Abstract

In this paper, we investigate the effects of a proposed reform that reduces high Danish vehicle registration taxes and instead introduces road user charging. We analyze how such a reform affects substitution between vehicles with different characteristics (vintage, fuel efficiency, engine size, quality, etc.), the lifetime of these vehicles, and the demand for kilometers and fuel. To do this, we formulate a tractable life-cycle model of vehicle ownership, type choice and usage. We explicitly model macroeconomic conditions, usage, aging, replacement, and scrappage to study the non-stationary equilibrium in the used car market, i.e., how equilibrium prices in the used car market and purchases of new as well as used cars vary over the business cycle. We structurally estimate this model using high-quality Danish register data that covers the entire population of cars and households in Denmark.

Keywords: Automobiles, emissions, carbon tax, dynamic programming, secondary market

JEL classification: D92, L11, L13, Q38
1 Introduction

Government policies that affect durable goods inherently influence equilibria in both the new and used markets. The presence of a secondary market may even lead to unintended consequences. This is particularly true in the automobile market. For example, Corporate Average Fuel Economy Standards in the United States can be expected to raise the price of new vehicles and delay scrappage of older and often more polluting vehicles (Jacobsen, 2013). In other countries, this effect is even more evident. In Denmark, the new vehicle registration tax nearly triples the price of vehicle, disincentivizing new vehicle purchases and leading to a much older fleet than would be expected given the high per capita income of the country.

There are important dynamic considerations in consumer decisions that mediate how policies affect the allocation of new and used durable goods. The stock of vehicles is persistent and vehicles depreciate in value over time. Moreover, transaction costs lead to inertia in consumer holdings due factors such as costly search or asymmetric information. These dynamic considerations are particularly important for the welfare consequences of policies addressed to both the primary and secondary markets.

This paper develops a tractable life-cycle model of vehicle ownership, vehicle choice, and usage to examine the effects of a proposed reform that reduces the exceptionally high Danish vehicle registration tax and replaces it with road user charging, in which drivers pay a tax based on the number of kilometers driven. We model the dynamic considerations of the consumer in a framework that includes macroeconomic conditions, aging, replacement, and scrappage. Using this framework, we study the non-stationary equilibrium in the secondary market and can replicate the waves of vehicle prices and ownership decisions corresponding to the business cycle that are observed in the data. We estimate our model using detailed data from the Danish register on all vehicles in Denmark and their subsequent odometer readings matched to individual and household-level demographics. These data contain longitudinal information on income, wealth, labour market status, household composition, distance to work, occupation, and family patterns, as well as information on all vehicle transactions and average sales prices at the make-model-vintage level.

This paper contributes to several strands of the literature. Economists have long been interested in the importance of secondary markets in the allocation of new and used durable goods (Rust, 1985b; Anderson and Ginsburgh, 1994; Hendel and Lizzeri, 1999a,b; Stolyarov, 2002; Gavazza, Lizzeri, and Roketskiy, 2014). Rather than examine the role of secondary markets alone, we study the role of secondary markets in influencing the welfare consequences of a major government policy. This policy affects the vehicle market, a well-studied market in the economics literature, with significant work on product differentiation and consumer choice of new vehicles (Bresnahan, 1981; Berry, Levinsohn, and Pakes, 1995; Goldberg, 1995; Petrin, 2002). These seminal papers allow for general patterns of substitution across differentiated products, but do not model secondary
markets or the dynamics of the consumer decision process. Several recent papers focus on the secondary market for vehicles and the influence of durability on the dynamics of vehicle demand (Adda and Cooper, 2000a; Stolyarov, 2002; Esteban and Shum, 2007; Chen, Esteban, and Shum, 2013). These papers set the stage for this work, but have more limited household heterogeneity and do not examine policy counterfactuals. Gavazza, Lizzerei, and Roketskiy (2014) focus more on the welfare and allocative role of secondary markets, with no particular policy context in mind. In contrast, Schiraldi (2011) models the consumer’s dynamic decision process to estimate transaction costs and run a counterfactual of a scrappage subsidy in Italy. We have developed our model with a particularly important policy counterfactual in mind and estimate it with impressively detailed disaggregated data.

Our paper also contributes to the economics literature examining environmental policies in vehicle market. For example, Bento, Goulder, Jacobsen, and von Haefen (2009) use a static model of consumer demand and a Bertrand oligopoly model for automobile supply to examine the welfare and distribution effects of vehicle taxes in the United States. (Jacobsen, 2013) builds on this modeling framework to examine the effects of Corporate Average Fuel Economy Standards in the United States. These papers model both vehicle choice and usage decisions to provide useful policy insight, but abstract from the intertemporal dependence of consumer decisions. Our paper also uses more comprehensive data that allows us to incorporate vehicle replacement due to accidents separately from scrappage, and to model the impact of macroeconomic conditions on the vehicle purchase decision. Gillingham (2012) develops a two-period vehicle choice and usage model to examine the effects of gasoline taxes and policies that change the price of new vehicles. The focus on this paper is on estimating the rebound effect, i.e., the additional driving in response to a policy that raises fuel economy. A major contribution of this paper is that it develops a tractable model of dynamic consumer choice to estimate primitives that allow us to simulate the counterfactual equilibrium and accordingly, the effects of an important policy reform that is actually being considered.

There are a number of attractive features of our approach for examining the effects of the proposed road user charging reform. First and most importantly, the structural parameters have a clear interpretation from the theoretical model, allowing for counterfactual simulations to examine the welfare effects of the proposed reform. Our data allow us to obtain aggregate demand for vehicle investments, fuel consumption, and usage by aggregating individual demands resulting from consumer dynamic optimizing behavior. Furthermore, our empirical setting and data contain several reforms that provide plausibly exogenous variation to identify our structural parameters or validate the model.

We find that our model converges and can not only replicate waves in the observed data due to business cycles, but can rationalize the vehicle choice and usage behavior in Denmark. We
estimate the impact of the policy on the total stock of cars in Denmark and on new and used car prices, total driving, and emissions.

The remainder of the paper is structured as follows. The next section provides background on the institutional setting and discusses our dataset. Section 3 develops our dynamic model of consumer purchase, vehicle type, replacement, and usage choices. Section 4 discusses our estimation approach and the data. Section 5 describes how we solve for the non-stationary equilibrium. Section 6 presents our results and Section 7 concludes.

2 Background and Data

2.1 Empirical Setting

Denmark provides a very useful empirical setting for examining policies that affect the new vehicle registration tax and the operating cost per kilometer driven. Vehicle taxation in Denmark currently is made up three components: a one-time registration tax when the vehicle first enters the Danish fleet, an annual tax, and fuel taxes. The registration tax is a very large proportional tax with a kink, where various deductions apply. For example, in 2010 the tax was 105 percent of the first DKK 79,000 and 180 percent of the portion of the price exceeding 79,000. Examples of deductions include a reduction of the taxable value of the vehicle of DKK 3,750 if ABS brakes are installed and a reduction of DKK 12,000 from the final tax if the vehicle drives 1920 km per litre of gas. Changes in the tax and the deductions of time can be useful for identification and to test the validity of our model if we identify our structural parameters based on variation in new vehicle prices.

The empirical setting has further interesting details relevant to our estimation. There has been three reforms from 1992 to the present with an increasing focus on creating incentives for households to purchase more fuel efficient vehicles. Data on the fuel efficiency of new vehicles is available from the first reform in 1997. This reform set the annual tax for all vehicles first registered prior to July 1, 1997 according to the weight of the vehicle. At the same time, it set the annual tax for all vehicles registered after July 1, 1997 according to the fuel economy of the vehicle (in kilometers per liter). The motivation behind this reform was to tax older vehicles for wear and tear on the road and incentivize households to purchase more fuel-efficient new cars.

In 2000, deductions in the registration tax were introduced for vehicles in the higher end of the fuel efficiency scale (above 25 km/l). Therefore, only a very limited fraction of the vehicles sold in that year were actually affected by the reform. In the 2007 reform, these deductions were expanded so that all vehicles have their registration tax depend on fuel efficiency according to a piecewise linear schedule. If the vehicle has a fuel efficiency (FE) of more than 16 km/l, it receives a deduction of 4,000(FE - 16), and if it has a fuel economy less than 16 km/l, the tax is increased by 1,000(16 - FE). Not surprisingly we see a very strong response at the extremes: The market share
of the most fuel efficient cars increases from 8.1 percent prior to the reform to 50.4 percent at the end of the period in 2011 whereas for cars driving 16.6 km/l or less it decreases from 71.3 percent to 19.4 percent. Figure 1 plots the distribution of fuel economy of newly registered cars in the period 1997-2011. It highlights the substitution towards more fuel efficient cars, with a particularly noticeable shift in response to the 2007 reform.

There are some interesting alternatives to the private personal car worth noting in the Danish context. Firstly, an employer may provide a car for his employee. If that is the case, the employee must pay taxes of the value of having the car at his disposal which we can then see in the tax registers. Secondly, there is the alternative of purchasing a car on yellow plates which implies a substantially reduced registration tax. Prior to 2007, this imposed no restrictions on the use of the vehicle but was only applicable for certain 2-seater vehicles (mainly vans but also some cars with back seats removed). The size of the registration tax for yellow plate vehicles changed considerably from 1997 to present, but most notably in 2007, where a considerable reform meant that the number of newly registered vans fell from a little over 62,774 in 2006 to about 15,212 in 2009 (see figure 2).

2.2 Data

The dataset used in this paper draws on many different Danish sources. At the core of the dataset is information on the fleet of vehicles registered in Denmark is available from Statistic Denmark in the database bildata. The main source for the database is the Central Register of Motor Vehicles. The database keeps track of nearly all vehicles in Denmark and in particular all private personal vehicles. For each vehicle we have the vehicle identification number (VIN) and the owner’s CPR number, which uniquely identifies all individuals in Denmark. This register not only contains basic vehicle information, but also allows us to track ownership over individual vehicles over time.

Socioeconomic data for the owners of vehicles comes from various Danish registers. These contain the full Danish population in each year with the exception of Danes living abroad. The CPR number is given to any individual taking residence for longer than 3 months in Denmark (6 months for Nordic or EU citizens) and is used in nearly all dealings with official authorities from education and taxation to the purchase of medicine and criminal records. Thus, the dataset includes detailed educational information, place of residence and time of movements, income and wealth information from the tax report (which for most employees is 3rd party reported). We merge in information on spouses and children to give an adequate picture of the household.

Another important vehicle register dataset contains information on the vehicle tests performed by the Danish Ministry of Transportation (MOT). There are three main types of tests, with the goal of ensuring that vehicles in Denmark are safe to drive. A registration test is performed when the

---

1Exceptions that are not included in the register include for example military vehicles.
vehicle is registered. Periodic tests are performed bi-annually from the fourth year since the car
was registered and the rest of its lifespan. Customs tests are performed on imported used vehicles
prior to their registration test when they are registered in Denmark. The most important variable
from the MOT tests is the odometer reading, which allows us to track the usage of individual
vehicles. Using the VIN, these odometer readings are merged with the vehicle register database.

Finally, the Danish Automobile Association (DAF) maintains a database of prices of vehicles
by make, model, variant, year and vintage, allowing us to follow the value of used cars as well.
The main limitation of this is that we do not observe what additional equipment was purchased
with the car but DAF provide an informed guess of what the typical price would be as well as a
high and low price, bounding the price range for that specific vehicle. Moreover, there are both
the price a professional car dealer would pay and the price he would demand for a given vehicle,
giving a proposed margin. The prices are highly reliable and are used by professional car dealers
in setting the price of a used vehicle.

3 The Model

We consider an infinite horizon model of both vertical and horizontal product differentiation. Let
\( \tau \) denote the “type” of vehicle. We will assume there are a finite number of possible types, \( \tau \in
\{1, \ldots, \tau\} \). These can be thought of as a make-model combination or simply a vehicle class (e.g.,
‘luxury’ ‘compact’ ‘economy’ ‘SUV’ ‘sport’ and ‘minivan’).

To capture vertical product differentiation, we also distinguish the age of the vehicle \( a, a \in
\{0, 1, \ldots, \bar{a}\} \) where \( a = 0 \) denotes a brand new vehicle, and \( a = 1 \) a one year old vehicle, and \( \bar{a} \) is
the oldest vehicle in the market. For simplicity, we let \( \bar{a} \) be a catchall class of all cars that are of
age \( \bar{a} = 20 \) or older. Thus, we index the set of cars that consumers in Denmark can choose from
by \( (\tau, a) \) where \( \tau \) specifies a particular type of car and \( a \) denotes its age.

This formulation is very useful for the tractability of the model, but does abstract from changes
in technology. We can note that changes in the real prices of cars are likely to be more attributable
to macroeconomic conditions than a particular technology innovation, but this is an area for future
work. There may also be a considerable degree of unobserved heterogeneity in used vehicles of a
given age. For example, some have been driven more than others, and some are in better condition
than others. However, Cho and Rust (2010) show that vehicle age and odometer readings are highly
correlated and that once age is included as a predictor of car prices, the incremental predictive value
of including the odometer is small. This still leaves unobserved heterogeneity in the condition
of the car and other dimensions. We account for such unobserved heterogeneity with random
variables that model the net effect of all of these factors.

Besides the variables \( (\tau, a) \) that we use to index the state and the choice set of cars that a
consumer can choose from at any given point in time, we introduce the key macro variables that we believe are relevant both for individual choices and for the equilibrium of the market as a whole, \((p,m)\) where \(p\) is the current price of gasoline and \(m\) is an indicator of the “macro state” of the Danish economy. We model \(m\) as a binary variable where \(m = 0\) indicates that the economy is in a recession period, and \(m = 1\) indicates a non-recession period.

Consumer expectations of the price that they could buy a typical car of type and age \((\tau, a)\) when the economy is in state \((p,m)\) are given by the function \(P(\tau, a, p, m)\). To account for heterogeneity in the prices of individual vehicles bought and sold, the transaction price of an individual vehicle is given by

\[
\tilde{P} = P(\tau, a, p, m)\tilde{\xi},
\]

where \(\tilde{\xi}\) is a lognormally distributed random variable that represents the effects of other characteristics besides the crude state of the vehicle \((\tau, a)\) and the current economic state \((p,m)\) that affect the actual transaction price. Let \(x\) denote a vector of individual-specific variables the most important of which include a) age, b) income, c) residential location, and d) distance to work. Age and income are treated as time-varying state variables, while all other variables are treated as time-invariant.

The characteristics of vehicle holdings in Denmark are especially convenient for reducing the curse of dimensionality in estimating a dynamic model. Nearly all households in Denmark own at most one vehicle. Thus, we restrict our analysis the the choice of owning a vehicle or not. At the start of each year a household makes a decision one whether to buy a new vehicle and or sell their existing vehicle. We will let \(\emptyset\) denote the state where a household does not own a car, and so a household that does not own a car can either continue not to own a car, or purchase a new or used car \((\tau, a)\). A household that does have a car \((\tau, a)\) can either decide to keep this car, or trade it for another car \((\tau', a')\) or just sell it without replacing it so its “car state” becomes \(\emptyset\). We will let \(u(\emptyset, x, \eta, p, m)\) be the utility a household with observed state \(x\) and unobserved type \(\eta\) gets from not owning a car when the economic state is \((p, m)\), whereas \(u(\tau, a, x, \eta, p, m)\) is the indirect utility the household recieves if it owns a car of type \((\tau, a)\) in the economic state \((p,m)\).

### 3.1 Household Dynamic Vehicle Choice Problem

Let \(d\) denote the household’s choice at time \(s\), where periods are indexed with the age of the head of the household.\(^2\) Denote \(D\) the full set of choices that households may have in different circumstances, and \(D_s\) the actual choice set faced at time \(s\) that is dependent of the state variables as explained below. The full set \(D\) is comprised of decisions to purchase a new car \((\tau', a')\) as well as the decision to keep the status quo \((d = \emptyset)\) or purge the existing car without replacing it \((d = \mathbb{P})\).

\(^2\)In the next section we’ll make a distinction between lifecycle time indexed with \(s\) and calendar time indexes with \(t\).
To keep the model tractable, we only control for the age of cars up to some upper bound $\bar{a}$, and from the point of view of the model car characteristics become fixed beyond that age. For the reasons explained in the next section, we assume that the oldest car which can be bought on the used car market is $\bar{a} - 1$ old, and there is no trading of $\bar{a}$-old cars other than selling to a scrap yard.

Therefore, a households which owns a car $(\tau, a)$ in the period $s$ (which we refer to as existing car) is choosing between whether to keep it, purge it or replace it with another car, new or used, i.e. $d \in D_s(\tau, a) = \{ \mathbb{K}, \mathbb{P} \} \cup \{(\tau', a'), a' < \bar{a} \}$. For the household without a car the choice set is one element smaller, namely it is not possible to get rid of non-existing car, i.e. $d \in D_s(\emptyset) = \{ \mathbb{K} \} \cup \{(\tau', a'), a' < \bar{a} \}$.

Our timing convention is the following. We assume that the existing car $(\tau, a)$, if any, is driven during current period $s$, and may be replaced by the other car $(\tau', a')$ (new if $a' = 0$) which becomes existing $\max\{\bar{a}, a' + 1\}$ years old car and is driven during the next period $s + 1$. If the household decides to keep their existing car ($d = \mathbb{K}$), it ages to become $(\tau, \max\{\bar{a}, a + 1\})$ in the period $s + 1$, and is driven in both period $s$ and period $s + 1$. If the household decides to get rid of their exiting car $(\tau, a)$, they still drive it in period $s$, but there is no car driven in period $s + 1$ which we denote $\emptyset$.

We allow for a possibility of an accident, which occurs with probability $\lambda_s$ to the existing car $(\tau, a)$ in time period $s$. In the event of an accident, the existing car can not be driven next period, and a monetary penalty is paid. The penalty is though of as a net payment that includes the scrap value of the broken car, fines, and out-of-pocket expenditures caused by an accident. From the modelling point of view, an even of the accident results in an additional restriction on the choice set $D_s$ in period $s$, namely the decision to keep the status quo $d = \mathbb{K}$ is removed from the available choices.

Let $v_s(\tau, a, x, \eta, p, m, d)$ be the lifetime expected present discounted utility of a household in state $(x, \eta)$ making the decision $d \in D_s$ at time $s$, when economy is in state $(p, m)$, and existing car is $(\tau, a)$.

We assume that there are additive IID Type 1 Extreme value components $\epsilon(d)$ to the indirect utilities of all choices $d \in D_s$ that this household faces at period $s$. We let $\epsilon = \{\epsilon(d) | d \in D_s\}$ be the vector of all of these additive components, which we assume capture the effect of idiosyncratic variability in the amount received for a car sold and the cost of a new car purchased and other factors, all of which are observed by the households at the time they make their decisions but not by the econometrician. Completely analogue assumptions are made for the households with no existing car. The Bellman equation for this household is

$$ V_s(\tau, a, x, \eta, p, m, \epsilon) = \max_{d \in D_s} \left[ v_s(\tau, a, x, \eta, p, m, d) + \epsilon(d) \right] $$

(2)

To simplify the exposition below, we disregard the accident penalty, which if different from zero requires one more argument in $v_s$ and $V_s$ below. Additional state variable indicating if an accident has happened in the current period would have to be added in the obvious way.
where the difference between $V_s$ and $v_s$ is that $V_s$ is the **maximized value of expected discounted utility over all of the discrete alternatives** $d \in D_s$, whereas $v_s$ is the expected discounted utility of a **particular** alternative $d \in D_s$. Thus, $V_s$ is the discrete-choice analog of the indirect utility function, and $v_s$ is the discrete choice analog of the direct utility function.

Applying the log-sum formula for the expectations of the maximum of the extreme value distributed random variables $V_{s+1}$, we can derive alternative specific Bellman equations to completely specify the model. To simplify the exposition below, denote

$$
\varphi(s, \tau, a, x, \eta, p, m, D_s) = \log \left( \sum_{d \in D_s} \exp \left\{ v_s(\tau, a, x, \eta, p, m, d) \right\} \right).
$$

For the households with existing car $(\tau, a)$ who decide to keep it $(d = \mathbb{K})$, we have

$$
v_s(\tau, a, x, \eta, p, m, \mathbb{K}) = u(\tau, a, x, \eta, p, m) + \beta_s \int_{x'} \int_{p'} \int_{m'} \left[ \lambda_s \varphi(s + 1, \tau, a + 1, x', \eta, p', m', \{\mathbb{K}\} \cup \{(\tau', a'), a' < \bar{a}\}) (1 - \lambda_s) \varphi(s + 1, \tau, a + 1, x', \eta, p', m', \{\mathbb{K}, \mathbb{P}\} \cup \{(\tau', a'), a' < \bar{a}\}) + \right] g_s(dx'|x, p', m') h(p', m'|p, m) dx' dp' dm',
$$

where $\beta_s$ is a mortality-adjusted discount factor for the household, $g_s(x'|x, p', m')$ is a transition probability density for the household’s state $x'$ at age $s + 1$ given their current state $x$ at age $s$ and the next period macro state $(p', m')$, and $h(p', m'|p, m)$ is a transition density for fuel prices and the macro state. If the household chooses to sell their existing car $(\tau, a)$ and not replace it with another one $(d = \mathbb{P})$, the Bellman equation is given by

$$
v_s(\tau, a, x, \eta, p, m, \mathbb{P}) = u(\tau, a, x, \eta, p, m) + \beta_s \int_{x'} \int_{p'} \int_{m'} \varphi(s + 1, \tau, 0, x', \eta, p', m', \{\mathbb{K}\} \cup \{(\tau', a'), a' < \bar{a}\}) g_s(dx'|x, p', m') h(p', m'|p, m) dx' dp' dm'.
$$

When the household chooses to replace their existing car $(\tau, a)$ with another car $(\tau', a')$ (new if
\( a' = 0 \) or used), the Bellman equation is given by

\[
v_s(\tau, a, \eta, p, m, \tau', a') = u_T(\tau, a, \tau', a', x, \eta, p, m) + \\
\beta_s \int_{x'} \int_{p'} \int_{m'} \left[ \\
\lambda_s \varphi(s + 1, \tau', a' + 1, x', \eta, p', m') \left\{ \mathbb{P} \cup \{(\tau'', a''), a'' < \bar{a}\}\right\} \\
(1 - \lambda_s) \varphi(s + 1, \tau, a' + 1, x', \eta, p', m') \left\{ \mathbb{K}, \mathbb{P} \cup \{(\tau'', a''), a'' < \bar{a}\}\right\} + \\
g_s(dx|x, p', m')h(p', m'|p, m)dx'dp'dm',
\]

where \( u_T(\tau, a, \tau', a', x, \eta, p, m) \) is the expected indirect utility to a consumer that decides to trade their vehicle \((\tau, a)\) for another one \((\tau', a')\) given by

\[
u_T(\tau, a, \tau', a', x, \eta, p, m) = \max_t \{w(t, y - f(t, \tau', a', p, m) + T(P(\tau', a', p, m) - P(\tau, a, p, m)), x, \eta),
\]

(4)

where \( T(P(\tau', a', p, m) - P(\tau, a, p, m)) \) is the total expected transactions cost involved in selling the existing car \((\tau, a)\) to buy another car \((\tau', a')\). If the other car is a new car \(a' = 0\), then the household will be subject to the very high Danish new car tax that can be up to 80% of the price of the new car \(P(\tau, a, p, m)\).

A modified form of this specification is to recognize that the transaction cost should reflect realized rather than expected prices of the vehicles bought and sold. If the household knows the realized transaction prices and is making its decision based on the realized rather than the expected transaction cost, the utility \( u_T(\tau, a, \tau', a', x, \eta, p, m) \) will be a random variable from the standpoint of the econometrician, since it reflects information on the realized transactions cost that the econometrician does not observe. Though there is an ad hoc and not a fully utility-consistent derivation of this realized utility, we assume that the extreme value error component (which is present for all alternatives) captures the difference between the ex ante formula for \( u_T \) given in equation (4) and the realized or ex post value that reflects the realized transaction cost that we do not observe as econometrician. So we make the assumption that the realized or ex post value of the indirect utility of trading the current vehicle \((\tau, a)\) for a new one \((\tau', a')\) is captured by the additive extreme value error component \( \varepsilon(d) \) where \( d = (\tau', a') \).

Now, consider a household without an existing car, which can continue to not own a vehicle \((d = \mathbb{K})\) or purchase a vehicle \((\tau', a')\). The value of the \( d = \mathbb{K} \) option is

\[
v_s(\emptyset, x, \eta, p, m, \mathbb{K}) = u(\emptyset, x, \eta, p, m) + \\
\beta_s \int_{x'} \int_{p'} \int_{m'} \varphi(s + 1, \emptyset, x', \eta, p', m') \left\{ \mathbb{K} \cup \{(\tau', a'), a' < \bar{a}\}\right\} g_s(dx|x, p', m')h(p', m'|p, m)dx'dp'dm'.
\]
The value for the decision to purchase a car \((\tau, a)\) is given by

\[
v_s(\emptyset, x, \eta, p, m, \tau, a) = u(\emptyset, x, \eta, p, m) + \beta_s \int_{x'} \int_{p'} \int_{m'} \left[ \lambda_s \phi(s + 1, \tau, a + 1, x', \eta, p', m', \{P\} \cup \{ (\tau', a') \} \} \right. \\
\left. (1 - \lambda_s) \phi(s + 1, \tau, a + 1, x', \eta, p', m', \{P, \bar{P}\} \cup \{ (\tau', a') \} \} \right] + \\
g_s(dx'|x, p', m') h(p', m'|p, m) dx'dp'dm',
\]

To close the model it is only left to specify the transition densities for the household characteristics \(x\) and macro state given by fuel prices \(p\) and macro indicator \(m\). Solving the model amounts to finding all choice specific values \(v_s\) in all points of the state space, which given our assumption about the IID extreme value distributed additively separable components, lead to the standard logit choice probabilities facilitating maximum likelihood estimation of the model.

### 3.2 Utility Specification

The approach here loosely follows that in Gillingham (2012) and Munk-Nielsen (2014). Let \(vkt\) be the total planned kilometers travelled by car over the coming year, and let \(f(vkt, \tau, a, p, c^o)\) be the total expected cost of this travel using the vehicle \((\tau, a)\) at fuel price \(p\) and operating costs per km of \(c^o\). The major component of this is the fuel cost which is proportional to the price of gasoline \(p\) (price per liter of fuel consumed, inclusive of taxes), and the distance travelled \(vkt\) (say measured in kilometers) times the fuel economy of the vehicle \(fe\) (measured in liters of fuel per kilometer travelled). However \(f(vkt, \tau, a, p, c^o)\) can also include road tolls and other expected maintenance costs that are a function of the distance travelled. Let \(u(vkt, y - f(vkt, \tau, a, p, c^o) - c^T, \eta)\) be the direct utility a household expects from owning a vehicle and driving a planned \(vkt\) kilometers, where \(c^T\) is the consumer “trade cost,” which is the total cost of a change in vehicle. \(c^T\) includes the difference between buying and selling costs as well as any transaction costs. We will model these transaction costs as a function of the macro state \(m\).

Let the utility of owning a vehicle and driving it be given by

\[
u(vkt, y - f(t, \tau, a, p, c^o) - c^T, \eta) = \gamma^{kt}(\tau, a, \eta, m) \log vkt + \gamma^{mu}(y) \left[ y - f(vkt, \tau, a, p, c^o) - c^T \right] - \gamma^{age} \sqrt{a}.
\]

This specification directly provides for diminishing marginal utility of driving. We could consider utility as a quadratic function of \(vkt\) to be more flexible, but the logarithm simplifies the
setting. The consumer’s optimal planned driving is then given by

\[ vkt^* = \arg\max_{vkt} u(vkt, y - f(vkt, \tau, a, p, c^0) - c^T, \eta) \]

Specify the driving costs for vehicle \((\tau, a)\) as

\[ f(vkt, \tau, a, p, c^0) = \left( \frac{p_f}{fe} + c^a \right) vkt \]

where we introduce \(p^{km} \equiv \frac{p_f}{fe} + c^a\) to simplify the exposition. The first-order condition for the optimal driving is then given by

\[ \gamma_{vkt} vkt^* = \gamma_{mu} p^{km} \]

\[ \Leftrightarrow vkt^* = \frac{\gamma_{vkt}}{\gamma_{mu} p^{km}}. \]

To allow the marginal utility of driving to be scaled by the type of the vehicle \((\mu_\tau)\), a quadratic in the age of the vehicle, and the unobserved consumer type \(\eta\), let

\[ \gamma_{vkt} = \gamma_{vkt0} + \gamma_{vkt1} a + \gamma_{vkt2} a^2 + \gamma_{vkt3} m + \mu_\tau + \eta + u_{vkt}. \]

This equation requires an assumption about the nature of \(u_{vkt}\). Since we are examining the planned driving for each consumer, we can assume that \(u_{vkt}\) is known to the consumer and thus is not stochastic in (5).

Now in order to allow for a diminishing marginal utility of income, let

\[ \gamma_{mu} = \frac{\gamma}{\sqrt{y}}. \]

Moreover, we allow for the macro state, \(m\), to affect the transaction costs, \(c^T\), in the following way,

\[ c^T = \gamma^0 + \gamma^m m + p^{new} - p^{old}, \]

where \(m \in \{0, 1\}\) so that the parameter \(\gamma^m\) shifts the sunk part of the transaction cost up and down.
Then the optimal $v_{kt}$ is given by

$$v_{kt}^* = \left( \frac{\gamma v_{kt0} + \gamma v_{kt1} a + \gamma v_{kt2} a^2 + \gamma v_{kt3} m + \mu + \eta + u_{vkt}}{\gamma m} \right) \sqrt{\frac{fe}{p^f - c^o_fe}}$$

$$= \frac{(\gamma v_{kt0} + \gamma v_{kt1} a + \gamma v_{kt2} a^2 + \gamma v_{kt3} m + \mu + \eta)fe \sqrt{p^f - c^o_fe}}{\gamma m} + u_{vkt1},$$

where $u_{vkt1} = u_{vkt} \left( \frac{\sqrt{fe}}{p^f - c^o_fe} \right)$. Note that if this equation is estimated separately, we cannot separately identify each of the coefficients. However, $\gamma m$ should also enter into the utility of the vehicle choice, so if we are estimating the model simultaneously, we will be able to identify all coefficients.

We estimate the driving decision model separately. In this case, we must pin down one of the coefficients. For example, we can normalize $\gamma m = 1$. In this case, we can perform a two-stage estimator and first estimate the following relationship

$$v_{kt}^* \left( \frac{p^f - c^o_fe}{\sqrt{fe}} \right) = \gamma v_{kt} + \gamma v_{kt1} a + \gamma v_{kt2} a^2 + \gamma v_{kt3} m + \mu + \eta + u_{vkt}$$

Appendix 1 derives the derivatives of the utility function, which are used for estimation.

### 4 Estimation of the Model

The detailed Danish register data enable us to identify the type of car and its age ($\tau, a$) for every Danish household that owns a car, and the type and age ($\tau', a'$) of a replacement vehicle for any household that trades a vehicle. So we construct a panel dataset $\{d_{i,t}, x_{i,t}\}$ for based on a large random sample from our data on all Danish households, $i = 1, \ldots, N$ over time periods $t$ where $d_{i,t}$ is the car holding/trading decision by household $i$ during year $t$, and $x_{i,t}$ are other household level variables we include in our dynamic programming model, the most important of which are the age of the household head $a_{i,t}$, the household’s income $y_{i,t}$ and its residential location (which we assume is a time invariant component of $x_{i,t}$).

Given the one year decision time intervals in our model, we fix a particular time at which decisions are assumed to take place for purposes of matching the model to the data. Specifically, we assume decisions are made on January 1 of each year. We also assume that income $y_{i,t}$ represents total income (after tax) in the previous year and the age variable $a_{i,t}$ is the age of the household head as of January 1. For the decision variable, we assume that a decision pertains to the coming year and so a household is recorded as trading its vehicle if we observe a sale between January 1st of the year and December 31st of that year.

To address attrition of older households (e.g., death of the household head, or reaching the maximum age where we assume the household head is too old to own a car or make further decisions...
about buying or selling vehicles), we can draw a sample of individuals over time so that the newly formed household are added to counter-balance attrition. Our sample thus remains approximately representative of the Danish population.

We propose a “semi-parametric two step” approach to estimation of the dynamic programming model. A key part of the model is consumers’ beliefs about the prices of vehicles, given by the price functions \( P(\tau, a, p, m) \). Based on the Danish used car price data, we can estimate a look-up table of this price function via regression. More specifically, we have a set of estimated price functions \( P(\tau, a, p, m) \) that can be used for our dynamic programming calculations of the form:

\[
P(\tau, a, p, m) = \exp \{ \gamma_0(\tau, p, m) + \gamma_1(\tau, p, m)a + \gamma_2(\tau, p, m)a^2 \},
\]

where the coefficients \( (\gamma_0, \gamma_1, \gamma_2) \) is estimated from a set of seemingly unrelated time series regressions with dependent variables \( \log(P_t(\tau, p)) \) for each vehicle class \( \tau \), for the set of all ages \( a \in \{0, \ldots, \bar{a}\} \) and including a flexible polynomial specification for how the \( (\gamma_0, \gamma_1, \gamma_2) \) coefficients depend on \( (p, m) \).

Another part of the “first stage” estimation involves estimating an income process for households to create the transition probability \( g_s(y'|y, p', m') \) predicting the probability distribution of income earned in the coming year, \( y' \) for a household whose head is age \( s \) and whose income in the previous year was \( y \) and when gas prices and the macroeconomic state of the economy in the coming year are \( p' \) and \( m' \), respectively. By including other household-specific variables and potentially also unobservable random effects \( \eta \), so we can use the Danish register data to estimate specifications of the form \( g_s(y'|y, x, \eta, p', m', p, m) \) where we include contemporaneous and lagged gas prices and macro variables to also help predict individual household income.

The log-likelihood for the full sample can be written as

\[
L(\theta) = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \log \left( \sum_{\eta} P(d_{it}|x_{it}, \eta, \theta) g_{a_{it}}(y_{it}|y_{it-1}, x_{it}, \eta, p_t, m_t, \theta) \lambda(\eta) \right)
\]

where \( \lambda(\eta) \) is the (discrete) probability that the household has unobserved preference type \( \eta \). The Type I extreme value distributional assumption leads to the standard multinomial logit expression for the choice probability \( P(d|x, \eta, \theta) \) where \( \eta \) represents the union of preference parameters entering the indirect utility function, the discount factor, and parameters of the income transition densities \( g_s(y',|y, x, \eta, p', m') \).

We then maximize the log-likelihood using several common algorithms, such as BHHH.
5 Solving for Equilibrium Prices

The following section contains our notes on how we are solving for equilibrium prices (and the motivation behind our choice) and has not been converted to paper form yet.

The first large scale numerical equilibrium calculation for new and used cars that we are aware of is Berkovec (1985). Berkovec’s model is static rather than dynamic, but he incorporates dynamic aspects by defining consumers as choosing the make/model and age of a used car as a static choice, but which also includes the expected price depreciation of the car as an addition to the maintenance and fuel costs. Berkovec estimated a nested logit model of car type and age choice using cross sectional data from the National Transportation Survey covering 1095 households. He merged in a data file on vehicle attributes (car weight, fuel economy, etc) from Cambridge Systematics, Inc. covering 831 different makes and models of vehicles for the 1967-1978 model years. “For the demand model estimation, vehicles were grouped into 131 classes based on 13 size/type and 11 vintage classifications with characteristics of each class defined as the weighted mean of the characteristics of vehicles in the class. These groupings, described in Table 1, are somewhat subjective, but are roughly aligned with industry marketing classifications.” (p. 201). Berkovec estimated auxiliary models of car scrappage using auto registration data from R.L. Polk, Inc. These data are also used to determine the total stock of cars in each of the 131 car type/age classes, thus determining the exogenous “supply” of vehicles in his analysis. He computed equilibrium prices by microaggregating his estimated nested logit model of vehicle demand using Census weighting to create a synthetic population representative of the US in 1978. He used a quasi-Newton algorithm to solve for the vector of 131 prices for the 131 type/age categories that sets the microaggregated demand for vehicles in each category to the supply.

Berkovec used population projections from the Current Population Survey and projected gas prices to make projections of market equilibrium for the period 1978 to 1990 under several scenarios. “The base-case simulation predicts a rapid increase in average new car fuel efficiency during 1978-1982. This increased efficiency is caused both by higher gasoline prices and by the rapid improvement in the fuel efficiency of the available new cars from 1978-1982 (caused partly by higher government fuel efficiency standards).” (p. 210). However Figure 1 in his paper shows that Berkovec’s base case forecasts substantially overpredict the sales of new cars in the early 1980s due to the model’s failure to incorporate macro shocks and the impact of the 1982 recession. The model does capture the increasing trend in fuel economy, “The base-case simulation predicts a rapid increase in average new car fuel efficiency during 1978-1982. This increased efficiency is caused both by higher gasoline prices and by the rapid improvement in the fuel efficiency of the available new cars from 1978-1982 (caused partly by higher government fuel efficiency standards).” (p. 210). Further, “The increased new car fuel efficiency along with the gradual retirements of older vehicles cause the predicted average MPG of the vehicle fleet to increase steadily over the
1978-1990 period as shown in Figure 4. Fleet fuel efficiency improvements are most rapid in the early part of the period as the inefficient early 1970s vintage vehicles are scrapped because of age and increased gasoline prices.” (p. 210). Thus our reading of the earliest work on numerically calculating equilibria in the automobile market confirms the importance of gas price shocks and consumer responses to these shocks by buying more fuel efficient vehicles. However it also confirms, via omission of macro shocks in the model and the model’s consequent failure to capture pronounced declines in new vehicle sales, the importance of cyclical macroeconomic fluctuations on equilibrium in the auto market. Berkovec’s paper did not present the calculated equilibrium prices, so it is hard to judge whether the model provided a sufficiently “realistic” model of the US auto market in other respects.

Around the same time as Berkovec’s study, a separate strand of the literature focused on the definition an calculation of auto market equilibrium accounting for the fact that consumers are making dynamic choices of when to buy and sell vehicles. However most previous studies that have calculated equilibria in new and used markets for vehicles have calculated stationary equilibria which are essentially steady state prices that equate supply and demand of new and used vehicles of all vintages/conditions. In these models the aggregate distribution of ages or odometer of the vehicles in the economy never changes since the economy is assumed to be in steady state. Examples of this approach include Rust (1985b), Stolyarov (2002), and Gavazza, Lizzeri, and Roketskiy (2014). In Rust’s model the odometer was treated as the relevant measure of the state of the car, so this model resulted in a continuum of goods and the equilibrium price is a function that assures that the distribution of demand for vehicles equals the distribution of the supply of vehicles. Further, stationarity implies that the market is in flow equilibrium so that the fraction of vehicles that are scrapped in any period are balanced by a corresponding fraction of new vehicles that are produced and sold to the market. In Rust’s model there are no accidents or reasons why vehicles would be scrapped “prematurely” and allow new vehicles are of the same type and vehicles only differ by their odometer values. An individual vehicle’s life history can be treated as a Markov process over its odometer readings, with a new vehicle starting its like with odometer value $O_0 = 0$, and then based on stochastic driving patterns over a sequence of owners until the vehicle is ultimately scrapped, the car will be in a series of odometer states $(O_0, O_1, O_2, \ldots, O_{T-1})$ until the first time $T$ at which $O_T \geq \gamma$ where $\gamma$ is a scrappage threshold then the car is scrapped. The steady state distribution of odometer values of the vehicles in the economy will be the unique invariant distribution of the regenerative Markov process representing the sequence of odometer values that a car “experiences” over its lifetime, until it is scrapped and replaced by a brand new vehicle (thereby assuring that flow equilibrium holds).

The invariant distribution of odometer values represents the “supply” of cars of different odometer values in steady state, and so it is only necessary to compute an equilibrium price function $P(O)$
that provides the price of a car as a function of its odometer $O$ that assures that the distribution of
the demand for cars of different odometer values equals the distribution of supply. Rust assumed
that there is an infinitely elastic supply of new cars at a price $P$ and an infinitely elastic demand
for scrapped vehicles at a price $P < P$. Thus, an equilibrium price function $P(O)$ is decreasing
function of the odometer value $O$ satisfying $P(0) = P$, and $P(O) = P$ for $O \geq \gamma$. In a market with
no transactions costs, things become much simpler since one can prove that it is optimal to trade
vehicles in every period and every consumer $\eta$ will have a most preferred odometer value $O(\eta)$
and will sell their old vehicle after one period of use and buy a replacement vehicle with their pre-
ferred odometer value $O^*(\eta)$ each period. The equilibrium price function $P(O)$ is constructed by
incenting a mass of consumers with the highest preference for vehicle newness to buy brand new
cars, and then the remaining consumers array themselves in a monotonic fashion based on their
preferred odometer value of vehicle to hold, $O^*(\eta) > 0$. So a new car is purchased by one of the
consumers whose $\eta$ value is in the highest category to induce them to buy brand new cars. Then
one period later, when the car’s odometer is some random state $O_1$ (represent a random amount
of usage of the vehicle during the period), the first owner sells the car in the secondhand market
to a consumer $\eta_1$ for which $O^*(\eta_1) = O_1$. Then a period later when the car’s odometer is $O_2$ the
second owner sells the car to a new owner $\eta_2$ for which $O^*(\eta_2) = O_2$ and so forth. Once there is
an owner $\eta_T$ whose optimal odometer is $O^*(\eta_T) < \gamma$ but who drives the car beyond the scrappage
threshold $\gamma$ (so that $O_T > \gamma$), then the car is sold for scrap value $P$, ending its life. If consumers
are homogeneous rather than heterogeneous, then the prices must adjust to make all consumers
indifferent about the odometer value of vehicle they hold: essentially if all consumers prefer newer
cars to older ones, in equilibrium the price depreciation on newer cars is higher to offset the higher
utility they offer consumers, and this generates a decreasing and convex equilibrium price function
$P(O)$. When consumers have quasi-linear preferences for cars, the optimal car type $O^*(\eta)$ turns
out to be the car for which the sum of maintenance/fuel costs plus expected price depreciation
is minimized. With homogeneous consumers, the price depreciation is a decreasing function
of the odometer value so that the higher price depreciation in new cars exactly counterbalances their
lower operating/maintenance costs relative to older cars, making consumers indifferent about the
odometer value since the sum of expected maintenance costs plus expected car price depreciation
is completely flat as a function of the odometer value in equilibrium.

Konishi and Sandfort (2002) proved the existence of a stationary equilibrium in an extension of
Rust’s stationary equilibrium model when there are transactions costs in addition to trading costs.
The distinction between these two costs is that trading costs are simply the net difference between
the price of the new car purchased and the resale price of the old car that is sold. Any additional
costs associated with trading above and beyond these trading costs are referred to as transactions
costs. This can include taxes, or fixed title fees or other payments made to an intermediary to
execute the trade. Rust (1985) showed that when there are no transactions costs, it will be optimal to trade the currently held vehicle in every period and buy the consumer’s optimal car $O^*(\eta)$ as discussed above — at least in a completely stationary environment. However when there are transactions costs in addition to trading costs, it is no longer optimal to trade vehicles every period. Instead, the optimal policy will typically take the form of an $(O, o)$ rule, the analog of the classic $(S, s)$ rule in inventory policy. That is, $O = O^*(\eta)$ corresponds to the optimal odometer that a consumer of type $\eta$ will choose conditional on it being optimal for the consumer to trade, whereas $o = o^*(\eta) > O$ is an optimal selling threshold for consumer of type $\eta$. Thus, after a consumer buys the optimal durable of odometer level $O^*(\eta)$ it will be optimal for the consumer to keep the durable under the current odometer exceeds the optimal sales threshold $o^*(\eta)$, at which point the consumer sells the existing car and repurchases another vehicle at the optimal odometer level $O^*(\eta)$. The challenge to proving the existence of equilibrium (as well as computing it) is to show that there exists an equilibrium price function that causes all of the different type of consumers $\eta$ to line up in the right way so that corresponding to each type of consumer $\eta$ whose optimal selling threshold $o^*(\eta)$ there is some other type of consumer $\eta'$ whose purchase threshold $O^*(\eta') > o^*(\eta)$. In particular, if there are consumers who have a strong preferences for newness and who find it optimal to purchase brand new cars and sell them relatively quickly (i.e. they have a relatively low optimal selling threshold $o^*(\eta)$, then there are other consumer types $\eta'$ who would be willing to buy these relatively new but not completely brand new cars. Further, in equilibrium we have to guarantee that the fraction of consumers of type $\eta'$ who a) are in a position to trade, and b) wish to buy a durable at a given odometer value $O(\eta)$ equals the fraction of consumers who are a) willing to sell, and b) whose cars have odometer values that match the desired odometer value $O^*(\eta)$ for all consumer types $\eta$ who are in the market to purchase a car of this odometer value. Konishi and Sandfort (2002) prove that a stationary equilibrium does exist under certain assumptions, which are naturally more restrictive than the conditions in Rust (1985) for existence of equilibrium when there are zero transactions costs.

Stolyarov (2002) and Gavazza, Lizzeri, and Roketskiy (2014) followed a stationary equilibrium approach similar to Rust (1985b), however they did not use vehicle odometer, but rather vehicle age as the key state variable determining the price of used car. If we also abstract from accidents and multiple different types of vehicles, in a steady state equilibrium there will be a maximum age $\bar{a}$ of the vehicle at which it will be sold for scrap, and the market will be in flow equilibrium which implies that the fraction $x$ of vehicles that are scrapped each period also equals to the fraction of new cars that are produced and sold to the market. It is easy to see that the invariant distribution in this case will be a discrete uniform distribution on the set of ages $\{0, 1, \ldots, \bar{a} - 1\}$ and $x = 1/\bar{a}$. This uniform distribution represents the steady state supply of cars. Equilibrium will be a set of prices $(P(0), P(1), \ldots, P(\bar{a} - 1))$ that equates demand and supply, i.e. it is a set of prices for
which the fraction of consumers who are willing to old age age $a$ also equals $x$. When there are no transactions costs, one can prove a result similar to Rust (1985a) that it is optimal for each individual (with heterogeneous preferences indexed by a one dimensional parameter $\eta$) to trade for their preferred age of car $a^*(\eta)$. The prices must adjust to insure that a fraction $x = 1/\overline{a}$ of consumers wish to hold cars of each age $a$, where $\overline{a}$ is the age under the given equilibrium at which the secondary market price equals the scrap value of the vehicle, $P(\overline{a}) = P$. When there are transactions costs, the trading strategy for a consumer will be given by two thresholds, $a(\eta)$ is the desired age of a replacement vehicle, and this vehicle is kept until it reaches age $\overline{a}(\eta)$ at which point the consumer sells the vehicle on the secondary market (or scraps it if $\overline{a}(\eta) \geq \overline{a}$ where $\overline{a} = \max_{\eta} \overline{a}(\eta)$). An equilibrium when there are transactions costs must cause the fraction of consumers holding each age of vehicle to equal the corresponding fraction of vehicles supplied of that age, $x$.

Stolyarov (2002) and Gavazza, Lizzeri, and Roketskiy (2014) compute equilibrium numerically in calibrated models designed to match data from the US car market. Even though the age distribution of cars is not uniformly distributed in the U.S. (the steady state age distribution predicted by both of their models when there are no accidents) and this is not consistent with their models, there are other aspects of the market for vehicles that their theory does do a better job of predicting. For example figure 1 in Gavazza et. al. shows a distinctly non-uniform distribution of the fraction of cars traded by age of vehicle and argues that their model provides a good overall explanation for this distribution and other features of the US car market.

“The model closely matches the statistics on aggregate trade in cars and activity in secondary markets, most notably the fraction of households acquiring a car and the fraction of cars traded. Figure 1 shows the fractions of used cars traded by age, generated by the model. The plot shows that the relationship between the fraction of used car traded and the age of the car is non-monotonic. These patterns of trade are the focus of Stolyarov (2002). The intuition for the spikes of trade is that, because of transaction costs, only a small fraction of households sells their car after one period, so the resale rates are (locally) increasing when cars are relatively new. Moreover, households are unlikely to purchase old cars that they would scrap after just one period, so the resale rates are (locally) decreasing when cars are old.” (p. 15). However the Gavazza et. al. paper did not compare many other dimensions in which it seems their model cannot match the data well. In particular, as we noted above, the distribution of ages of vehicles is definitely not a uniform distribution as their model presumes, and as we show below, this distribution is not stationary over time: various gas price, macro and other shocks create “waves” in the age distribution of vehicles that slowly evolve in predictable ways as the auto stock ages.

However the stationary equilibrium perspective is not an unreasonable point of departure. Rust (1985a) showed that the continuous equilibrium framework he developed results in a stationary
equilibrium that provides a reasonably good approximation to both the steady state age distribution
and the the distribution of odometer values of vehicles in the U.S. However there is considerable
evidence that the distribution of vehicles is not stationary, especially in the aftermath of significa-
cant oil price shocks, or after recessions. Large increases in oil prices cause consumers to trade
and scrap older fuel-inefficient vehicles and buy newer fuel-efficient ones. This tends to thin and
truncate the upper tail of older vehicles and add more mass to the the younger ages in the age distri-
bution of vehicles. Recessions can have the opposite effect, causing consumers to delay purchases
of new vehicles and continue to hold their existing vehicles or delay scrapping them. This tends
to thin out the distribution at the younger ages of vehicles and thicken and extend the upper tail of
older age vehicles.

Adda and Cooper (2000b) present a micro-aggregated model of auto holdings that accounts for
macro shocks and provides clear evidence that as a result of these shocks and time series variation
in gas prices and other factors, the age distribution of vehicles is definitely not time-stationary.
Figure 3 repeats a figure from their paper, where we see the “wave” in the age distribution slowly
travelling to the right over time. That is, a peak in new car sales in a given year, becomes a peak
in 1 year old cars in the next year, a peak in 2 year old cars two years later, and so forth. Figure 7
suggest that indeed, a lot of the variation in the age distribution of cars over time is due to cyclical
variation in new car sales, and we see substantial variations in sales over time with the troughs in
sales corresponding to historical recession periods.

Figure 7 plots a similar evolution of the age distribution of the vehicle stock, but for Denmark.
We also see similar “waves” in this distribution, which shows that the car market in Denmark also
appears to affected by various shocks that create reasonably big cycles in the number of new cars
purchased, which are then reflected in the waves in the age distributions plotted in the figure.

Adda and Cooper (2000b) also point out “echo effects” on booms in new vehicle sales in one
period of time leading to “echo booms” in scrappages and new car sales as this peak in the age
distribution works its way to the right into the region of the age distribution (approximately at
15 years of age) where there is a high incidence of scrappage. “Using these CDFs, one can also
visualize echo effects. Suppose that in some period, such as 1989 in France, there is a burst of
sales. In subsequent years, as this cars age, there is a bulge in the CDF. Eventually, these cars
reach the optimal (state dependent) scrapping age and there is a new burst of sales. Of course, this
process is tempered by the endogenous and exogenous scrapping that occurs at earlier ages.” (p.
6).

Adda and Cooper (2000b) argue that consumers are more likely to postpone decisions to replace
old vehicles with new ones during recession periods, and these delays in replacement decisions is
a major factor driving the cycles in new car sales. They use a micro-aggregated model of optim-
mal replacement of vehicles which assumes that a secondary market for vehicles do not exist: all
consumers purchase brand new cars and hold the cars until they are scrapped in their model. The model contains a “macro state variable” in addition to idiosyncratic (conditionally independently distributed incomes given the macro shock) that induces correlation in individual agent replacement decisions. In a recession period, the adverse macro shock causes many agents with older vehicles to postpone their decision to scrap their vehicles and buy a new one, and this causes a trough in new car sales. On the other hand in a boom period, a larger than usual number of individuals with older cars choose to scrap them and buy new cars, and this leads to peaks in the time series of sales of new cars. Over time these booms and peaks in new car sales create the appearance of a wave travelling in the distribution of vehicle ages until enough time passes and all the vehicles have “died off” via accidents and scrappage.

We would like to build a model that relaxes the stationarity assumption and allows for the effects of various macro shocks or changes in fuel prices, however to do so, we would ordinarily need to carry the age distribution of vehicles as part of the “state variable” of the model since as we observed in the existing literature, the current age distribution of cars represents the currently available supply of vehicles, and this distribution of supply evolves rather slowly over time as new vehicles are produced, the oldest vehicles are scrapped, some younger vehicles disappear or “die” due to accidents, and the remaining younger vehicles become one year older. However especially when we are considering a model where there are many different types of vehicles, it becomes very problematic to have to carry the entire distribution of cars around as a state variable in individual dynamic programming problems, since this is a high dimensional continuous state variable that would make it infeasible to solve the dynamic programming due to the curse of dimensionality.

Therefore we follow the pragmatic approach of Krusell and Smith (1998) and assume that equilibrium car prices can be well predicted using a much smaller-dimensional set of “sufficient statistics”. In this case we will assume that the sufficient statistics are just the current price of gas and the macro state, \((p, m)\). In Krusell and Smith, they also used the average value of the individual specific savings as an additional part of the sufficient statistic. We could follow this and use the average vehicle age as part of the vector of sufficient statistics to help do a better job of summarizing the current level of “supply” of vehicles. However for the time being, we will start by trying to see if gas prices and the macro state are sufficient to provide good approximations, on the grounds that even relatively severe recessions or large increases in gas prices do not change individual behavior by a sufficient amount to radically alter the age distribution of vehicles in the short run. If we find that restricting the price function to depend only on \((p, m)\) hampers our ability to find an approximate equilibrium, or results in solutions that do not seem to match well the observed evolution of prices and age distributions for vehicles in Denmark, then we may be faced to look at more complicated versions of this model that adds various moments of the multivariate distribution of the vehicle stock (such as including mean vehicle age by type of vehicle), but we
hope to avoid this extra complexity.

We outline the strategy for finding equilibrium prices numerically. Similar to previous papers surveyed, we will be searching for a vector of prices \( P(\tau, a, p, m) \) that approximately equates supply and demand for vehicles for all car types and ages and in all of the relevant states \((p, m)\). How do we do this in practice? Imagine that we start with an initial guess of the equilibrium prices \( P \).

We solve the dynamic programming problem to determine the optimal vehicle holding and trading strategies. Then starting with an initial “configuration” which we denote by \( S_0(\tau, a, p, m) \) which is the total number of cars of each type \( \tau \) and age \( a \) owned by a hypothetical Danish population of simulated individuals at the start of year 0. We can take the actual holdings of vehicles or a scaled version of the age distribution of vehicle holdings to scale to the size of the “economy” we desire to simulate for purposes of calculating an equilibrium (it will be faster to simulate for a smaller number of individuals).

Then given these initial holdings \( S_0 \) and the guess at the equilibrium price \( P \) we simulate the dynamic programming solutions for our simulation sample of household and this result in some consumers scrapping their vehicles and others selling their vehicles to buy other new or used vehicles, and still others selling and not buying a vehicle to replace the one they sold, and others who have vehicle buying a car, and finally another group (which will probably be the largest segment of the population) choosing to keep their existing cars one more year. Once these decisions are made we will get an implied micro-aggregated “excess demand” for vehicles that we can denote by \( ED_0(\tau, a, S_0, P, p_0, m_0) \) where this is the difference between the supply of vehicles for sale of type/age \((\tau, a)\) and the demand for vehicles of this same type/age group. By our assumption of infinitely elastic supply for all new vehicles (since all cars sold in Denmark are imports) the excess demand for new vehicles is zero: \( ED_0(\tau, 0, S_0, P, p_0, m_0) = 0 \) and the price of these vehicles is fixed exogenously outside of our model, i.e. we will treat \( P(\tau, 0, p, m) \) as “data” from the standpoint of solving for an equilibrium of the model.

Similarly, we assume all vehicles that are in the maximum age category \( \bar{a} \) (say \( \bar{a} = 20 \)) are scrapped for a value \( P(\tau, \bar{a}, p, m) \) that we also treat as “data” since it is not determined as part of the equilibrium of the model. Thus, we only need to determine equilibrium prices at ages \( a \in \{1, \ldots, \bar{a} - 1\} \). In our calculation of excess demand, we exclude the fraction of households who choose to keep their cars. So the excess demand for type/age category \((\tau, a)\) is the difference between the number of households who have a car in this group that wish to sell their cars, and the number of households that wish to buy a car of this type/age category. Let \( ED_0(S_0, P, p_0, m_0) \) be the vector of these excess demands for all \( \tau \) and all \( a \in \{1, \ldots, \bar{a} - 1\} \), and let \( ||ED_0(S_0, P, p_0, m_0)|| \) be some norm of this vector of excess demands, such as the sum of absolute values (\( L_1 \) norm), or sum of squared excess demands (\( L_2 \) norm) or the maximum absolute excess demand over all \( \tau \) and \( a \in \{1, \ldots, \bar{a} - 1\} \) (\( L_\infty \) norm). Then we see to minimize the norm of the excess demands. However
note that this norm assumes a particular initial condition $S_0$ and $(p_0,m_0)$ and we are seeking a price function that minimizes the norm of excess demands for all possible initial conditions.

One convenient way to do this is to conduct time series simulations of the model. To do this, we ignore the possibility that the model may not be in equilibrium at time 0, and thus we assume (counterfactually) that the supply in the initial period $t = 0$ equals the micro-aggregated demand, i.e. $D_0(S_0,P_0,p_0,m_0)$ represents the total demand for cars of all types and ages at time $t = 0$ but just after all households have conducted their desired trades to obtain their updated optimal holdings of vehicles. This becomes the new stock of cars under the (counterfactual) hypothesis that supply equals demand at this initial hypothetical equilibrium price vector $P_0$, the vector of prices in $P(\tau,a,p_0,m_0)$. Let $S_1 = \Gamma(D_0(S_0,P_0,p_0,m_0))$ be the “aged” supply of cars at the start of simulation period 1, but just before households have made their decisions/trades to update their vehicle holdings according to their dynamic programming strategies. Essentially this is a deterministic mapping that just involves incrementing the age of every type of car by 1, but allowing for accidents to deplete the supplies of available vehicles. Note that we treat accidents as exogenous stochastic events in the simulation and so there is a small stochastic element in this updating rule and we could write $\hat{S}_1 = \Gamma(D_0(S_0,P_0,p_0,m_0))$ to reflect the randomness in the aged stock of vehicles due to stochastic accidents.

Given $S_1$ at $t = 1$ we then micro-aggregate the optimal holding/trading decisions from the dynamic programming problem, where we will also update the population in our simulated sample to account for “deaths” of older households and “births” (or entry) of new households and the period $t = 1$ excess demand vector $ED_1(S_1,P_1,p_1,m_1)$ reflects both the updating of the the simulated population (including the aging of the households that have not died off by one year) and the realized values of the period $t = 1$ gas price $p_1$ and macro shock $m_1$, and their impact on the prices of vehicles via the posited equilibrium price function $P(\tau,a,p_1,m_1)$. Then continuing in a recursive fashion we assume counterfactually that the period $t = 1$ is in equilibrium and take the supply of vehicles in period $t = 1$ as equal to the number of vehicles demanded $D_1(S_1,P_1,p_1,m_1)$ and then using this we age this stock of vehicles via simulation of deterministic aging and accidents to get $\hat{S}_2 = \Gamma(D_1(S_1,P_1,p_1,m_1))$ where $\Gamma$ denotes this simulated accident/aging operator.

Suppose we run this simulation for a total of $T$ periods, where $T$ is sufficiently large that our simulation has covered “most” of the possible macro states and gas prices (or represents a good random sample from the ergodic distribution of these values). Then we propose to form the average of the norms of the excess demands over this $T$ period simulation to an average excess demand we denote by $\nabla(P)$ given by

$$\nabla(P) = \frac{1}{T} \sum_{t=0}^{T} \|ED_t(S_t,P,p_t,m_t)\|$$

(8)

where $S_{t+1} = \Gamma(D_t(S_t,P,p_t,m_t))$. We propose to search for an approximate equilibrium price
function $P(\tau,a,p,m)$ by minimizing $\nabla(P)$ numerically over the relevant “free” components of $P(\tau,a,p,m)$ (i.e. taking into account our assumption that the price of new cars of all models and scrap prices of all models $\tau$ is exogenously specified outside of the model).

6 Results

We have preliminary results so far, but we have achieved convergence and have been able to replicate the wave patterns with our macroeconomic state variable. We have achieved this with both simulated data and most recently, with a very small subsample of actual Danish register data.

The following two figures show some of our preliminary results graphically. First, Figure 7 shows the equilibrium prices, non-parametrically estimated. These are used in our expected price function. Second, Figure 8 shows the estimated waves of the vehicle stock in the car age distribution over time. This result is very promising, for it indicates that our estimation approach can rationalize key patterns in the observed data.

We anticipate having a more complete set of results within the next month or so.

7 Conclusion

To come soon...

Appendix 1: Derivatives

To write up the derivatives of the utility function, it is useful to first insert optimal driving back into the utility function.

$$u = \gamma^{\text{kt}}(\tau,a,\eta,m) \log(vkt) + \gamma^{\text{mu}}_i(y) \left[ y - \left( \frac{p^{\text{fuel}}}{fe} + c^o \right) vkt - c^T \right] - \gamma^{\text{age}} \sqrt{a}.$$ 

To re-iterate, optimal driving is found from the FOC,

$$\gamma^{\text{kt}}(\tau,a,\eta,m) \frac{1}{vkt} = \frac{\gamma^{\text{mu}}_i(y) p^{km}}{\gamma^{\text{kt}}(\tau,a,\eta,m) \gamma^{\text{mu}}_i(y) p^{km}}.$$ 

23
Note that it is important to insert the expression of $vkt^*$ as a function of parameters for taking the derivatives to take into account that perturbing parameters will change utility directly but also “indirectly” by changing optimal driving.

Inserting this back, the utility becomes,

$$u = vkt(t, a, \eta, m) \log \frac{vkt(t, a, \eta, m)}{\mu(y)p^{km}} - \gamma^{age}\sqrt{a}$$

Now we are ready to consider specific derivatives. Let us first consider one of the coefficients entering into $\gamma^{mu}$ which at this stage is only $\gamma^m$, since $\gamma^{mu} \equiv \gamma^m / \sqrt{y}$.

$$\frac{\partial u}{\partial \gamma^m} = vkt(t, a, \eta, m) \frac{1}{\gamma^{mu}(y)p^{km}} \left\{ - \frac{vkt(t, a, \eta, m)}{p^{km}} \frac{1}{[\gamma^{mu}(y)]^2} \frac{\partial \gamma^{mu}(y)}{\partial \gamma^m} + (y - c^T) \frac{\partial \gamma^{mu}(y)}{\partial \gamma^m} \right\}$$

$$= - \frac{vkt(t, a, \eta, m)}{\gamma^{mu}(y)} \frac{\partial \gamma^{mu}(y)}{\partial \gamma^m} + (y - c^T) \frac{\partial \gamma^{mu}(y)}{\partial \gamma^m}$$

$$= \left[ (y - c^T) - \frac{vkt(t, a, \eta, m)}{\gamma^{mu}(y)} \right] \frac{\partial \gamma^{mu}(y)}{\partial \gamma^m}.$$ 

Finally, note that

$$\frac{\partial \gamma^{mu}(y)}{\partial \gamma^m} = \frac{1}{\sqrt{y}}.$$ 

Now, consider instead one of the coefficients entering into $\gamma^{vkt}$. Specifically, for $\gamma^v \in \{\gamma^{vkt,0}, \gamma^{vkt,1}, \gamma^{vkt,2}, \gamma^{vkt,3}\},$
\[
\frac{\partial u}{\partial \gamma^v} = \frac{\partial \gamma^{kt}(\tau, a, \eta, m)}{\partial \gamma^v} \log \left[ \frac{\gamma^{kt}(\tau, a, \eta, m)}{\gamma^v} \right] + \gamma^{kt}(\tau, a, \eta, m) \left[ \frac{1}{\gamma^{kt}(\tau, a, \eta, m)} \right] \frac{1}{\gamma^v} - \frac{\partial \gamma^{kt}(\tau, a, \eta, m)}{\partial \gamma^v} \frac{1}{\gamma^{kt}(\tau, a, \eta, m)} \left[ \frac{1}{\gamma^v} - 1 \right] \left\{ \log \left[ \frac{\gamma^{kt}(\tau, a, \eta, m)}{\gamma^v} \right] + 1 - 1 \right\} \] 

More conveniently, we can write this as

\[
\frac{\partial u}{\partial \gamma^v} = \log(vkt^*) \frac{\partial \gamma^{kt}(\tau, a, \eta, m)}{\partial \gamma^v}.
\]

Therefore, since \(\gamma^{kt}(\tau, a, \eta, m) \equiv \gamma^{kt,0} + \gamma^{kt,1}a + \gamma^{kt,2}a^2 + \gamma^{kt,3}m + \eta + u^{kt}\), we have that

\[
\frac{\partial u}{\partial \gamma^{kt,0}} = \log(vkt^*),
\]

\[
\frac{\partial u}{\partial \gamma^{kt,1}} = \log(vkt^*)a,
\]

\[
\frac{\partial u}{\partial \gamma^{kt,2}} = \log(vkt^*)a^2,
\]

\[
\frac{\partial u}{\partial \gamma^{kt,3}} = \log(vkt^*)m.
\]

These derivatives are used in the code where the analytical derivatives are calculated.
References


Figure 1: Fuel economy of new vehicles in Denmark over time.

Figure 2: Time series of vans in Denmark shows the effect of the 2007 reform.
Figure 3: Evolution of the Age Distribution of Cars in the United States

Figure 1: Sales of New Cars, France and US

Figure 4: Time Series of New Vehicle Sales in US and France
Figure 5: Time Series of New Vehicle Sales in US, 1951 to 2012

Figure 6: Evolution of the Age Distribution of Cars in Denmark
Figure 7: Equilibrium prices, non-parametrically estimated.

Figure 8: Simulated waves of the vehicle stock over time.