they to a certain extent make EVs come as an addition to an existing CV rather than as a replacement for it. So while the new EV would not contribute to global pollution, it still results in local externalities like congestion, accidents, and noise. Such externalities are sizeable for the sub-population of marginal EV adopters, which tend to live in or around urban areas. Moreover, complementarity will increase the regressivity of EV policies, since two-car ownership is increasing in income.

Counterfactual simulations indicate that if portfolio effects are not taken into account, the abatement cost per tonne of $\mathrm{CO}_{2}$ will be under-estimated by around $30 \%$. In contrast, simple fuel taxes discourages driving in CVs, which simultaneously promotes EVs and targets both local and global externalities and is furthermore progressive.

Table 5: Elasticities (computed for the 2017 cross-section)

|  | Baseline | Demographics |  | Fuel prices |  |  | Purchase prices |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Income | WD | CV | EV | Both | CV | EV | Both |
| Car Portfolio Shares |  |  |  |  |  |  |  |  |  |
| $\operatorname{Pr}(0$ cars $)$ | 0.272 | -0.79 | -0.03 | 0.21 | 0.00 | 0.21 | 0.35 | 0.01 | 0.36 |
| $\operatorname{Pr}(1$ car $)$ | 0.416 | -0.01 | -0.04 | -0.00 | 0.00 | -0.00 | 0.01 | 0.01 | 0.01 |
| $\operatorname{Pr}(2$ cars $)$ | 0.312 | 0.70 | 0.07 | -0.18 | -0.00 | -0.18 | -0.31 | -0.02 | -0.33 |
| Number of Cars by Fuel Type |  |  |  |  |  |  |  |  |  |
| Number of cars | 1.040 | 0.42 | 0.03 | -0.11 | -0.00 | -0.11 | -0.18 | -0.01 | -0.19 |
| - EV | 0.025 | 0.45 | 0.08 | 0.21 | -0.07 | 0.14 | 0.35 | -0.97 | -0.62 |
| - diesel | 0.474 | 0.53 | 0.06 | -0.13 | 0.00 | -0.13 | -0.31 | 0.02 | -0.29 |
| - gasoline | 0.542 | 0.31 | 0.00 | -0.11 | 0.00 | -0.10 | -0.10 | 0.02 | -0.09 |
| Driving |  |  |  |  |  |  |  |  |  |
| Driving (km/year) | 14,501 | 0.50 | 0.06 | -0.13 | -0.00 | -0.14 | -0.23 | -0.01 | -0.24 |
| - EV | 378 | 0.47 | 0.11 | 0.23 | -0.08 | 0.15 | 0.40 | -1.03 | -0.63 |
| - Diesel | 7,880 | 0.59 | 0.08 | -0.15 | 0.00 | -0.15 | -0.34 | 0.02 | -0.32 |
| - Gasoline | 6,242 | 0.39 | 0.03 | -0.14 | 0.00 | -0.14 | -0.14 | 0.02 | -0.13 |
| Taxes and Costs (NOK) |  |  |  |  |  |  |  |  |  |
| Total tax revenue | 22,979 | 0.84 | -0.01 | 0.42 | 0.00 | 0.42 | 0.56 | 0.09 | 0.64 |
| - Fuel | 6,566 | 0.47 | 0.05 | 1.61 | 0.01 | 1.62 | -0.19 | 0.01 | -0.18 |
| - Registrations | 14,224 | 1.10 | -0.04 | -0.05 | 0.00 | -0.05 | 1.01 | 0.12 | 1.14 |
| - Toll | 1,280 | 0.19 | 0.00 | -0.08 | 0.01 | -0.07 | -0.13 | 0.09 | -0.04 |
| - Annual | 909 | 0.41 | 0.03 | -0.11 | 0.00 | -0.11 | -0.20 | 0.01 | -0.18 |
| Total ownership cost ${ }^{a}$ | 32,740 | 0.53 | 0.02 | -0.10 | -0.00 | -0.10 | 0.24 | 0.01 | 0.24 |
| Externalities |  |  |  |  |  |  |  |  |  |
| Local externalities (NOK) ${ }^{\text {b }}$ | 10,820 | 0.56 | 0.05 | -0.13 | -0.00 | -0.13 | -0.23 | -0.01 | -0.24 |
| CO2 emissions (tonnes) | 2.209 | 0.49 | 0.06 | -0.16 | 0.00 | -0.16 | -0.20 | 0.02 | -0.19 |

Notes: All outcomes are per household per year, and monetary outcomes are in 2015 NOK.
${ }^{a}$ : Total ownership cost includes all expected monetary spending except fuel, i.e. the "rental price" (purchase less depreciated resale price), annual tax, and toll payments.
${ }^{b}$ : Local externalities include noise, congestion, accidents, infrastructure, and local air pollution. See Appendix A.4.3.

Table 6: Portfolio Synergies, 2017


Notes: The matrix displays the average portfolio synergy as measured by the average supermodularity of utility, $\Delta_{i j k}$ in Equation (12). A higher value (closer to zero) means stronger synergy between the two cars.

Table 7: Counterfactual simulations

|  | (1) | (2) | (3) |
| :---: | :---: | :---: | :---: |
| Exogenous policy variables |  |  |  |
| Targeted segment | EV | CV | CV |
| Tax instrument | Purchase | Fuel | Purchase |
| Effect on consumer price | -20.0\% | +2.06\% | +1.62\% |
| Welfare effects (annual, per household) |  |  |  |
| $\Delta \mathrm{CO}_{2}(\mathrm{~kg})$ | -7.266 | -7.266 | -7.266 |
| $\Delta$ Consumer surplus (NOK) | 172.365 | -215.051 | -323.603 |
| $\Delta$ Taxes (NOK) | -337.859 | 179.273 | 185.206 |
| $\Delta$ Local externalities (NOK) | 27.437 | -29.261 | -40.443 |
| Abatement cost ( NOK per kg CO 2 ) | -26.552 | -0.897 | -13.481 |
| Number of cars |  |  |  |
| Cars | 3,769.3 | -5,025.7 | -6,713.4 |
| - EV | 11,068.3 | 229.1 | 303.7 |
| - CV | -7,299.0 | -5,254.9 | -7,017.1 |
| Households by Portfolio Choice |  |  |  |
| No car | -1,189.2 | 2,545.0 | 3,310.4 |
| EV | 2,772.5 | 103.5 | 143.2 |
| EV,EV | 117.6 | 3.9 | 5.3 |
| CV | -4,163.5 | -167.7 | -50.6 |
| CV,CV | -5,598.2 | -2,602.5 | -3,558.2 |
| CV,EV | 8,060.7 | 117.9 | 149.9 |
| Driving (expected average percentage changes) |  |  |  |
| Total driving | 0.253 | -0.278 | -0.376 |
| EV driving | 27.944 | 0.478 | 0.644 |
| Diesel driving | -0.313 | -0.304 | -0.542 |
| Gasoline driving | -0.341 | -0.289 | -0.228 |

Note: Evaluated on the 2017 cross-section of our dataset.

Table 8: Counterfactual Results With and Without Portfolio Effects

| Targeted segment | Full model ( $\Gamma_{i j k} \neq 0$ ) |  |  | Restricted model ( $\Gamma_{i j k}=0$ ) |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | EV | CV | CV | EV | CV | CV |
| Tax instrument | Purchase | Fuel | Purchase | Purchase | Fuel | Purchase |
| Abatement cost per tonne | -26,552 | -897 | -13,481 | -16,856 | -785 | 16,520 |
| Number of cars (change) | 3,769 | -5,026 | -6,713 | 3,636 | -6,633 | -8,818 |
| Household with EV | 2,773 | 104 | 143 | 4,562 | 185 | 254 |
| Households with EV,CV | 8,061 | 118 | 150 | 8,304 | 163 | 214 |

Note: Showing excerpts from Tables 7 and D.5.

## References

Adamou, Adamos, Sofronis Clerides, and Theodoros Zachariadis, "Welfare Implications of Car Feebates:A Simulation Analysis," The Economic Journal, 2013.

Anderson, Soren T and James M Sallee, "Designing policies to make cars greener," Annual Review of Resource Economics, 2016, 8, 157-180.

Archsmith, James, Alissa Kendall, and David Rapson, "From cradle to junkyard: assessing the life cycle greenhouse gas benefits of electric vehicles," Research in Transportation Economics, 2015, 52, 72-90.
_ , _, and _ , "From cradle to junkyard: assessing the life cycle greenhouse gas benefits of electric vehicles," Research in Transportation Economics, 2015, 52, 72-90.
_ , Kenneth Gillingham, Christopher R Knittel, and David S Rapson, "Attribute Substitution in Household Vehicle Portfolios," Technical Report 2020.

Bento, Antonio M, Lawrence H Goulder, Mark R Jacobsen, and Roger H Von Haefen, "Distributional and efficiency impacts of increased US gasoline taxes," The American Economic Review, 2009, pp. 667-699.

Borger, B. De, I. Mulalic, and J. Rouwendal, "Substitution between cars within the household," Transportation Research A, 2016, 85, 135-156.

Busse, Meghan R, Christopher R Knittel, and Florian Zettelmeyer, "Are Consumers Myopic? Evidence from New and Used Car Purchases," The American Economic Review, 2013, 103 (1), 220-256.

Chandra, Ambarish, Sumeet Gulati, and Milind Kandlikar, "Green drivers or free riders? An analysis of tax rebates for hybrid vehicles," Journal of Environmental Economics and Management, 2010, 60 (2), 78 - 93.

Clinton, Bentley C. and Daniel C. Steinberg, "Providing the Spark: Impact of financial incentives on battery electric vehicle adoption," Journal of Environmental Economics and Management, 2019, 98.

DeShazo, J.R., Tamara L. Sheldon, and Richard T. Carson, "Designing policy incentives for cleaner technologies: Lessons from California's plug-in electric vehicle rebate program," Journal of Environmental Economics and Management, 2017, 84, 18-43.

Dubin, J.A. and D.L. McFadden, "An econometric analysis of residential electric appliance holdings and consumption," Econometrica: Journal of the Econometric Society, 1984, pp. 345-362.

Durrmeyer, Isis and Mario Samano, "To rebate or not to rebate: Fuel economy standards versus feebates," The Economic Journal, 2018, 128 (616), 3076-3116.

Ershov, Daniel, Jean-William Laliberté, Mathieu Marcoux, and Scott Orr, "Estimating complementarity with large choice sets: An application to mergers," Working Paper, 2021.

Figenbaum, Erik and Marika Kolbenstvedt, "Learning from Norwegian battery electric and plug-in hybrid vehicle users: Results from a survey of vehicle owners," TØI report 1492/2016, 2016.

Fridstrøm, Lasse, "Dagens og morgendagens bilavgifter," TØI report 1708/2019, 2019. Norwegian only (English summary).

Gallagher, Kelly Sims and Erich Muehlegger, "Giving green to get green? Incentives and consumer adoption of hybrid vehicle technology," Journal of Environmental Economics and Management, 2011, 61 (1), 1 - 15.

Gentzkow, Matthew, "Valuing new goods in a model with complementarity: Online newspapers," American Economic Review, 2007, 97 (3), 713-744.

Gillingham, K., "Selection on Anticipated Driving and the Consumer Response to Changing Gasoline Prices," Working paper, 2012.

Gillingham, Kenneth and Anders Munk-Nielsen, "A tale of two tails: Commuting and the fuel price response in driving," Journal of Urban Economics, 2019, 109, 27-40.
_ and James H. Stock, "The Cost of Reducing Greenhouse Gas Emissions," Journal of Economic Perspectives, November 2018, 32 (4), 53-72.
_, Fedor Iskhakov, Anders Munk-Nielsen, John P Rust, and Bertel Schjerning, "Equilibrium trade in automobiles," Journal of Political Economy, 2022, 130 (10).
_, _, _, John Rust, and Bertel Schjerning, "A Dynamic Model of Vehicle Ownership, Type Choice, and Usage," Working Paper, 2015.

Gillingham, Kenneth T, Sébastien Houde, and Arthur A Van Benthem, "Consumer myopia in vehicle purchases: Evidence from a natural experiment," American Economic Journal: Economic Policy, 2021, 13 (3), 207-38.

Grigolon, Laura, Mathias Reynaert, and Frank Verboven, "Consumer valuation of fuel costs and tax policy: Evidence from the European car market," American Economic Journal: Economic Policy, 2018, 10 (3), 193-225.

Grzybowski, Lukasz and Frank Verboven, "Substitution between fixed-line and mobile access: the role of complementarities," Journal of Regulatory Economics, 2016, 49 (2), 113-151.

Hang, Derrick, Daniel McFadden, Kenneth Train, and Ken Wise, "Is vehicle depreciation a component of marginal travel cost?: a literature review and empirical analysis," Journal of Transport Economics and Policy (JTEP), 2016, 50 (2), 132-150.

Heckman, James J, "Sample selection bias as a specification error," Econometrica: Journal of the econometric society, 1979, pp. 153-161.

Holland, Stephen P, Erin T Mansur, Nicholas Z Muller, and Andrew J Yates, "Are there environmental benefits from driving electric vehicles? The importance of local factors," American Economic Review, 2016, 106 (12), 37003729.

Iaria, Alessandro and Ao Wang, "A note on stochastic complementarity for the applied researcher," Economics Letters, 2021, 199, 109731.

Ito, Koichiro and James M Sallee, "The economics of attribute-based regulation: Theory and evidence from fuel economy standards," Review of Economics and Statistics, 2018, 100 (2), 319-336.

Jacobsen, Mark, "Evaluating U.S. Fuel Economy Standards in a Model with Producer and Household Heterogeneity," American Economic Journal: Economic Policy, 2013, 5(2), 148-187.

Jenn, Alan, Katalin Springel, and Anand R Gopal, "Effectiveness of electric vehicle incentives in the United States," Energy Policy, 2018, 119, 349-356.

Johansen, Bjørn Gjerde, "Measuring Substitution Patterns for Differentiated Products - Demand for Electric Vehicles," Working Paper, available at request, 2020.

Knittel, Christopher R. and Ryan Sandler, "The Welfare Impact of SecondBest Uniform-Pigouvian Taxation: Evidence from Transportation," American Economic Journal: Economic Policy, November 2018, 10 (4), 211-42.

Kverndokk, Snorre, Erik Figenbaum, and Jon Hovi, "Would my driving pattern change if my neighbor were to buy an emission-free car?," Resource and Energy Economics, 2020, 60, 101-153.

Li, Jing, "Compatibility and investment in the us electric vehicle market," Job Market Paper, 2019.

Manski, Charles F and Leonard Sherman, "An empirical analysis of household choice among motor vehicles," Transportation Research Part A: General, 1980, 14 (5-6), 349-366.

Manzini, Paola, Marco Mariotti, and Levent Ülkü, "Stochastic complementarity," The Economic Journal, 2019, 129 (619), 1343-1363.

Muehlegger, Erich and David S Rapson, "Subsidizing mass adoption of electric vehicles: Quasi-experimental evidence from California," Technical Report, National Bureau of Economic Research 2018.

Munk-Nielsen, Anders, "Diesel Cars and Environmental Policy," Working Paper, 2014.

Reynaert, Mathias, "Abatement strategies and the cost of environmental regulation: Emission standards on the European car market," The Review of Economic Studies, 2021, 88 (1), 454-488.

Rødseth, Kenneth Løvold, Paal Brevik Wangsness, Knut Veisten, Katharina Alena Høye, Rune Elvik, Ronny Kløboe, Harald Thune-Larsen, Lasse Fridstrøm, Elizabeth Lindstad, Agathe Rialland, Kristofer Odolinski, and Jan-Eric Nilsson, "Eksterne kostnader ved transport i Norge," TØI report 1704/2019, 2019. Norwegian only (English summary).

Small, K.A. and H.S. Rosen, "Applied Welfare Economics with Discrete Choice Models," Econometrica, 1981, 49 (1), 105-130.

Springel, Katalin, "Network externality and subsidy structure in two-sided markets: Evidence from electric vehicle incentives," American Economic Journal: Economic Policy, 2021, 13 (4), 393-432.

Thomassen, Øyvind, Howard Smith, Stephan Seiler, and Pasquale Schiraldi, "Multi-category competition and market power: a model of supermarket pricing," American Economic Review, 2017, 107 (8), 2308-51.

Wakamori, Naoki, "Portfolio considerations in differentiated product purchases: An application to the Japanese automobile market," University of Mannheim Discussion Papers, 2015, (499).

West, S.E., "Distributional effects of alternative vehicle pollution control policies," Journal of Public Economics, 2004, 88 (3), 735-757.

Xing, Jianwei, Benjamin Leard, and Shanjun Li, "What does an electric vehicle replace?," Journal of Environmental Economics and Management, 2021, 107, 102432.

Yan, Shiyu, "The economic and environmental impacts of tax incentives for battery electric vehicles in Europe," Energy Policy, 2018, 123, 53-63.

## Online Appendix

## A Data

This chapter describes the data and the dataset creation in more detail. Section A. 1 describes the main dataset: the household data including car ownership at the household level. Section A. 2 describes how we define car types in the model. Section A. 3 describes the driving data, and finally Section A. 4 gives an overview of other data sources such as fuel prices, purchase taxes and local external costs.

## A. 1 Car ownership and demographics

## A.1.1 Car ownership

The central motor vehicle register in Norway is an annual data set that contains information about all vehicles registered in the country, including a unique car identifier and unique individual identifiers of the owners that can be matched with other administrative data sources. It also includes several attributes of the cars (further described in Section A.2) as well as dates for important events such as first registration, de-registrations, re-registrations and potentially the scrap date.

First, we select a subset of the vehicles. We omit vehicles that run on alternative fuel types (e.g. hydrogen and kerosene; these are less than 100 vehicles in total). We also omit all vehicles that are not passenger cars with the exception of small vans, such as motorcycles, scooters, tractors, buses, trucks, etc. We only keep vehicles that are 25 years or younger, since the data on ownership and driving patterns for older cars is too irregular. ${ }^{42}$

We omit all vehicles where the owner is a firm rather than an individual (slightly less than 10 percent of the car fleet). Some of these cars are leasing vehicles used by households, but as the household using the car is unobserved, it is not possible to include these vehicles in the data. Leasing cars are typically sold second-hand to households when they are 2-4 years old, and at this point they will reappear in our sample. See Figure B. 1 and related text for more information. Finally, the motor vehicle register also contains information about type of driving each vehicle is

[^24]registered for. We only keep vehicles registered for private transport (meaning that taxis, emergency vehicles, embassy vehicles, military vehicles, hearses, etc. will be omitted, even if the registered owner is an individual rather than a firm).

To create annual observations of car ownership at the household level we stock sample the owners of the cars from the motor vehicle register at a given date at the end of the year (December 31st), and allocate the cars to the corresponding household. Stock sampling has the advantage that we do not have to make assumptions about how to allocate a car between individuals if the car changed owner within a given year.

When defining car ownership, we only consider the two youngest cars in three (or more) vehicle households. Thus, we assume that the third (and fourth, etc.) vehicle is irrelevant for the household's decision of the two-car portfolio. Three-car ownership is at most $0.9 \%$ and often likely to be due to a transition from one two-car portfolio to another. Regardless, this simplification is necessary for the tractability of our model.

## A.1.2 Socio-economic and location data

Our final data set is defined at the household $\times$ year level, where car ownership at the household level is obtained by linking individuals with the same unique household identifier. This is further linked to other socio-economic information from various Norwegian registers, such as the national population register and tax records. In particular, we define the following variables: "couple" is a dummy variable for whether the household consists of one or two adult members; "age" is the average age of the two spuses; and "income" is defined as the sum of labor and capital income net of taxes and transfers for both spouses.

We observe the geographical location of the households' residence at the "basic statistical unit" level (henceforth referred to as "neighborhood") each year. With more than 14,000 neighborhoods in Norway, and an average population of less than 200 households, this is the most detailed geographical classification available. These neighborhoods are typically small in densely populated areas and significantly larger in rural areas.

By linking employers to employees, we also know the neighborhood of individuals' workplaces. For individuals with several employers within a given year, we choose the workplace associated with the highest labor income, given that the neighborhood of the workplace is non-missing.

This information allows us to associate each individual with a work route. We use the road network Elveg, a publicly available data set maintained by The Nor-
wegian Public Roads Administration (NPRA) that includes all drivable roads that are either longer than 50 meters or part of a network. ${ }^{43}$ This allows us to calculate the fastest route along the road network between the road links that are closest to the centroids of each neighborhood; i.e. the route that minimizes driving time according to the speed limit on each road link. We have also obtained information on toll gates in Norway from the NPRA, including coordinates and annual rates. This information is added to the appropriate links in the road network, allowing us to calculate associated toll payments with each work route. ${ }^{44}$ For toll gates with time-differentiated charges we use the price during rush hours, as we are looking at tolls associated with the work trips.

We use the information on routes between neighborhoods to create variables for work distance and toll payments. Work distance is defined as the average of the spouses' work distances, where the work distance is set to zero for individuals that are unemployed. In cases where one of the spouse's workplace is lacking information on neighborhood, we use the other spouse's work distance instead. If all working individuals in a household have missing information on the workplace neighborhood, the work distance variable is set to missing. The variable for toll payments is defined similarly, but we use the sum of spouses' tolls instead of the average. We then create a measure of potential annual toll payments, assuming that spouses drive to work 220 days each year. ${ }^{45}$

We also use the information on neighborhoods to create variables for "living in a city" as well as "working in a city". The first either takes the value zero or one, while the latter will take the value 0.5 if the household is a couple and only one of the spouses work in a city. In this context, "city" is defined as within the boundaries of Oslo, Bergen, Trondheim or Stavanger, the four largest cities in Norway.

The "average toll payment" variable is the cost associated with a one-way work trip. We observe that the use of tolls in Norway is significantly expanded over the years for which we have data; the share of households exposed to tolls on their work trip has increased from 8.5 percent in 2005 to almost 20 percent in 2017.

[^25]Furthermore, the average one-way cost for the exposed households has increased from about 16 NOK in 2005 to 28 NOK in 2017, measured in 2015-NOK. ${ }^{46}$ This is partly due to a general increase in toll levels, and partly due to the introduction of rush hour charges in toll cordons of several major cities, making driving into the city center during congested hours more expensive. Taken together, this means that the average household's work trip exposure to tolls increased from 1.37 NOK in 2005 to 5.63 NOK in 2017 (one way, measured in 2015 kroners). It should be noted that part of the increase in toll levels between 2010 and 2017 is offset by the expansion of electric vehicles, and the fact that they can drive through toll gates at zero cost.

## A.1.3 Sample selection

We impose three main sample selection criteria.
Criterion 1: Some firms have missing information on neighborhoods, meaning that we are not able to create work distance and toll measures for the individuals working there. These households are removed.

Criterion 2: We remove households where the average one-way work distance is above 150 kilometers. These individuals are likely to either work from home, or commute by other modes of transport (several individuals are for instance living in one of the major cities in Norway and working in another about 500 kilometers away; these individuals are more likely to commute by plane and less often than every day). ${ }^{47}$

Criterion 3: As we use the natural logarithm of income as an explanatory variable in the model, we remove all households where the net income is negative.

Table A. 1 shows how the sample gradually drops as these selection criteria are imposed. We are left with almost 26.5 million household-year observations, which is $89 \%$ of the raw data.

[^26]Table A.1: Car ownership sample selection

|  | Raw data | Criterion 1 | Criterion 2 | Criterion 3 |
| ---: | ---: | ---: | ---: | ---: |
| $N$ | $29,731,353$ | $27,355,078$ | $26,634,996$ | $26,470,192$ |
| Share | $100 \%$ | $92.0 \%$ | $89.6 \%$ | $89.0 \%$ |

Note: Criterion 1: work distance observed. Criterion 2: Work distance not greater than 150 km . Criterion 3: Non-negative income.

## A. 2 Car Types and Attributes

The central motor vehicle register in Norway contains information about all vehicles registered in the country, including first registration date, make, model and car attributes such as fuel type, fuel efficiency, engine effect, weight, etc. The vast amount of detail means that we observe several hundreds of thousands of unique car types (i.e. unique bundles of car attributes).

New-car prices: The motor vehicle register does not contain car prices. We have obtained MSRPs ("Manufacturer's Suggested Retail Prices") for November each year from OFV ("Opplysningsrådet for Vegtrafikk") at a similarly detailed level. ${ }^{48}$ Prices are merged to the motor vehicle register using a fuzzy matching procedure that compares car attributes as well as strings for make and model names across the two data sets. Prices are then added manually for the most common vehicles where the fuzzy matching fails. This procedure found a matching price for about 96 percent of all cars. As the number of electric vehicle models available between 2011 and 2017 was relatively low, we went through all the models manually to ensure that all electric vehicle prices are correct. We have also added "range" for electric vehicles manually, as this attribute is not included in the motor vehicle register. ${ }^{49}$

Used-car prices: Unfortunately, we do not have any information on transaction prices for used cars. Therefore we use a depreciated value of the new car price, where the depreciation rates are based on The Norwegian Tax Authority's evaluation of used cars. ${ }^{50}$ These rates are used to calculate the registration tax for an imported used car. This rate is on average $12.5 \%$ annually. This is very close to the average rates observed by Gillingham et al. (2015) in suggested used-car prices from the Danish Automobile Dealer Association, and those rates have very low dispersion

[^27]Table A.2: Car attributes by car type, 2017

| Car type | Fuel type | Age | Weight | Effect | Range | $\mathrm{CO}_{2}$ | Ownership cost | Driving cost |
| :--- | :--- | ---: | :--- | ---: | ---: | ---: | ---: | :--- |
| 1: Small, new | Gasoline | 0.52 | 1.27 | 99.9 | - | 114.3 | 138,499 | 0.684 |
| 2: Small, new | Diesel | 0.56 | 1.48 | 105.4 | - | 126.7 | 178,809 | 0.619 |
| 3: Large, new | Gasoline | 0.51 | 1.80 | 144.3 | - | 81.7 | 236,803 | 0.489 |
| 4: Large, new | Diesel | 0.53 | 1.84 | 140.1 | - | 143.1 | 273,454 | 0.702 |
| 5: Small, 1-4 years | Gasoline | 3.29 | 1.12 | 78.7 | - | 115.5 | 82,327 | 0.683 |
| 6: Small, 1-4 years | Diesel | 3.50 | 1.40 | 90.0 | - | 122.7 | 106,544 | 0.595 |
| 7: Large, 1-4 years | Gasoline | 2.82 | 1.51 | 115.9 | - | 114.0 | 132,223 | 0.682 |
| 8: Large, 1-4 years | Diesel | 3.28 | 1.75 | 120.6 | - | 156.7 | 166,340 | 0.763 |
| 9: Small, 5-11 years | Gasoline | 8.51 | 1.05 | 65.8 | - | 137.4 | 62,780 | 0.802 |
| 10: Small, 5-11 years | Diesel | 8.32 | 1.37 | 79.4 | - | 137.3 | 74,863 | 0.654 |
| 11: Large, 5-11 years | Gasoline | 8.69 | 1.41 | 105.2 | - | 170.0 | 83,933 | 0.991 |
| 12: Large, 5-11 years | Diesel | 8.56 | 1.76 | 112.2 | - | 188.0 | 102,529 | 0.903 |
| 13: Small, 12+ years | Gasoline | 17.06 | 1.09 | 68.3 | - | 171.0 | 55,120 | 0.977 |
| 14: Small, 12+ years | Diesel | 15.00 | 1.42 | 83.1 | - | 166.7 | 63,814 | 0.790 |
| 15: Large, 12+ years | Gasoline | 16.75 | 1.45 | 109.6 | - | 221.0 | 62,954 | 1.271 |
| 16: Large, 12+ years | Diesel | 16.03 | 1.90 | 98.3 | - | 234.2 | 70,616 | 1.126 |
| 17: Small EV, new | Electric | 0.51 | 1.46 | 96.3 | 268.1 | 0.0 | 118,818 | 0.155 |
| 18: Small EV, used | Electric | 3.02 | 1.40 | 84.5 | 213.7 | 0.0 | 80,625 | 0.155 |
| 19: Large EV, new | Electric | 0.30 | 2.26 | 301.2 | 494.4 | 0.0 | 296,614 | 0.183 |
| 20: Large EV, used | Electric | 2.89 | 2.13 | 273.3 | 450.7 | 0.0 | 162,841 | 0.183 |

Notes: Weight is in tonnes, engine effect is in kW , range in $\mathrm{km}, \mathrm{CO}_{2}$ emissions in $\mathrm{g} / \mathrm{km}$, and driving cost in $2015 \mathrm{NOK} / \mathrm{km}$ (based on fuel efficiency and fuel price). Ownership costs are computed for three years of ownership and discounted (at annual rate 0.95), comprising "rental cost" (purchase price minus depreciated resale value), annual taxes, repair and maintenance costs.
across car types.
Car attributes: The attributes from the motor vehicle register we use in the model include "engine effect", measured in 100 kW ; "weight", measured in tonnes; "fuel efficiency", measured in liter/kWh per kilometer; and "car age", measured in years and calculated as the time interval from the first registration date of the vehicle until the date where we stock sample car ownership (December 31st each year). As "weight" tends to work as a proxy for car quality in discrete choice models, and electric vehicles are heavier due to the battery, we subtract the weight of the battery for electric vehicles ( $30 \%$ of the weight).

Table A. 2 shows the attributes of each of the 20 car types for 2017. Note that the characteristics vary over time depending on which cars falls into the categories in each year. Furthermore, the EV only started becoming available from 2011 and onwards.

## A. 3 The Driving Dataset

## A.3.1 Periodic vehicle inspections

Car owners in Norway are required to conduct a periodic vehicle inspection before four years have passed since the date of the first registration. For subsequent vehicle inspections, not more than two years can have passed from the previous inspection. For cars that are second-hand imported and more than three years old, the deadline for the first inspection is within one year since the car was first registered in Norway. ${ }^{51}$ It is illegal to drive in a car that has not been to the periodic vehicle inspection within the required deadline: if the driver of such a car is stopped in a police control, the car's license plates are removed.

The control has two purposes: first, to check that the car is safe to use. Second, to check that noise and local pollutants are not above the allowed threshold values. If the car is not approved, the owner is responsible for conducting required repairs and do a second inspection within two months. If the car does not pass a new inspection within the two months, it is illegal to use.

For our purpose, the most important aspect of the periodic inspections is that odometer readings are recorded in a register that is linked to the motor vehicle register through a unique car identifier. This means that we are able to merge car and owner characteristics to each odometer reading.

## A.3.2 Driving periods

We use odometer readings to create a separate data set containing car specific driving periods. The length of a driving period is the time between two subsequent periodic inspections ( $\approx 2$ years), and the mileage during a driving period is the difference between odometer readings. For each car's first driving period, the length will be the time interval between the first registration and the first periodic inspection of the car ( $\approx 4$ years), and the mileage will be the odometer reading at the first inspection.

We only keep driving periods from the cars in our main sample (see Section A.1.1) where the car is owned by households from our main sample (see Section A.1.3) for at least a part of the driving period. This leaves about 12.6 million driving periods.

We then drop all driving periods with start dates earlier than January 1st 2005 ( $\approx 2.1$ million observations) for two reasons. First, our main data set on car ownership starts in 2005. Second, odometer readings were not recorded in Norway prior

[^28]to 2005 , meaning that the first driving period for the 2005 car fleet will span all years from the car is first registered in Norway.

The driving data is slightly more noisy than the car ownership data and includes some odometer readings that are obviously wrong, resulting in some cars being reported as driven tens of thousands of kilometers each day. To remove these outliers, we drop all driving periods where the first difference of odometer readings is either negative or above 200 kilometers per day. This removes about 80,000 driving periods $(\approx 0.8$ percent), and leaves a sample of about 10.5 million observations.

## A.3.3 Driving sub-periods and sample selection

Definition (driving sub-period): We define driving sub-periods as the part of a driving period where (a) the household that owns the car do not change and (b) the ownership status for other vehicles owned by the same household do not change.

Sub-periods are thus characterized by unique owners with unique car portfolios. As in Section A.1, the car portfolio for households owning more than two cars will consist of the two cars of the most recent vintage. The driving periods are split in 21.8 million unique sub-periods, corresponding to 2.1 sub-periods per driving period in average.

The household data described in section A.1.2 is annual, while sub-periods can be of arbitrary length, from a single day to multiple years. Thus, we use weighted averages of the annual household data, where the weights correspond to the amount of the driving sub-period that falls within each year. For example, if a sub-period spans from the middle of 2010 to the middle of 2012, the household specific demographic data for 2011 is weighted by 0.5, while the 2010 and 2012 values are weighted by 0.25 each.

While the car attributes described in Section A. 2 needs to be grouped by car types as they will be used in a discrete choice model, the observed driving is conditional on car ownership. This means that we can use the actual car attributes in the driving sub-likelihood rather than the car type specific averages to preserve more of the heterogeneity in the data. All car attributes except "age" are time-invariant. For the age of the car, we use the age at the middle point of the driving sub-period.

Finally, we add the fuel/energy price of the type that corresponds to the car (diesel, gasoline or electricity, see Section A.4). Diesel and gasoline prices are observed at a monthly level, while electricity prices are observed quarterly. As with the household data we use the weighed average of the prices that are contained in the driving sub-period. Note that the randomness of the timing of periodic vehicle
inspections will introduce random variation in the average fuel price a household is faced with during a driving sub-period.

Next, we impose as set of sample selection criteria that reduce the number of driving periods from $10,532,123$ to $6,088,227$, or by about 42 percent. In the following sections, the sample selection criteria will be described more in detail.

Criterion 1: As we are using actual car attributes rather than car type averages, we require that all car attributes are observed in the motor vehicle register. Moreover, for sub-periods with two-car portfolios, we require that the attributes of the additional car the household owns must be observed. The most problematic attribute is "fuel efficiency", which is missing for a significant share of the older cars in first years of the data.

Criterion 2: We remove driving periods where the length is not within the interval [ 2 years minus 150 days, 4 years plus 150 days] as we consider these to be observations of an unnatural length.

Criterion 3: We remove driving periods where less than 70 percent of the driving period is covered by valid sub-periods. This may happen because the owner of the car is unobserved for a significant part of the driving period, or excluded from our sample. It is for instance common for company-owned cars to be bought second-hand by households.

For each driving period, we calculate driving in kilometers per day. For each sub-period we calculate the percent of the total permissible days of driving that this sub-period makes up, so that the sum over weights across sub-periods belonging to the same driving period is $100 \%$.

Table A.3: Driving sub-periods, sample selection

|  | Raw data | Criterion 1 | Criterion2 | Criterion 3 |
| :--- | ---: | ---: | ---: | ---: |
| Sub-periods | $21,763,490$ | $15,932,834$ | $14,905,140$ | $13,403,807$ |
| Share of raw data | $100 \%$ | $73.2 \%$ | $68.5 \%$ | $61.6 \%$ |
| Avg. km/day | 36.5 | 38.4 | 38.1 | 37.3 |

Note: Criterion 1: period startes before 1 Jan 2005 or car attributes missing. Criterion 2: unnatural inspection timing. Criterion 3: more than $70 \%$ of the overall driving period has been dropped to preceding criteria or is unaccounted for.

Table A. 3 shows how the number of valid sub-periods is reduced as the three sample selection criteria are imposed. The last rows contain all vehicle types, while the rows above separates between gasoline, diesel and electric vehicle sub-periods. For each of these groups, three statistics are displayed: the number of sub-periods, the share they constitute of the raw data as well as the average kilometers driven per day within each group. As the first periodic inspection is conducted when the
car is approximately four years old, we have a relatively small sample of electric vehicle observations. ${ }^{52}$

## A. 4 Other data sources

## A.4.1 Purchase taxes

Purchase taxes apply to new vehicles with diesel or gasoline driven internal combustion engines (including hybrids), and consist of VAT ( 25 percent) as well as a one-off registration tax. The latter has several components that depend on the attributes of the car: weight, engine effect, cylinder volume, $\mathrm{CO}_{2}$ emissions and $\mathrm{NO}_{x}$ emissions. The registration tax is a piece-wise linear function of these attributes.

Using the annual tax rates set by the government and car specific attribute values from the central motor vehicle register, we can calculate the registration tax for each car. Using the MSRPs, we also calculate the VAT.

There has been significant changes in which attributes that are included in the registration tax, as well as changes in the attribute specific rates and location of kinks in our sample period 2005-2017. Changes are politically motivated by wanting to increase the weight of the $\mathrm{CO}_{2}$ component relative to the other components, and this provides price variation over time as well as across vehicles with different attributes. Tax rates as a function of a car's attributes are displayed in Figure A.1.

In 2005, the registration tax included components for weight, engine effect and cylinder volume. The cylinder volume component was replaced by a $\mathrm{CO}_{2}$ component in 2007. In 2009, the $\mathrm{CO}_{2}$ component was changed to a feebate component, where rebates were given to cars with type approved $\mathrm{CO}_{2}$ emissions of less than 120 grams per kilometer, while rates for emissions above this threshold increased. The $\mathrm{NO}_{x}$ component was introduced in 2010, but is small in magnitude compared to the other components. The $\mathrm{CO}_{2}$ component has been gradually increased each year, with comparable reductions in the engine effect component, and in 2017 the latter was completely phased out.

For hybrid vehicles, some additional clauses apply. First, the engine effect component is only calculated based on the internal combustion engine. Second, the weight component is calculated without the weight of the battery. The deduction for battery weight was set to 10 percent of the vehicle weight before 2010. From

[^29]Figure A.1: Components of the registration tax, selected years.


Notes: All tax components are displayed in 2015-kroners. Technical attributes are displayed on the same scale, the units are displayed in the legend. The registration tax for a new vehicle is the sum of all components.

2010, it was increased to 26 percent for plugin hybrids. In 2017, it was reduced from 10 percent to 5 percent for non-plugin hybrids.

On average, taxes constitute almost 50 percent of the final consumer price for new diesel and gasoline cars. The progressive nature of the registration tax means that large cars with high fuel consumption are taxed more heavily. The change over time in the $\mathrm{CO}_{2}$ component relative to engine effect and cylinder volume has favored diesel vehicles over gasoline vehicles. The feebate aspect of the $\mathrm{CO}_{2}$ component has also reduced the registration tax for most hybrids and some small diesel vehicles to zero or close to zero - this will happen in cases where the $\mathrm{CO}_{2}$ rebate completely offsets the other components. New car prices and tax components for new small/large diesel/gasoline vehicles are displayed in Figure A.2. The decline in the registration tax for large gasoline cars partly reflects that this group has an increasing share of hybrids.

Finally, we want to show that although price endogeneity is a common problem when working with disaggregate data, it is less problematic in our setting when

Figure A.2: Prices and tax components for new ICEV car types over time.


Notes: All tax components are displayed in 2015 NOK. Car types correspond to car types 1-4 from Section A.2.
considering aggregate car types. In Table A.4, new car prices excluding taxes and the registration tax for ICEVs are regressed on the relevant car attributes included in the structural model (EVs are excluded here, as their tax rate is zero). The first columns display results for new cars on the car model level, while the last columns display results for the four groups of new ICEVs included in the model. All regressions are non-weighted by sales and include all available new ICEVs during the period 2005-2017. The first columns show, as expected, that a large portion of price variation is unaccounted for in the disaggregate. The latter columns however show that while weight, engine effect and the diesel dummy explain $96 \%$ of the variation in producer prices, they only explain $71 \%$ of the variation in the registration tax. This has two major implications: First, by aggregating cars into car types, virtually all unobserved variation in producer prices are removed. Second, even though the registration tax is a function of the attributes of the vehicle there is enough variation stemming from changes over time in the tax rates (see Figure A.1) to provide residual price variation for aggregate car models. ${ }^{53}$

[^30]Table A.4: Regressions of Car Price or Tax on Characteristics

| Weight (tonnes) | Disaggregate car models |  |  |  | Aggregate car models |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Price ex tax |  | Registration tax |  | Price ex tax |  | Registration tax |  |
|  | $-24.45^{* * *}$ | (2.138) | 231.80 *** | (1.571) | 148.70*** | (24.97) | 391.40*** | (72.51) |
| Effect (100 kw) | 239.04*** | (1.334) | 203.34*** | (0.981) | $98.13^{* *}$ | (25.47) | $-211.52^{* * *}$ | (73.95) |
| Diesel (dummy) | 41.61*** | (1.812) | $-14.84^{* * *}$ | (1.332) | -6.214 | (5.741) | -25.003 | (16.67) |
| Constant | $-29.28^{* * *}$ | (2.838) | -377.04*** | (2.086) | -100.20*** | (12.78) | -177.38*** | (28.05) |
| Observations | 23,777 |  | 23,777 |  | 52 |  | 52 |  |
| Mean of $y$ | 255.10 |  | 263.16 |  | 200.08 |  | 127.64 |  |
| Adjusted $\mathrm{R}^{2}$ | 0.661 |  | 0.863 |  | 0.956 |  | 0.709 |  |
| RMSE | 105.74 |  | 77.73 |  | 11.43 |  | 33.20 |  |

Notes: Standard errors in parantheses. Outcome variables are measured in 1000 NOK (2015-kroners). Observations are new ICEV car types available in the period 20052017. Left: Before aggregation. Right: after aggregation (car types 1-4; see Section A.2). Regressions are not weighted by sales, explaining why the average values of the dependent variables does not match when comparing the disaggregate car models to the aggregate car models.

## A.4.2 Fuel and energy prices

There are three main tax components for diesel and gasoline in Norway. First, a $\mathrm{CO}_{2}$ component. Second, a road use component, with the purpose of capturing all other externalities than $\mathrm{CO}_{2}$, such as noise and local pollutants, wear and tear of roads, congestion and accidents. ${ }^{54}$ These two components are diesel- and gasoline specific flat rates that are adjusted annually.

We do not have access to fuel prices exhibiting geographical variation. However, the fact that we observe fuel prices that vary by month, and the fact that the timing of driving periods vary almost arbitrarily means that we have individual level variation in fuel price exposure. Note also that variation in fuel efficiency across car models means that the driving cost per kilometer will be different for different cars even within the same time period.

Figure A. 3 displays how the CPI-adjusted gasoline and diesel prices as well as the different tax components have evolved over the sample period in 2015-kroners. The tax components for gasoline has been higher than the tax components for diesel for all months in the data set, resulting in a gasoline price that is higher per liter.

In addition to the 25 percent VAT, there is a flat tax rate on electricity as well. This is meant to incentivize households to be more energy efficient, but is significantly lower than the rates on diesel and gasoline measured as percent of the final consumer price. For electricity we only have access to prices on quarterly

[^31]Figure A.3: Diesel and gasoline price components

Panel A: Gasoline price


Panel B: Diesel price


Figure A.4: Electricity and fuel prices
Panel A: Electricity price
Panel B: Fuel prices


intervals, displayed in the left panel of Figure A.4. The right panel contrasts this to the final diesel and gasoline price during the same period.

National diesel, gasoline and electricity prices are publicly available through Statistics Norway, and can be accessed at their website.

## A.4.3 Local external costs of driving

To predict the local externalities of driving, we use marginal damage costs quantified by The Institute of Transport Economics in Norway in 2019 for Norwegian conditions. Damage costs are differentiated by vehicle type and the population density in the area where driving occurs, and reported separately for $\mathrm{CO}_{2}$ emissions, local emissions, noise, congestion, accidents and wear-and-tear of road infrastructure. We use all components except $\mathrm{CO}_{2}$ emissions, and the resulting external costs per kilometer are reported in Table A.5.

Table A.5: Local externalities of driving, 2015 NOK per kilometer

| Car type | Area | Local <br> pollutants | Noise | Congestion | Accident | Infra- <br> structure | Total |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| CV | Large urban area | 0.396 | 0.297 | 1.332 | 0.108 | 0.027 | 2.160 |
| CV | Other | 0.041 | 0.174 | 0.112 | 0.108 | 0.027 | 0.461 |
| EV | Large urban area | 0.206 | 0.297 | 1.332 | 0.108 | 0.027 | 1.980 |
| EV | Other | 0.016 | 0.174 | 0.112 | 0.108 | 0.027 | 0.437 |

Notes: Local externalities are obtained from (Rødseth et al., 2019) and converted to 2015kroners. We assume that households located in Oslo, Bergen, Trondheim or Stavanger live in large urban areas ( $25 \%$ of households). For the "other" category, we assume that $59 \%$ of the remaining population live in areas with 15,000 inhabitants or more, while the rest of the population live in rural areas. This is in accordance with Statistics Norway's numbers on population density.

## B Additional descriptives

Section B. 1 presents additional descriptives in the form of various figures and linear regressions. Section B. 2 presents linear regressions to illustrate the conditional relationship between toll payments and EV ownership.

## B. 1 Car ownership and driving

Figure B.1: New car purchases by year, propulsion system and owner type.


Figure B. 1 is based on aggregate new car sales records, and illustrates two things. First, the increasing share of hybrid vehicles over time. For the car types used in the structural model, hybrid vehicles are included in the diesel and gasoline segments. Second, the increase in the number of new cars that are bought by companies. Company-owned cars are typically sold second-hand to households when they are 2-4 years old, implying that even though the share of new cars bought by companies is high, the share of the car fleet that is owned by companies is still low (less than

Figure B.2: Car ownership by income and work distance deciles, 2017


$\square 0$ cars $-\quad 1$ car $-\quad 2$ cars


Notes: The first work distance group contains all households with zero work distance ( $40 \%$ ). Remaining groups are equally large. The y-axis indicates the average of different variables within each bin. The top row displays the share of households owning 0,1 and 2 cars. The bottom row displays average number of cars of by fuel type owned by households within each bin.
ten percent in 2017).
Figures B. 2 and B. 3 show raw means of car ownership (extensive margin) and driving (intensive margin) respectively, by deciles of income (left) and work distance (right), while Figure B. 4 illustrates the distribution of driving across all households and years.

Table B. 1 shows the market shares for each of the 231 car portfolios in 2017. Table B. 2 illustrates what the market shares of two-car portfolios would have been under independence. Finally, Tables B. 3 and B. 4 present results from linear regressions on driving and car ownership respectively, to illustrate conditional correlations in the data.

Figure B.3: Driving by income and work distance deciles.


Note: The first work distance group contains all households with zero work distance (40\%). Remaining groups are equally large. The y-axis indicates kilometers per day, household and car. The top row splits driving between 1 and 2 car households. The bottom row displays driving by fuel type. Driving is driving period weighted averages within each group, where the weights correspond to the number of days of the driving period that the car is owned by each household. Note that some of the bins in the figure will have few observations: there are for instance few electric vehicles and two-car households in the lower income and work distance groups.

Figure B.4: Distributions of driving.
A: Driving by number of cars owned
B: Driving by car type, new cars


Notes: Both panels show driving per car per day. The left panel displays a histogram of driving for one- and two-car households. The right panel shows kernel densities of driving per day by fuel type. As driving decreases with car age, the right panel only includes driving periods between the first registration and the first safety inspection (i.e. $\approx 0-4$ year old cars). Both panels are weighted by the length of the driving period.

Table B.1: Portfolio Market Shares in 2017


Notes: Each cell shows the share of households in 2017 that own the corresponding car portfolio. This uses the Ownership Dataset.

Table B.2: Predicted Portfolio Market Shares Under Independence


Notes: Each cell shows the predicted probability of the bundle, $(j, k)$, denoted $\hat{s}_{j k}$, computed as $\hat{s}_{j j}=s_{j}^{2}$, where $s_{j \mid 1} \equiv \frac{s_{j}}{\sum_{k \in \mathcal{J}} s_{k}}$ is the observed frequency of one-car purchase, and $\hat{s}_{j k}=2 s_{j} s_{k}$ for $j \neq k$ (since we do not distinguish between $(j, k)$ and $(k, j)$ ).

Table B.3: Linear regression coefficients, driving.

| Dependent variable: | (1) | (2) | (3) | (4) |
| :---: | :---: | :---: | :---: | :---: |
| Twocar dummy | -0.0111 | $0.173^{* * *}$ | $0.188^{* * *}$ | $5.545^{* * *}$ |
|  | (0.0146) | (0.0141) | (0.0141) | (0.0389) |
| Household income (log) | 1.782*** | -0.604*** | -0.594*** | $0.224^{* *}$ |
|  | (0.0181) | (0.0171) | (0.0171) | (0.0174) |
| Work distance | $0.143^{* *}$ | 0.160*** | 0.161*** | $0.161^{* * *}$ |
|  | (0.000432) | (0.000469) | (0.000469) | (0.000468) |
| Diesel dummy |  | $4.450{ }^{* * *}$ | $3.334^{* *}$ | $3.334^{* * *}$ |
|  |  | (0.0294) | (0.0208) | (0.0207) |
| EV dummy |  | -0.153 | -5.633*** | -5.438*** |
|  |  | (0.153) | (0.178) | (0.178) |
| Toll (NOK per worktrip) |  | -0.0446*** | -0.0447*** | -0.0446*** |
|  |  | (0.000382) | (0.000382) | (0.000380) |
| Toll times EV dummy |  | $0.0261^{* *}$ | $0.0261^{* * *}$ | $0.0251^{* * *}$ |
|  |  | (0.00335) | (0.00335) | (0.00334) |
| Price per km |  | $1.535^{* * *}$ |  |  |
|  |  | (0.0533) |  |  |
| Fuel/energy price |  |  | -0.328*** | -0.352*** |
|  |  |  | (0.00734) | (0.00730) |
| Car age difference |  |  |  | -0.234*** |
|  |  |  |  | (0.00134) |
| Engine effect difference |  |  |  | 0.0344*** |
|  |  |  |  | (0.000357) |
| Weight difference |  |  |  | 0.00197*** |
|  |  |  |  | (0.0000345) |
| Constant | $21.43^{* * *}$ | $41.64{ }^{* * *}$ | $46.33^{* * *}$ | $34.96{ }^{* * *}$ |
|  | (0.234) | (0.223) | (0.242) | (0.247) |
| Demographic controls | $\checkmark$ | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Car controls |  | $\checkmark$ | $\checkmark$ | $\checkmark$ |
| Mean depvar: | 35.61 | 35.61 | 35.61 | 35.61 |
| Observations: | 13,403,807 | 13,403,807 | 13,403,807 | 13,403,807 |
| Sum of weights: | 6,088,227 | 6,088,227 | 6,088,227 | 6,088,227 |

* $\mathrm{p}<0.10,{ }^{* *} \mathrm{p}<0.05,{ }^{* * *} \mathrm{p}<0.01$. Robust standard errors in parentheses.

Notes: Dependent variable is driving per car per day. Observations are sub-periods (unique combinations of driving periods, households and car portfolios). Regressions are weighted with the share of the corresponding driving period a sub-period constitutes. Prices are measured in 2015-kroners. The first column regresses driving on demographics. The second column includes controls for car attributes and price per kilometer (fuel price over fuel efficiency). The third column controls for fuel price instead of price per kilometer. Fuel price is $\mathbf{1}\{$ gasoline $\} \times$ gasoline price $+\mathbf{1}\{$ diesel $\} \times$ diesel price $+\mathbf{1}\{E V\} \times$ electricity price (in liters $/ \mathrm{kWh}$ ). Additional demographic controls include "age", "couple" and "city". Additional car controls include "car age", "engine effect" and "car weight".

Table B.4: Linear regression coefficients, car ownership 2017

| Dependent variable: | Gasoline | Diesel | EV | Cars |
| :--- | :---: | :---: | :---: | :---: |
| Household income (log) | $0.0449^{* * *}$ | $0.0792^{* * *}$ | $0.0147^{* * *}$ | $0.139^{* * *}$ |
|  | $(0.000293)$ | $(0.000466)$ | $(0.000125)$ | $(0.000694)$ |
| Work distance | $-0.0000573^{* *}$ | $0.00635^{* * *}$ | $-0.000230^{* * *}$ | $0.00606^{* * *}$ |
|  | $(0.0000291)$ | $(0.0000360)$ | $(0.0000138)$ | $(0.0000338)$ |
| Age | $0.00281^{* * *}$ | $-0.00215^{* * *}$ | $-0.000627^{* * *}$ | 0.0000287 |
|  | $(0.000194)$ | $(0.0000195)$ | $(0.00000595)$ | $(0.0000227)$ |
| Couple dummy | $0.187^{* * *}$ | $0.419^{* * *}$ | $0.0363^{* * *}$ | $0.641^{* * *}$ |
|  | $(0.000853)$ | $(0.000940)$ | $(0.000266)$ | $(0.00109)$ |
| City dummy | $-0.0516^{* * *}$ | $-0.226^{* * *}$ | $0.00643^{* * *}$ | $-0.272^{* * *}$ |
|  | $(0.000881)$ | $(0.000869)$ | $(0.000380)$ | $(0.00101)$ |
| Toll (NOK per worktrip) | $0.000211^{* * *}$ | $-0.00217^{* * *}$ | $0.00163^{* * *}$ | $-0.000333^{* * *}$ |
|  | $(0.0000225)$ | $(0.0000250)$ | $(0.0000147)$ | $(0.0000234)$ |
| Constant | $-0.406^{* * *}$ | $-0.590^{* * *}$ | $-0.143^{* * *}$ | $-1.138^{* * *}$ |
|  | $(0.00340)$ | $(0.00549)$ | $(0.00143)$ | $(0.00832)$ |
|  |  |  |  |  |
| Mean depvar: | 0.4085 | 0.5224 | 0.0470 | 0.9779 |
| Observations: | $2,169,769$ | $2,169,769$ | $2,169,769$ | $2,169,769$ |

* $\mathrm{p}<0.10,{ }^{* *} \mathrm{p}<0.05,{ }^{* * *} \mathrm{p}<0.01$. Robust standard errors in parentheses.

Notes: Dependent variable is number of cars owned per household, by propulsion system (columns 1-3) and in total (column 4). Income and toll payments are measured in 2015kroners.

## B. 2 EV ownership and toll payments

Figure B. 5 displays the geographical variation in EV ownership and toll exposure by presenting neighborhood-level averages in 2014 and 2017. The panels indicate (1) a growth in both outcomes over time and (2) a high degree of spatial correlation, in particular around the main cities.

One concern is that there is spurious correlation between toll payments and EVs: i.e., that tolls are more common in areas where the share of EV owners would be high for other reasons. To test this, we run linear regressions on EV ownership with a varying degree of fixed effects. Results are displayed in Table B.5. The coefficients on toll payments are remarkably similar between specifications, indicating that the within-neighborhood variation in tolls has a similar impact on EV ownership as toll variation across Municipalities.

Table B.5: Effect of tolls on EV ownership, 2017

|  | $(1)$ | $(1)$ | $(1)$ |
| :--- | :---: | :---: | :---: |
| Toll (NOK per worktrip) | $0.00124^{* * *}$ | $0.00105^{* * *}$ | $0.000964^{* * *}$ |
|  | $(0.00000816)$ | $(0.00000849)$ | $(0.00000874)$ |
| Work distance (kms) | $-0.000439^{* * *}$ | $-0.000295^{* * *}$ | $-0.000259^{* * *}$ |
|  | $(0.0000106)$ | $(0.0000108)$ | $(0.0000110)$ |
| Household income (log) | $0.00698^{* * *}$ | $0.00672^{* * *}$ | $0.00670^{* * *}$ |
|  | $(0.000156)$ | $(0.000156)$ | $(0.000159)$ |
| Employment (dummy) | $0.0491^{* * *}$ | $0.0516^{* * *}$ | $0.0501^{* * *}$ |
|  | $(0.000534)$ | $(0.000534)$ | $(0.000538)$ |
| Age (years) | $-0.000167^{* * *}$ | $-0.000131^{* * *}$ | $-0.000182^{* * *}$ |
|  | $(0.00000856)$ | $(0.00000856)$ | $(0.00000875)$ |
| Couple (dummy) | $0.0186^{* * *}$ | $0.0187^{* * *}$ | $0.0159^{* * *}$ |
|  | $(0.000352)$ | $(0.000351)$ | $(0.000355)$ |
| City (dummy) | $-0.00657^{* * *}$ |  |  |
| Work in city (dummy) | $(0.000376)$ |  | $0.0301^{* * *}$ |
|  | $0.0447^{* * *}$ | $0.0333^{* * *}$ | $0.000704)$ |


| Municipality FE: |  | $\checkmark$ |  |
| :--- | :---: | :---: | :---: |
| Neighborhood FE: |  |  | $\checkmark$ |
| Observations: | $2,169,769$ | $2,169,747$ | $2,169,585$ |

* $\mathrm{p}<0.10,{ }^{* *} \mathrm{p}<0.05,{ }^{* * *} \mathrm{p}<0.01$. Robust standard errors in parentheses.

Notes: Dependent variable is whether a household owns an EV (0/1). Income and toll payments are measured in 2015-kroners. Note that the city dummy will be absorbed by the fixed effects, and that Municipality level fixed effects will be absorbed by neighborhood level fixed effects as neighborhoods never cross Municipality borders.

Figure B.5: Geographical variation in EV share and toll exposure


Notes: Left panels display share of households that owns EVs by neighborhood, while right panels display average neighborhood-level toll exposure where tolls are measured in 2015 kroners for a one-way work trip. Top panels are 2014, while bottom panels are 2017. Note that the scale of all panels are different. For confidentiality reasons, neighborhoods with less than 50 households have their values replaced by the Municipality level average. Note that densely populated areas where EV ownership and toll exposure tends to be higher have neighborhoods too small to be visible on the maps.

## C Mathematical Appendix

## C. 1 Complementarities

Proposition: In the logit model with binary demand, $\mathcal{D}=\{0, j, k,(j, k)\}$, the two goods are Hicksian complements if and only if $\Gamma>0$.

Proof: Expected demand is $Q_{j}=\operatorname{Pr}(j)+\operatorname{Pr}(j, k)$, and

$$
\begin{align*}
\frac{\partial Q_{j}}{\partial U_{k}} & =\operatorname{Pr}(j)\left[\frac{\partial U_{j}}{\partial U_{k}}-\sum_{d \in \mathcal{D}} \operatorname{Pr}(d) \frac{\partial U_{d}}{\partial U_{k}}\right]+\operatorname{Pr}(j, k)\left[\frac{\partial U_{j k}}{\partial U_{k}}-\sum_{d \in \mathcal{D}} \operatorname{Pr}(d) \frac{\partial U_{d}}{\partial U_{k}}\right] \\
& =\operatorname{Pr}(j, k) \operatorname{Pr}(0)-\operatorname{Pr}(j) \operatorname{Pr}(k) \tag{13}
\end{align*}
$$

since $\sum_{d \in \mathcal{D}} \operatorname{Pr}(d) \frac{\partial U_{d}}{\partial U_{k}}=\operatorname{Pr}(k)+\operatorname{Pr}(j, k)$.
Next, we leverage the logit choice probability functional form,

$$
\operatorname{Pr}(d)=\frac{\exp \left(U_{d}\right)}{1+\exp \left(U_{j}\right)+\exp \left(U_{k}\right)+\exp \left(U_{j}+U_{k}+\Gamma\right)},
$$

in writing

$$
\begin{aligned}
\frac{\partial Q_{j}}{\partial U_{k}} & >0 \\
\Leftrightarrow \operatorname{Pr}(j, k) \operatorname{Pr}(0) & >\operatorname{Pr}(j) \operatorname{Pr}(k) \\
\Leftrightarrow \frac{\operatorname{Pr}(j, k)}{\operatorname{Pr}(0)} & >\frac{\operatorname{Pr}(j)}{\operatorname{Pr}(0)} \frac{\operatorname{Pr}(k)}{\operatorname{Pr}(0)} \\
\exp \left(U_{k}+U_{j}+\Gamma\right) & >\exp \left(U_{k}\right) \exp \left(U_{j}\right) \\
\Gamma & >0 .
\end{aligned}
$$

Proposition: In the binary model, C. 2 holds if and only if C. 3 holds.
Proof: Note first, that $\mathcal{D}=\{0, j, k,(j, k)\}$. From (13), we see that

$$
\frac{\partial Q_{j}}{\partial U_{k}}=\operatorname{Pr}(j, k) \operatorname{Pr}(0)-\operatorname{Pr}(j) \operatorname{Pr}(k)
$$

so Hicksian complementarity (C.3) holds iff.

$$
\begin{gathered}
\operatorname{Pr}(j, k) \operatorname{Pr}(0)>\operatorname{Pr}(j) \operatorname{Pr}(k) \\
\Leftrightarrow s_{j k} s_{0}>s_{j} s_{k},
\end{gathered}
$$

where $s_{d}$ denotes the share of households choosing $d \in \mathcal{D}$. Next, note that the Correlation Criterion (C.2) can be written as

$$
\begin{aligned}
\frac{s_{j k}}{s_{j k}+s_{k}} & >s_{j k}+s_{j} \\
\Leftrightarrow s_{j k} & >s_{j} s_{k}+s_{j k}\left(s_{k}+s_{j}+s_{j k}\right) \\
\Leftrightarrow s_{0} s_{j k} & >s_{j} s_{k} .
\end{aligned}
$$

since $s_{0}+s_{j}+s_{k}+s_{j k}=1$.

Proposition: In the model with binary demand, $\mathcal{D}=\{0, j, k,(j, k)\}$, the crossderivative of demand is

$$
\frac{\partial Q_{j}}{\partial U_{k}}=\operatorname{Pr}(j, k)-Q_{j} Q_{k} .
$$

## Proof:

$$
\begin{aligned}
\frac{\partial Q_{j}}{\partial U_{k}} & =\operatorname{Pr}(j)\left[\frac{\partial U_{j}}{\partial U_{k}}-\sum_{d \in \mathcal{D}} \operatorname{Pr}(d) \frac{\partial U_{d}}{\partial U_{k}}\right]+\operatorname{Pr}(j, k)\left[\frac{\partial U_{j k}}{\partial U_{k}}-\sum_{d \in \mathcal{D}} \operatorname{Pr}(d) \frac{\partial U_{d}}{\partial U_{k}}\right] \\
& =\operatorname{Pr}(j, k)-[\operatorname{Pr}(j)+\operatorname{Pr}(j, k)][\operatorname{Pr}(k)+\operatorname{Pr}(j, k)] \\
& =\operatorname{Pr}(j, k)-Q_{j} Q_{k}
\end{aligned}
$$

Proposition: In the full model with symmetric bundles, $\mathcal{D}=\{0, j, k,(j, k),(j, j),(k, k)\}$, the cross-derivative of demand is

$$
\frac{\partial Q_{j}}{\partial U_{k}}=\operatorname{Pr}(j, k)-Q_{j} Q_{k} .
$$

Proof: Using the definition of $Q_{j}$,

$$
\begin{aligned}
\frac{\partial Q_{j}}{\partial U_{k}} & =\frac{\partial}{\partial U_{k}}[\operatorname{Pr}(j)+\operatorname{Pr}(j, k)+2 \operatorname{Pr}(j, j)] \\
& =\operatorname{Pr}(j)\left[0-\frac{\partial \Lambda}{\partial U_{k}}\right]+\operatorname{Pr}(j, k)\left[1-\frac{\partial \Lambda}{\partial U_{k}}\right]+2 \operatorname{Pr}(j, j)\left[1-\frac{\partial \Lambda}{\partial U_{k}}\right] \\
& =\operatorname{Pr}(j, k)-Q_{j} \frac{\partial \Lambda}{\partial U_{k}},
\end{aligned}
$$

where $\Lambda$ is the logsum,

$$
\begin{aligned}
\Lambda \equiv & \log \left[1+\exp \left(U_{j}\right)+\exp \left(U_{k}\right)+\exp \left(U_{j}+U_{k}+\Gamma_{j k}\right)\right. \\
& \left.+\exp \left(2 U_{j}+\Gamma_{j j}\right)+\exp \left(2 U_{k}+\Gamma_{k k}\right)\right] .
\end{aligned}
$$

Now all that remains is to show that $\frac{\partial \Lambda}{\partial U_{k}}=Q_{k}$ :

$$
\begin{aligned}
\frac{\partial \Lambda}{\partial U_{k}} & =\operatorname{Pr}(k)+\operatorname{Pr}(j, k)+2 \operatorname{Pr}(k, k) \\
& =Q_{k}
\end{aligned}
$$

## C. 2 Estimation

In this section, we describe details relating to our econometric estimator. Specifically, how we match the sub-periods (where our model predicts driving) to the driving periods (where we observe driving), and how we conduct inference.

First, it is useful to consider an example to fix ideas. Suppose a household buys a car in July of 2015, then adds a second car to its portfolio in July of 2016, before the first car has an inspection in July of 2017. Focus on the first car: its driving period covers two years, composed of two sub-periods: the first sub-period (2015-06
to 2016-06) covers $50 \%$ of the driving period and the owner is a one-car household. The second sub-period (2016-06 to 2017-06) also covers $50 \%$ but the owner is now a two-car household. Our predicted total for the driving period is $50 \%$ times the one-car driving plus $50 \%$ times the prediction from the two-car household.

Formally, let $t \in\{1, \ldots, T\}$ denote driving periods, and $s_{t} \in\left\{1, \ldots, S_{t}\right\}$ denote sub-periods, where $w_{t s}$ denotes the fraction of period $t$ covered by sub-period $s$, i.e. $\sum_{s=1}^{S_{t}} w_{t s}=1$ for all $t$. We have on average $\frac{1}{T} \sum_{t=1}^{T} S_{t}=2.2$ sub-periods per driving period. During a sub-period, we can identify a unique household, $i$, with a stable car portfolio, $d_{i}$, throughout. Let $x_{t s}^{*}(\theta)$ denote the predicted driving for that household, given model parameters $\theta$. The vector of demographics, $\mathbf{z}_{i}$, are weighted averages of the years covered by the sub-period, and car attributes, $\mathbf{q}_{d_{i}}$, are taken directly from the car without aggregating to the 20-type level (so we have more variation in car attributes in the driving sub-likelihood than in the discrete choice sub-likelihood).

The normality assumption that we make is thus

$$
x_{t}^{\mathrm{data}}=\sum_{s=1}^{S} w_{t s} x_{t s_{t}}^{*}(\theta)+\eta_{t}, \quad \eta_{i} \sim \operatorname{IID\mathcal {N}}\left(0, \sigma_{x}^{2}\right) .
$$

Thus, the likelihood contribution from driving period $t$ is

$$
\log \frac{1}{\sigma_{x}} \phi\left(\frac{x_{t}^{\mathrm{data}}-\sum_{s=1}^{S} w_{t s} x_{t s}^{*}(\theta)}{\sigma_{x}}\right)
$$

The full likelihood function thus becomes

$$
\mathcal{L}(\theta)=\frac{1}{N} \sum_{i=1}^{N} \log \operatorname{Pr}\left(d_{i} \mid \theta\right)-\frac{1}{T} \frac{1}{2} \sum_{t=1}^{T}\left(\frac{x_{t}^{\mathrm{data}}-\sum_{s=1}^{S_{t}} w_{t s} x_{t s}^{*}(\theta)}{\sigma_{x}}\right)^{2}-\log \sigma_{x}
$$

Next, we are faced with the issue that we use a different number of observations in our ownership dataset $(N=52,739)$ and driving dataset $(T=2,588,591)$. The number of discrete choice observations is set low for computational reasons, but we choose to retain more observations for the driving sub-likelihood, which is computationally inexpensive. We could have reduced $T$ to be the same observations that we have discrete choices for in order to use standard inference. Instead, we opt to add additional driving observations for increased precision, but conduct inference based on the lower of the two observation counts, $N$, to be conservative.

Thus, we estimate standard errors using the following "Sandwich" formula:

$$
\begin{align*}
\operatorname{Cov}(\hat{\theta}) & =\frac{1}{N} A^{-1} B A^{-1}  \tag{14}\\
A & =\sum_{i=1}^{N} \nabla \ell_{i}(\theta)^{\prime} \nabla \ell_{i}(\theta) \\
B & =\sum_{i=1}^{N} \nabla^{2} \ell_{i}(\theta) \\
\ell_{i} & =\log \operatorname{Pr}\left(d_{i} \mid \theta\right)-\frac{1}{T} \frac{1}{2} \sum_{t=1}^{T}\left(\frac{x_{t}^{\mathrm{data}}-\sum_{s=1}^{S_{t}} w_{t s} x_{t s}^{*}(\theta)}{\sigma_{x}}\right)^{2}-\log \sigma_{x}
\end{align*}
$$

## D Additional results

Figure D.1: Fit of driving for 1-car households


Notes: This figure displays average driving per day for one-car households by deciles of car attributes and household characteristics. x-axes also denote averages within each decile. The fit is evaluated on the full $100 \%$ driving sample.

Figure D.2: Car characteristics by income and work distance


Figure D.3: Number of cars over time



Figure D.4: Car ownership by car type


Table D.1: Cross-price elasticities, 2017


Notes: The matrix displays the price elasticity of demand for the alternative in the $j$ th row, $Q_{i j}$ from Equation (2), when the price of the car in the $k$ th column is increased by one percent.

Table D.2: Share of households with negative cross-price elasticity, 2017

| Small gasoline, new |  | 1.0 | 1.1 | 1.1 | 1.0 | 1.0 | 1.0 | 1.0 | 1.1 | 1.0 | 1.0 | 1.0 | 1.6 | 1.1 | 1.2 | 1.1 | 3.3 | 3.6 | 1.9 | 2.0 | 16 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| small diesel, new | 1.0 |  | 1.0 | 1.0 | 1.1 | 1.0 | 1.0 | 1.0 | 1.2 | 1.0 | 1.0 | 1.0 | 2.5 | 1.3 | 1.4 | 1.2 | 6.0 | 6.9 | 2.1 | 2.4 |  |
| Large gasoline, new | 1.1 | 1.0 |  | 1.0 | 1.2 | 1.0 | 1.0 | 1.0 | 1.4 | 1.1 | 1.1 | 1.0 | 4.2 | 1.4 | 2.0 | 1.3 | 6.6 | 7.8 | 2.0 | 2.3 |  |
| Large diesel, new | 1.1 | 1.0 | 1.0 |  | 1.1 | 1.0 | 1.0 | 1.0 | 1.4 | 1.1 | 1.1 | 1.0 | 5.0 | 1.5 | 2.2 | 1.4 | 11.8 | 16.4 | 2.3 | 2.6 |  |
| Small gasoline, 1-4 years | 1.0 | 1.1 | 1.2 | 1.1 |  | 1.0 | 1.1 | 1.1 | 1.0 | 1.0 | 1.0 | 1.0 | 1.2 | 1.0 | 1.1 | 1.0 | 2.6 | 2.8 | 1.9 | 2.0 | 12 |
| small diesel, 1-4 years | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |  | 1.0 | 1.0 | 1.1 | 1.0 | 1.0 | 1.0 | 1.9 | 1.2 | 1.3 | 1.1 | 4.9 | 5.6 | 2.1 | 2.3 |  |
| Large gasoline, 1-4 years | 1.0 | 1.0 | 1.0 | 1.0 | 1.1 | 1.0 |  | 1.0 | 1.2 | 1.0 | 1.1 | 1.0 | 2.1 | 1.2 | 1.4 | 1.1 | 4.1 | 4.6 | 1.9 | 2.0 | 10 |
| Large diesel, 1-4 years | 1.0 | 1.0 | 1.0 | 1.0 | 1.1 | 1.0 | 1.0 |  | 1.3 | 1.1 | 1.1 | 1.0 | 3.4 | 1.3 | 1.7 | 1.3 | 7.8 | 9.5 | 2.2 | 2.5 |  |
| Small gasoline, 5-11 years | 1.1 | 1.2 | 1.4 | 1.4 | 1.0 | 1.1 | 1.2 | 1.3 |  | 1.0 | 1.0 | 1.1 | 1.0 | 1.0 | 0.9 | 1.0 | 2.0 | 2.0 | 2.1 | 2.0 |  |
| small diesel, 5-11 years | 1.0 | 1.0 | 1.1 | 1.1 | 1.0 | 1.0 | 1.0 | 1.1 | 1.0 |  | 1.0 | 1.0 | 1.4 | 1.1 | 1.1 | 1.1 | 3.9 | 4.2 | 2.2 | 2.3 | 8 |
| Large gasoline, 5-11 years | 1.0 | 1.0 | 1.1 | 1.1 | 1.0 | 1.0 | 1.1 | 1.1 | 1.0 | 1.0 |  | 1.0 | 1.2 | 1.0 | 1.1 | 1.0 | 2.5 | 2.7 | 1.8 | 1.9 |  |
| Large diesel, 5-11 years | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.1 | 1.0 | 1.0 |  | 2.1 | 1.2 | 1.3 | 1.2 | 5.2 | 5.9 | 2.1 | 2.3 |  |
| Small gasoline, 12-25 years | 1.6 | 2.5 | 4.2 | 5.0 | 1.2 | 1.9 | 2.1 | 3.4 | 1.0 | 1.4 | 1.2 | 2.1 |  | 1.0 | 0.8 | 1.0 | 1.5 | 1.5 | 3.8 | 3.2 | 6 |
| small diesel, 12-25 years | 1.1 | 1.3 | 1.4 | 1.5 | 1.0 | 1.2 | 1.2 | 1.3 | 1.0 | 1.1 | 1.0 | 1.2 | 1.0 |  | 0.9 | 1.0 | 2.6 | 2.7 | 2.6 | 2.5 |  |
| Large gasoline, 12-25 years | 1.2 | 1.4 | 2.0 | 2.2 | 1.1 | 1.3 | 1.4 | 1.7 | 0.9 | 1.1 | 1.1 | 1.3 | 0.8 | 0.9 |  | 1.0 | 1.5 | 1.5 | 2.4 | 2.2 | 4 |
| Large diesel, 12-25 years | 1.1 | 1.2 | 1.3 | 1.4 | 1.0 | 1.1 | 1.1 | 1.3 | 1.0 | 1.1 | 1.0 | 1.2 | 1.0 | 1.0 | 1.0 |  | 2.8 | 3.1 | 2.5 | 2.5 |  |
| Small electric, new | 3.3 | 6.0 | 6.6 | 11.8 | 2.6 | 4.9 | 4.1 | 7.8 | 2.0 | 3.9 | 2.5 | 5.2 | 1.5 | 2.6 | 1.5 | 2.8 |  | 2.6 | 3.3 | 3.1 |  |
| small electric, used | 3.6 | 6.9 | 7.8 | 16.4 | 2.8 | 5.6 | 4.6 | 9.5 | 2.0 | 4.2 | 2.7 | 5.9 | 1.5 | 2.7 | 1.5 | 3.1 | 2.6 |  | 3.6 | 3.3 | 2 |
| Large electric, new | 1.9 | 2.1 | 2.0 | 2.3 | 1.9 | 2.1 | 1.9 | 2.2 | 2.1 | 2.2 | 1.8 | 2.1 | 3.8 | 2.6 | 2.4 | 2.5 | 3.3 | 3.6 |  | 2.0 |  |
| Large electric, used | 2.0 | 2.4 | 2.3 | 2.6 | 2.0 | 2.3 | 2.0 | 2.5 | 2.0 | 2.3 | 1.9 | 2.3 | 3.2 | 2.5 | 2.2 | 2.5 | 3.1 | 3.3 | 2.0 |  |  |
| Outside option | 0.0 | 0.0 | 0.0 | 0.2 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.4 | 0.0 |  |
| $\operatorname{sn} \operatorname{mal}^{\operatorname{ll} 9^{2 a^{\circ}} \sin ^{\ln e}}$ | $2 \mathrm{ar}^{6}$ | $s^{m^{2}}$ |  | $2^{2 x 0}$ |  |  |  |  |  |  | $\mathrm{sm}^{2}$ |  | $220$ |  |  |  |  |  |  |  |  |

Notes: The matrix displays the share of households where the cross-price elasticity of demand $\partial Q_{i j} / \partial p_{k}$ is negative for the alternative in the $j$ th row when the price of the car in the $k$ th column is increased, where $Q_{i j}$ is given in Equation (2).

Table D.3: Average kms/day per car as function of car portfolio, 2017.


Notes: The matrix displays average driving (kms/day) for the car in the $j$ th row, when it is combined in a portfolio with the car in the $k$ th column. The first column displays average driving for one-car households. Averages across households are unconditional on choice probabilities, meaning that households receive equal weight even though some are unlikely to end up with a particular car portfolio.

Table D.4: Parameter Estimates, No Portfolio Effects

| Demographics: | Driving ( $\alpha_{z}$ ) |  | Outside option $\left(\omega_{z}\right)$ |  | Utility of money $\left(\gamma_{z}\right)$ |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\ln$ (household income) | 0.00767 | (8.67) | -0.612 | (-2.62) | -0.636 | (-8.57) |
| work distance (kms) | 0.000652 | (5.10) | -0.00102 | (-1.88) | 0.0104 | (1.61) |
| age (avg) | -0.00112 | (-6.11) | -0.0100 | (-1.41) | -0.0240 | (-3.48) |
| 1 \{city\} | -0.0346 | (-2.66) | 0.271 | (0.965) | -0.416 | (-1.69) |
| 1 \{couple\} | 0.0190 | (3.68) | 0.266 | (1.49) | -0.329 | (-1.77) |
| cons | 0.0233 | (3.69) | 6.85 | (2.44) | 11.3 | (13.6) |
| Car attributes: | Driving ( $\alpha_{q}$ ) |  | Car ownership ( $\xi_{q}$ ) |  | Portfolio shifting $\left(\Gamma_{4}\right)$ |  |
| car age | 0.00118 | ( 1.86) | 0.297 | (10.6) | 0.00 | - |
| car age ${ }^{2}$ | -0.000282 | ( -4.67) | -0.00651 | (-2.75) | 0.00 | - |
| engine effect ( $100 \mathrm{~kW} \mathrm{)}$ | 0.0127 | ( 1.47) | 4.30 | (9.07) | 0.00 | - |
| engine effect $\times 1\{\mathrm{EV}\}$ | -0.0213 | (-6.16) | -3.25 | (-11.4) | 0.00 | - |
| total weight (tonnes) | 0.145 | ( 34.8) | -9.12 | (-25.5) | 0.00 | - |
| 1 \{diesel\} | 0.0317 | ( 1.65) | -0.791 | (-0.955) | 0.00 | - |
| 1 \{EV\} | 0.0896 | ( 7.96) | -3.44 | (-7.85) | 0.00 | - |
| Other variables |  |  | Portfolio effects |  |  |  |
| Range by work distance ( $\varphi_{1 i j}$ ) | $4.61 \mathrm{e}-05$ | ( 3.24) | Ownersh | satiation ( $\Gamma_{1}$ ) | 0.00 | - |
| Local EV incentives ( $\varphi_{2 i j}$ ) | 1.75 | ( 8.06) | Range an | xiety ( $\Gamma_{2}$ ) | 0.00 | - |
| Driving squared ( $\alpha_{2}$ ) | -0.00331 | (-16.1) | Driving | atiation ( $\Gamma_{3}$ ) | 0.00 | - |
| S.D. of error term, driving ( $\sigma_{x}$ ) | 29.7 | ( 201.3) | Driving | bstitution ( $\Gamma_{5}$ ) | 0.00 | - |
| Realised toll payment share ( $\theta_{\text {toll }}$ ) | 0.388 | ( 0.801) |  |  |  |  |

Notes: These estimates are similar to those in Table 4, except that that portfolio effects are disabled by forcing $\Gamma_{i j k}=0$. This forces 11 parameters to zero but the remaining 37 are still estimated. Parentheses show the $t$-statistics corresponding to each parameter value using the standard errors from the Sandwich formula in equation (14). The results are based on a random sub-sample of $N=52,739$ discrete choice observations. The driving dataset is a $30 \%$ random sample of all driving periods, yielding $S=2,588,591$ odometer readings.

Table D.5: Counterfactual simulations, two-car model without portfolio effects

|  | (1) | (2) | (3) |
| :---: | :---: | :---: | :---: |
| Exogenous policy variables |  |  |  |
| Targeted segment | EV | CV | CV |
| Tax instrument | Purchase | Fuel | Purchase |
| Effect on consumer price | -20.0\% | +2.06\% | +1.62\% |
| Welfare effects (annual, per household) |  |  |  |
| $\Delta \mathrm{CO}_{2}(\mathrm{~kg})$ | -10.966 | -10.716 | -10.176 |
| $\Delta$ Consumer surplus (NOK) | 159.014 | -228.425 | -342.170 |
| $\Delta$ Taxes (NOK) | -316.820 | 177.989 | 117.153 |
| $\Delta$ Local externalities (NOK) | 27.026 | -42.022 | -56.905 |
| Abatement cost (NOK per kg CO2 ${ }^{\text {) }}$ | -16.856 | -0.785 | -16.520 |
| Number of cars |  |  |  |
| Cars | 3,636.3 | -6,633.1 | -8,817.5 |
| - EV | 13,199.5 | 361.6 | 484.4 |
| - ICEV | -9,562.9 | -6,994.6 | -9,301.9 |
| Households by Portfolio Choice |  |  |  |
| No car | -1,674.3 | 3,146.6 | 4,096.5 |
| EV | 4,562.2 | 185.3 | 253.6 |
| EV,EV | 166.6 | 6.6 | 8.5 |
| ICEV | -4,850.3 | 154.6 | 370.9 |
| ICEV,ICEV | -6,508.4 | -3,656.2 | -4,943.3 |
| ICEV,EV | 8,304.2 | 163.1 | 213.8 |
| Driving (expected average percentage changes) |  |  |  |
| Total driving | 0.235 | -0.377 | -0.494 |
| EV driving | 36.479 | 0.741 | 0.983 |
| Diesel driving | -0.494 | -0.444 | -0.706 |
| Gasoline driving | -0.414 | -0.358 | -0.310 |

Note: See Table 7 - the difference is that these results are based on the model with $\Gamma_{i j k}=0$ (estimates are in Table D.4).


[^0]:    *We are thankful for comments and suggestions by Erich Muehlegger, Mogens Fosgerau, Bertel Schjerning, Andreas Moxnes and Katinka Holtsmark. Anders gratefully acknowledges funding from the Innovation Fund Denmark under the URBAN grant and from the Danish Economic Councils. Bjørn acknowledges funding from the Norwegian Research Council ("Driving towards the low-emission society" project \# 267942). Data on loan from Statistics Norway have been essential.
    ${ }^{\dagger}$ Institute of Transport Economics and University of Oslo, Department of Economics, Norway. E-mail: bgj@toi.no
    $\ddagger$ University of Copenhagen, Department of Economics, Denmark. E-mail: amn@econ.ku.dk

[^1]:    ${ }^{1}$ Kverndokk et al. (2020) report that $89 \%$ of households that own a CV and buy an EV report that they drive their CV no less than before.

[^2]:    ${ }^{2}$ We assume that households are rationally forward-looking and internalizing future fuel cost savings. Recent research has found evidence in support of this (e.g. Busse et al. 2013; Grigolon et al. 2018).
    ${ }^{3}$ Of course, a general trend towards EVs has been present in most countries, but the timing and nature of those policies are highly heterogenous.

[^3]:    ${ }^{4}$ We emphasize that a major reason for this is the coarse aggregation of cars into just 20 types. Thus, cars of competing firms have been bunched together, so the price variation across firms and idiosyncratic car quality within a segment of attribute space gets averaged out. This identification strategy could not be used if one is interested in cars at the firm level e.g. to study market power.

[^4]:    ${ }^{5}$ West (2004); Bento et al. (2009); Grigolon et al. (2018); Munk-Nielsen (2014).
    ${ }^{6}$ Empirically, see e.g. Grzybowski and Verboven (2016); Thomassen et al. (2017); Ershov et al. (2021).
    ${ }^{7}$ In their setting, complementarity is between the continuous choices, e.g. soda and chips.
    ${ }^{8}$ E.g. Chandra et al. (2010); DeShazo et al. (2017); Jenn et al. (2018); Muehlegger and Rapson (2018); Yan (2018); Clinton and Steinberg (2019).

[^5]:    ${ }^{9}$ In Norway, $95 \%$ of EV owners report to charge at home (Figenbaum and Kolbenstvedt, 2016). So while access to a charging network matters, it is not necessarily first-order important for EV adoption.
    ${ }^{10}$ In Norway, $95 \%$ of electricity is produced by hydro power so it is less relevant for our specific empirical setting.
    ${ }^{11} \mathrm{~A}$ wider literature has studied fuel economy standards imposed directly on car manufacturers, including Durrmeyer and Samano (2018); Reynaert (2021) and earlier work reviewed in Anderson and Sallee (2016).

[^6]:    ${ }^{12}$ During the period 2005-2017, the following attributes have affected the registration tax: weight, cylinder volume, engine effect, type approved $\mathrm{CO}_{2}$ emissions and type approved $\mathrm{NO}_{x}$ emissions.
    ${ }^{13}$ While charging network is of course important for adoption, the causality might run even more strongly in the other direction, with entry of charging stations responding to local demand,

[^7]:    possibly ahead of time if firms are forward-looking. Springel (2021) pursued a such a joint model, but in a one-car setting.
    ${ }^{14}$ The data does not include leased cars, which are listed as owned by a company and therefore cannot be tracked to the users. In practice, this means that we are misclassifying some households as having fewer cars at their disposal than they actually do. Note that EVs are under-represented among company owned cars - see Appendix Figure B. 1 for more information.
    ${ }^{15}$ The road network is from 2015. We do not have time-series variation in the road network.

[^8]:    ${ }^{16}$ For CVs, this segmentation is based on the average vehicle weight within each age and fuel category each year. EVs available in Norway during our sample period naturally fall in two categories, where the "large" category consists of Tesla Models S and X.
    ${ }^{17}$ For a CV, the age groups are $[0 ; 1),[1 ; 4),[5 ; 11),[12 ; \infty)$ years, while for an EV we just have new and used, $[0 ; 1)$ and $[1 ; \infty)$, since there is no EV in our sample prior to 2011.
    ${ }^{18}$ The share of hybrid vehicles is increasing over time (see Appendix Figure B.1), and in 2017, the car type "large, new gasoline" consists mainly of hybrids (see Table A.2, third row). Hybrid vehicles have higher weight and ownership cost, relatively high engine efficiency, but lower driving cost and emissions per kilometer. This shift makes a lot of sense due to the favorable tax treatment relative to more polluting traditional cars.

[^9]:    ${ }^{19}$ See Appendix Table B.3, which regresses driving in km/day on household and car characteristics. The 0.2 is the coefficient on a dummy for whether the household owns two cars.

[^10]:    ${ }^{20}$ The fact that preferences are quasi-linear implies that consumers cannot be (that) risk averse. Therefore, allowing non-degeneracy (e.g. a unit root expectation process) is unlikely to affect

[^11]:    results much.
    ${ }^{21}$ Theoretically, this means that the marginal utility of driving can become negative for high values of $x$, but we do not find this to be an issue in practice.
    ${ }^{22} \mathrm{We}$ have also experimented with heterogeneity in $\alpha_{2}$, to allow for differences in the curvature of the driving function, but found that this did not improve the fit visibly.
    ${ }^{23}$ In practice, we do not find issues with negative predicted driving to be an issue. This is unlike supermarket shopping applications, where zero expenditures are common and so non-negativity constraints must be taken seriously (Thomassen et al., 2017).

[^12]:    ${ }^{24}$ More specifically, we use a discount rate of $\left(0.95+0.95^{2}+0.95^{3}\right) / 3=0.903$ since the "second period" has a three year duration. The only exception to this is the resale price of the car: as

[^13]:    the car is sold at the end of the third year, the resale value is discounted by $0.95^{3}$. This subtle point is omitted from Section 4 where, to ease the exposition, we only considered two periods. An alternative to Equation 3 would have been $U_{i d}=u_{i d 1}+\sum_{t=2}^{4} \beta^{t-1} \mathbb{E}\left(u_{i d t}\right)$, which is what we approximate with our discount factor.
    ${ }^{25}$ The large EV has three times the engine effect of the small EV, whereas the corresponding difference between a large and small CV is roughly 1.5 times (see Table A.2).
    ${ }^{26} \mathrm{We}$ encountered a small number of very low values of $\mathrm{WD}_{i}$, resulting in very large $\varphi_{1 i j}$. To deal with this, we use $\max \left(\mathrm{WD}_{i}, 5\right)$ (and have experimented with other floors than 5 km , finding that it makes little difference).

[^14]:    ${ }^{27} 1\{i$ works in major city $\}$ is not strictly and indicator variable, as we let it take the value 0.5 in cases where the household consists of two adults but only one of them works in a major city.
    ${ }^{28} \mathrm{An}$ alternative would be to let expected maintenance costs affect the variable cost of driving. However, assuming that maintenance costs are considered as fixes is consistent with the literature; see e.g. Hang et al. (2016).

[^15]:    ${ }^{29}$ To explore whether tolls should be a variable or fixed cost, we examined the relationship between driving and tolls and found that households responded much less to tolls than to other monetary variation in the cost of driving (in a richer model of driving, one might imagine that households have luxury trips and necessity trips and that commuting is perhaps a necessity, explaining the weak intensive margin response to toll prices). Furthermore, assuming tolls to be fixed circumvents the issue of where households conduct marginal leisure trips (not all households go to their office in weekends) and which tolls roads such trips might cross.

[^16]:    ${ }^{30}$ This is simplest to see for one-car households: because $\alpha_{2}$ enters in the denominator of optimal driving $x_{i j}^{*}$ in (4), the two sets of coefficients ( $\boldsymbol{\alpha}_{z}, \boldsymbol{\alpha}_{q}, \gamma_{z}, \alpha_{2}$ ) and ( $\left.\lambda \boldsymbol{\alpha}_{z}, \lambda \boldsymbol{\alpha}_{q}, \lambda \gamma_{z}, \lambda^{-1} \alpha_{2}\right)$ result in identical $x_{i j}^{*}$ for any $\lambda \neq 0$. Conversely, choice probabilities will be different.

[^17]:    ${ }^{31}$ Note that there is also an EV-dummy in $\xi_{i j}$ to control the relative baseline market shares of EV and CV.
    ${ }^{32}$ This assumes 220 days of commuting by a CV.

[^18]:    ${ }^{33}$ Note that since the registration tax is a deterministic function of the producer price, we can achieve a perfect fit with piece-wise linear functions in a given year. The remaining $30 \%$ residual variation is due to the reforms that change this schedule from year to year. Johansen (2020) exploits the same variation and provides a more detailed decomposition and analysis with a more fine-grained choice set.

[^19]:    ${ }^{34}$ This is consistent with Norway being a small, open economy. Furthermore, in related work Johansen (2020) provides evidence that the pass-through of tax changes as a result of changes over time in rates of the registration tax in Norway is insignificantly different from $100 \%$. Gallagher and Muehlegger (2011) finds a pass-through of fuel taxes to consumer prices in the US of approximately $100 \%$, and Adamou et al. (2013) find that there is little difference between assuming a $100 \%$ passthrough of purchase taxes and estimating supply side responses for the car market in Europe. With regards to EVs, Muehlegger and Rapson (2018) find the pass-through of subsidies in California to consumer prices to be indistinguishable from $100 \%$.
    ${ }^{35}$ Modeling the interaction between the primary and secondary market requires an equilibrium model along the lines of Gillingham et al. (2022).

[^20]:    ${ }^{36}$ Memory constraints on the computational server restrict us from using the full sample.

[^21]:    ${ }^{37}$ We note that the $1 \%$ of households that display complementarity frequently are often households with a high probability of choosing no car. It may seem puzzling to find complementarity among households unlikely to own two cars but it is a property of the logit model. Intuitively, it is related to a violation of the monotonicity property discussed by Manzini et al. (2019).
    ${ }^{38}$ More generally, $\Delta_{i j k}$ captures the notion of a "discrete double derivative" mentioned by Gentzkow (2007). In Gentzkow's model, $\Delta_{i j k}>0$ iff. $j$ and $k$ are complements. But since we have non-binary demand, $\Delta_{i j k}>0$ is necessary but not sufficient (see Appendix C.1).

[^22]:    ${ }^{39}$ Naturally, the VAT-based policy would be improved by using an attribute-based tax rather than a uniform proportional tax. We chose the uniform tax to have a parsimonious presentation.

[^23]:    ${ }^{40}$ The full model has $+11,068$ EVs but $-7,299 \mathrm{CVs}$, while the restricted model has $+13,200 \mathrm{EVs}$ and $-9,563 \mathrm{CVs}$.
    ${ }^{41}$ One-car models will completely miss this point, because cars will by definition be substitutes: i.e. once a household purchases an EV in a one-car model, they are required to reduce their CV driving to zero.

[^24]:    ${ }^{42}$ For the older cars, the data on registrations and re-registrations tend to be noisy, meaning that it is difficult to know whether a car is actually in use or not. For instance, several of the older cars are labeled as de-registered or scrapped even though they have a subsequent EU control. In addition, cars obtain veteran status when they are 30 years old, reducing the ownership tax to almost zero. Thus, the owner has less incentive to de-register a car even if it is not in use. Furthermore, owners of veteran cars are not required to undertake EU controls as often. This means that we do not observe the mileage of these cars.

[^25]:    ${ }^{43}$ See e.g. https://kartverket.no/globalassets/standard/ horinger-standarder/vegnett-5.0-elveg-2.0/produktspesifikasjon-elveg-2_
    0 -hoeringsutgave-oktober-2018.pdf (Norwegian only).
    ${ }^{44}$ We use a static road network from 2015. This means that all variation in work distances within households over time will come from either households moving, workplaces moving or individuals changing jobs. The cost of passing toll gates however vary by year according to the rates.
    ${ }^{45}$ As we do not know the share of individuals that drive to work (or alternatively the share of days each year that individuals drive to work), this will be estimated in the model. Thus, the number of days chosen has no impact on the solution of the model. If we were to specify that individuals drove to work 110 days each year instead, it would simply result in the "toll" parameter being twice as large.

[^26]:    ${ }^{46}$ If we assume that both spouses drive to work 220 days each year, this corresponds to 7,000 NOK/year and 12,500 NOK/year respectively.
    ${ }^{47}$ Extremely long work distances can also be due to errors in the data. First, the location of the firm may be wrong. This can for instance happen if the firm has moved but the change has not been recorded in the registers yet. Second, if the household has recently moved there may be a discrepancy in the data, since residential locations are reported at a given date while employeremployee relationships are recorded throughout the year. Third, some individuals may have moved but not updated their residential location yet. It is for instance common for young people to move to other cities to study while still being registered with their parent(s) address.

[^27]:    ${ }^{48}$ By the contract with OFV, we are not permitted to share this data. However, the MSRPs each year in PDF format can be found here: https://www.skatteetaten.no/en/rates/ car-prices---list-prices-as-new/ Note that the MSRPs include purchase taxes.
    ${ }^{49}$ Range, estimated for Norwegian conditions for most EV models, can be found here: https: //elbil.no/om-elbil/elbiler-idag/.
    ${ }^{50}$ See https://www.skatteetaten.no/en/rates/car-rates---deduction-for-use-table/.

[^28]:    ${ }^{51}$ The time intervals are different for veteran cars, passenger vehicles not used for private transport such as taxis etc., as well as vehicles with a total weight above 7.5 tonnes. However, these intervals apply to all vehicles in our data set.

[^29]:    ${ }^{52}$ The odometer readings data ends on December 31st 2017, meaning that four year old vehicles would be bought new at December 31st, 2013. As the sale of EVs in Norway has increased exponentially, the majority of the cars are of a younger vintage than this. We do utilize odometer readings for some cars that were registered later, but this is provided that the owner has conducted a periodic vehicle inspection earlier than she is required to.

[^30]:    ${ }^{53}$ The root mean squared error indicates that conditional on weight, engine effect and fuel type, the remaining variation in kroners is three times larger for the registration tax than for the price excluding taxes.

[^31]:    ${ }^{54}$ According to the Government; see e.g. https://www.regjeringen.no/no/tema/okonomi-og-budsjett/skatter-og-avgifter/veibruksavgift-pa-drivstoff/id2603482/ (Norwegian only).

