Diesel Cars and Environmental Policy

Anders Munk-Nielsen*

September 2, 2015

Abstract

In this paper, I measure the costs of environmental taxation of car ownership and usage in Denmark. Using full population Danish register data covering 1997–2006, I estimate a discrete-continuous model of car choice and usage that explicitly allows households to select cars based on expected usage conditional on observed and unobserved heterogeneity. I validate the model using a major Danish reform in 2007 which prompted a substantial shift in the characteristics of purchased cars unique to the Danish setting compared to the rest of Europe. Through counterfactual simulations, I find that both Danish reforms in 1997 and 2007 were cost-ineffective at reducing CO_2 emissions compared to a fuel tax. Moreover, I find that the diesel market share responds strongly to taxation but that environmental goals can be reached both with and without a large diesel share in the fleet.

Keywords: Car taxation, fuel taxation, environmental policy, discrete/continuous choice estimation.
JEL codes: D12, H23, Q53, Q58, L98.

*Department of Economics, University of Copenhagen, anders.munk-nielsen@econ.ku.dk. I would like to thank David Brownstone, Hamish Lowe, Bo Honoré, Aureo de Paula, John Rust, Kenneth Gillingham, Mogens Fosgerau, Jesse Burkhardt, Bertel Schjerning, Søren Leth-Petersen, Jeppe Druedahl and Thomas Jørgensen for comments and feedback. This paper is part of the IRUC research project financed by the Danish Council for Strategic Research (DSF). Financial support is gratefully acknowledged. All remaining errors are my own.

1 Introduction

Since 1980, greenhouse gas emissions from the Danish transport sector have increased from 10 to 15 mio tons CO_2 annually while all remaining sectors together have reduced emissions from 55 to 30 mio tons. In Denmark as well as the rest of the developed world, a consensus is emerging that emissions from the transport sector must be decreased if environmental goals are to be reached. The goal of this paper is to measure the cost-effectiveness of environmentally motivated tax policies that have targeted car choice and use.

Towards this end, I estimate a structural 2-period discrete-continuous model of new car purchase and subsequent usage by Danish households. My dataset covers all new car purchases for the period 1997–2006 as well as subsequent driving over a 4-year period and detailed demographics from the Danish registers. In 2007, a major Danish reform was implemented, followed by substantial changes in the characteristics of newly purchased cars. In particular, the diesel share of new cars in Denmark increased remarkably compared with other European countries at that time. My sample period stops before the reform but I know the response from external sources and can use this to validate my model.

My results contribute to the understanding of the costs of environmental car taxation. The model gives predictions on car choices and subsequent driving, allowing me to analyze the impact of counterfactual policy scenarios on tax revenue, substitutions in the new car market, total driving, fuel demand and CO_2 emissions. I find that a simple fuel tax would have been more efficient per ton of CO_2 than both the 1997 and 2007 reforms were. Other studies found fuel taxes to be more effective compared with taxes that target car characteristics (Grigolon, Reynaert, and Verboven, 2015) and with emissions standards (Jacobsen, 2013).

I also contribute with new insights regarding the increasing diesel share. This has received attention by policy makers as awareness has increased about the negative health effects of local air pollution from diesel cars.[1] A key descriptive fact is that diesel car drivers tend to drive on average 60.0% more than gasoline car drivers. I therefore estimate a high-dimensional discrete-continuous model that explicitly accounts for selection based on observed and unobserved heterogeneity in driving. To my knowledge, I am the first to empirically explore the rise in the diesel share accounting for endogeneous selection. My findings indicate that the diesel share is highly sensitive to the way that car taxes discriminate between gasoline and diesel cars. Environmentally motivated car taxes tend to target the fuel efficiency but must correct for the inherently higher efficiency of diesel cars. I show that the diesel cars are neither necessary nor sufficient for environmental goals. To shed light on what the diesel would be in absence of discriminatory taxation based on fuel type, I counterfactually equalize car taxes and fuel taxes for gasoline and diesel cars and find a level slightly higher than that in 2006, but lower than for most other European countries.

My findings complement existing knowledge on car choice and usage due to the unique nature of my setting; by studying a small open economy without domestic car production and using a reform to explore the validity of the model, I can address some of the issues that are inherent in many of the classic studies of car taxation. Firms respond to car tax policies for example by

[1] In 2012, the World Health Organization moved diesel exhaust to their list of carcinogens — substances that are definitely known to cause lung cancer.

changing their portfolios (Reynaert, 2014). Policies that affect a small market such as the Danish will tend to provide a smaller incentive for automakers to change their portfolios, reducing this supply side concern. Similarly, the market is too small for shocks that are unique to Denmark to affect global fuel prices.

There may, however, still issues with common demand shocks across countries, such as increasing urbanization. Therefore, another strength is the access to full population detailed register data, including demographic information on work distance, income. In addition to accounting for changing urbanization patterns, this allows me to model household driving very precisely. Thereby, I can also give an accurate estimate of the response in driving to an exogenous increase in fuel efficiency (the so-called *rebound effect*), which has been widely debated in the literature (e.g. Small and Van Dender, 2007; Bento et al., 2009; Gillingham, 2012; Hymel and Small, 2015). I estimate the rebound effect for Denmark to be –0.30.

The rest of the paper is organized as follows; Section 1.1 discusses the contributions from this paper in the context of related literature. Section 2 presents the institutional setting and the data and presents some preliminary descriptive evidence. Section 3 lays out the theoretical model while Section 4 gives the empirical strategy for estimation and discusses identification. Section 5 presents the estimates and structural elasticities. Section 6 contains the counterfactual policy simulations and section 7 concludes. Appendix A contains a list of the notation used throughout the paper as well as the core equations of the structural model for easy reference.

1.1 Related Literature

I mainly contribute to the literature on the cost of environmental policies in the car market. Recently, a number of papers have emphasized European settings. D'Haultfæuille, Givord, and Boutin (2013) study the French *Bonus/Malus* reform of 2008 which is a feebate similar to the Danish one. They find that the reform had a negative environmental impact, mainly because it led to more cars being sold at the extensive margin. My model conditions on entry into the new car market so I make no claims on the extensive margin results. Adamou, Clerides, and Zachariadis (2013) counterfactually study the impact of a feebate, finding that the reform needs to look more like a fee than a rebate in order to be optimal. Grigolon, Reynaert, and Verboven (2015) find that fuel taxes are more efficient than vehicle taxes in reducing fuel usage than taxes working through the fuel efficiency of cars. Using cross-country market-level data, they find that discriminatory fuel taxes and differences in fuel efficiency alone explain 40% of the differences across countries. My results indicate that discriminatory ownership and purchase taxes may well account for a substantial part of the remaining 60%. Mabit (2014) also uses Danish data and analyzes the 2007 reform that is also under study in this paper and finds the changes in car characteristics occurring in the period to be as important as the reform.

A number of other studies consider more small-scale reforms, typically affecting smaller segments. These are generally found to be cost-ineffective. Huse and Lucinda (2013) consider a Swedish reform affecting only highly efficient green cars using a BLP model. They find that the implicit price of CO2-emissions from that reform was far above the social cost of carbon in Sweden. Beresteanu and Li (2011) and Chandra, Gulati, and Kandlikar (2010) study incentive schemes aimed at hybrid cars in the US and Canada and both find them to be cost-ineffective.

The papers cited above all target the demand side of the market but a large American literature focuses on supply side instruments, primarily the Corporate Average Fuel Economy (CAFE) standards. These require car makers to reach a certain weighted average fuel economy across their sold cars, subject to a number of technical details. Goldberg (1998) is one early study of CAFE standards utilizing joint modeling of car choice and usage, finding that policies targeting the car choice are favorable to fuel taxes. Building on the framework by Bento et al. (2009), recent work by Jacobsen (2013) compares the cost-effectiveness of CAFE standards and fuel taxes, finding the latter to be the more effective. Reynaert (2014) and Clerides and Zachariadis (2008) are among the few papers studying the effects of the European fuel economy standards, announced in 2007 and to be fully binding by 2015. Reynaert (2014) focuses on the responses of the European automakers, finding that they primarily respond by technology adoption.

A different strand of literature looks at the fuel type of the purchased cars, focusing on the choice of diesel vs. gasoline. This is a much more prevalent option in the European than the American context and the diesel market share increased substantially up through the early 1990's, following the introduction of the direct injection or common rail technology. Miravete, Moral, and Thurk (2014) study this in the Spanish setting, finding that the policy treatment of diesel vs. gasoline in Europe functioned in effect as a subsidy to European car makers. On the methodological side, Verboven (2002) uses within-model variation between car models that only differ in using gasoline or diesel fuel for identification in a BLP framework. Grigolon, Reynaert, and Verboven (2015) also consider heterogeneity in driving but assume a zero fuel price elasticity of driving. My paper is the first to my knowledge to study the dieselization while estimating the driving decision simultaneously.

Endogenous selection of consumers into car types based on individual driving demand has been emphasized in recent work. This paper builds on Gillingham (2012) who introduces endogenous selection both based on observables, unobservables and explicitly on expectations about future fuel prices. The model builds on Dubin and McFadden (1984). Some work has used 2-step approaches to integrating type choice and usage (e.g. Goldberg 1998; West 2004; D'Haultfæuille, Givord, and Boutin 2013), while more recent work has promoted simultaneous estimation (e.g. Bento et al. 2009; Feng, Fullerton, and Gan 2013; Jacobsen 2013 and in particular Gillingham, 2012). The model explicitly accounts for the selection effect required to identify the so-called *rebound effect*, namely the effect on driving of increasing fuel efficiency (see e.g. Small and Van Dender, 2007).

In terms of the data used, this paper is novel in applying micro data on car choice and usage matched with household-level demographics for the full Danish population over a long period of 9 years. Many papers in the car demand literature have only used market-level data (e.g. Berry, Levinsohn, and Pakes 1995; Miravete, Moral, and Thurk 2014; Reynaert 2014; Verboven 2002). The papers using micro-level data either use survey data (West 2004; Bento et al. 2009; Jacobsen 2013), often with only a limited number of years, or do not observe household demographics at the micro level (e.g. Gillingham, 2012).

Two major aspects of car demand that I do not tackle in this paper are multi-car households, dynamics and myopia. Even though the data would allow it, I choose not to include 2-car households in this study.[2] This is to make sure the choiceset in the model remains tractable. Since only 12.1% of Danish households between 18 and 65 years own 2 or more cars, I capture the largest segment this way (see Figure B.8).

A recent literature has looked at the question of whether consumers correctly take into account future savings in fuel cost when making a car purchase.[3] I make no claims to answering this question but will follow the empirical work indicating that that consumers are rational and time-consistent when they make their vehicle and driving decisions. However, I will allow some flexibility in consumer expectations about future fuel prices.

Finally, some authors have emphasized the dynamics of vehicle ownership decisions, opting for a fully dynamic structural model.[4] While this facilitates the study of important aspects such as the used-car market, scrappage and ownership durations one must trade off complexity elsewhere in the model and it is central to maintain a high-dimensional choiceset to accurately fit in the effects of the policies considered. As most other non-dynamic papers, the model presented in this paper conditions on entry into the new car market. If the reforms change substitutions between the used and new car market, such effects will be ignored. In that sense, the focus of this paper is purely on the substitution patterns in the car market.

2 Background and Data

In this section, I will first describe the institutional setting in Denmark, focusing on the taxation of cars in the period. I then discuss the data, explaining the different data sources and the construction of the final dataset. Finally, I present descriptive evidence on car choice and driving.

2.1 Institutional Setting

Car taxation in Denmark consists of three elements; a registration tax, a bi-annual ownership tax and fuel taxes. The registration tax is paid at the time of purchase and is a linear function of the purchase price with a kink,

$$\tau_t^{\text{reg}}(p^{\text{gross}}) = 1.05 \cdot \min(K_t, p^{\text{gross}}) + 1.80 \cdot \max(0, p^{\text{gross}} - K_t),$$

[2]Some of the only studies focused on modeling multi-car households are Spiller (2012); Borger, Mulalic, and Rouwendal (2013); Wakamori (2011). Bento et al. (2009) take a different approach, considering each car as a *choice occasion*. An alternative approach in my setting would be to ignore knowledge about other cars and consider the two instances as independent or to add a control.

[3]The findings have been mixed with some support for myopia (Allcott and Wozny, 2012) and some against (Busse, Knittel, and Zettelmeyer (2013); Sallee, West, and Fan (2010); Grigolon, Reynaert, and Verboven (2015)). The interested reader is referred to the literature review by Greene (2010) which documents that there has been extremely mixed evidence in the empirical literature. Another strand of literature emphasizes certain behavioral aspects that I will not consider in this paper; Gallagher and Muehlegger (2011) find that tax incentives working through the purchase price are more effective than ones working through income tax deductions, and Li, Linn, and Muehlegger (2014) find that driving responds more strongly to fuel taxes than to changes in the fuel product price.

[4]Many recent dynamic models build on the optimal replacement model by Rust (1987). These models are much better suited to looking at issues like vehicle scrappage (Adda and Cooper, 2000; Schiraldi, 2011), and the used car market (Adda and Cooper (2000); Schiraldi (2011); Chen, Esteban, and Shum (2010); Gavazza, Lizzeri, and Rokestkiy (2014); Gillingham et al. (2013); Stolyarov (2002)Chen, Esteban, and Shum, 2010; Gavazza, Lizzeri, and Rokestkiy, 2014; Stolyarov, 2002; Gillingham et al., 2013). Such issues are beyond the scope of this paper.

where K_t is a politically set kink, $\tau_t^{\text{reg}}(\cdot)$ denotes the registration tax and p^{gross} is the raw car price including VAT (25%) but net of deductions.[5] Consequently, taxes make up just over 160% of the purchase price of the average Danish car. The second tax, ownership tax, is paid twice a year and depends on the fuel efficiency (in kilometers per liter, km/l) of the car according to a schedule that is updated irregularly over the period and accounted for in the estimation. There is a separate schedule for diesel cars where the tax rate is higher for any given level of fuel efficiency. This balances the fact that diesel cars on average have higher fuel efficiency than gasoline cars. The third tax element, fuel taxes, are comprised of a fixed and a proportional component and the total fuel tax amounts to 68.0% of the gasoline price, averaged over my sample period (58.5% for diesel). The composition of taxes and product price for the gasoline and diesel prices are shown in Figure B.4.

There were two major reforms of interest in the sample period; A change in the bi-annual tax in 1997 and a change in the registration tax in 2007. My data does not cover both before and after either of these reforms. All cars first registered before July 1st 1997 have their bi-annual tax rate set according to the weight (and still follow that scheme) while those first registered after that date follow the fuel efficiency. The 2007 reform was a so-called *feebate*, working through the registration tax and giving a rebate to green cars and added a fee to inefficient cars. The rebate was DKK 4,000 per unit of km/l over the pivot (16 km/l for gasoline cars and 18 km/l for diesel cars). The corresponding fee was slightly lower, at 1,000 DKK per km/l. Figure 2.1 shows the prompt change in the fuel efficiency of new cars after the reform is introduced and 2.2 shows an even greater change in the diesel share of newly purchased cars. From the European Automobil Dealer Association, I have access to the diesel share in other European countries, which is also shown in 2.2, highlighting that the response was unique to Denmark.[6]

2.2 Data

The dataset contains all new cars purchased between July 1st 1997 and December 31st 2006 and is based on matched Danish administrative data. The car ownership information comes from The Central Motor Register, which holds license plate ownership information. Driving information comes from the mandatory safety inspection which all cars must attend four years after purchase. At this test, it is evaluated whether the car is in safe condition and the odometer is measured and recorded. Therefore, the driving data comes from a 4-year period following purchase. Demographic informations on the car owners and the remainder of their household is obtained by matching the personal identifier (CPR number) with the Danish registers. The most important variable is the computed work distance measure (described in appendix B.3). This measure captures the product of the work distance and the number of days that the individual goes to work, regardless of the mode choice. Households are only eligible for the deduction if they are working and their private address is further than 12 km from the address of their

[5] Deductions are given for example for installed safety equipment which are not observed in the data and therefore ignored in this paper. Anecdotally, some deductions are larger than the cost of installing the equipment, meaning that the equipment is universally adopted.

[6] I have been unable to get the similar fuel efficiency numbers for other European countries. I expect that the Danish response is unique in relation to the timing but that the general trend is certainly shared across countries. The source for the Danish diesel share and average fuel efficiency post-2007 is Statistics Denmark's aggregate statistics (statistikbanken.dk), but I do not have this information in my micro data.

primary work place, which is the case for a little under half of the individuals. Appendix B.3.1 provides details on this unique variable.

A car type in the data is defined as a unique Vehicle Type Approval number. These are identifiers assigned by the Ministry of Transportation when a car is approved for import and sale in Denmark. They vary at a finer level than the traditional (make-model-year) in some respects, since any change in the vehicle that might alter safety aspects of the car in operation require a new approval for import. The identifier does not contain information on the make year, however. Car characteristics are merged using this identifier. An important limitation of the data is that I do not observe the age of the car; instead, I observe the year the car was first registered in Denmark and use this to construct the age, assuming that the car is not an imported used car. Imports of used cars are not a big problem for my setting because the high Danish car taxes imply that the used-car prices are generally very high. I have access to new car prices and depreciation rates are available from a dataset maintained by the Danish Automobile Dealer Association (DAF). The depreciation rates are used by used car dealers in Denmark when they make an offer on a used car of a given age in normal condition and the new car prices are merchant suggested retail prices (MSRPs). Fuel prices are available at the daily level from the Danish Oil Industry Association (EOF; www.eof.dk). These prices are recommended retail prices for the entire country so local variations and price wars do not show up in the data.[7] In Appendix B.3.2, I show that the product prices of both types of fuel track international oil prices very closely (Figure B.5). All tax rates are taken directly from the law texts using www.retsinformation.dk with the exception of fuel taxes, that come from EOF.

As many of the classic car choice papers, the emphasis of this paper is on the new car market. While car ownership is observed for used cars, prices and characteristics are only available for cars purchased from 1997 and forward.

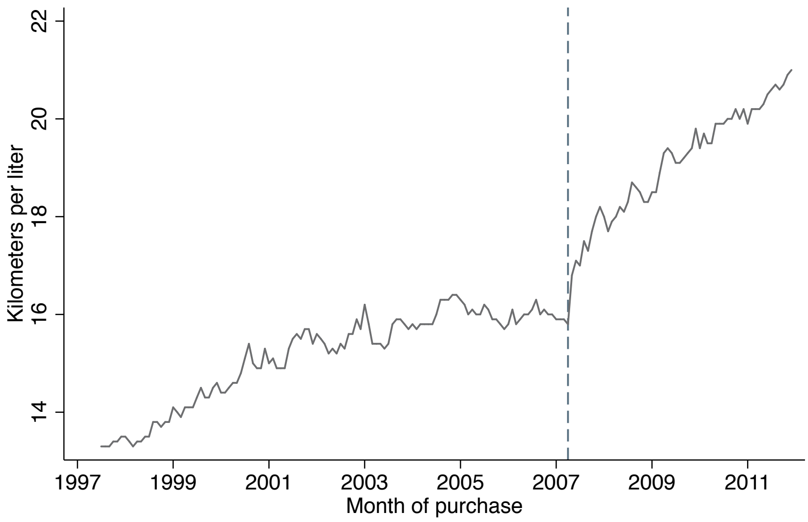
In order to evaluate the welfare consequences of the counterfactual policies, one needs a measure of the marginal external costs of driving. These are taken from DTU Transport (2010) and shown in Appendix B.2.[8] The key thing to note about externalities is that the per-kilometer externality of congestion and accidents are far larger than environmental externalities (this has been emphasized by e.g. De Borger and Mayeres, 2007).

The final estimation sample contains $N = 128,910$ new car purchases by Danish couples in 1997–2006. The sample selection is described in details in Appendix B.1. To ensure demographic heterogeneity, I have selected only households consisting of couples. Adding singles could easily be done but would require many additional parameters and they account for less than 20% of all new purchases. I also deselect cars with missing observations as well as car types that are purchased fewer than 30 times. The final dataset has a total of $J = 1,177$ different cars to choose from. Even so, the choiceset facing a single household is much smaller than this because no car was available in all sample years. Working with a choiceset of this high dimensionality in a discrete choice setting is challenging but it allows me to implement and explore the tax system very precisely.

[7]In the literature estimating the demand for driving, many papers rely on spatial variation in fuel prices for identification. This would not be appropriate for Denmark, however, since the country is so small that it would be hard to establish regions that would avoid trading across markets.

[8]I have recalculated from a per kilometer to per liter externality in terms of air pollution from CO_2 and other particle emissions.

Figure 2.1: Fuel Efficiency of Newly Purchased Cars in Denmark



Source: Statistics Denmark, avg. fuel efficiency of newly purchased cars.
The horizontal line marks the 2007 reform in April.
The dots show yearly averages over the middle date of the year.

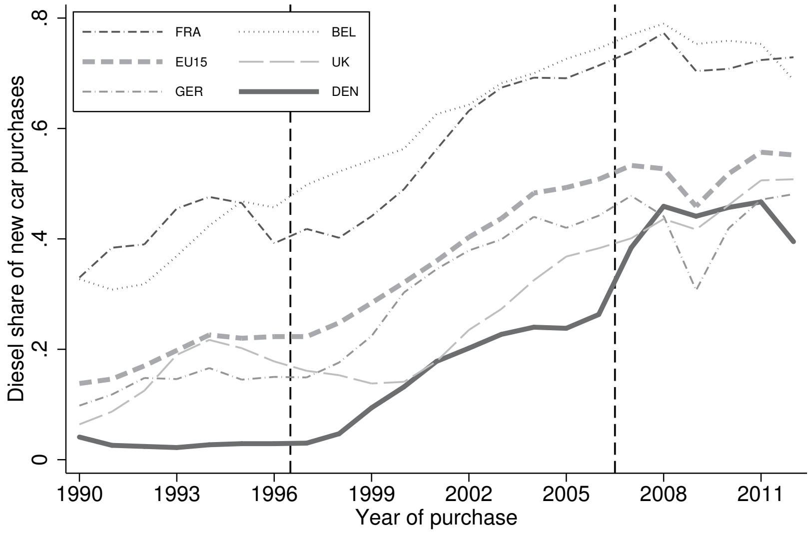
2.3 Descriptive Evidence

In the period 1997–2006, the fuel efficiency of newly purchased cars increased from just over 13 kilometers per liter (km/l) to 16 km/l, as shown in Figure 2.1. The figure furthermore shows a sharp change occurring right when the Danish feebate of 2007 was implemented in 2007. However, in the same period there was a drastic increase in diesel car sales, which made up 3.0% of all new cars sold in 1997 but had increased to 26.3% by 2006. Furthermore, this number increased to 38.4% in 2007. While the increasing trend over the period was shared by many other European countries, where the average diesel share increased from 22.3% to 50.8%, the jump in 2007 is absent for those countries. Figure 2.2 shows the diesel share of new purchases for Denmark together with 4 other countries and the Western European average. The common trend naturally opens the question of how much of the changes in characteristics was driven by changes in demand, supply and policy.

Table 2.1 shows summary statistics for the estimation sample both in terms of cars and households. Regarding average work distance variable, this is zero if the household has less than 12 km to work. The reported averages of 11.8 km for males and 8.12 for females are therefore the averages of this censored variable.

To get a first grasp of the conditional correlations in the data, Figure B.13 shows the distribution of driving for gasoline car drivers and for diesel car drivers. The average gasoline car drives 49.2 km per day while the average diesel car drives 78.8 km per day. This is confirmed in Table B.4; the table shows regressions where car characteristics of the purchased vehicles are regressed on the demographic variables of the households purchasing them. The estimates indicate that an increase in the male's work distance of one standard deviation is associated with the probability that the household buys a diesel car by 5.1 percentage points. The coefficient on real household income is positive for weight, engine power (kW) and size (cc) and the real price. This means that richer households tend to buy larger and more powerful cars. Figure

Figure 2.2: Diesel Cars — Fraction of Total New Car Sales in European Countries



The horizontal lines mark the Danish reforms in July 1997 and April 2007
Source: ACEA (http://www.acea.be/collection/statistics).

B.12 visualizes the spatial dimension of this and shows that the urban regions of Denmark have low work distances and low diesel shares while some of the regions with the longest work distances also have a higher prevalence of diesel cars. Moreover, the figure shows that there are rural regions in the eastern part of Denmark where diesel cars are very popular in spite of work distances being lower. Appendix B.3 contains more descriptive statistics.

In Appendix B.4, I present detailed descriptives for the fuel price development over time. Fuel prices have increased by 23.0% and 33.7% for gasoline and diesel fuel respectively. This has mainly been driven by changes in the product price as Danish fuel taxes rates have been largely unchanged over the period (cf. Figure B.4). The fact that the diesel share has increased in spite of diesel fuel prices growing faster than gasoline prices indicates that either the characteristics or the differential tax rates of diesel cars have changed even faster in a favorable direction. Finally, even though the relative price of diesel to gasoline has increased from 80.6% to 89.2% over the period, there has substantial gyrations in the relative price year to year (Figure B.6).

More detailed descriptives are presented in appendix B.3 but to paraphrase, the only household demographic that appears to predict diesel purchase is the home-work distances of each of the spouses. This variable is also an important predictor of the household's vehicle kilometers travelled (VKT) and elasticity of driving with respect to the price per kilometer (PPK). The variable is rarely available in empirical studies and often considered to be the main component of household fixed effects in driving equations.

3 Model

In this section, I outline the decision model of the households. I first present the functional form of the two-period utility function. I then solve for optimal planned driving in the second period, before inserting this back into the first-period utility to derive the expected utility of choosing a given car.

Table 2.1: Summary Statistics — Shortened Names in Parentheses

Car Variables			
	N	Mean	Std.
Fuel efficiency (km/l, e)	128,910	14.68	2.56
Weight (tons, q^{weight})	128,910	1,660.80	201.63
Horsepower (kW, q^{kw})	128,910	70.71	16.94
Displacement (cc, q^{displace})	128,910	1,580.08	265.40
Diesel (%)	128,910	0.1108	0.31
Price (2005 DKK, p^{car})	128,910	219,284.20	66,522.11
Depreciation factor (δ)	128,910	0.8741	0.0118
Units Sold	128,910	228.20	213.48
Demographic Variables			
	N	Mean	Std.
Work distance, male (WDm)	128,910	11.80	19.63
Work distance, female (WDf)	128,910	8.12	14.84
Gross income (2005 DKK, inc)	128,910	701,058.5	456,223.5
Number of kids (nkids)	128,910	0.9866	1.07
Unemployment, male (unempm)	128,910	0.0859	0.28
Unemployment, female (unempf)	128,910	0.1616	0.37
Age, male (agem)	128,910	43.99	10.12
Age, female (agef)	128,910	42.00	10.27
Male income %	128,910	0.5894	0.13
Urban area (bigcity)	128,910	0.2084	0.41

The model builds on the discrete-continuous selection model literature going back to Dubin and McFadden (1984). The idea is that the usage in the second period comes out of Roy's identity. This type of framework was applied to car choice and usage by Bento et al. (2009) and Gillingham (2012). The model presented below is based closely on the latter but with the extension of allowing household demographics to affect driving not only through the price sensitivity parameter but also through the mean driving.

3.1 Household Utility

The model is a two-period model; in the first period, t_1, the household purchases a car of type j at the price p_j^{car} under uncertainty about fuel prices in the future. In the second period, t_2, fuel prices are realized and the household makes its driving decision. Households enter the new market at different points in time and thus face different sets of available cars, $\mathcal{J}_{t_1}$, and different fuel prices. In the implementation, t_1 is the calendar year in which the household enters the new car market, i.e. $t_1 \in \{1997, ..., 2006\}$. The driving period length is four years, because the first mandatory safety inspection at which the odometer is measured in the data occurs after four years. At the end of the second period, four years later, the car is sold at the used-car price given by $\delta_j^4 p_j^{\text{car}}$, where δ_j is a car-specific annual depreciation factor obtained from the Danish Automobile Dealer Association (the δ_j is 0.874). There is no outside option of not owning a car and there are no used cars in the choiceset.[9] In that sense, the model conditions on entry into

[9]The main reason for not having an outside option because this simple quasi-linear two-period model is not well-suited to deal with the inherently dynamic problem of purchasing a car, which represents a major investment

the new car market but remains agnostic about why and when this entry occurs.[10]

The utility function takes the form

$$u_{ij} = u_{ij1} + \beta^4 \mathbb{E}(u_{ij2}),$$

where β is the annual discount factor (fixed at 0.95) which is raised to the power four because there are four years between purchase and driving period. Both of the period-utilities are quasi-linear in the consumption of the composite outside good. First-period utility takes the form

$$u_{ij1} = \gamma_i \left(y_{it_{1i}} - p_j^{\text{car}} - 4\tau_j \right) + u^{\text{own}}(j),$$

where $u^{\text{own}}(j)$ is utility from owning a car but not related to the driving, τ_j is the annual tax and y_{it} denotes household income in period t. The parameter γ_i scales the utility of money relative to that of driving and it varies across households according to $\gamma_i \equiv \gamma_z' z_i$, where z_i is a vector of household demographics. For the primary results, I let $u^{\text{own}}(j) = \alpha_0' q_j$, where q_j is a vector of observable characteristics for the car such as weight, engine power but not including fuel efficiency, e_j, which is restricted to enter the model through the cost structure.[11] This term shifts mean utilities of buying a given car in a way that is unrelated to the driving utility so as to better fit market shares.

In the second period, the household must choose how many kilometers to drive, x. The second-period utility is given by

$$u_{ij2} = \gamma_i \left(y_{it_2} + \delta_j^4 p_j^{\text{car}} - \frac{p_{jt_2}^{\text{fuel}}}{e_j} x \right) + \alpha_{1ij} x + \alpha_2 x^2,$$

where e_j is the fuel efficiency of car j in kilometers per liter, $p_{jt_2}^{\text{fuel}}$ is the price of fuel (gasoline or diesel depending on the fuel type of car j), and α_{1ij} is a parameter that affects the utility of driving an extra kilometer. This parameter is heterogeneous and correlated with demographics and car characteristics as follows:

$$\alpha_{1ij} \equiv \alpha_{10} + \alpha_{1z}' z_i + \alpha_{1q}' q_j + c_i.$$

The variable c_i is a time-constant random effect that is independent of z_i and q_j and captures heterogeneity in the utility of driving that is unobserved by the econometrician but observed by the household. The assumption that utility from driving is quadratic yields a computationally attractive form for optimal driving as we shall see. It implies theoretically a bliss point in driving but in the application, all households were far below this point. The coefficient α_2 has also been allowed to vary over i and j but the additional parameterization did not improve model fit so I

in Denmark on account of the large taxes. I ignore the used-car market partly due to missing data on car characteristics, which would heavily skew my sample over time, and partly due to the dimensionality; including that many more car types would force be to reduce the dimensionality of the choiceset.

[10] One could imagine a fully dynamic optimal stopping problem where the consumer in each period considers replacing his current car, e.g. Schiraldi (2011). However, then it would be computationally very challenging to have a choiceset of $J = 1,177$ cars.

[11] In future work, it would also be interesting to include information on parents' automobile choice where available in the registers as persistence in brand preference within a family has been documented in the literature (Anderson et al., 2013).

chose the more parsimonious specification.

3.2 Solving the Consumer's Problem

In period t_2 when the household makes its VKT choice, x, it conditions on the purchased car. Thus, optimal driving maximizes u_{ij2} conditional on j. Interior solutions must therefore satisfy the first-order condition;[12]

$$x = -\frac{1}{2\alpha_2}\left(\alpha_{1ij} - \gamma_i \frac{p^{\text{fuel}}_{jt_2}}{e_j}\right) \equiv x^*_{ij}(p^{\text{fuel}}_{jt_2}). \tag{3.1}$$

Thus, optimal driving is characterized by a linear equation, where car characteristics shift the level of driving and household demographics shift both the level and the price sensitivity of driving. In particular, note that the unobserved driving type, c_i, shifts the level of driving. The linear form conveniently allows me to relate the structural parameters to reduced-form regressions of VKT on the price per kilometer, defined as the fuel price divided by the fuel efficiency, $p^{\text{fuel}}_{jt_2}/e_j$, since the scaled parameters, $-\frac{\alpha_{1ij}}{2\alpha_2}$ and $\frac{\gamma_i}{2\alpha_2}$, are identified by the driving equation. This is also useful for finding good starting values.

When I insert the optimal driving rule from (3.1) back into the full utility function I obtain an expression that can be computed based on data. Due to the quasi-linearity, the income term, $\gamma_i(y_{it_{1i}} + \beta^4 y_{it_{2i}})$, does not vary over j and so can be dropped from the specification. Instead, income is allowed to enter through both the heterogeneous parameters, α_{1ij} and γ_i, to capture correlations with taste patterns and leisure activities. The expected utility of choosing car j is

$$\begin{aligned} u_{ij} \quad = \quad & -\gamma_i 4\tau_j + \gamma_i\left[1 - (\beta\delta_j)^4\right] p^{\text{car}}_j + u^{\text{own}}(j) \\ & + \beta^4 \mathbb{E}\left\{ -\gamma_i \frac{p^{\text{fuel}}_{jt_2}}{e_j} x^*_{ij}(p^{\text{fuel}}_{jt_2}) + \alpha_{1ij} x^*_{ij}(p^{\text{fuel}}_{jt_2}) + \alpha_2 \left[x^*_{ij}(p^{\text{fuel}}_{jt_2})\right]^2 \middle| p^{\text{fuel}}_{jt_1} \right\}. \end{aligned} \tag{3.2}$$

All that remains is to specify the household's expectations at time t_1 about fuel prices at time t_2 conditional on fuel prices at time t_1s. In the literature, many implementations have used static expectations, whereby the expectation in (3.2) collapses to a single number. Gillingham (2012) uses a unit root and also allows consumers to use prices of futures on fuel in their forecast. He finds that it makes little difference to his results. I have implemented both static expectations, perfect foresight and a unit root with a drift. For the unit root, the expectation in equation (3.2) must be solved by numerical integration. I do this using Gauss-Hermite quadrature, which performs extremely well for univariate integrals. As it turns out, the specification of the fuel price expectations do not greatly impact my main results. There are two intuitive reasons for this; firstly, the variation in fuel efficiency in the choiceset is larger than the variation in fuel prices over time. Secondly, the quasi-linear utility function implies that consumers are risk neutral. In a model with diminishing marginal utility of money or credit constraints, concerns about fuel prices rising too much might push the household down to low levels of consumption and high curvature. The non-linearity of such a model could yield much greater differences depending on the expected fuel prices.

[12] At the estimated parameter values, the model only predicts strictly positive VKT for all households.

One simplification that the fuel price expectation structure has imposed in all cases is that the gasoline and diesel price processes are not modeled jointly by the households. The price of diesel has moved from 80.6% to 89.2% of the gasoline price over the period, but there are substantial fluctuation year to year (cf. Figure B.6). The forecasts from a bivariate time-series process of the two fuel prices, possibly including oil prices, might yield an improvement but I leave this for future work.

I will conclude the model section with a brief discussion of the assumptions imposed by the model. The quasi-linearity of the model affords a lot in terms of simplifying the model solution. An alternative interpretation, due to Bento et al. (2009), is that the model considers the problem of a household *renting* a car for four years; since the household pre-commits to selling the car again and there is no uncertainty about future car prices, the analogy is very clear. This simplification admits more complexity elsewhere. Moreover, curvature is more likely to make a difference for the decision about *when* to go on the new car market; households might choose to postpone car purchases simply due to the fear of becoming unemployed and receiving a large negative income shock. Since this is beyond the scope of this paper, I choose to focus on having a highly detailed model of the car choice conditional on entry. Instead, I rely on capturing some of these effects by allowing income to change the marginal utility of money and driving by including it in γ_i and α_{1ij} to capture some of these effects. This is similar to how many papers in the literature following Berry, Levinsohn, and Pakes (1995) have done it.

Computationally, the main challenge with the implementation of the model is the dimensionality of the choiceset, $\mathcal{J}$. Avoiding aggregating cars has the advantage of clarity as well as precision in terms of calculating tax revenue and other counterfactual outcomes that rely on the precise characteristics of individual cars; such details might get lost in aggregation. The model has been implemented in `c`, which has yielded a considerable speedup over Matlab in particular due to parallelization and explicit utilization of the sparsity structure of $\mathcal{J}$ due to some car models not being available in all years.

4 Empirical Strategy

In this section, I first outline the econometric methodology and derive the likelihood function. I then discuss where the identifying variation is coming from in the data and comment on the implementation of the estimator. Finally, I outline how I simulate from the model and calculate counterfactual outcomes based on the estimated parameters.

4.1 Econometric Methodology

The econometric methodology follows Gillingham (2012). The dataset contains for each household the discrete car choice, d_i, and the continuous driving choice, x_i. Furthermore, it contains the realized average fuel price over the household's driving period, $p^{\text{fuel}}_{jt_{2i}}$, and finally the vector of demographic variables, z_i. The subscript i in period t_{2i} is to remind the reader that there is cross sectional variation in the fuel price insofar as two households' periods do not perfectly overlap. Fuel prices also vary with j depending on the fuel type of the car. Other than that, the year of purchase gives the annual fuel prices that year and the choiceset and characteristics of the cars available in that year.

To obtain non-degeneracy of the model, an error term is added to both choice margins; an IID Gaussian measurement error to the optimal driving equation and an IID Extreme Value term to the conditional utility, u_{ij}. The observed driving for household i, x_i, is therefore written as

$$x_i = x^*_{id_i}(p^{\text{fuel}}_{d_i t_{2i}}) + \eta_i, \quad \eta_i \sim \mathcal{N}(0, \sigma_x^2),$$

This means that the partial likelihood contribution for the observed driving is given by

$$f_x(x_i|\theta) = \frac{1}{\sigma_x\sqrt{2\pi}} \exp\left\{ -\frac{\left[x_i - x^*_{id_i}(p^{\text{fuel}}_{d_i t_{2i}})\right]^2}{2\sigma_x^2} \right\}, \tag{4.1}$$

where the dependence of predicted driving, $x^*_{id}(\cdot)$, on the unobserved type, c_i, is subsumed.

For the type choice, the full utility for household i from choosing car $j \in \mathcal{J}_{t_1}$ becomes

$$\tilde{u}_{ij} = u_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ij} \equiv \frac{1}{\lambda}\tilde{\varepsilon}_{ij} \quad \tilde{\varepsilon}_{ij} \sim \text{IID Extreme Value.}$$

I will discuss the scale parameter, λ, in greater detail below. The probability that car j maximizes household i's utility is therefore given by

$$\Pr_i(j|\theta) = \frac{\exp(u_{ij}/\lambda)}{\sum_{j' \in \mathcal{J}_{t_1}} \exp(u_{ij'}/\lambda)}.$$

I will estimate one version of the model where $c_i = 0$ for all i. In that model, the full log-likelihood contribution for for household i becomes

$$\ell_i^{\text{full}}(\theta) = f_x(x_i|\theta)\Pr_i(d_i|\theta).$$

In the general case, I will assume that $c_i \sim \mathcal{N}(0, \sigma_c^2)$ and the likelihood gets the typical integrated likelihood form similar to the mixed logit:

$$\ell_i^{\text{sim}}(\theta) = \int f_x(x_i|\sigma_c c; \theta)\Pr_i(d_i|\sigma_c c; \theta)\, \mathrm{d}\Phi(c),$$

where Φ is the Gaussian cdf. The conditioning on the individual effect $c = \sigma_c c_i$ is made explicit in both $f_x(\cdot)$ and $\Pr_i(\cdot)$ in the equation as a reminder that it enters into α_{1ij} and thus in both optimal driving and choice-specific utilities. In this sense, the c_i variable has the interpretation of a random effect. The univariate integral will be computed using Gauss-Hermite quadrature.[13]

The model has been implemented in the programming language c using Matlab's interface, Mex. For optimizing the likelihood function, I have alternated between using a gradient based (quasi-Newton) and a gradient-free (Nelder-Mead) solver with semi-analytic numerical gradients (exploiting the linear structure of the random coefficients) and BHHH approximation of the

[13] For the results presented here, only 8 nodes were used. Future work is under way using more nodes. Comparing quadrature with simulation using simple, smooth functions and univariate integrals, it was found that quadrature attains the same level of precision as simulation using five to ten times more evaluations of the integrand. This point was also highlighted by Dubé, Fox, and Su (2012) and Judd and Skrainka (2011).

Hessian due to Berndt et al. (1974).[14]

The logit scale parameter, λ, is not identified in the outset because the scale of utility can be moved up and down by α_2. However, I found that the likelihood was more easy to manage numerically with a re-normalization setting $\alpha_2 := -1$ and estimating λ instead. Unfortunately, the likelihood function turned out to be extremely flat in the direction of λ. Instead, I estimated the model over a grid of λ-values and picked the λ that produced the best fit for the data while also giving sensible elasticities. If I allow the optimizer to choose λ freely, the optimizer terminates without convergence at a λ value of just over 100,000, at which point the model produces zero elasticities (to the fifth decimal) on all margins. I discuss this issue in greater detail in Appendix C and outline a potential model extension that would allow me to estimate the scale parameter jointly with the remaining parameters. This approach involves estimating car type fixed effects vis-a-vis Berry, Levinsohn, and Pakes (1995).

4.2 Identification

The model relies on both cross-sectional and time-series variation as well as within-household variation. The variation in fuel prices and the choiceset over time identifies how households substitute between available cars under different circumstances. The parameters in the utility function are moreover tied down by there being two observed outcomes for each household; the discrete car choice and the continuous driving choice. In that sense, the model intuition is not far from a Heckman selection model; the exclusion restrictions are the fuel prices at the time of purchase, the choiceset available at the time of purchase as well as the structure of the model. In essence, the model imposes the strict cross-sectional restriction that consumers value money in a similar fashion when making car purchase decisions and driving decisions. The driving decisions should be thought of as covering several years and not the daily driving decisions, where households can switch purchases over the week days in response to daily variation in prices.

There has been a considerable increase in fuel prices in my sample period which, as discussed in section B.3, has arguably been driven by world-market factors. To leverage variation from changes in the tax rates over the period, I have explicitly coded the annual tax rates, τ_j, and included those in the model. The characteristics of available cars have also changed substantially due to technological progress over the ten-year period, which has made cars more fuel efficient for any given level of car weight. These sources of variation are fine to the extent that the changes in car makers' portfolios is driven by tax policies or demand side effects in other, bigger markets. However, there may of course be common trends in demand across countries leading to this. For example, urbanization patterns across many developed countries have followed similar patterns with more households moving to the urban areas. My work distance variable will capture such trends, so in terms of the driving equation, I am more worried about correlated trends in leisure driving. In related work, Gillingham and Munk-Nielsen (2015) explore many different sources of variation to estimate the medium-run, 1-year elasticity of driving with respect to fuel prices and find a central elasticity of -0.30 with household fixed effects. This is very close to what I

[14] Whenever the gradient-based solver would get stuck, unable to improve the likelihood along the gradient direction, the Nelder-Mead solver proved useful in breaking free of the local optimum.

find when I take into account selection, even though I don't include fixed effects.

In terms of the discrete car choice, the model can be thought of as a mixed logit with a particular functional form imposed on the choice-specific utilities. In much of the literature on car choice the driving equation is not considered but there will often be either sophisticated nesting structures on the logit errors or car specific fixed effects in the Berry, Levinsohn, and Pakes (1995) or both (Grigolon, Reynaert, and Verboven, 2015). In future research, it would be interesting to see these features integrated in a discrete-continuous choice model. I propose such an extension in Appendix C but leave the estimation to future research.

4.3 Simulating From the Model

As with most structural models, it is essential to be able to simulate counterfactual behavior from the model. Essentially, we want to compute simple statistics characterizing the final market outcome of making changes to taxes, prices or the characteristics of cars. These outcomes might be the CO_2 emitted, tax revenue, the average fuel efficiency, etc. Formally, suppose we are interested in some outcome ω_{ij}. Then define the average expected outcome as

$$\tilde{\mathbb{E}}(\omega|\theta) \equiv \frac{1}{N}\sum_{i=1}^{N}\sum_{j\in\mathcal{J}_i}\Pr_i(j|\theta)\omega_{ij}. \tag{4.2}$$

This is the average (over households) weighted average (over available choices weighted with conditional choice probabilities) outcome.

Note that in the computation of (4.2), I need to take a stand on the stochastic variables in the model; η_i, ε_{ij} and c_i. The measurement error is set to zero, $\eta_i := 0$. Since I am weighting by conditional choice probabilities, the expression is implicitly an expectation over ε_{ij}. Lastly, c_i is set to zero for all households; instead, one could integrate out the random effect unconditionally, but given the quasi-linearity and the linear driving equation, it is unlikely that such efforts would yield very different results.[15] Standard errors have not been computed for the expected outcomes.

Two examples of outcomes of particular interest require an extra comment. Firstly, the CO_2 emissions; These are calculated using the kg of CO_2 that is emitted by the combustion of a liter of each fuel,[16] yielding the following CO_2 emissions (in kg) conditional on choosing car j and realized fuel price $p^{\text{fuel}}_{jt_{2i}}$,

$$CO_{2,ij} \equiv \left(\mathbf{1}_{\{j\text{ is gas}\}}2.392^{\text{kg}}/_{\text{l}} + \mathbf{1}_{\{j\text{ is diesel}\}}2.64^{\text{kg}}/_{\text{l}}\right)\frac{x^*_{ij}(p^{\text{fuel}}_{jt_{2i}})}{e_j}.$$

Setting $\omega_{ij} := CO_{2,ij}$ in (4.2) gives the average expected CO_2 emissions. The analysis emphasizes CO_2 emissions to focus on the environmental aspects but might as well have emphasized fuel

[15] The reason why the random effect makes a difference in estimation is that here, information from both periods are employed simultaneously and thus the simulated likelihood will apply the highest weight to the region of the support of c_i that best rationalize household i's two decisions. An alternative approach that might yield different results would be to integrate out c_i *conditional on choices*; this is in line with the approach outlined in Train (2009, ch. 11). That strategy has some similarities with a latent class model where one can compute the probability that $c_i = c_q$ for some quadrature node q, and use these weights in counterfactual simulations.

[16] These numbers come from www.ecoscore.be (and are confirmed by www.environment.gov and www.epa.gov).

use; the two are proportional.

Secondly, the tax revenue can be calculated conditional on car purchase and subsequent usage. The conditional total tax revenue, τ_{ij}^{total}, is given by

$$\tau_{ij}^{\text{total}} \equiv \tau_j^{\text{fuel}} \frac{p_{jt_{2i}}^{\text{fuel}}}{e_j} x_{ij}^*(p_{jt_{2i}}^{\text{fuel}}) + \tau^{\text{reg}}(p_{tj}^{\text{car}}) + 4\tau^{\text{annual}},$$

where $\tau^{\text{reg}}(\cdot)$ gives the registration tax and τ_j^{fuel} is the fuel taxes in pct. of the total fuel price. Setting $\omega_{ij} := \tau^{\text{total}}$ in (4.2) gives the average expected tax revenue for the government.

Lastly, following Small and Rosen (1981) and Gillingham (2012), the model yields the usual "logsum" welfare measure defined as

$$\text{CS} \equiv \frac{1}{N} \sum_{i=1}^{N} \log \left[\sum_{j \in \mathcal{J}_i} \exp(u_{ij}) \right], \tag{4.3}$$

which can be used to evaluate the welfare impacts on consumers from changing parameters of the choiceset such as car characteristics or tax rates. It should be noted though that since there is no outside option, this welfare measure does not take into account that households may choose not to own a car.

5 Results

In this section, I present the estimation results. I start by presenting the structural parameter estimates and discussing these. To assess the validity, I discuss the driving equation and relate it to a partial estimation of the driving parameters alone as well as to what has been found in the literature. To get a better intuitive grasp of the model behavior at the estimated parameters, I compute a number of relevant outcomes and calculate the elasticities of these with respect to exogenous variables. Finally, I discuss robustness and consider alternative specifications of the fuel price expectations process.

Table D.1 shows the structural estimates from the preferred specification allowing random effects ($c_i \neq 0$) and where consumers have perfect foresight with respect to fuel prices. I will discuss the fuel price expectations later. The coefficients have the expected signs; households with higher work distances tend to drive more (α_{1z}-coefficients are positive) and be more price-responsive in their driving (γ_z-coefficients also positive, increasing the magnitude of the utility of money).[17] Urban households tend to drive their cars less and older households also drive less. Heavy cars tend to be driven more intensively as indicated by $\alpha_{1,\text{weight}}$ and $\alpha_{1,\text{weight}^2}$ both being positive; this is consistent with car weight proxying for unobserved luxury characteristics. The term, α_0, captures utility from the car ownership that are unrelated to driving. The parameters entering into α_0 tend to be very large, but recall that they should be divided by the λ-value of 10,000. The diesel coefficient ($\alpha_{0,\text{diesel}}$) is negative; this indicates that there is some characteristic about diesel cars that keeps households from buying them even though their other characteristics

[17] Gillingham and Munk-Nielsen (2015) explore precisely this feature of the data, finding that it high-driving households switch from driving to car to using other modes of transport when fuel prices increase. The behavior is consistent with a model of switching costs in changing transport mode to work from private car to public transportation.

make them more attractive than a given gasoline car. Finally, note that the dispersion in the unobserved driving type, c_i, is estimated to be 16.09, while the standard error on the driving equation measurement error is 21.95. This indicates that the endogenous selection of car type based on other factors still play a considerable role even though work distance is accounted for.

Table 5.1: Estimated parameters

Fixed Parameters				
Parameter				Value
β				0.95
λ				10000
α_2				-1
Model: Perfect foresight, random effects.				
General Parameters				
Parameter			Estimate	t
σ_x			16.093	(69.12)
σ_α			21.951	(31.77)
Demographics				
	γ_z		α_{1z}	
Parameter	Estimate	t	Estimate	t
Constant	47.596	(35.22)	74.927	(14.88)
Age	-8.447	(-18.97)	8.901	(8.71)
Age^2	7.363	(15.88)	-15.168	(-14.39)
Work distance, male	8.170	(18.95)	17.889	(69.45)
Work distance, female	1.079	(19.46)	9.684	(108.20)
Income	-9.457	(-31.44)	-8.768	(-39.94)
Number of kids	1.453	(11.65)	-0.458	(-2.93)
Urban dummy	-0.210	(-1.48)	-1.412	(-10.09)
Car Parameters				
Parameter			Estimate	t
$\alpha_{0,\text{weight}}$			124074.734	(41.91)
$\alpha_{0,\text{weight}^2}$			-5009.689	(-5.67)
$\alpha_{0,\text{kw}}$			-413.653	(-25.53)
α_{0,kw^2}			5.114	(46.83)
$\alpha_{0,\text{displace}}$			-194.172	(-0.15)
$\alpha_{0,\text{displace}^2}$			4976.559	(13.12)
$\alpha_{0,\text{diesel}}$			-4235.595	(-24.99)
$\alpha_{1,\text{weight}}$			18.876	(3.12)
$\alpha_{1,\text{weight}^2}$			10.189	(5.64)

Recall from section 4 that the VKT equation can be estimated separately, using the partial likelihood function from equation (4.1). Table 5.2 shows the elasticities of VKT with respect to the fuel efficiency, the weight of the car, and the fuel price.[18] Elasticities are computed numer-

[18] The estimated linear equation regresses VKT on demographics and car characteristics as well as demographics interacted with the price per kilometer, defined as the fuel price (gasoline or diesel depending on the car) divided by the fuel efficiency.

Table 5.2: Elasticities of VKT From the Structural Model — Partial Estimates and Preferred Specification

	Fuel efficiency	Weight	Fuel Prices
Partial Likelihood (using $f_x(\cdot)$)			
Mean	0.718	1.323	–0.725
Std.	0.426	0.339	0.431
Preferred Specification (using $\ell^{\text{sim}}(\cdot)$)			
Mean	0.279	0.858	–0.282
Std.	0.085	0.171	0.086

ically for each observation using finite differences and reports both the average and standard deviation of the elasticity across observations.[19] The elasticity of driving with respect to the fuel efficiency from the partial model is –72.5%. This central elasticity, when properly identified, is what Small and Van Dender (2007) refer to as the *rebound effect*. This estimate is fairly close to the approximately –80% that Frondel, Peters, and Vance (2008); Frondel, Ritter, and Vance (2012) find using German data. The estimate from the full model accounting for selection, however, is –28.2%. Gillingham (2012) finds a bias in the same direction but smaller in magnitude with a rebound effect of –21% dropping to –15% when selection is accounted for. Bento et al. (2009) find a mean elasticity of –35% which also controls for selection. Note that the elasticity with respect to fuel price and fuel efficiency are the same (except for the sign and direction) since they only enter the model together in the price per kilometer.[20] Finally, the estimates in Table 5.2 indicate that weight has a large effect on driving with an increase in weight of 1% being associated with an increase in driving of 0.858%. This implies that to understand the impacts of a car reform on driving and thereby emissions, it is not enough to just focus on the fuel efficiency; the weight of the chosen vehicles can also have strong effects on the final driving.

To get a better grasp of the implied behavior by the structural elasticities, Table 5.3 shows a range of economic outcomes simulated from the model in column (1) by the method described in Section 4.3. The table also shows elasticities of these outcomes with respect to four different variables in columns (2)–(5), computed using finite differences.

Column (2) shows the relative change in each expected outcome when the fuel efficiency of each car in the choiceset is increased by 1%. For the fuel efficiency of the chosen vehicles, this has an elasticity of 0.90 so that the average expected fuel efficiency is 0.9% higher. This implies that when technological progress makes cars more fuel efficient, households substitute away some of this for other characteristics; the weight increases by 0.09%, the engine power (kW) by 0.24% and the diesel share falls by 0.18%. More interesting, the CO_2 elasticity is –57%, so that a 1% improvement in fuel efficiency does not give a full 1% improvement in CO_2 emissions. This is partly due to consumers switching away from efficient cars and partly due to consumers driving

[19] The dispersion in the elasticity is driven by the dispersion in the computed coefficient $\hat{\gamma}_i \equiv \hat{\gamma}_z' z_i$.

[20] Gillingham (2012) allows e_j to shift the mean u_{ij} by putting it in the term $\alpha_0' q_j$ in (3.2).

Table 5.3: Structural Elasticities — Quasi, Perfect Foresight, Random effect

	Levels	Elasticities			
	(1) Baseline	(2) Fuel efficiency	(3) Weight	(4) Fuel prices	(5) O95 prices
Consumer welfare					
CS	114970.09	0.25	1.29	-0.25	-0.20
Total taxes					
E(total taxes)	146623.83	0.08	0.40	-0.08	-0.06
Ownership tax					
E(Regtax revenue)	106556.44	0.23	0.27	-0.23	-0.14
E(Owntax revenue)	11093.62	0.29	0.31	-0.28	-0.16
Fuel tax					
E(O95 revenue)	25115.85	-0.49	0.63	0.50	-0.02
E(Diesel revenue)	3857.92	-0.89	2.98	0.88	2.33
Driving/fuel use					
E(VKT)	79663.89	0.30	1.01	-0.30	-0.19
E(litre O95)	4340.92	-0.49	0.63	-0.50	-1.01
E(litre D)	891.32	-0.89	2.98	-0.12	2.33
E(litre D|urban)	188.02	-0.86	3.04	-0.14	2.24
E(kg CO2)	12736.56	-0.57	1.06	-0.43	-0.39
Characteristics					
E(fe)	15.92	0.90	-0.00	0.10	0.20
E(we)	1.70	0.09	1.15	-0.09	-0.04
E(kw)	77.08	0.24	0.13	-0.23	-0.25
E(displace)	1.65	0.18	0.12	-0.18	-0.16
E(% diesel)	18.49	-0.16	1.86	0.15	2.33
E(% diesel|urban)	3.89	-0.14	1.88	0.13	2.24

The model is quasi-linear with perfect foresight and
and random effects (σ_α is estimated).
The baseline column is expected outcomes, all other are elasticities.
(1): baseline 2006 scenario, (2) fuel efficiency up by 1%, (3): weight up by 1%,
(4): all fuel prices up by 1%, (5): only O95 up by 1%.
Counterfactuals are run on 2006 data.

longer since the cost of driving an extra km is now lower. This result has huge implications for climate policy since it means that in order to reduce CO_2 emissions by 1%, the required improvement in fuel efficiency is approximately 1.75%.

Column (3) shows the effects of increasing the weight of all cars by 1%. This increases VKT by 1.01% and CO_2 by 1.06%. Note that the elasticity of driving with respect to car weight was even stronger in the partial equation, indicating that selection is at play.

Column (4) shows the effects of increasing the real fuel price at the pump by 1%.[21] The most notable result here is that tax revenue *falls*, indicating that the Danish taxes are at the wrong side of the Laffer curve's top; While revenue from fuel taxes increase, revenue from the registration and the ownership tax fall by much more because households buy different types of cars. CO_2 emissions fall by 0.41%, which should be compared to the intensive-margin response of 0.28% implied by Table 5.2.[22]

Finally, column (5) increases gasoline prices by 1% but keeps diesel prices constant. The result is a 2.33% change in the probability of purchasing a diesel car (and thus of the diesel market share). This gives a first indication that the diesel market share is highly sensitive to cost differences.

Based on the elasticities of CO_2, tax revenues and welfare with respect to fuel prices, it is possible to compute the marginal cost of CO_2 reductions from a fuel tax. Back of the envelope calculations indicate, that a reduction of one ton of CO_2 would cost society 7389.90 DKK.[23] This number is far above the Social Cost of Carbon of 260 DKK per ton as suggested by the US Environmental Protection Agency. The high cost is perhaps not surprising given how high the tax level is in Denmark.

To examine robustness, the model has also been estimated assuming static expectations and a unit root as described in section 3.2. These different specifications gave quite similar results in terms of elasticities and implications for the counterfactual simulations so the perfect foresight model was chosen. The results with static expectations are shown in Appendix D.1; the elasticities of the relevant quantities relatively unchanged compared the corresponding ones for the model with perfect foresight, although the driving response is –0.39 instead of –0.30. This implies a higher reduction in driving, and the cost per ton of CO_2 for the fuel tax implied by these estimates is correspondingly lower: 5843.02 DKK. The key to understanding the difference between the parameters estimated under the two sets of assumed price expectation formation is the realized movements in fuel prices (see Figure B.3); prices have been increasing throughout

[21]Note that to obtain this using taxes, one would have to take into account supplier responses. For the US, Marion and Muehlegger (2011) find a pass-through to consumers of almost 100% but given the substantially higher taxes in Denmark, that conclusion might not be valid here. Nonetheless, I abstract from the question of passthrough.

[22]Table 5.2 conditions on car choice so any given relative change in driving will produce the same relative change in fuel consumption and thus in CO_2 emissions.

[23]The required change in fuel prices to reduce CO_2 by 1 ton is approximately $\Delta p = (\mathcal{E}_{\mathrm{CO}_2,p}\mathrm{CO}_2/p)^{-1} = (0.43\frac{12.7\,\mathrm{ton}}{8.5\,\mathrm{DKK/l}})^{-1} \cong 1.55$DKK/l. This implies an approximate change in consumer surplus and taxes of

$$\begin{aligned}\Delta\mathrm{CS} &= \mathrm{CS} \times \mathcal{E}_{\mathrm{CS},p} \times \frac{\Delta p}{p} = 114,970.09 \times -0.25 \times \frac{1.55}{8.5} \cong -5248,13\,\mathrm{DKK} \\ \Delta\mathrm{Taxes} &= 146,623.83 \times -0.08 \times \frac{1.55}{8.5} = -2141,78\,\mathrm{DKK}.\end{aligned}$$

the period and the likelihood conditions on the same car and driving choices. Therefore, if consumers knew that prices would increase yet did not choose an even more fuel efficient car to curb fuel costs, it must be because they valued the fuel savings less relative to the other car characteristics. I have chosen the perfect foresight model as the preferred specification because the in-sample fit of the diesel share is better (Figure D.3). However, I think that for future research it might be more fruitful to focus on modeling the time-series development in the relative price of gasoline and diesel; Figure B.6 shows that the relative price of diesel to gasoline has gyrated around an increasing trend and gyrations appear to show up in the predicted diesel share. Figure D.2 illustrates that the model's over- and under-predictions seem to be correlation with the gyrations.

6 Counterfactual Policy Simulations

In this section, I present a sequence of counterfactual policy simulations. I start with a discussion of the model structure and assumptions and what they imply for the applicability of the counterfactuals. I then present a counterfactual simulation, implementing the out-of-sample 2007 reform in-sample. Next, I assess the role of the 1997 reform in driving the increase in diesel cars in Denmark. Finally, I present a counterfactual exploring the diesel share in absence of discriminatory ownership and fuel taxes.

6.1 External Validity

The strength of the model is in understanding how households trade off between available cars in the choiceset in characteristic space and how this interacts with driving decisions. In that sense, the model is well-suited for understanding how car tax policies feed into driving behavior and the related externalities. The high-dimensional choiceset makes the model precise in terms of modeling the tax system and leveraging policy variation. However, the computational cost of this dimensionality is that the model conditions on entry into the new car market. This means that all the simulated effects are for the average household; the model is uninformative as to changes in the number of households (or cars per household). Moreover, restricting the model to new cars effectively eliminates real-world substitution alternatives in the form of the used-car market and the outside option of not owning any car. Ignoring substitution options for the consumers will inflate my estimate of the consumer loss related to increasing taxation.[24]

The second main restriction of the model is that supply side responses to the proposed reforms are ignored, i.e. assuming a 100% passthrough of taxes. In reality, profit maximizing car sellers in oligopolistic competition will likely change the relative prices of cars in their portfolios. In defense of this assumption, Adamou, Clerides, and Zachariadis (2013) find little difference between their simulation results when they use their estimated supply side pricing function or simply assuming 100% passthrough in a European context. For fuel taxes, Gallagher and Muehlegger (2011) find that passthrough in the US to consumers is approximately 100%. Moreover, given the small size

[24]This is because in the model, the consumer has no choice but to shift around in the choiceset. In a more realistic model, the consumer can also choose the outside option or used cars. Instead of being forced to absorb higher taxes, the consumer has the option of not owning a car. Since this alternative is unaffected by fuel taxes, the consumer surplus measure in (4.3) will drop less when that alternative is available.

of the Danish market relative to other European countries, auto makers are unlikely to change their production to cater to Denmark.

6.2 The 2007 Reform: Model Validation and Policy Evaluation

As described in 2.1, the 2007 reform was a feebate, meaning that it gives a rebate to green cars and puts a fee on inefficient cars. The *pivot point* of the reform, differentiating green cars from dirty ones, was set to 16 km/l for gasoline cars and 18 km/l for diesel cars. Recall that 2007 is not in the estimation sample because driving information is only available for a small number of cars purchased in this year.

Table 6.1 shows the implications of implementing the 2007 feebate in 2006. Most importantly, the diesel market share goes up from 18.5% to 24.5%, an increase of 32.3%.[25] The true response to the 2007 reform was an increase in the diesel share of 46.0%. In other words, the model can explain two-thirds of the relative shift in the diesel share. Similarly, the model predicts the average fuel efficiency to increase by 7.04% whereas the actual response to the reform was 5.73%. In this case, the model overshoots but as Figure (2.1) illustrates, the fuel efficiency continues to increase in the following years, increasing by an additional 7.63% in 2008. I view these as good out-of-sample fits.[26]

Regarding the predicted environmental impact of this reform, the average expected CO_2 emissions fall by 892.2kg or 7.0%. Some of this comes through the intended channel of improved fuel efficiency which increases by 7.0%, but recall from table 5.3 that this only translates into approximately $0.57 \cdot 7.0\% = 4.0\%$ reductions in CO_2. In particular, the reform as a by-product reduces weight by 3.5% which translates into less driving, yielding an additional $1.01 \cdot 3.5\% = 3.5\%$ in CO_2 reductions. In other words, the reform's impact on the weight of the chosen vehicles is almost as important as the intended impact via fuel efficiency.

In terms of welfare, the 2007 reform increased consumer surplus but decreased taxes by much more. Even accounting for the lowered driving and thus lower non-CO_2 externalities, the societal cost of the predicted reduction in CO_2 was 11,886.99 DKK/ton. This is a 60.9% higher cost per ton of CO_2 than that of the fuel tax, cf. section 5, and even further from Social Cost of Carbon of 260 DKK/ton. It is not uncommon to find high implied costs of CO_2 savings in the literature, e.g Beresteanu and Li (2011) and Huse and Lucinda (2013), although my estimates are exceptionally high. However, the feebate is asymmetric with a higher rebate than fee; in light of Adamou, Clerides, and Zachariadis (2013) it is not surprising that it is in-effective.

Given that the model fits the shift to diesels, the next question is which part of the policy design led to this shift. The pivot point of 16 km/l for gasoline and 18 km/l is an obvious candidate given that the median difference between gasoline and diesel cars is higher than 4 km/l. I therefore implement a counterfactual where the pivots instead are set to 16 km/l and 20 km/l. The results of this counterfactual are shown in Table D.2; here, diesel share only

[25] One important note to make in this regard is that the diesel share in the sample in 2006 is 18.5% whereas in the full population it is 21.8%. As discussed in appendix B.1, this is due to diesel cars being over represented in the car types that are only purchased by very few households and therefore dropped from the sample. I expect that these niche cars would be hard to fit in this model framework.

[26] I have been unable to find data to produce a graph comparing fuel efficiencies across European countries similarly to how Figure 2.2 shows diesel penetration rates. My impression is that change in fuel efficiency in 2007 for Denmark is still uniquely large but not as different from the rest of Europe as is the case for the diesel share.

Table 6.1: Counterfactual Simulations — The 1997 and 2007 Reforms

	(1) Baseline	(2) 1997	(3) 2007	(4) Internalization
Consumer welfare				
CS	114,970.09	99,607.67	115,989.89	115,569.79
Total taxes				
E(total taxes)	146,623.83	176,422.24	134,398.55	146,854.69
Ownership tax				
E(Regtax revenue)	106,556.44	117,238.77	98,175.80	107,363.31
E(Owntax revenue)	11,093.62	26,613.14	9,813.59	9,131.02
Fuel tax				
E(O95 revenue)	25,115.85	31,519.34	21,695.96	23,698.46
E(Diesel revenue)	3,857.92	1,050.99	47,13.20	6,661.89
Driving/fuel use				
E(VKT)	79,663.89	78,391.96	78,740.62	79,518.56
E(litre O95)	4,340.92	5,447.67	3,749.84	4,095.94
E(litre D)	891.32	242.82	1,088.92	1069.70
E(litre D|urban)	188.02	50.40	230.59	230.87
E(kg CO2)	12,736.56	13,671.87	11,844.36	12,621.51
Characteristics				
E(fe)	15.92	14.69	17.04	16.12
E(we)	1.70	1.73	1.64	1.71
E(kw)	77.08	89.93	70.07	76.64
E(displace)	1.65	1.86	1.54	1.65
E(% diesel)	18.49	4.97	24.48	23.28
E(% diesel|urban)	3.89	1.03	5.18	5.03

The counterfactuals are run on data for 2006.

1997: The green ownership tax is replaced with the weight based annual tax.

2007: The 2007 feebate reform is implemented on 2006 data.

Internalization: Annual and registration taxes for diesels are set in the same way as gasoline cars but the diesel price is increased by 1.923 DKK/l.

increases marginally by 6.2%. Moreover, this alternative version of the reform yields 91% of the CO_2 reductions of the actual reform with almost identical consumer surplus and tax revenue. This provides evidence that the CO_2 reductions achieved by the feebate were not simply due to a shift to diesel cars.

6.3 The 1997 Reform: The Role of Taxation in the Dieselization

The 1997 reform changed the annual tax from being based on the weight of the car to being based on the fuel efficiency (see section 2.1). However, cars first registered before July 1st 1997 still follow a weight-based scheme. In this counterfactual, I compute the annual tax for all cars based on that scheme instead of the actual, fuel efficiency based scheme.[27] The average expected outcomes in 2006 under this counterfactual are shown in column (2) of Table 6.1. Figure 6.1 shows the predicted diesel share year by year in the sample.[28] The results show that while the diesel share would still have increased, the increase would have been substantially lower. In 2006, the predicted share is 4.97%, which is substantially below the baseline of 18.49%. The reason for this difference is that the post-1997 tax regime rewards high fuel efficiency while the pre-1997 regime punished heavy cars. Since diesel cars are inherently more fuel efficient and tend to be heavier, they are likely to benefit from this. Already in 1997, it is clear that the new tax scheme favors diesel cars; for the average car in the choiceset in 1997, the actual annual tax of a diesel car was 8.0% higher than the average annual tax for a gasoline car. However, under the counterfactual, pre-1997 regime the diesel car would be paying an 18.8% higher annual tax. Under the counterfactual, the predicted diesel share is substantially below the predicted under the actual tax regime (Figure 6.1) and this is driven mainly by this difference moving even further in favor of gasoline cars over the period.[29] The fact that the difference in the average annual tax increases over time also helps to explain why the response in the diesel share following the 1997 reform in Figure 2.2 is not a drastic shift as is the case for the 2007 reform.

6.4 The "Optimal" Diesel Share

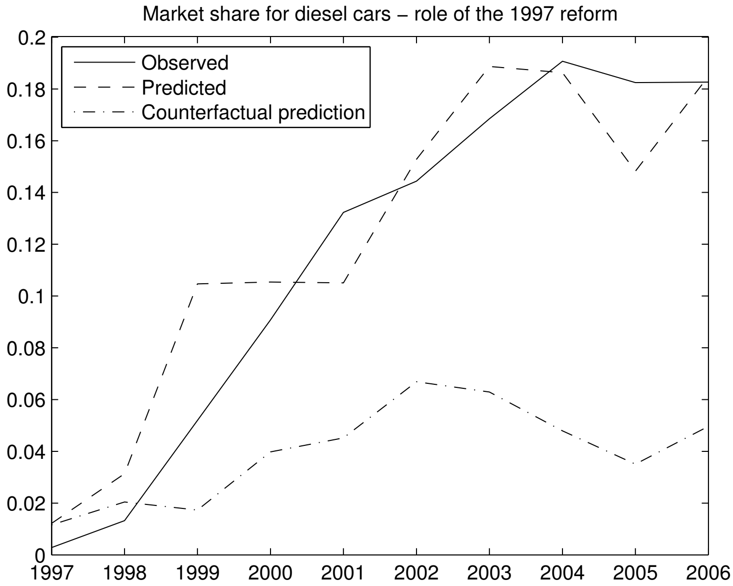
For both the 1997 and 2007 reforms, I have found that the reforms were both misaligned in their differential treatment of diesel and gasoline cars, causing a change in the status-quo diesel share. Given that there are discriminatory elements both in fuel taxes and ownership taxes, I explore the question: what would be the free-market outcome if the only discrimination in taxes was due to differences in externalities? To answer this question, I implement a counterfactual on the 2006 data. The only source of differences in externalities between diesel and gasoline cars is related to the fuel usage; the burning of diesel fuel emits slightly more CO_2 and emits harmful

[27] There might be many other counterfactuals equally interesting as the alternative to the 1997-reform; if for example the rates were changed over time to encourage scrapping of vintages from before 1997. That does not, however, appear to be the case.

[28] The in-sample fit of the diesel share (the "Predicted" curve in Figure 6.1) fluctuates around the observed diesel share. The deviations are timed along with the movements in the relative price of diesel to gasoline (Figure B.6). One way to improve the fit might be to add a bivariate forecast, since the relative deviations appear to be strongly mean-reverting around a trend.

[29] By the end, however, the average diesel car in the choiceset would have paid 61.8% more, had it followed the old scheme, while the actual, post-1997 annual tax only imposed a 23.4% higher annual tax on diesel owners.

Figure 6.1: Predicted Dieselization From the Baseline Model vs. the Weight Tax Counterfactual.



local air pollutants that gasoline does not.[30] Therefore, I first equalize ownership taxes, setting those for diesel cars equal to those for gasoline cars. Fuel taxes are not equal in the outset since gasoline has a higher fixed component of the taxes (see Appendix B.3.2). Therefore, I first equate fuel taxes by increasing the fixed component for diesel fuel up to the level gasoline and then add an additional per-liter tax equal to the per-liter external cost. The estimates of marginal external costs are taken from DTU Transport (2010) (see Appendix B.2). Assuming a 100% passthrough to consumers, I can simulate whether the diesel market share would be above or below the baseline level for Denmark in absence of discrimination — this exercise is similar to internalizing an externality using a Pigovian tax, except that the baseline gasoline tax is not necessarily optimal. In this sense, I do not claim to find the optimal diesel share but rather an improvement over the status quo.

The results are shown in column (4) of table 6.1. The central conclusion is that the predicted diesel share increases by 25.9% based on the 2006 diesel share (from 18.49% to 23.28%). This puts the predicted counterfactual diesel share between the 2006 and 2007 levels. Note that any incomplete passthrough would directly dampen this effect. An interesting additional conclusion that can be drawn from this counterfactual is that the proposed policy appears to represent an unambiguous improvement; Consumer surplus and tax revenue go up, CO_2 emissions go down and VKT also goes down (so externalities from congestion and accidents also decrease). However, these improvements are very small economically. This counterfactual indicates that when the added externalities of diesel cars are priced (subject to the externality prices that I have used), the added value of those cars (in terms of efficiency, for example) relative to their price makes them a valuable part of the car fleet.

[30]In 2012, the World Health Organization moved diesel fumes to the list of substances that are known to cause lung cancer. There is regulation in place, effectively requiring diesel cars to be fitted with particle filters to reduce this type of pollution. These are taken into account by the external cost estimates.

7 Conclusion

In this paper, I estimate a structural discrete-continuous model of car choice and usage, allowing endogenous selection into car types based on expected future driving. The model is estimated using high quality full population register data for Denmark covering 1997–2006. To validate the estimates, I exploit the Danish car taxation reform of 2007 which prompted clear changes in new car type decisions immediately, unique to Denmark, in particular in the diesel market share. Implementing the 2007 reform counterfactually in 2006, I find that the model is able replicate the strong responses to the reform in terms of the diesel share and the fuel efficiency.

A consistent finding is that Danish households have responded very strongly to the tax incentives given by the 1997 and the 2007 reform. The implication is that both reforms were highly cost-ineffective ways of obtaining CO_2 reductions compared to a fuel tax, mainly due to foregone tax revenue. A central mechanism behind this is that according to simulations from the model, a 1% technological increase in the fuel efficiency of all cars only translates into a 0.57% reduction in CO_2 emissions; this is partly due to households substituting these fuel savings away for larger, more luxurious cars and partly due to the *rebound effect*, whereby households being pushed towards more efficient cars in turn drive them more intensively (at an elasticity of –30%). This greatly limits the effectiveness of environmental policies. Additionally, my results indicate that the effects of car taxes on driving that work through the weight of the chosen car may be at least as important as those working through the fuel efficiency.

To evaluate the two tax reforms of the period, I compare their cost-effectiveness to a fuel tax. I find that fuel taxes are much more effective. However, the cost per ton of CO_2 is still many times larger than the social cost of carbon, possibly due to the high level of taxes in Denmark in the outset. In particular, I find that increasing fuel taxes may *lower* tax revenue if they are increased; while they do increase fuel tax revenue, this is offset by an even larger drop in car taxes as consumers shift away from the luxury segment.

Another finding is that the reforms were responsible for most of the increase in the diesel share that occurred in my sample period. In particular, the Danish feebate reform in 2007 could have been designed differently to yield 91% of the CO_2 reductions but with only a minor increase in the diesel share. Nevertheless, I also show that the societal gains from diesel cars outweigh their negative aspects and that the diesel share in 2006 is close to the optimal level for the Danish setting.

Appendix

A Notation and Core Equations

This section is meant as a quick reference to give an overview of the model and the notation used in this paper.

The notation is as follows,

$$
\begin{aligned}
j &\quad - \quad \text{car type (e.g. 2003 Volvo V70 Turbo Diesel)},\\
d_i &\quad - \quad \text{the chosen car type by household } i,\\
x &\quad - \quad \text{vehicle kilometers travelled (VKT, a generic decision variable)},\\
x_i &\quad - \quad \text{the observed driving for household } i \text{ (conditioning on } d_i),\\
x^*_{ij}(p^{\text{fuel}}) &\quad - \quad \text{the optimal driving rule},\\
e_j &\quad - \quad \text{fuel efficiency of a car of type } j \text{in } {}^{\text{km}}/_{\text{l}},\\
p^{\text{car}}_{tj} &\quad - \quad \text{price of a new car of type } j \text{ in year } t,\\
p^{\text{fuel}}_{tj} &\quad - \quad \text{fuel price (the subscript } j \text{ is there to distinguish diesel or octane)},\\
\gamma_i &\quad - \quad \text{utility of driving relative to outside consumption (household-specific)},\\
z_i &\quad - \quad \text{household attributes correlated with driving utility},\\
y_{it} &\quad - \quad \text{household income in period } t,\\
\beta &\quad - \quad \text{discount factor (fixed at 0.95)},\\
\delta_j &\quad - \quad \text{vehicle-specific depreciation rate (e.g. 0.8)},\\
\alpha_{1ij}, \alpha_2 &\quad - \quad \text{utility from driving is quadratic in VKT with these coefficients},\\
\alpha_0 &\quad - \quad \text{coefficients on } q_j\text{; Utility from car } j \text{ that is not related to driving},\\
\varepsilon_{ij} &\quad - \quad \text{IID extreme value type II shock (to the car type choice utility)},\\
\eta_i &\quad - \quad \text{measurement error in the VKT equation},\\
\zeta &\quad - \quad \text{coefficients used in the linear interpretation of optimal driving.}
\end{aligned}
$$

The full utility can be written as

$$
\begin{aligned}
u_{ij} \quad = \quad & \gamma_i \left[1 - (\beta\delta_j)^4\right] p^{\text{car}}_{jt_{1i}} - 4\gamma_i\tau_j + u^{\text{own}}(j) \\
& + \beta^4 \mathbb{E}\left\{-\gamma_i \frac{p^{\text{fuel}}_{jt_{2i}}}{e_j} x^*_{ij}(p^{\text{fuel}}_{jt_{2i}}) + \alpha_{1ij} x^*_{ij}(p^{\text{fuel}}_{jt_{2i}}) + \alpha_2 \left[x^*_{ij}(p^{\text{fuel}}_{jt_{2i}})\right]^2\right\}.
\end{aligned}
$$

where

$$
\begin{aligned}
\gamma_i &= \gamma'_z z_i,\\
u^{\text{drive}}_{ij}(x) &= \alpha_{1ij} x + \alpha_2 x^2,\\
\alpha_{1ij} &= \alpha_{10} + \alpha'_{1z} z_i + \alpha'_{1q} q_j + c_i, \quad c_i \sim \mathcal{N}(0, \sigma^2_c).
\end{aligned}
$$

The driving rule, $x^*_{ij}(p^{\text{fuel}}_{jt})$, is given by

$$x^*_{ij}(p^{\text{fuel}}_{jt}) = -\frac{1}{2\alpha_2}\left(\alpha_{1ij} - \gamma_i \frac{p^{\text{fuel}}_{jt}}{e_j}\right).$$

In the estimation, z_i contains mean spouse age, age squared, work distance for both spouses, real gross income, the number of kids and a dummy for living in a major urban area (Copenhagen, Odense, Aarhus or Aalborg). The characteristics, q_j, are vehicle total weight, engine displacement in cc, engine horsepower in kW and squares of all these variables and a dummy for diesel. To keep the number of parameters down, only the total weight and its square was used in α_{1ij} — the remaining were close to insignificant and greatly increased estimation running time.

B Data

B.1 Sample Selection

Table B.1 shows how the sample size (new car purchases) gradually drops from the initial 311,057 cars to 128,910 as different sample selection criteria are imposed. The first criterion states that the household purchasing the car must own it for at least 90% of the 4-year driving period. This causes the most dramatic reduction in sample size because many households sell the car within this period. Figure B.1 shows a histogram of the fraction of the 4-year period that the purchasing household owns the car for the full sample of 311,057 purchases (disregarding the mass point at 100%). This shows that the share declines steadily down from 90% to 0%. The choice of 90% is to emphasize the need for accurate data on the driving to ensure that the selection on anticipated driving is pinned down by the data. Future work should look checking the sensitivity of the results to reducing the 90%.

Table B.1: Sample Selection

	(1) New cars	(2) Owns>90%	(3) Ncars<1.5	(4) #sold > 30	(5) Final sample
1997	14,500	8,866	8,252	6,453	6,019
1998	45,075	27,986	24,895	22,248	21,374
1999	42,260	25,846	22,540	20,165	19,525
2000	30,070	17,699	15,350	12,764	12,461
2001	23,774	12,182	10,389	8,057	7,893
2002	28,648	16,305	14,035	11,611	11,016
2003	22,733	12,516	10,774	8,961	8,600
2004	29,535	16,552	14,095	11,901	11,548
2005	36,722	22,794	18,999	15,863	15,490
2006	37,740	24,670	19,793	15,458	14,984
N	311,057	185,416	159,122	133,481	128,910

(2): The family owns the car at least 90% of driving period,
(3): The family may own another car but no more than 50%
of the driving period of this car,
(4): At least 30 of this car sold in full sample, (5): final sample.

Figure B.1: Fraction of the Driving Period Where the Original Owner Still Owns the Car

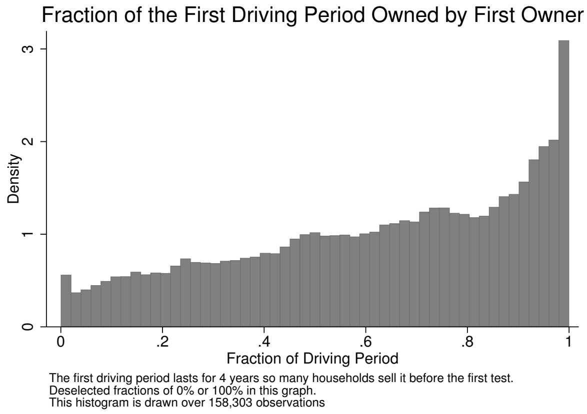


Table B.2: Deselecting Cars That are Rarely Sold and the Resulting Diesel Share

Threshold	Diesel % in 06	$\lvert\mathcal{J}\rvert$	N
30	18.5%	1,177	128,007
20	19.6%	1,518	136,977
10	20.6%	2,105	144,820
5	21.0%	2,783	149,112
0	21.8%	7,572	154,089

The second criterion deselects 2-car households but allows a second car to be present for up to 50% of the period.

The third criterion deselects certain car types from the choice set by deleting purchases of cars that were purchased fewer than 30 times in the period 1997–2006. This has a very unfortunate implication in that diesel cars are heavily over represented in this group. Table B.2 shows the implications on the sample size (N), the number of cars ($\lvert\mathcal{J}\rvert$) and the diesel market share in 2006 of setting this limit to 20, 10, 5 and 0 respectively. The true market share in 2006 was 21.8% but the restriction on the choice set results in a share of just 18.5%. However, bringing this up towards the truth increases the size of the choice set immensely, making estimation computationally very burdensome.

The final criterion makes routine checks such as dropping extreme observations (outside of the 0.1th or 99.9th percentiles) or rows with missing or senseless values.

B.2 Marginal External Costs of Driving

In this subsection, the marginal external cost estimates used for welfare calculations and for the construction of the diesel internalization counterfactual in section 6.4 are described. The cost estimates are taken from DTU Transport (2010) and they are provided by a major Danish research institution and used by Danish policy makers. The external costs of driving a km in a gasoline and diesel car respectively are reproduced in table B.3.

Two things are worth noting; Firstly, pollution and climate change costs are dwarfed by the congestion and accident externalities. While this particular externality is not well addressed with the model applied in this paper because it depends critically on when and where the driving

Table B.3: Marginal External Costs per Km Travelled by Fuel Type[a]

	Gas	Diesel	Unit
Noise	0.0478	0.0478	DKK/km
Accident	0.2095	0.2095	DKK/km
Congestion	0.3368	0.3368	DKK/km
Infrastructure	0.0097	0.0097	DKK/km
Air pollution	0.1352126	0.668475	DKK/liter
Climate	0.19764	0.21000	DKK/liter

[a] Source: DTU Transport (2010). Note that only air pollution and climate depend on the fuel type.

takes place, it does mean that an increased traffic flow should be highly discouraged.

Secondly, the only place where diesel car externalities are different from those of gasoline cars is in terms of air pollution and climate change. The difference in climate change externalities stem from the fact that diesel cars typically drive farther per litre of fuel (a sales-weighted average of 18.1 versus 13.5 km/l for in 2006) while diesel only contains 10.4% more CO_2 per litre than gasoline does (2.640 kg/l 2.392 kg/l). The difference in air pollution comes primarily from particulate matter. For the Belgian context, Mayeres and Proost (2013) report that particulate matter makes up 85.0% of all emissions-related externalities per ton of diesel, far more than the externalities from SO_2 and NO_x. In fact, the marginal externality of diesel air pollution depends crucially on the population density. Since a dummy for living in one of the four largest Danish cities is already in the model, the expected diesel use and diesel market share has been calculated conditional on urban residence. It turned out that urban diesel use and purchases followed the overall numbers quite closely for the reforms considered here.

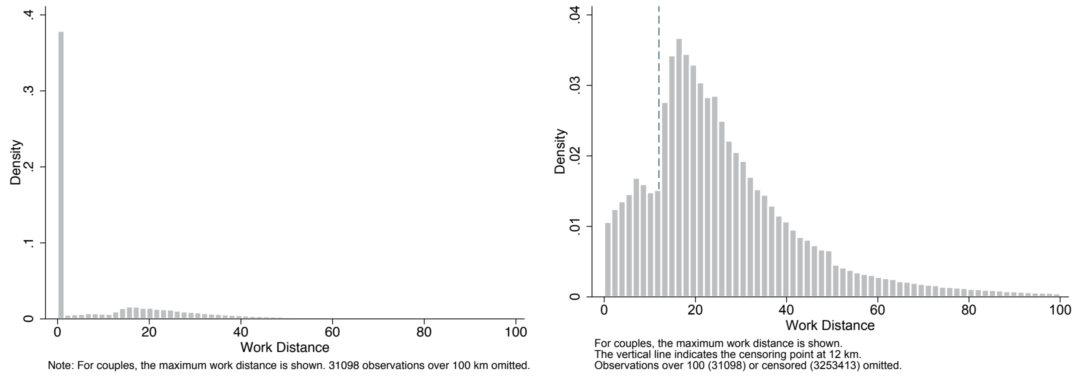
B.3 Descriptives

B.3.1 Work Distance

The work distance variable is the only one that is not taken directly from the register data. This one is calculated based on the travel tax deduction which comes from the personal tax registers. In Denmark, anyone living further than 12 km from their work place is eligible for a deduction depending on the distance times the number of days worked. The deduction is regardless of the number of hours worked and regardless of the type of transportation actually used by the worker. The deduction is a linear function of the km travelled above 24 (to and from work) but the rate drops to half after 100 km. In 2005, for example, it was DKK 1.68 for each km above 24 but below 100 and 0.84 for each km above 100. The rate was changed each year and twice in 2000. Moreover, as a part of a larger Danish reform in 1998 dubbed the Whitsun package, there was an adjustment to give a lift for the low-paid.

Note that in order to construct a work distance measure, one needs to know the number of days worked which is not observed. Therefore, it is assumed that everyone work 225 days a

Figure B.2: Work Distance Distribution



year.[31] Note, however, that this only means that the work distance variable may be imprecise for the actual distance to the work but still precise about the variable of interest that is the annual km commuted to work. Figure B.2 shows the distribution of the constructed work distance measure for the *larger dataset* from which I take my estimation sample; in the left panel, the full distribution is showed. This clearly shows the censoring with a large mass point at zero. The right panel removes these zeros and shows the remaining distribution. There is a clear discontinuity at a work distance of 12 km, consistent with the fact that this is the threshold for eligibility. For all the observations below 12 km, we know that their actual work distance is larger than 12 km but that they must have worked fewer than 225 days. For example, individuals with part time employment can be expected to fall there.

B.3.2 Fuel Prices

Figure B.3 shows the development in gasoline and diesel prices in Denmark in 2005 DKK. Prices have generally been increasing and moreover, it appears that diesel prices were converging on gasoline prices up towards 2008. Figure B.4 shows the price composition for both types of fuel; the fixed tax rate (dubbed the “Energy Tax”, which is split up into a CO2 tax in 2005) is fairly constant over the period with the exception of 1999 for gasoline and 2000 for diesel, where it is increased by 12% for both fuel types. In other words, most of the variation in fuel prices in Denmark comes from the product price. Figure B.5 shows the product price for Octane 95 and Diesel fuel together with the Western Texas Intermediate crude oil price. This figure shows that the prices have tracked the oil price very closely over the period, in particular for diesel fuel.

B.3.3 Car Characteristics

Figure B.7 shows the fraction of diesel cars in the register data (i.e. also data not included in my estimation sample). It shows the increase in the diesel market share that appears to really start increasing after 1997. The larger share of diesel cars with vintages in the 1980s can either be due to higher market share there or due to a different scrappage pattern for diesel cars then.

Figure B.8 shows the number of cars owned per household by year. The graph is based on

[31] The official numbers for public sector employees in 2007–2010 were 224, 226, 225 and 228.

Figure B.3: Real Price of Octane 95, 1980–2011

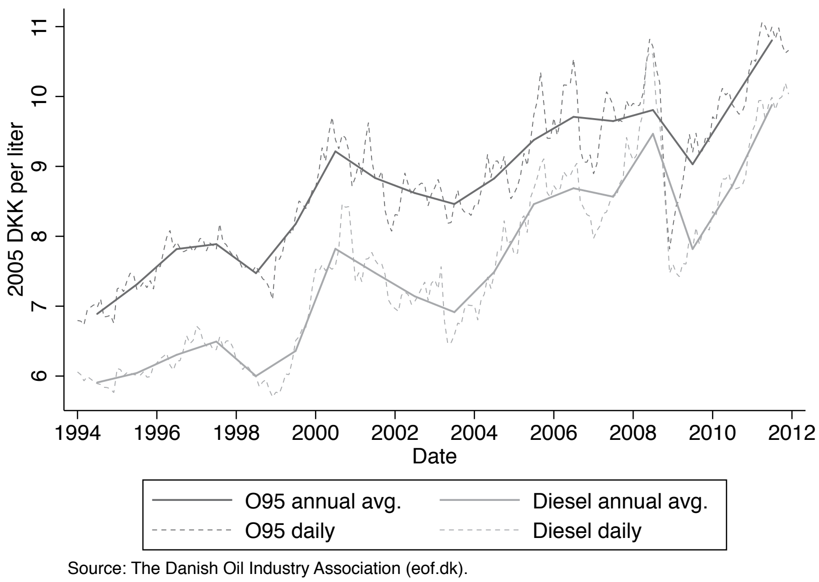


Figure B.4: The Composition of the Price of Gasoline and Diesel Fuel

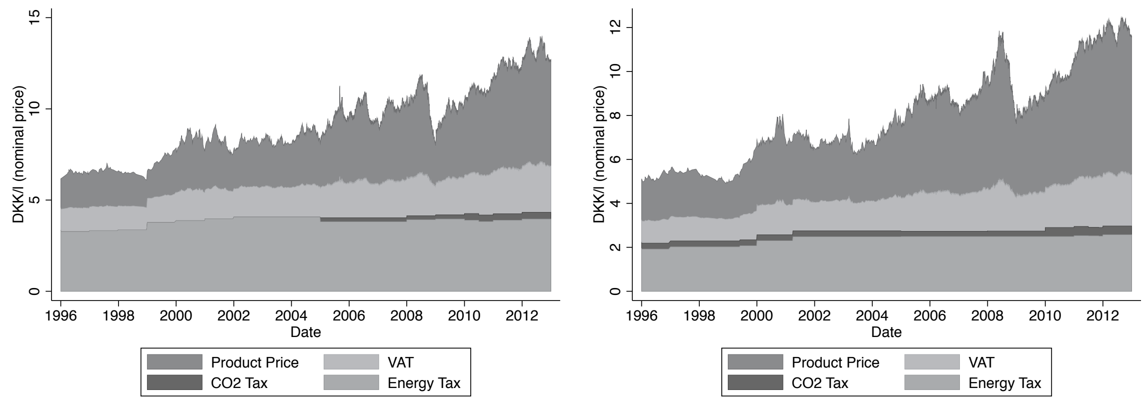


Figure B.5: Gasoline and Diesel Product Price And Crude Oil Prices

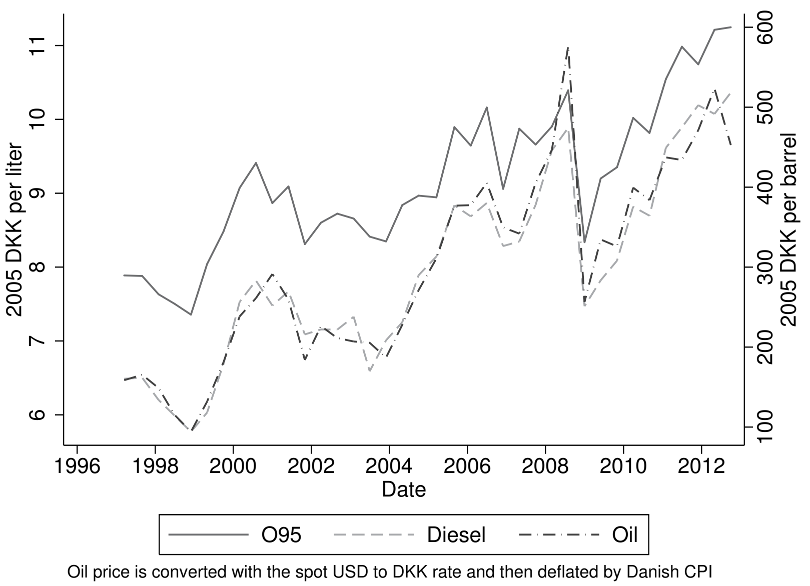


Figure B.6: Relative Fuel Prices

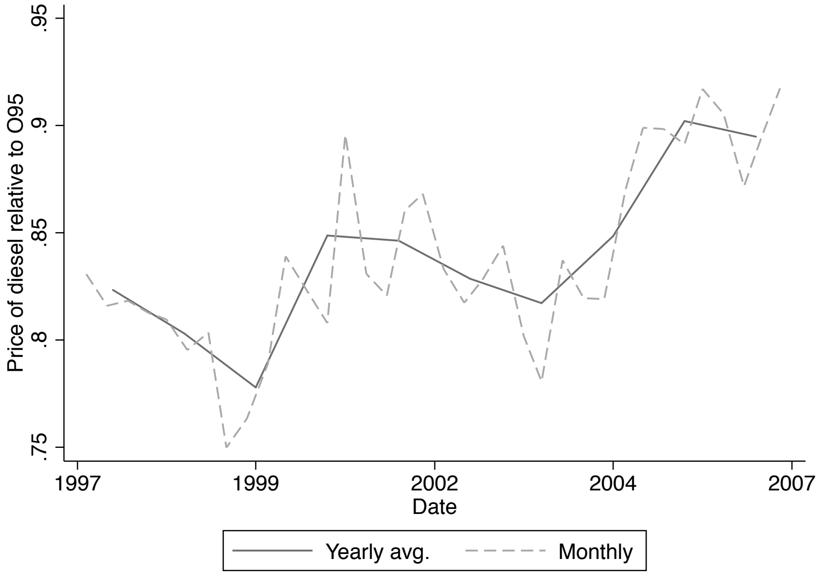


Figure B.7: Diesel Share in Denmark by Vintage

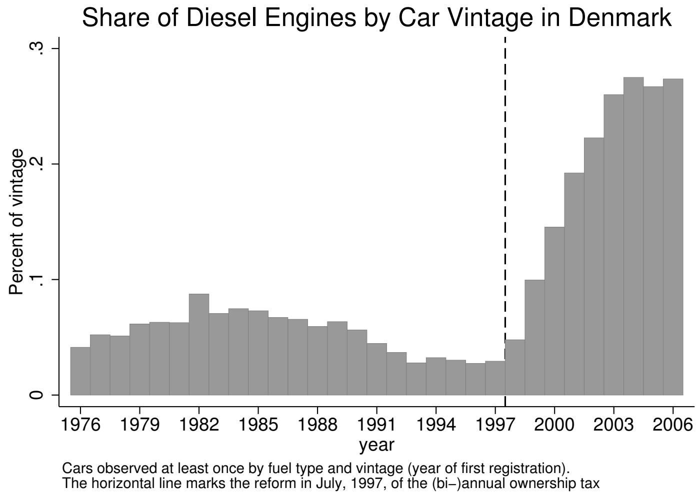
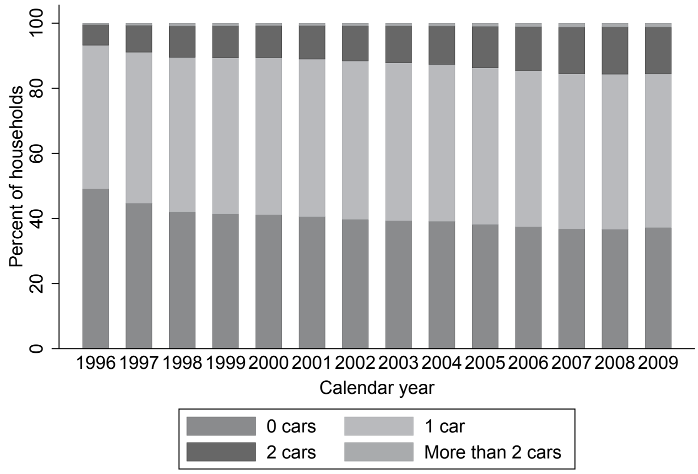


Figure B.8: Number of Cars Owned per Household



Note: each bar shows the share of households in a given year that own the particular number of cars.

a dataset containing all households and cars. The figure indicates that even though there has been an increasing share of households owning more than 1 car, the share is still extremely small compared to for example the US.

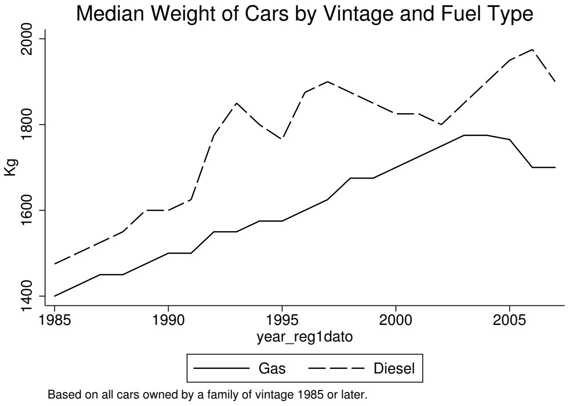
Figures B.9–B.11 show the development in median characteristics of sold cars. The most notable development is the increasing trend in weight for both types of fuel that has occured all the way back to the 80's. In this paper, weight proxies for the quality of the car by measuring comfort and the carrying capacity of the car. Similarly, fuel efficiency has gone up dramatically but here we see that while it has been somewhat monotone for gasoline cars, almost all the growth for diesel cars occured in 1997–99. Two things are worth noting there; Firstly, only 17 diesel cars are in the sample in 1997 so we are talking about very small numbers. Secondly, the advent of the Common Rail injection technology which quickly became standard in all diesel engines was the main reason for this. Apart from improving performance in terms of fuel efficiency, it also greatly improved the torque of the cars (which is not in my data) and changed the sound signature, making it more appealing to many consumers (according to an car salesman I have talked to).

The development in engine displacement, horse power and purchase price are much more erratic. This underlines the advantage of the chosen empirical model where all these characteristics are used in the household's comparison across cars, rather than focusing on each characteristic separately.

To better grasp the overall patterns in what car characteristics certain households end up with, table B.4 shows the estimates from regressing each car characteristic on household demographics. The results are much as one would expect with for example richer households purchasing more powerful and luxurious cars. It also shows some ambiguity in the effect of work distance — if males have a long work distance, they tend to prefer having a more comfortable ride whereas females tend to go for a more fuel efficient, smaller car.

The patterns shown in Table B.4 also show up clearly in the spatial patterns; Figure B.12 shows two maps of Denmark where the municipalities have been colored according to the average of the work distance (the maximum within the household) of Danish car-owning households in

Figure B.9: Median Characteristics Over Time — Weight and Fuel Efficiency



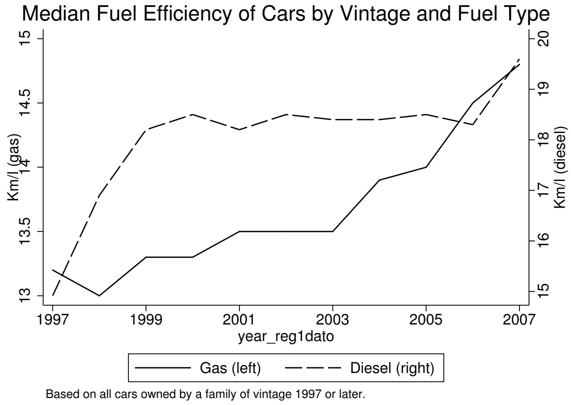
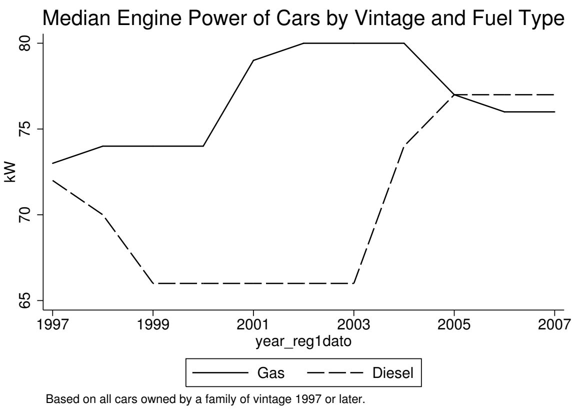


Figure B.10: Median Characteristics Over Time— Engine Power and Displacement



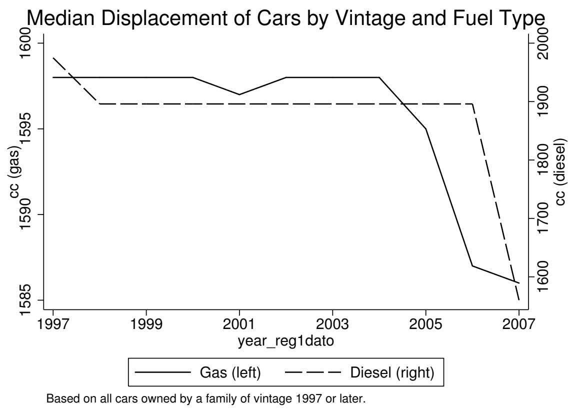


Figure B.11: Median Characteristics Over Time— Real Price (2005 DKK)

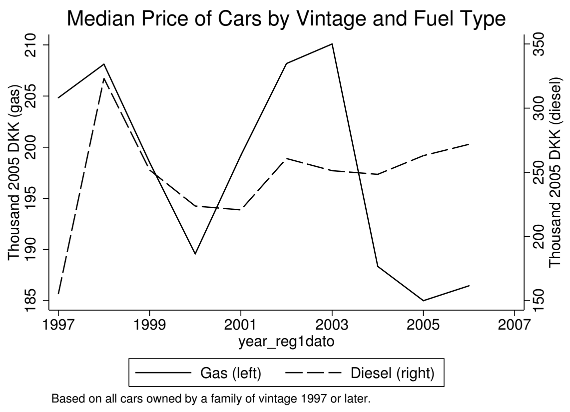


Table B.4: Car characteristics of new cars

	(1)	(2)	(3)	(4)	(5)	(6)
	Km/l	Weight	Diesel	kW	Displace	Real price
$p^{\text{fuel (O95)}}$	0.415***	-0.0193***	0.0410***	-1.923***	-4.645	-8612.1***
	(19.33)	(-10.84)	(14.93)	(-12.53)	(-1.92)	(-14.37)
GDP (2005=1)	-11.39***	-0.165**	-1.771***	-3.053	-978.8***	-244123.3***
	(-17.05)	(-2.98)	(-20.71)	(-0.64)	(-13.01)	(-13.11)
Age (m)	-0.0136	0.00373***	0.000654	0.473***	4.907***	1520.1***
	(-1.48)	(4.88)	(0.55)	(7.18)	(4.72)	(5.90)
Age squared (m)	0.0000800	-0.0000430***	-0.0000207	-0.00549***	-0.0593***	-17.83***
	(0.78)	(-5.03)	(-1.57)	(-7.45)	(-5.10)	(-6.20)
Age (f)	-0.0400***	0.00491***	-0.00118	0.335***	4.149***	1516.1***
	(-4.89)	(7.23)	(-1.13)	(5.73)	(4.50)	(6.64)
Age squared (f)	0.000306***	-0.0000445***	-2.68e-08	-0.00325***	-0.0430***	-14.48***
	(3.30)	(-5.80)	(-0.00)	(-4.92)	(-4.13)	(-5.62)
Work dist. (m)	0.0150***	0.000136***	0.00262***	-0.00892***	0.555***	84.66***
	(44.34)	(4.86)	(60.59)	(-3.70)	(14.59)	(9.00)
Work dist. (f)	0.0178***	-0.000444***	0.00238***	-0.0512***	-0.160**	-59.57***
	(39.57)	(-11.91)	(41.42)	(-15.95)	(-3.16)	(-4.76)
Income	-0.000000245***	2.45e-08***	-6.92e-09***	0.00000296***	0.0000434***	0.0166***
	(-16.83)	(20.33)	(-3.71)	(28.43)	(26.49)	(40.84)
Male inc %	0.00797	-0.000305	0.000721	0.00768	0.287	121.5
	(1.42)	(-0.66)	(1.00)	(0.19)	(0.45)	(0.78)
# kids	-0.303***	0.0417***	0.0000115	1.244***	24.10***	7462.6***
	(-37.23)	(61.88)	(0.01)	(21.43)	(26.33)	(32.94)
Urban dummy	0.0198	-0.00707***	-0.00564**	-0.754***	-8.729***	-2115.9***
	(1.22)	(-5.26)	(-2.72)	(-6.51)	(-4.78)	(-4.68)
Unemployed (m)	0.260***	-0.0407***	-0.00916**	-3.412***	-53.62***	-15345.8***
	(11.16)	(-21.06)	(-3.07)	(-20.47)	(-20.41)	(-23.59)
Unemployed (f)	0.170***	-0.0146***	0.00574*	-1.426***	-21.18***	-5694.4***
	(9.50)	(-9.86)	(2.51)	(-11.14)	(-10.49)	(-11.40)
Linear time trend	0.434***	0.0181***	0.0449***	1.082***	14.93***	7397.0***
	(43.37)	(21.76)	(35.06)	(15.11)	(13.23)	(26.49)
Constant	21.45***	1.653***	1.231***	65.59***	2230.9***	404939.1***
	(42.70)	(39.69)	(19.16)	(18.28)	(39.43)	(28.91)
N	128910	128910	128910	128910	128910	128910

For variable labels, m denotes male and f denotes female.
Same sample as the one used for the two-period model.
(1) Fuel efficiency in km/l, (2) weight in tons, (3) LPM for diesel,
(4) engine power in kW and (5) displacement in cc.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure B.12: Spatial Illustration: Municipality-averages of Work Distance and Diesel Frequency

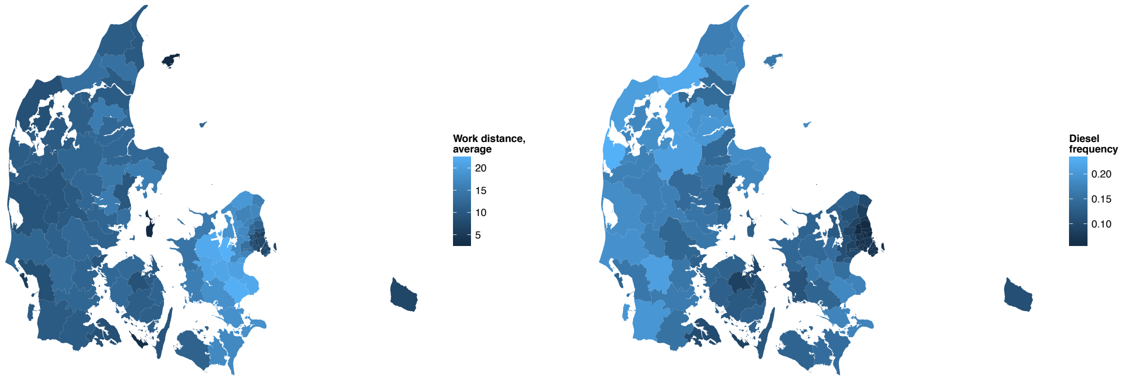
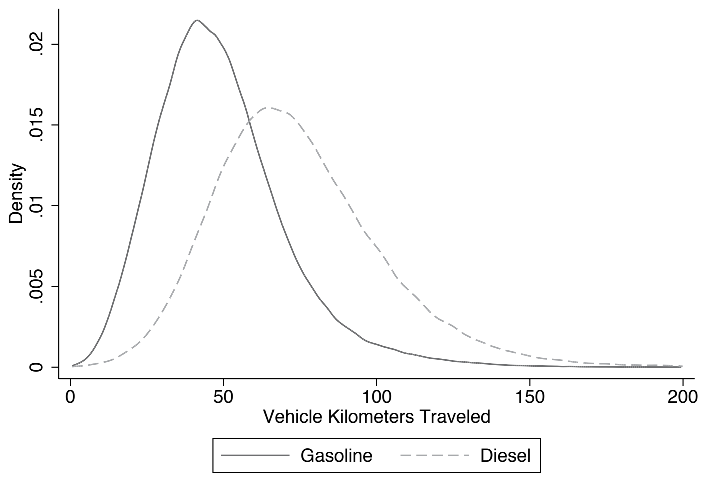


Figure B.13: Driving by Fuel Type



the left panel and the frequency of diesel cars in the right panel. These maps are drawn for a larger sample containing all car-owning households. The patterns show two interesting aspects. Firstly, there positive relationship between work distance and diesel ownership is in line with both urban areas and rural areas of Sealand (the big western Island); urban areas have low work distances and low diesel shares and vice versa for the rural areas of Sealand. However, in the Eastern part of the country, there appears to be low work distances and high diesel frequencies. These areas have very different employment patterns from the greater Copenhagen region, which is most likely a part of the explanation.

B.3.4 Descriptive Evidence on Driving

Figure B.13 shows the driving distribution for diesel car drivers and gasoline car drivers. The distribution for diesel car drivers is shifted strongly towards higher driving.

Figure B.14 shows median vehicle kilometers travelled (VKT) against median fuel price over time for gasoline cars (left panel) and diesel cars (right panel). Both figures show that the typical car purchased in later years ends up driving less than in earlier years and that fuel prices have been increasing. This is consistent with a negative fuel price elasticity.

Figure B.14: Median VKT vs Fuel Price Over Time for Gas and Diesel

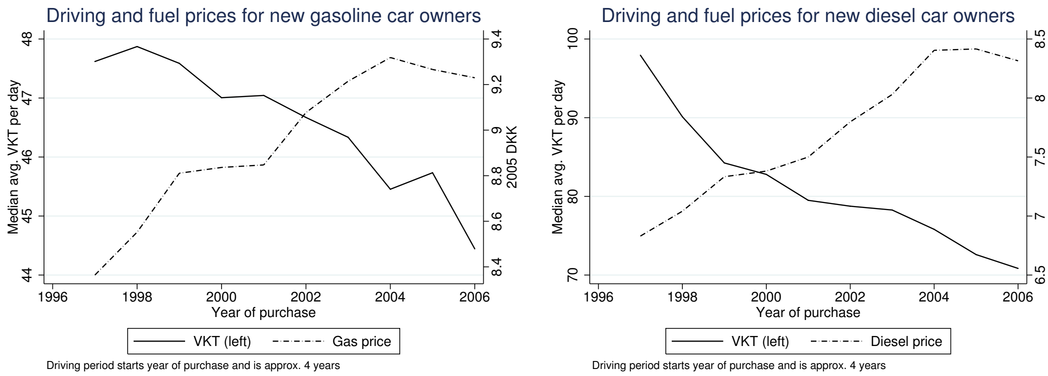
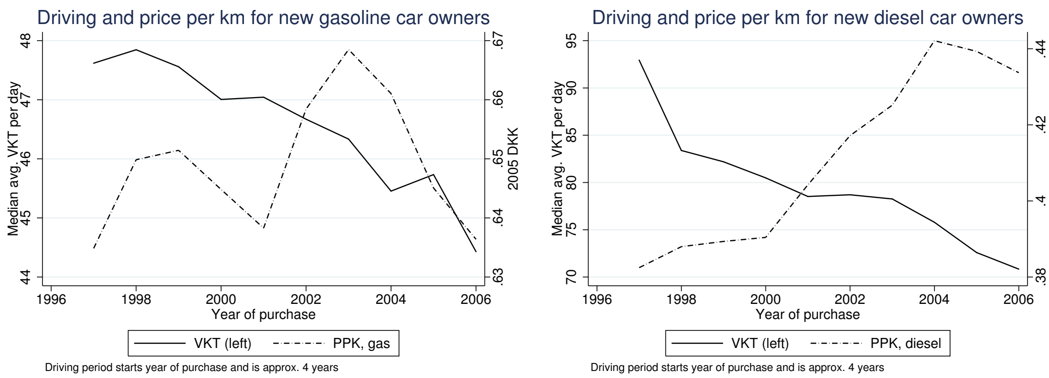


Figure B.15: Median VKT vs Price Per Kilometer (PPK) Over Time for Gas and Diesel



The corresponding figures where the price per kilometer (PPK, p_{jt}^{fuel}/e_j) is used are shown in figure B.15 and here the picture is much less clear picture because fuel efficiency also increases over time. This is precisely the selection effect at play where consumers are moving to more fuel efficient cars to counteract the increasing fuel prices.

Table B.5 shows the results from regressing VKT on PPK, car characteristics and household demographics. The most central result is that the mean estimated PPK-elasticity depends very strongly on whether a different mean driving is allowed for diesel car households (which decreases the mean elasticity from –.74 to –.30). This is further emphasized by the fact that the elasticity is –0.16 when estimated on the gasoline sample only and –.39 on the diesel subsample. Gillingham and Munk-Nielsen (2015) explore the heterogeneity in the fuel price elasticity on household demographics and the interested reader is referred to that paper.

C Joint Estimation of the λ-parameter

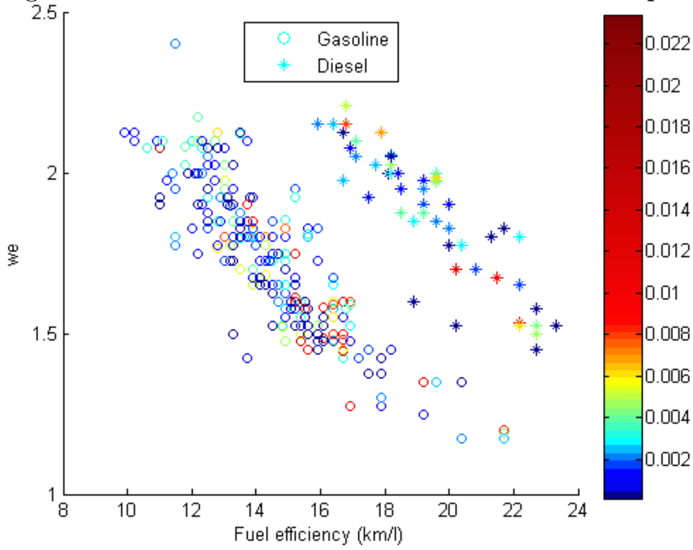
In this section, I discuss the issue with the estimation of the logit error term scaling parameter, λ, and present an idea for estimating a more flexible extension of the model that might facilitate joint estimation. I first discuss the problem, providing intuition about the λ parameter and why the maximum likelihood estimate is so high. I then argue that car fixed effects can be the cause of the problem and that controlling for these may solve the issue of the high λ. In light of this, I conclude with an outline a strategy for incorporating car type fixed effects into the model in

Table B.5: VKT Regressions — Price per Kilometer (PPK) Elasticity

	(1) Simple	(2) Diesel dummy	(3) Year FE	(4) Only gas	(5) Only diesel
Price per km	-50.50***	-20.32***	-16.82***	-10.01**	-65.73
	(-21.52)	(-8.29)	(-6.17)	(-2.83)	(-1.94)
GDP	-41.98***	-35.10***	-57.10***	-50.23***	-82.38***
	(-29.57)	(-24.69)	(-12.38)	(-10.68)	(-3.88)
Age (m)	0.518***	0.538***	0.536***	0.415***	1.256***
	(10.62)	(11.12)	(11.09)	(8.51)	(6.15)
Age squared (m)	-0.00777***	-0.00792***	-0.00789***	-0.00677***	-0.0138***
	(-13.57)	(-13.91)	(-13.87)	(-11.85)	(-5.59)
Work dist. (m)	0.353***	0.348***	0.348***	0.333***	0.381***
	(119.98)	(118.89)	(118.99)	(104.54)	(47.49)
Work dist. (f)	0.340***	0.334***	0.334***	0.356***	0.260***
	(87.05)	(85.83)	(85.92)	(84.04)	(24.27)
Income	-0.00000250***	-0.00000239***	-0.00000234***	-0.00000231***	-0.00000367***
	(-19.67)	(-18.89)	(-18.54)	(-18.68)	(-4.55)
# kids	0.236***	0.223***	0.219***	0.128	0.515*
	(3.61)	(3.43)	(3.37)	(1.93)	(2.05)
Urban dummy	-1.131***	-1.106***	-1.092***	-1.262***	0.973
	(-8.84)	(-8.70)	(-8.60)	(-9.88)	(1.80)
Unemployed (m)	0.492**	0.438*	0.459*	0.565**	-0.499
	(2.64)	(2.36)	(2.48)	(3.03)	(-0.64)
Unemmployed (f)	-0.0700	-0.0930	-0.0784	-0.0474	-0.410
	(-0.49)	(-0.65)	(-0.55)	(-0.33)	(-0.74)
Km/l	0.610***	-0.257*	-0.118	0.117	-1.475
	(6.07)	(-2.51)	(-1.02)	(0.77)	(-1.90)
Weight	0.0329***	0.0209***	0.0210***	0.0229***	0.0116***
	(67.23)	(36.60)	(36.45)	(38.60)	(5.36)
Engine power	-0.0368***	0.0426***	0.0428***	0.0483***	0.101***
	(-5.52)	(6.18)	(6.19)	(6.67)	(3.82)
Engine size	0.0140***	0.00367***	0.00352***	0.00180***	0.00178
	(33.71)	(7.55)	(7.23)	(3.36)	(1.14)
Diesel dummy		17.99***	18.09***		
		(40.22)	(40.37)		
Constant	28.02***	43.76***	66.88***	54.31***	157.5***
	(10.49)	(16.30)	(11.82)	(8.78)	(5.85)
Year FE	No	No	Yes	Yes	Yes
N	128007	128007	128007	114623	13384
R^2	0.348	0.356	0.357	0.235	0.216
Avg. elasticity	-0.744	-0.300	-0.248	-0.158	-0.392

In variable names, m denotes male and f denotes female.
Column 4 contains only gasoline cars and 5 only diesels.
Year FE: for each year, a dummy for whether the driving period covers the year.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure C.1: Market Shares in the Characteristics Space



the spirit of Berry, Levinsohn, and Pakes (1995) for future research.

As mentioned in Section (4), I have chosen to normalize $\alpha_2 = -1$ and estimate λ. When I estimate the model, I first estimate the reduced-form driving parameters subject to the normalization α_2. I then use these parameters as the starting values for the full, joint optimization. However, the likelihood function is increasing in λ up to the point where λ is so high that the model just predicts uniform choice probabilities for all choices. Recall that in logit models where the choice-specific utilities are non-linear (for example the present model or dynamic discrete choice models), the λ is sometimes identified and then it acts as a *smoothing* parameter. In some sense, it is analogous to the bandwidth in a Nadaraya-Watson kernel density estimator; in one extreme, when $\lambda \to 0$, the choice probabilities converge to an indicator function for the highest utility choice, $\Pr(j) = \mathbf{1}\{j = \arg\max_{j'} u_{j'}\}$. In the other extreme, when $\lambda \to \infty$, we choice probabilities become uniform, $\Pr(j) = 1/|\mathcal{J}|\forall j$, where $|\mathcal{J}|$ is the number of choices available. In intuitive terms, λ indicates how *precise* the model is, since it does not alter the ordering of the conditional utilities of the alternatives.

With this intuition at hand, it is easier to understand why the likelihood function is maximized for such a high value of λ. For given values of the remaining structural parameters, the model will tend to assign similar choice probabilities to cars that are in the same region of the choiceset. However, Figure C.1 illustrates that this is not the case in the data. The figure shows a scatter plot of the cars that are available in the 2006 choiceset. The x-axis denotes fuel efficiency in km/l and the y-axis denotes weight in metric tons while the coloring of the dots indicates the market share of each car in 2006. The figure shows that there are cars that are very close in characteristics with very dissimilar market shares. This will all else equal point towards characteristics not being important for determining the market shares of cars. However, the cross-equational restrictions implied by the model structure are such that to reduce the importance of the characteristics, the driving predictions will be altered. Therefore, I conjecture that the high value of λ is a way for the optimizer to reduce the importance of characteristics in predicting market shares without resulting in a bad fit of the driving equation.

If the explanation I have outlined above is correct, then controlling for car fixed effects should solve the problem. However, the big question for future work on this is whether they will suck up all the variation and result in a model where car choice becomes equally unresponsive to changes in policy; something which a priori must be wrong in light of the stark changes in the fuel efficiency and the diesel share following the 2007 reform.

For future research, I will now outline a potential strategy for estimating an extension of the model presented in this paper that allows for fully flexible car type fixed effects, $u^{\text{own}}(j) = \xi_j$. This is in line with the agenda of the Berry, Levinsohn, and Pakes (1995) literature, emphasizing the importance of unobserved car characteristics correlated with price (and possibly other characteristics).

This model with product-level fixed effects may be estimated in two ways; A direct approach would be to simply estimate all the $J-1 = 1,176$ dummies with maximum likelihood. Estimating such a large number of parameters would not be feasible using numerical derivatives, but with analytic derivatives and the BHHH approximation of the Hessian, complexity only increases linearly in the number of parameters.

An alternative approach is to apply a fixed point like that proposed by Berry (1994). Let $\Gamma : \mathbb{R}^{J-1} \to \mathbb{R}^{J-1}$ be the operator defined by $\Gamma(\xi^{[i]}) = (\Gamma_1(\xi^{[i]}), ..., \Gamma_{J-1}(\xi^{[i]}))$, where

$$\Gamma_j(\xi^{[i]}) = \xi_j^{[i]} + \sum_{t \in \mathcal{T}_j} \varpi_{jt} \left[\log s_{jt}^{\text{data}} - \log s_{jt}^{\text{pred}}(\xi^{[i-1]}) \right],$$

where s_{jt} is the market share for car j in year t, $\mathcal{T}_j$ is the set of years where car j was available and

$$\varpi_{jt} = \frac{N_t}{\sum_{t \in \mathcal{T}_j} N_t},$$

where N_t is the number of households going on the market in year t. Letting $\tilde{u}_{ij} \equiv u_{ij} - u^{\text{own}}(j)$, the predicted market share is given by

$$s_{jt}^{\text{pred}}(\xi^{[i-1]}) = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{\exp\left[(\tilde{u}_{ij} + \xi_j^{[i-1]})/\lambda\right]}{\sum_{k \in \mathcal{J}_t} \exp\left[(\tilde{u}_{ik} + \xi_k^{[i-1]})/\lambda\right]}.$$

This gives rise to the following algorithm;

Algorithm: A Berry (1994) fixed point.

Initialization: Set $\xi_j^{[0]} := \sum_{t \in \mathcal{T}_j} \varpi_{jt} \log s_{jt}^{\text{data}}$ and pick a reference car, j_0, for which $\xi_{j_0} := 0$.

Iteration: Given $\xi^{[i-1]}$, let $\xi^{[i]} = \Gamma(\xi^{[i-1]})$. Continue until $\|\xi^{[i]} - \xi^{[i-1]}\| < \epsilon$.

Recently, there has been some debate about numerical concerns with the implementations of algorithms using nested fixed points such as Berry (1994); Berry, Levinsohn, and Pakes (1995); Rust (1987). Dubé, Fox, and Su (2012) have emphasized the importance of using a tight inner-loop tolerance (ϵ) to avoid numerical noise spilling out into the outer loop. They suggest using the MPEC approach (Judd and Su, 2012). Instead, I follow the approach by Iskhakov et al. (2015) and use analytic derivatives for the inner loop, replacing the fixed-point iteration shown above with a root-finding solver for the quadratic system of non-linear equations,

$$\xi - \Gamma(\xi) = 0.$$

By using the analytic Jacobian of the operator Γ, which has a computationally simple form, I find that the solver converges in 13 iterations to machine precision.

D Additional Results

Table D.1 shows the structural elasticities from the preferred specification. The results are estimated based on a model with perfect foresight that allows random effects ($c_i \neq 0$). For the presented set of estimates, α_2 was fixed to -1, but very recently, I have successfully estimated that coefficient as well without it significantly changing the results.

Table D.2 shows the results from the baseline on the 2006 data as well as the 2007 counterfactual implemented in 2006 (same as column (3) of table 6.1) and an additional simulation of the 2007 reform where the pivot point of diesel cars is moved from 18km/l to 20km/l. The motivation is that the pivot point for gasoline cars is 16km/l but a typical diesel car drives about 4 km further per liter of fuel than a gasoline car. In that sense, the pivot of 20km/l should provide a better balance in the incentives.

In figure D.1 is shown the observed diesel share, the simulated diesel share from the model and a counterfactual simulation where both fuel price time series are kept at the 1997 level. The figure shows that the diesel share would have been higher in the later years if fuel prices had not changed. Two important points should be noted; Firstly, since the model conditions on entry into the new car market, raising or lowering fuel prices, for all cars will not change results as drastically as if more households were allowed to switch into car ownership. Nonetheless, raising fuel prices will lower expected driving and utility so given the convex utility in driving, some consumers will move towards more fuel efficient vehicles and therefore also diesel cars. This is also why, in the structural elasticities in table 5.3 we saw that when all fuel prices go up by 1%, the diesel share grows by 0.15%.

Secondly, the more important implication of holding fuel prices at the 1997 level is that the *relative* price of gasoline to diesel is kept constant. Figure D.2 plots two time series. On the left axis is the expected price of gasoline divided by the expected price of diesel (under perfect foresight — i.e. the fuel prices that are driving expectations) for a household going on the market in the given year and on the right axis is the predicted diesel market share for the year divided by the observed share. The figure shows that the tendency of the model to over or under-predict the diesel share is systematically related to the relative fuel prices. For example, the predicted share has two particularly striking periods; In 99–00, the prediction moves from over to under the observed share, coinciding with a sharp jump down in the relative price (diesel caught up with gasoline). In 05, the model has a kink down, under-predicting the diesel share. This coincides with a sharp jump down in the relative price from 117.9% to 110.9%, making diesels less favorable. Note that the predicted to observed share is not shown for 1997 because it is 432%. This extreme number is due to the observed share being quite close to zero in that year.

Table D.1: Estimated parameters

Fixed Parameters				
Parameter				Value
β				0.95
ψ				1
λ				10000
Model: Perfect foresight, quasi-linear, random effects.				
General Parameters				
Parameter			Estimate	t
σ_x			16.093	(69.12)
σ_α			21.951	(31.77)
Demographics				
	γ_z		α_{1z}	
Parameter	Estimate	t	Estimate	t
Constant	47.596	(35.22)	–	(–)
age	-8.447	(-18.97)	8.901	(8.71)
agesq	7.363	(15.88)	-15.168	(-14.39)
WDm	8.170	(18.95)	17.889	(69.45)
WDf	1.079	(19.46)	9.684	(108.20)
inc	-9.457	(-31.44)	-8.768	(-39.94)
nkids	1.453	(11.65)	-0.458	(-2.93)
city	-0.210	(-1.48)	-1.412	(-10.09)
Car Parameters				
Parameter			Estimate	t
α_{10}			74.927	(14.88)
α_{20}			-1.000	†
$\alpha_{0,\text{weight}}$			124074.734	(41.91)
$\alpha_{0,\text{weight}^2}$			-5009.689	(-5.67)
$\alpha_{0,\text{kw}}$			-413.653	(-25.53)
α_{0,kw^2}			5.114	(46.83)
$\alpha_{0,\text{displace}}$			-194.172	(-0.15)
$\alpha_{0,\text{displace}^2}$			4976.559	(13.12)
$\alpha_{0,\text{diesel}}$			-4235.595	(-24.99)
$\alpha_{1,\text{weight}}$			18.876	(3.12)
$\alpha_{1,\text{weight}^2}$			10.189	(5.64)

†: Fixed parameter, see section **??**.

Table D.2: Simulation of the 2007 Feebate Reform — The Role of the Diesel Pivot

	(1) Baseline	(2) 2007	(3) 2007 alt.
Consumer welfare			
E(CS	114970.09	115989.89	115363.51
Total taxes			
E(total taxes)	146623.83	134398.55	134482.53
Ownership tax			
E(Regtax revenue)	106556.44	98175.80	97702.93
E(Owntax revenue)	11093.62	9813.59	9779.74
Fuel tax			
E(O95 revenue)	25115.85	21695.96	23122.16
E(Diesel revenue)	3857.92	4713.20	3877.70
Driving/fuel use			
E(VKT)	79663.89	78740.62	78323.44
E(litre O95)	4340.92	3749.84	3996.34
E(litre D)	891.32	1088.92	895.89
E(litre D|urban)	188.02	230.59	189.60
E(kg CO2)	12736.56	11844.36	11924.39
Characteristics			
E(fe)	15.92	17.04	16.75
E(we)	1.70	1.64	1.64
E(kw)	77.08	70.07	70.72
E(displace)	1.65	1.54	1.54
E(% diesel)	18.49	24.48	19.63
E(% diesel|urban)	3.89	5.18	4.15

2007: The feebate reform of 2007 is implemented in 2006.

2007 alt.: As 2007 but the diesel pivot is 20 km/l instead of 18 km/l.

Figure D.1: Counterfactual Simulation: The Diesel Share Under Constant Fuel Prices

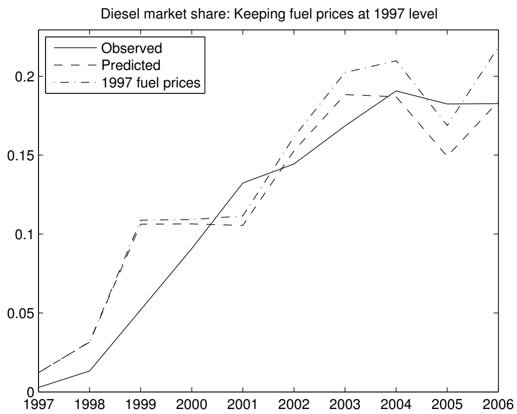
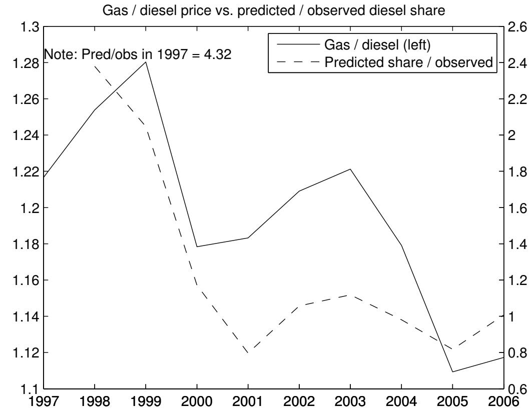


Figure D.2: Relative Fuel Prices and Relative Market Share Error



D.1 Static Expectations

Table D.3 shows the structural elasticities from the model estimated imposing the assumption of static expectations. The elasticity of driving with respect to PPK is now –39% as opposed to –30% for the perfect foresight estimates, indicating that to fit the data, the estimates must emphasize monetary costs more in this version of the model. Similarly, when the fuel efficiency of all cars in the choice set go up by 1%, the expected fuel efficiency goes up by 0.93% as opposed to 0.90% with perfect foresight. In other words, consumers are still substituting away some technological gains in fuel efficiency for other engine characteristics but not as much as earlier. And in particular, as PPK rises, the expected diesel share now falls. Finally, as the weight of all cars goes up by 1%, the expected weight now goes up by 1.58%, as opposed to just 1.15% earlier and the expected driving response (allowing for changes on the extensive margin) goes up by 1.71% as compared to 1.01% under static expectations.

In short, the estimates from the model imposing static expectations imply that money matters more to consumers and that the weight of the car also matters more for how much it is driven.

Figure D.3 compares the diesel share predictions from the models that impose perfect foresight and static expectations respectively with the observed diesel share. The movements in the two are highly similar but there is a slight tendency in the later years for the static expectations prediction to be slightly below the other.

Figure D.4 shows the 1997 counterfactual simulation using the estimates imposing static expectations. It shows that the conclusion from the perfect foresight model still holds; The counterfactual simulation where the 1997 reform was never imposed show a dramatically smaller diesel share in all years (but still an increase over time).

Table D.3: Structural Elasticities — Static Expectations

	(1) Baseline	(2) Fuel efficiency	(3) Weight	(4) Fuel prices	(5) O95 prices
		Consumer welfare			
CS	64412.21	0.43	2.32	-0.43	-0.33
		Total taxes			
E(total taxes)	139431.76	0.05	1.24	-0.05	0.05
		Ownership and registration tax			
E(Regtax revenue)	101066.75	0.19	1.11	-0.19	-0.02
E(Owntax revenue)	9999.10	0.23	1.37	-0.23	0.03
		Fuel/RUC tax			
E(O95 revenue)	23801.00	-0.55	0.85	0.55	-0.04
E(Diesel revenue)	4564.90	-0.43	6.32	0.41	2.08
		Driving/fuel use			
E(VKT)	81183.20	0.39	1.74	-0.39	-0.21
E(litre O95)	4113.67	-0.55	0.85	-0.44	-1.03
E(litre D)	1054.66	-0.43	6.32	-0.58	2.08
E(litre D|urban)	225.43	-0.40	6.40	-0.61	1.99
E(kg CO2)	12624.18	-0.52	2.04	-0.47	-0.34
		Characteristics			
E(fe)	16.18	0.93	-0.18	0.06	0.15
E(we)	1.70	0.10	1.58	-0.10	-0.01
E(kw)	72.07	0.14	0.71	-0.14	-0.06
E(displace)	1.53	0.10	0.58	-0.10	-0.00
E(% diesel)	19.77	0.25	4.36	-0.26	2.08
E(% diesel|urban)	4.20	0.27	4.39	-0.29	1.98

Elasticities based on estimates imposing static expectations

(2): Relative changes when e_j increases by 1% for all j.

(3): Relative changes when weight$_j$ increases by 1% for all j.

(4): Relative changes when fuel prices increase by 1%.

(4): Relative changes when gasoline prices increase by 1%.

All numbers are averages weighted with CCPs.

Figure D.3: Diesel Share Predictions — Comparing the Perfect Foresight and Static Expectations Predictions

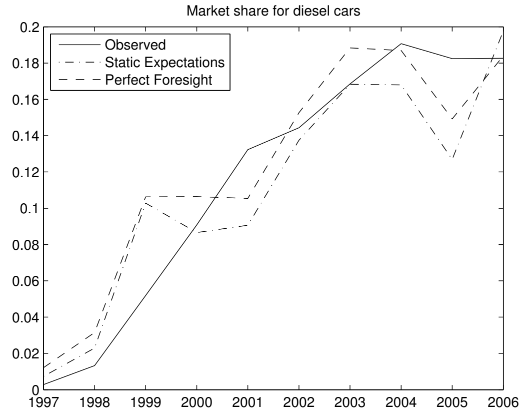
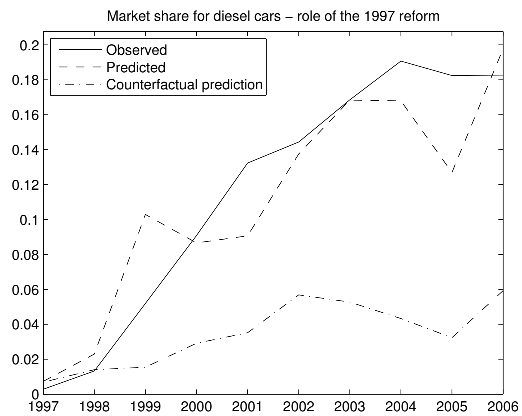


Figure D.4: 1997 Counterfactual — Static Expectations



References

Adamou, Adamos, Sofronis Clerides, and Theodoros Zachariadis. 2013. "Welfare Implications of Car Feebates:A Simulation Analysis." *The Economic Journal* URL `http://dx.doi.org/10.1111/ecoj.12094`.

Adda, Jérôme and Russell Cooper. 2000. "Balladurette and Juppette: A Discrete Analysis of Scrapping Subsidies." *Journal of Political Economy* 108 (4):778–806. URL `http://www.jstor.org/stable/10.1086/316096`.

Allcott, Hunt and Nathan Wozny. 2012. "Gasoline prices, fuel economy, and the energy paradox." *NBER Working Paper* .

Anderson, Soren T., Ryan Kellogg, Ashley Langer, and James M. Sallee. 2013. "The Intergenerational Transmission of Automobile Brand Preferences: Empirical Evidence and Implications for Firm Strategy." Working Paper 19535, National Bureau of Economic Research. URL `http://www.nber.org/papers/w19535`.

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

Beresteanu, Arie and Shanjun Li. 2011. "Gasoline Prices, Government Support, And The Demand For Hybrid Vehicles In The United States." *International Economic Review* 52 (1):161–182. URL `http://dx.doi.org/10.1111/j.1468-2354.2010.00623.x`.

Berndt, Ernst R, Bronwyn H Hall, Robert E Hall, and Jerry A Hausman. 1974. "Estimation and inference in nonlinear structural models." In *Annals of Economic and Social Measurement, Volume 3, number 4*. NBER, 653–665.

Berry, S., J. Levinsohn, and A. Pakes. 1995. "Automobile prices in market equilibrium." *Econometrica: Journal of the Econometric Society* 63 (4):841–890.

Berry, Steven T. 1994. "Estimating discrete-choice models of product differentiation." *The RAND Journal of Economics* :242–262.

Borger, Bruno De, Ismir Mulalic, and Jan Rouwendal. 2013. "Substitution between Cars within the Household." *Tinbergen Institute Discussion Paper* .

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

Chandra, Ambarish, Sumeet Gulati, and Milind Kandlikar. 2010. "Green drivers or free riders? An analysis of tax rebates for hybrid vehicles." *Journal of Environmental Economics and Management* 60 (2):78 – 93. URL `http://www.sciencedirect.com/science/article/pii/S0095069610000598`.

Chen, Jiawei, Susanna Esteban, and Matthew Shum. 2010. "How much competition is a secondary market?" Working Papers 2010-06. URL `http://ideas.repec.org/p/imd/wpaper/wp2010-06.html`.

Clerides, Sofronis and Theodoros Zachariadis. 2008. "The effect of standards and fuel prices on automobile fuel economy: an international analysis." *Energy Economics* 30 (5):2657–2672.

De Borger, Bruno and Inge Mayeres. 2007. "Optimal taxation of car ownership, car use and public transport: Insights derived from a discrete choice numerical optimization model." *European Economic Review* 51 (5):1177–1204.

D'Haultfæuille, X., P. Givord, and X. Boutin. 2013. *The Ecoomic Journal* .

DTU Transport. 2010. Tech. rep., DTU. URL `http://www.modelcenter.transport.dtu.dk/Publikationer/Transportoekonomiske-Enhedspriser`. Version 1.1.

Dubé, Jean-Pierre, Jeremy T Fox, and Che-Lin Su. 2012. "Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation." *Econometrica* 80 (5):2231–2267.

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

Feng, Ye, Don Fullerton, and Li Gan. 2013. "Vehicle choices, miles driven, and pollution policies." *Journal of Regulatory Economics* 44 (1):4–29. URL `http://dx.doi.org/10.1007/s11149-013-9221-z`.

Frondel, Manuel, Jorg Peters, and Colin Vance. 2008. "Identifying the Rebound: Evidence from a German Household Panel." *Energy Journal* 29 (4):154–163.

Frondel, Manuel, Nolan Ritter, and Colin Vance. 2012. "Heterogeneity in the rebound effect: Further evidence for Germany." *Energy Economics* 34 (2):461–467.

Gallagher, Kelly Sims and Erich Muehlegger. 2011. "Giving green to get green? Incentives and consumer adoption of hybrid vehicle technology." *Journal of Environmental Economics and Management* 61 (1):1 – 15. URL `http://www.sciencedirect.com/science/article/pii/S0095069610000768`.

Gavazza, Alessandro, Alessandro Lizzeri, and Nikita Rokestkiy. 2014. "A quantitative analysis of the used-car market." *American Economic Review* URL `http://mpra.ub.uni-muenchen.de/id/eprint/38414`.

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

Gillingham, Kenneth, Fedor Iskhakov, Anders Munk-Nielsen, John Rust, and Bertel Schjerning. 2013. "A Dynamic Model of Vehicle Ownership, Type Choice, and Usage." *Working Paper* .

Gillingham, Kenneth and Anders Munk-Nielsen. 2015. "The Tail Wagging the Dog: Commuting and the Fuel Price Response in Driving." *Working Paper* .

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

Greene, David L. 2010. “How Consumers Value Fuel Economy: A Literature Review.” Tech. Rep. EPA-420-R-10-008, U.S. Environmental Protection Agency.

Grigolon, Laura, Mathias Reynaert, and Frank Verboven. 2015. “Consumer Valuation of Fuel Costs and the Effectiveness of Tax Policy: Evidence from the European Car Market.” *Working Paper* .

Huse, Cristian and Claudio Lucinda. 2013. “The Market Impact and the Cost of Environmental Policy: Evidence from the Swedish Green Car Rebate.” *The Economic Journal* URL `http://dx.doi.org/10.1111/ecoj.12060`.

Hymel, Kent M. and Kenneth A. Small. 2015. “The Rebound Effect for Automobile Travel: Asymmetric Response to Price Changes and Novel Features of the 2000s.” *Energy Economics* forthcoming.

Iskhakov, Fedor, John Rust, Bertel Schjerning, and Jinhyuk Lee. 2015. “Constrained Optimization Approaches to Estimation of Structural Models: Comment.” *Working Paper* URL `http://ssrn.com/abstract=2583655`.

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

Judd, Kenneth and Che-Lin Su. 2012. “Constrained Optimization Approaches to Estimation of Structural Models.” *Econometrica* 80 (5):2213–2230.

Judd, Kenneth L and Benjamin S Skrainka. 2011. “High performance quadrature rules: How numerical integration affects a popular model of product differentiation.” *Working Paper* URL `http://ssrn.com/abstract=1870703`.

Li, Shanjun, Joshua Linn, and Erich Muehlegger. 2014. “Gasoline Taxes and Consumer Behavior.” *American Economic Journal: Economic Policy* 6 (4):302–42. URL `http://www.aeaweb.org/articles.php?doi=10.1257/pol.6.4.302`.

Mabit, Stefan L. 2014. “Vehicle type choice under the influence of a tax reform and rising fuel prices.” *Transportation Research Part A: Policy and Practice* 64:32–42.

Marion, Justin and Erich Muehlegger. 2011. “Fuel tax incidence and supply conditions.” *Journal of Public Economics* 95 (9–10):1202–1212. URL `http://www.sciencedirect.com/science/article/pii/S0047272711000545`.

Mayeres, Inge and Stef Proost. 2013. “The taxation of diesel cars in Belgium - revisited.” *Energy Policy* 54 (0):33–41. URL `http://www.sciencedirect.com/science/article/pii/S0301421511009670`.

Miravete, Eugenio J., Maria J. Moral, and Jeff Thurk. 2014. “Protecting the European Automobile Industry through Environmental Regulation: Adoption of Diesel Engines.” .

Reynaert, Mathias. 2014. "Abatement strategies and the cost of environmental regulation: Emission standards on the European car market." *KU Leuven Center for Economic Studies Discussion Paper Series DPS14* 31.

Rust, John. 1987. "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." *Econometrica* :999–1033.

Sallee, James M, Sarah E West, and Wei Fan. 2010. "The effect of gasoline prices on the demand for fuel economy in used vehicles: Empirical evidence and policy implications." *Working Paper* URL `http://www.ntanet.org/images/stories/pdf/proceedings/09/032.pdf`.

Schiraldi, P. 2011. "Automobile replacement: a dynamic structural approach." *The RAND journal of economics* 42 (2):266–291.

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

Small, K.A. and K. Van Dender. 2007. "Fuel Efficiency And Motor Vehicle Travel: The Declining Rebound Effect." *Energy Journal* 28 (1):25.

Spiller, Elisheba. 2012. "Household Vehicle Bundle Choice and Gasoline Demand: A Discrete-Continuous Approach." *Working Paper* .

Stolyarov, Dmitriy. 2002. "Turnover of used durables in a stationary equilibrium: Are older goods traded more?" *Journal of Political Economy* 110 (6):1390–1413.

Train, K.E. 2009. *Discrete choice methods with simulation.* Cambridge University Press, second ed.

Verboven, Frank. 2002. "Quality-based price discrimination and tax incidence: evidence from gasoline and diesel cars." *Rand Journal of Economics* :275–297.

Wakamori, Naoki. 2011. "Portfolio considerations in differentiated product purchases: An application to the Japanese automobile market." *Bank of Canada Working Paper* (27).

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