CIRRELT

Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation

A Dynamic Discrete-Continuous Choice Model of Car Ownership, Usage and Fuel Type

Aurélie Glerum
Oskar Blom Västberg
Emma Frejinger
Anders Karlström
Muriel Beser Hugosson
Michel Bierlaire

June 2015

CIRRELT-2015-23

Bureaux de Montréal :
Université de Montréal
Pavillon André-Aisenstadt
C.P. 6128, succursale Centre-ville
Montréal (Québec)
Canada H3C 3J7
Téléphone : 514 343-7575
Télécopie : 514 343-7121

Bureaux de Québec :
Université Laval
Pavillon Palasis-Prince
2325, de la Terrasse, bureau 2642
Québec (Québec)
Canada G1V 0A6
Téléphone : 418 656-2073
Télécopie : 418 656-2624

www.cirrelt.ca

A Dynamic Discrete-Continuous Choice Model of Car Ownership, Usage and Fuel Type

Aurélie Glerum[1], Oskar Blom Västberg[2], Emma Frejinger[3,*], Anders Karlström[2], Muriel Beser Hugosson[2], Michel Bierlaire[1]

[1] École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory (TRANSP-OR)

[2] Royal Institute of Technology (KTH), School of Architecture and the Built Environment, Department of Transport Science, SE-100 44 Stockholm Sweden

[3] Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT) and Department of Computer Science and Operations Research, Université de Montréal, P.O. Box 6128, Station Centre-Ville, Montréal, Canada H3C 3J7

Abstract. This paper presents a dynamic discrete-continuous choice model (DDCCM) of car ownership, usage, and fuel type that embeds a discrete-continuous choice model (DCCM) into a dynamic programming (DP) framework to account for the forward looking behavior of households in the context of car acquisition. More specifically, we model the transaction type, the choice of fuel type, and the annual driving distance for up to two cars in the household. We present estimation and cross validation results based a subsample of the Swedish population that is obtained from combining the population and car registers. Finally we apply the model to analyze a hypothetical policy; of a subsidy that reduces the annual cost of diesel cars.

Keywords: Dynamic discrete-continuous choice model, dynamic programming, dynamic discrete choice models, car ownership and usage, constant elasticity of substitution.

---

* Corresponding author: Emma.Frejinger@cirrelt.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2015

1 Introduction

Numerous governments have in the past implemented policies aiming at influencing the composition of the car fleet, motivated by goals of reducing green house gas emissions. In this context, quantitative models play an important role in understanding and predicting the changes in demand in response to policy changes. The literature on car related choice models is vast, but there appears to be an agreement that it is important to consider car ownership (number of cars) and car usage (distance driven with each car) simultaneously at the household level. For instance, increased fuel efficiency has a direct effect of reduced green house gas emissions, but at the same time usage of the car may become less costly which may increase driving distance, resulting in increased emissions. Another aspect is that a car is a costly and durable good that can be used over a long period of time and sold at a second hand market. Fuel cost expectations and the dynamics of replacing cars are important aspects of household car ownership decisions. In this paper we present a dynamic discrete choice model (DDCM) (Rust, 1987; Aguirregabiria and Mira, 2010), that accounts for the forward looking behaviour of households when making car transactions (increasing/decreasing the household fleet or changing its composition).

The nature of a discrete choice (car ownership and fuel type) and continuous choice (car usage) is a defining feature of our proposed model. In this paper we specify a *dynamic discrete-continuous choice model (DDCCM)* that jointly models car replacement decisions, choice of fuel type and usage of each car within a household. To the best of our knowledge, this is the first paper that brings together the two methodologies of discrete-continuous choice and dynamic discrete choice models, to jointly model a household's decisions regarding ownership and usage of up to two cars. As we explain in the following section, this is not the first dynamic discrete choice model for car ownership. Existing models are however limited to owning at most one car which is not realistic in the Swedish market where a large share of the households own two cars (the share of three car households is however very small and is neglected in this model).

In this context, one of the main issues is how to model the continuous choice variables capturing the annual driving distances for each car in a household. Depending on variations in the fuel prices, households owning both a gasoline and a diesel car can choose use one car more than the other. In this paper we address this important aspect with a constant elasticity of substitution (CES) utility function.

This paper contributes to the literature by proposing a DDCCM model that models households' forward looking behaviour related to car transaction decisions and choice of fuel type. We model the annual driving distance as a myopic choice with a CES utility function that can capture the substitution between the cars for two-car households. Moreover, we present estimation and cross-validation results based on a combination of the Swedish population and car registers.

The paper is structured as follows. Section 2 reviews the literature on dynamic car ownership models and joint models of car type and usage and Section 3 presents the methodological framework of the DDCCM. The empirical results are presented in Section 4, and finally, Section 5 concludes.

2 Literature review

A comprehensive review of the vast literature related to car ownership, type and usage choices is out of the scope of this paper (for reviews we refer the reader to e.g. de Jong et al., 2004; de Jong and Kitamura, 2009). Even though many sophisticated studies have been published on static models, we focus here on the fairly scarce literature directly related to our work, which deals with dynamic models taking into account the forward-looking behavior of decision-makers. One of the most notable studies is Schiraldi (2011). It focuses on the estimation of transaction costs in a dynamic framework based on aggregate data and analyzes the effect of a scrappage policy in Italy. Moreover Schiraldi (2011) models the price on the second-hand market and provides an excellent review of related literature. Similarly to Schiraldi (2011) we assume that a decision-maker maximizes the expected discounted lifetime utility modeled by a value function that is the solution to the Bellman equation. Moreover we make the same assumptions (actually dating back to Rust, 1987) to deal with the 'curse of dimensionality' and obtain an operational model. As opposed to Schiraldi (2011) where a consumer holding a car a given year decides whether to hold, sell or scrap the car (if the consumer does not hold a car he/she decides to continue that way or to buy one), we have a more complex choice setting because a household can hold more than one car and we also model the usage of each car. Unlike Schiraldi (2011) who allows for endogenous price in second-hand market, we propose model for the demand side hence assuming that supply variables are exogenous.

Related to our work is also the one by Xu (2011) who develops a dynamic discrete choice model to explain car acquisition decisions and choice of fuel type. Car usage is however not considered so there are no continuous choice variables. The model is applied to stated preferences data collected in Maryland and corresponds to a optimal stopping model similar to the one by Rust (1987). de Lapparent and Cernicchiaro (2012) and Cernicchiaro and de Lapparent (2015) present a DDCM to explain choices regarding car acquisition and usage. In their research. Unlike in our model, they treat the cars in a two car-household as independent and they have a simpler structure of the action space. These studies highlight the benefits of dynamic models over traditional static ones: de Lapparent and Cernicchiaro (2012) and Cernicchiaro and de Lapparent (2015) show that DDCMs have a better in-sample fit than static discrete choice models and Xu (2011) shows that they outperform static models in terms of recovering market trends.

Other modeling approaches have been considered in order to jointly model car ownership and usage. For instance, a very interesting study is presented by Gillingham (2012) who models cars' monthly mileage conditional on vehicle type. He integrates consumers' expectations about the cars' future resale prices and future gasoline prices after a six-year period.

It is in particular important to mention an ongoing research project at University of Copenhagen (Kenneth Gillingham, Fedor Iskhakov, Anders Munk-Nielsen, John Rust and Bertel Schjerning) with an objective similar to ours, namely a discrete-continuous dynamic choice model for transaction decisions and usage. Their presentation at the IRUC seminar (Copenhagen, December 2012) inspired us to view the continuous choice variable of car usage as a myopic choice conditional on the discrete choice variables. To the best of our knowledge of this so-far unpublished work, the models have important differences because they consider the supply side (endogenous prices) but limits the households can own at most one car. The latter is a reasonable assumption in the Danish market but not the Swedish one.

There is also literature on duration models that model the time elapsed between two car transactions. For instance, de Jong (1996) presents an interesting study based on a system of models including a duration model for the time between car replacement decisions and a regression model for annual car usage. The main difference between the dynamic model presented here and a duration model is that households are assumed to be forward-looking. This means that they optimize their choices taking expected future utility into account. Moreover, socio-economic characteristics are not assumed constant between transactions and we can model several choices jointly.

3 The dynamic discrete-continuous choice modeling framework

In this section we present the DDCCM framework. We start by stating the main assumptions on which the model is based. Then we describe the model structure, from the base components to the specification of the full model. One of the key elements of the choice variable is the annual mileage of each car and we explain in detail its specification. We end the section by discussing maximum likelihood estimation of the model.

3.1 Main assumptions

The DDCCM is formulated as a discrete-continuous choice model that is embedded into a dynamic programming (DP) framework. We model the joint decision of vehicle transactions, mileage and fuel type, based on the following assumptions.

Decisions are made at a household level. In addition, we assume that each household can have at most two cars. Larger household fleets may also be considered but at the cost of increased complexity. As pointed out by de Jong and Kitamura (2009), it may be relevant to consider three car households for prediction even though the current share in several markets (typically European markets) is low.

The choice of vehicle transaction and fuel type(s) is *strategic*, that is, we assume that households take into account the future utility of the choice of these variables in their decision process.

We consider an infinite-horizon problem to account for the fact that households make long-term decisions in terms of car transactions and fuel type. For example, individuals are assumed to strategically choose the fuel type of the car they purchase according to their expectation of fuel prices in the next years, or they decide to purchase only one car at present knowing that they might add another car in the future years.

We make the simplifying assumption that when households decide how much they will drive their car for the upcoming year, they only consider the utility of this choice for that particular year without accounting for whether the residual value of their car is affected by usage. In other words, the choice of mileage(s) is *myopic*, that is, households do not take into account the future utility of the choice of the current annual driving distance(s) in their decision process.

Similarly to de Jong (1996) we make the reasonable assumption that the choice of mileage(s) is conditional on the choice of the discrete decision variables (i.e. the transaction type and the fuel type).

3.2 Definition of model components

The DP framework is based on four fundamental elements: the *state space*, the *action space*, the *transition function* and the *instantaneous utility*. In this section, we describe each of these in detail.

The *state space* S is constructed based on the following variables.

- The *age* y_{ctn} of car c of household n in year t. We set an upper bound for the age $\bar{Y}$, assuming that above this upper bound, changes in age do not affect the utility or transition from one state to another. This implies that we have $y_{ctn} \in Y = \{0, 1, \ldots, \bar{Y}\}$.
- The *fuel type* f_{ctn} of car c of household n in year t. A car c can have any fuel type $f_{ctn} \in F = \{0, 1, \ldots, \bar{F}\}$, where $1, \ldots, \bar{F}$ is the list of available fuel types in the market of interest. The level 0 indicates the absence of a car.

As described in Section 3.1, each household can have at most two cars. Each state $s_{tn} \in S$ can hence be represented as

$$s_{tn} = (y_{1tn}, f_{1tn}, y_{2tn}, f_{2tn}), \tag{1}$$

where the car denoted by the index 1 is the car which has been in a household n's fleet for the longest time, and the car denoted by the index 2 is the car which entered the household in a later stage.

For households who have access to company cars, the size of the state space can be computed as

$$
\begin{aligned}
|S| &= (|Y| \times (|F|-1)+1)^2 && (2)\\
&+ (|Y| \times (|F|-1)+1) && (3)\\
&+ 1. && (4)
\end{aligned}
$$

The first term (2) consists of the number of possible states for two-car households (the exponent 2 stands for the two cars in the household). The second term (3) is the number of possible states for one-car households and the last term (4) represents the absence of cars in a household. It is important to keep the size as low as possible since we need to solve the DP problem repeatedly when estimating the model parameters. To show that the definition of the above state space can be small, we provide a small numerical example; assuming that cars can be at maximum 9 years old and that the market is composed of gasoline and diesel cars only, the size of the state space reaches the reasonable size of 463.

The *action space* A is constructed based on the following variables.

- The *transaction* $h_{tn} \in H$ in household n's composition of the car fleet in year t. Every year, the household can choose to increase, decrease or replace all or part of the fleet, or do nothing. We additionally make the simplifying assumption that a household cannot purchase more than one car per time period. The enumeration (see Figure 1) leads to nine possible transactions.
- The *annual mileage* $\tilde{m}_{ctn} \in \mathbb{R}^+$ of each car c chosen by household n.
- The *fuel type* $\tilde{f}_{ctn} \in F$ of each car c chosen by household n.

Each action $a_{tn} \in A$ can be represented as

$$a_{tn} = (h_{tn}, \tilde{m}_{1tn}, \tilde{f}_{1tn}, \tilde{m}_{2tn}, \tilde{f}_{2tn}). \qquad (5)$$

It is worth noting that we have a completely discrete state space, while the action space is discrete-continuous. Moreover, all actions are not available from all states. Hence, we have $a_{tn} \in A(s_{tn})$ and the total number of discrete actions are obtained by enumerating all possible actions from each particular state. Table 1 summarizes the number of discrete actions that can be attained for households with 0, 1 or 2 cars, depending on the type of transaction which is chosen. To give an example, we assume that households can choose between two fuel types (gasoline and diesel). In this case, a 1-car household that decides to increase the fleet of 1 car has the choice between 2 possible actions, i.e. a gasoline or a diesel engine. In the row 'Sum', the total number of possible discrete actions for households with respectively 0, 1 or 2 cars are reported.

Given that a household n is in a state s_{tn} and has chosen an action a_{tn}, the *transition function* $f(s_{t+1,n}|s_t, a_{tn})$ is defines the probability of ending up in next state $s_{t+1,n}$. In our case the transition probability is assumed to be degenerate.

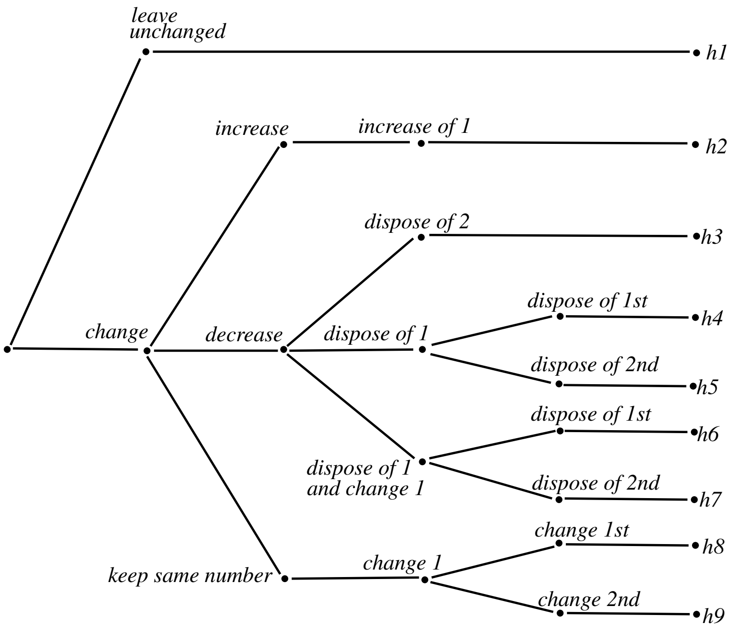

Figure 1: The nine possible transactions in a household fleet

Assuming that $a_{tn}^D = (h_{tn}, \tilde{f}_{1tn}, \tilde{f}_{2tn})$ gathers the discrete components of an action a_{tn} and $a_{tn}^C = (\tilde{m}_{1tn}, \tilde{m}_{2tn})$ gathers the continuous components, the *instantaneous utility* is defined as

$$u(s_{tn}, a_{tn}^C, a_{tn}^D, x_{tn}, \theta) = v(s_{tn}, a_{tn}^C, a_{tn}^D, x_{tn}, \varepsilon_C(a_{tn}^C), \theta) + \varepsilon_D(a_{tn}^D), \tag{6}$$

where variable x_{tn} contains socio-economic information relative to the household, θ is a vector of parameters to be estimated. Expression $v(s_{tn}, a_{tn}^C, a_{tn}^D, x_{tn}, \varepsilon_C(a_{tn}^C), \theta)$ is a deterministic term, $\varepsilon_D(a_{tn}^D)$ and $\varepsilon_C(a_{tn}^C)$ are the random error term for the discrete and continuous actions, respectively. Similarly as proposed by Rust (1987), the instantaneous utility has an additive-separable form. We note that we include $\varepsilon_C(a_{tn}^C)$ in the general description of the model but we remove it in Section 3.4 because we assume deterministic utility for the continuous choice in this study.

Transaction name	0 car	1 car	2 cars
h_1: leave unchanged	1	1	1
h_2: increase 1	2	2	-
h_3: dispose 2	-	-	1
h_4: dispose 1st	-	1	1
h_5: dispose 2nd	-	-	1
h_6: dispose 1st and change 2nd	-	-	2
h_7: dispose 2nd and change 1st	-	-	2
h_8: change 1st	-	2	2
h_9: change 2nd	-	-	2
Sum	3	6	12

Table 1: Number of possible actions for households with 0, 1 or 2 cars (in the action space generated by the discrete components of the choice variable).

3.3 Value function

As in a DDCM case (see e.g. Aguirregabiria and Mira, 2010), the value function of the DDCCM is defined as

$$V(s_{tn},x_{tn},\theta)=\max_{a_{tn}\in A}\{u(s_{tn},a_{tn},x_{tn},\theta)+\beta\sum_{s_{t+1,n}\in S}\bar{V}(s_{t+1,n},x_{t+1,n},\theta)f(s_{t+1,n}|s_{tn},a_{tn})\}$$
$$=\max_{a_{tn}\in A}\{v(s_{tn},a^C_{tn},a^D_{tn},x_{tn},\varepsilon_C(a^C_{tn}),\theta)+\varepsilon_D(a^D_{tn})$$
$$+\beta\sum_{s_{t+1,n}\in S}\bar{V}(s_{t+1,n},x_{t+1,n},\theta)f(s_{t+1,n}|s_{tn},a_{tn})\} \quad (7)$$

In order to obtain a version of the Bellman equation that does not depend on the random utility error term $\varepsilon_D(a^D_{tn})$, we consider the *integrated value function* $\bar{V}(s_{tn},x_{tn},\theta)$, given as follows.

$$\bar{V}(s_{tn},x_{tn},\theta)=\int V(s_{tn},x_{tn},\theta)dG_{\varepsilon_D}(\varepsilon_D(a^D_{tn})) \quad (8)$$

where G_{ε_D} is the CDF of ε_D.

In the case where all actions are discrete and the random terms $\varepsilon_D(a^D_{tn})$ are i.i.d. extreme value, it corresponds to the logsum (see e.g. Aguirregabiria and Mira, 2010). We aim at finding a closed-form formula in the case where the choices are both discrete and continuous. In fact, a closed-form formula is possible in the special case where the choice of mileage of each car in the household is assumed myopic. This implies that individuals choose how much they wish to drive their car(s) every year, without accounting for the expected discounted utility of this choice for the following years.

Under the hypothesis of myopicity of the choice of annual driving distance(s), the integrated value function is obtained as follows.

$$\begin{aligned}
\bar{V}(s_{tn},x_{tn},\theta) &= \int V(s_{tn},x_{tn},\theta)dG_{\varepsilon}(\varepsilon_D(a^D_{tn})) \\
&= \int \max_{a_{tn}\in A}\{u(s_{tn},a_{tn},x_{tn},\theta,\varepsilon(a_{tn}))+ \\
&\quad \beta \sum_{s_{t+1,n}\in S} \bar{V}(s_{t+1,n},x_{t+1,n},\theta)f(s_{t+1,n}|s_{tn},a_{tn})\}dG_{\varepsilon}(\varepsilon_D(a^D_{tn})) \\
&= \int \max_{a^D_{tn}}\{\max_{a^C_{tn}}\{v(s_{tn},a^C_{tn},a^D_{tn},x_{tn},\varepsilon_C(a^C_{tn}),\theta)\}+\varepsilon_D(a^D_{tn}) \\
&+ \beta \sum_{s_{t+1,n}\in S} \bar{V}(s_{t+1,n},x_{t+1,n},\theta)f(s_{t+1,n}|s_{tn},a^D_{tn})\}dG_{\varepsilon}(\varepsilon_D(a^D_{tn})) \\
&= \log\sum_{a^D_{tn}}\exp\{\max_{a^C_{tn}}\{v(s_{tn},a^C_{tn},a^D_{tn},x_{tn},\varepsilon_C(a^C_{tn}),\theta)\} \\
&+ \beta \sum_{s_{t+1,n}\in S} \bar{V}(s_{t+1,n},x_{t+1,n},\theta)f(s_{t+1,n}|s_{tn},a^D_{tn})\} \qquad (9)
\end{aligned}$$

Similarly as in the case of a DDCM, the value function can be solved by iterating on (9).

Since we model both acquisition and usage, we assume that $v(\cdot)$ of (9) is the sum of a utility linked with the acquisition the vehicles v^D_{tn} and a utility linked with the usage of the car v^C_{tn}

$$v(s_{tn},a^C_{tn},a^D_{tn},x_{tn},\varepsilon_C(a^C_{tn}),\theta) = v^D_{tn}(s_{tn},a^D_{tn},x_{tn},\theta)+v^C_{tn}(s_{tn},a^D_{tn},a^C_{tn},x_{tn},\varepsilon_C(a^C_{tn}),\theta) \qquad (10)$$

thus decoposing the utility specification into a discrete choice and a continiuous choice component. We will in the following consider these in turn.

3.4 Utility specification of the continuous choice component

3.4.1 Optimal mileage for households owning two cars with different fuel types

By assumption, each household can have at most two cars. This implies that for two-car households, the annual mileage of each car must be decided every year. That is, $v(s_{tn},a^C_{tn},a^D_{tn},x_{tn},\varepsilon_C(a^C_{tn}),\theta)$ of (9) is maximized with respect to the two annual driving distances. Given the additive form of (10), we only need to maximize expression $v^C_{tn}(s_{tn},a^D_{tn},a^C_{tn},x_{tn},\varepsilon_C(a^C_{tn}),\theta)$ with respect to a^C_{tn}. For the sake of simplicity, we assume in this model formulation that there is no random term related to the continuous choice variable and we omit $\varepsilon_C(a^C_{tn})$.

If a household owns two cars, one car is generally driven more than the other one, i.e. one is used for long distances while the other is used for shorter trips.

We therefore make the assumption that the households do not choose how much to drive each car independently, but rather the repartition of the total mileage that it plans to drive across the two cars. Moreover, the use of both cars in the household is highly dependent on fuel prices. Hence in the common case of a two-car household owning both a car of a fuel f_1 and a car of a fuel f_2 (e.g. a diesel and a gasoline car), the repartition of mileages might fluctuate depending on this economic feature.

In the abovementioned case, this motivates the use of a CES utility function for the choice of mileages for cars of different fuel types within a same household, since it allows to evaluate how likely households substitute the use of one car with the other, when the difference between the fuel prices is changing.

Let us denote the mileages of the chosen cars with fuels f_1 and f_2 as $\tilde{m}_{f_1tn}$ and $\tilde{m}_{f_2tn}$, respectively. They are defined as follows.

$$\tilde{m}_{f_1tn} := \tilde{m}_{1tn} \cdot I_{1f_1tn} + \tilde{m}_{2tn} \cdot I_{2f_1tn} \tag{11}$$

and

$$\tilde{m}_{f_2tn} := \tilde{m}_{1tn} \cdot I_{1f_2tn} + \tilde{m}_{2tn} \cdot I_{2f_2tn}, \tag{12}$$

where I_{cf_1tn} is equal to c if car $c \in \{1,2\}$ is a car driven with fuel f_1 (e.g. gasoline), 0 otherwise, and $I_{cf_2tn} = 1 - I_{cf_2tn}$ is an indicator of whether the car c is driven with fuel f_2 (e.g. diesel).

The deterministic utility of driving is given by the following CES function

$$v^C_{tn}(s_{tn}, a^D_{tn}, a^C_{tn}, x_{tn}, , \theta) = \theta_v(\tilde{m}^\rho_{f_1tn} + \tilde{m}^\rho_{f_2tn})^{\frac{1}{\rho}}. \tag{13}$$

Parameter ρ with $\rho \leq 1$ and $\rho \neq 0$, is related to the elasticity of substitution σ between the two cars, given by

$$\sigma = \frac{1}{1-\rho} \tag{14}$$

and the parameter θ_v with $\theta_v \geq 0$ sets the scale of the utility.

The choice of $\tilde{m}_{f_1tn}$ and $\tilde{m}_{f_2tn}$ must be made such that the budget constraint of the household holds

$$p_{f_1tn}\tilde{m}_{f_1tn} + p_{f_2tn}\tilde{m}_{f_2tn} = \mathrm{Inc}_{tn}, \tag{15}$$

where $p_{ftn} := \mathrm{cons}_{ftn} \cdot \mathrm{pl}_{ft}$ is the cost per km of driving a car with fuel $\tilde{f} \in \{f_1, f_2\}$, that is the product of the car consumption cons_{ftn} and the price of a liter of fuel pl_{ft} for that car. Variable Inc_{tn} is the (scaled) share of the household's annual income which is used for expenses related to car fueling. It is based on a fixed budget share of 8 % of the household disposable income[1], adjusted with a tax subsidy available in the presence of diesel cars

$$\mathrm{Inc}_{tn} = \frac{0.08 \cdot \mathrm{Disp\ Inc}_{tn} + N_{\mathrm{diesel},t,n}\theta_{CESdiesel} \cdot 1\,000}{100\,000} \tag{16}$$

[1]This value is obtained from the population and car register data.

where $N_{\text{diesel},t,n} \in \{0,1,2\}$ is the number of diesel cars owned by household n at time t.

The optimal value of mileages for both cars is obtained by solving the following maximization problem

$$\max_{\tilde{m}_{f_1tn},\tilde{m}_{f_2tn}} v^C_{tn}, \text{ such that } p_{f_1tn}\tilde{m}_{f_1tn} + p_{f_2tn}\tilde{m}_{f_2tn} = \text{Inc}_{tn} \tag{17}$$

The above formulation of the CES utility fonction with the budget constraint has the following advantages. First, the constraint enables us to solve the maximization problem according to one dimension only. Such an approach has been considered by Zabalza (1983), in a context of trade-off between leisure and income. Second, the use of a CES function is also convenient because the elasticity of substitution is directly obtained from the estimate of parameter ρ.

The following analytical solution for $\tilde{m}_{f_2tn}$ can be obtained

$$\tilde{m}^*_{f_2tn} = \frac{\text{Inc}_{tn} \cdot p_{f_2tn}^{(1/(1-\rho))}}{p_{f_1tn}^{(\rho/(\rho-1))} + p_{f_2tn}^{(\rho/(1-\rho))}} \tag{18}$$

which allows us to infer the value of the optimal mileage for the car with the other fuel type

$$\begin{aligned}\tilde{m}^*_{f_1tn} &= \frac{\text{Inc}_{tn}}{p_{f_1tn}} - \frac{p_{f_2tn}}{p_{f_1tn}}\tilde{m}^*_{f_2tn} \\ &= \frac{\text{Inc}_{tn}}{p_{f_1tn}} - \frac{p_{f_1tn}}{p_{f_1tn}} \cdot \frac{\text{Inc}_{tn} \cdot p_{f_2tn}^{(1/(1-\rho))}}{p_{f_1tn}^{(\rho/(\rho-1))} + p_{f_2tn}^{(\rho/(1-\rho))}} \end{aligned} \tag{19}$$

Consequently, we obtain the optimal value for the deterministic utility of the continuous actions:

$$v^{C*}_{tn} = \theta_v \left(\left(\frac{\text{Inc}_{tn}}{p_{f_1tn}} - \frac{p_{f_2tn}}{p_{f_1tn}} \cdot \frac{p_{f_2tn}^{(1/(\rho-1))} \cdot \text{Inc}_{tn}}{p_{f_1tn}^{(\rho/(\rho-1))} + p_{f_2tn}^{(\rho/(\rho-1))}} \right)^{\rho} + \left(\frac{\text{Inc}_{tn} \cdot p_{f_2tn}^{(1/(\rho-1))}}{p_{f_1tn}^{(\rho/(\rho-1))} + p_{f_2tn}^{(\rho/(\rho-1))}} \right)^{\rho} \right)^{\frac{1}{\rho}} \tag{20}$$

Then v^{C*}_{tn} can be inserted back in (10) and the Bellman equation (9) becomes

$$\begin{aligned}\bar{V}(s_{tn},x_{tn},\theta) &= \log\sum_{a^D_{tn}} \exp\{v^D_{tn}(s_{tn},a^D_{tn},x_{tn},\theta) + v^{C*}_{tn}(s_{tn},a^D_{tn},a^{C*}_{tn},x_{tn},\theta) \\ &+ \beta \sum_{s_{t+1,n}\in S} \bar{V}(s_{t+1,n},x_{t+1,n},\theta) f(s_{t+1,n}|s_{tn},a^D_{tn})\}, \end{aligned} \tag{21}$$

where $a^{C*}_t = (\tilde{m}^*_{1,t},\tilde{m}^*_{2,t})$. The integrated value function $\bar{V}$ can then be computed by value iteration. Let us note that the optimal mileage(s) for car 1 and car 2 are obtained by the following mappings.

$$\tilde{m}^*_{1tn} = \tilde{m}^*_{f_1tn} \cdot I_{1f_1tn} + \tilde{m}^*_{f_2tn} \cdot I_{1f_2tn} \tag{22}$$

$$\tilde{m}^*_{2tn} = \tilde{m}^*_{f_1tn} \cdot I_{2f_1tn} + \tilde{m}^*_{f_2tn} \cdot I_{2f_2tn} \tag{23}$$

3.4.2 Optimal mileage for households with two cars of the same fuel

In the case when both cars in a household have the same fuel type, (19) and (18) give

$$\tilde{m}^*_{1tn} = \tilde{m}^*_{2tn} = \frac{\text{Inc}_{tn}}{2 \cdot p_{f_1 tn}}. \tag{24}$$

We assume that the parameter related to the elasticity between the two cars is different for this case and denote it by ρ_s and the optimal utility of driving v^{C*}_{tn} is

$$v^{C*}_{tn} = \theta_v ({\tilde{m}^*_{1tn}}^{\rho_s} + {\tilde{m}^*_{2tn}}^{\rho_s})^{1/\rho_s} = \theta_v 2^{1/\rho_s} \frac{\text{Inc}_{tn}}{2 \cdot p_{f_1 tn}}. \tag{25}$$

In the special case when $\rho_s = 1$, i.e., if there were perfect substition between the two cars, any combination of milage that satisfies (15) is an optimal solution to the optimization problem (17), but the optimal utility is still be given by (25).

3.4.3 Optimal mileage for one-car householdsl

For one-car households with a car of fuel f_1, the optimization problem reduces

$$\max_{\tilde{m}_{f_1 tn}} v^C_{tn}, \text{ such that } p_{f_1 tn} \tilde{m}_{f_1 tn} = \text{Inc}_{tn}. \tag{26}$$

The optimal mileage $\tilde{m}^*_{1tn}$ for the only car in the household is hence given by

$$\tilde{m}^*_{1tn} = \tilde{m}^*_{f_1 tn} = \frac{\text{Inc}_{tn}}{p_{f_1 tn}} \tag{27}$$

and consequently the optimal utility of driving is

$$v^{C*}_{tn} = \theta_v \tilde{m}^*_{1tn}. \tag{28}$$

3.5 Utility specification of the discrete choice component

We now turn our attention to the discrete choices and we present the deterministic part of the instantaneous utilities. The utility is divided into the CES utility function, transaction costs for buying, changing and disposing of cars, and ownership costs.

CES utility functions

The parameters of the CES utility functions are

ρ in CES function when owning two cars of different fuel

θ_v CES scale when two cars with different fuel or one car

$\theta_{CESdiesel}$ Parameter for additional cost associated with having a diesel car. Multiplied with 1000 and the number of diesel cars and subtracted from the amount of money available for fuel in the CES utility

θ_0 Parameter for CES utility when having two cars with the same fuel

As estimating the substitution parameter ρ is rather complicated due to the highly non-linear form of the utility, we define $\theta_0 = \theta_v 2^{1/\rho_s}/2$ rather than estimating ρ_s in (25). This means that the substitution parameter for two cars with the same fuel is given by:

$$\rho_s = \frac{\log(2)}{\log(2\theta_0) - \log(\theta_v)}. \tag{29}$$

Transaction costs

We assume that the transaction cost is different for different actions and in the following we present the specification for (i) disposing, (ii) buying and (iii) changing a car.

Disposing of a car comes with a fixed transaction cost θ_2 and an age dependent transaction cost θ_1

θ_1 Transaction cost for disposing or changing a car, dependent on 1/age of the car.

θ_2 Constant transaction cost of disposing or changing a car in any state.

$$\begin{aligned} u_{dispose}(s_{tn}) =& \theta_1 (d_1/y_{1tn} + d_2/y_{1tn}) \\ &+ (d_1 + d_2)\theta_2 \end{aligned} \tag{30}$$

where $d_1 = 1$ if car 1 is disposed or changed and $d_2 = 1$ if car 2 is disposed or changed.

Buying a car comes with a constant transaction cost as well as utilities dependent on the fuel type of the new car.

θ_3 Utility of buying car (transaction cost). Transaction $h2$, $h6$, $h7$, $h8$ and $h9$.

θ_4 Constant utility for buying second car with different fuel from the first car.

θ_5 Constant utility for buying second car with the same fuel from the first car.

θ_{10} Constant for buying a diesel car

$$
\begin{aligned}
u_{buy} = & \theta_3 + \theta_4 \cdot \text{buy second car with different fuel from first} \\
& + \theta_5 \cdot \text{buy second car with same fuel as first} \\
& + \theta_{10} \cdot \text{buy diesel car}
\end{aligned} \tag{31}
$$

Finally, we assume that changing a car incurs an additional transaction cost given by a constant

θ_6 Utility of changing car (disposing and selling a car during same year). Transaction $h2$, $h6$, $h7$, $h8$ and $h9$.

$$
u_{change} = \theta_6 \cdot \text{change one car} \tag{32}
$$

Utility of owning/not owning a car

The utility of not owning any car is, as reference, fixed to zero. The cost for owning one or two cars are given by

θ_7 Constant for owning two cars with different fuel

θ_8 Cost of owning a car, dependent on log of age of the car

θ_9 Dummy for keeping a car of age 5 or larger.

$$
\begin{aligned}
u_{\text{own two cars}} = & \theta_8(\log(y_{1tn}+1)+\log(y_{1tn}+1))+ \\
& \theta_7(\text{own two cars of different fuel})+ \\
& \theta_9(k_1 \cdot (y_{1tn} \geq 5) + k_2(y_{2tn} \geq 5))
\end{aligned} \tag{33}
$$

where $k_1 = 1$ if car 1 is kept and k_2 is one if car 2 is kept.

3.6 Maximum likelihood estimation and validation

The parameters of the DDCCM are obtained by maximizing the log of the likelihood function

$$
\mathscr{L}(\theta) = \prod_{n=1}^{N} \prod_{t=1}^{T_n} P(a_{tn}^D | s_{tn}, x_{tn}, \theta), \tag{34}
$$

where N is the total population size, T_n is the number of years household n is observed and $P(a_{tn}^D | s_{tn}, x_{tn}, \theta)$ is the probability that household n chooses a particular discrete action a_{tn}^D at time t

$$
P(a_{tn}^D | s_{tn}, x_{tn}, \theta) = \frac{v_{tn}^D(s_{tn}, a_{tn}^D, x_{tn}, \theta) + v_{tn}^{C*}(s_{tn}, a_{tn}^D, a_{tn}^{C*}, x_{tn}, \theta) + \beta \sum_{s_{t+1,n} \in S} \bar{V}(s_{t+1,n}, x_{t+1,n}, \theta) f(s_{t+1,n} | s_{tn}, a_{tn}^D)}{\sum_{\tilde{a}_{tn}^D} \left\{ v_{tn}^D(s_{tn}, \tilde{a}_{tn}^D, x_{tn}, \theta) + v_{tn}^{C*}(s_{tn}, \tilde{a}_{tn}^D, \tilde{a}_{tn}^{C*}, x_{tn}, \theta) + \beta \sum_{s_{t+1,n} \in S} \bar{V}(s_{t+1,n}, x_{t+1,n}, \theta) f(s_{t+1,n} | s_{tn}, \tilde{a}_{tn}^D) \right\}} \tag{35}
$$

The simplest way to estimate this type of model is using the nested fixed point (NXFP) algorithm proposed by Rust (1987) where the DP problem is solved for each iteration of the non-linear optimization algorithm searching of the parameter space. Given the model assumptions in this paper the DP problems can be solved in short computational time which makes it possible to use NXFP and obtain parameter estimates in a reasonable computational time.

Before presenting empirical results in the following section, we note that we have validated the maximum likelihood estimation of the model using several samples of simulated data (for which the true model is known) and we show that the parameter estimates are not significantly different from their true values (the detailed results of the validation study are presented in Glerum et al., 2014).

4 Application on Swedish car fleet data

As an example of application, we consider the case of the evolution of private car ownership within Swedish households from 1999 to 2008. For that purpose we use data from two large registers: the Swedish population register and the register of the whole Swedish car fleet.

The present section first describes how the households were selected among the Swedish population. The model specification and estimation results are subsequently presented. The last part of the section presents a policy scenario simulating the effect of the dieselization of the car fleet of 2009 (see Hugosson and Algers, 2012; Kageson, 2013).

4.1 Selection of the sample of Swedish households

We use a random sample of a stratum of the Swedish population for the sake of model estimation. The subpoulation is defined such that it only contains households that (i) performed one type of actions as defined in Section 3.2 from 1999 to 2008 and that (ii) exclusively own private cars. The latter excludes households that have access to a company car that can be used for private purposes. We exclude these since we do not have any information about the company cars, such as fuel type and odometer readings. A household is defined on the information available in the population register which enables us to identify single individuals, married couples with or without children and unmarried couples with children in common. A household in our data set hence gathers the information relative to the cars belonging to each of its members, including children above 18 years living with their parents.

One of the features of the DDCCM is its capability to capture substitution between cars of different fuel types. In the application, we focus on substitution between gasoline and diesel vehicles only, due to the very small share of alternative-fuel and hybrid vehicles in the Swedish car fleet. We present some descriptive statistics for this subpopulation in Table 2.

For model estimation we sample 5,000 households from this subpopulation, out of these 5,000 we excluded households that were only partially observed over the time period (1999-2008) and the resulting sample consists of 4,447 households and 33,254 observations.

4.2 Estimation results

The estimated results are reported in Table 3 which contains two models, one where we estimate the discount factor and one where it is fixed to zero (myopic decision-maker). All parameter estimates have their expected signed and are significant. It is interesting to note that the data allows us to identify the discount factor which has been fixed in other studies. The value is 0.92 (significantly different from 1) which supports the hypothesis of forward looking decision makers. The in-sample fit is significantly better for Model 1 compared to Model 2. We can also note that the parameter ratios are different for the two models. θ_0 and θ_v from Table 3 yield $\rho_s = 0.758$ which is within the feasible interval between 0 and 1, more over it is greater than ρ. This means that the substitution effect between two cars with the same fuel type is smaller than that between two cars with different fuel, as expected.

To interpret some of the other variables in terms of monetary value, let us first study how the marginal utility changes when additional income become available in the current year. The income of the current year enters the utility through the CES function $v_{tn}^{C}*$. This utility is linear in income and the slope depends on (i) the fuel prices and (ii) the number of cars the household own and their fuel type, as can be seen from Equations (20), (25) and (28). Calculating the derivative of this utility with respect to income spent on fuel is trivial. As an example, consider a one car household:

$$v_{tn}^{C*} = \theta_v \cdot 2 \cdot \tilde{m}_{1tn}^{*} = \theta_v \frac{\mathrm{Inc}_{tn}}{p_{f_1 tn}}. \tag{36}$$

The derivative of the utility for household n in time t with respect to income in the current time step t then becomes:

$$\frac{\partial v_{tn}^{C*}}{\partial \mathrm{Inc}_{tn}} = \frac{\theta_v}{p_{f_1 tn}}. \tag{37}$$

Remember that $p_{f_1 tn}$ is the cost per kilometer of driving a car with fuel f_1. It has been assumed throughout this paper that the fuel consumption is 0.081/km. For a household owning a gasoline car in year 2004 (see Table 2 for gasoline prices) we therefore get

$$\frac{\partial v_{tn}^{C*}}{\partial \mathrm{Inc}_{tn}} = 1.12/(100,000\,\mathrm{kr}). \tag{38}$$

4.3 Out of sample validation

To test the performance of the model we do a cross-validation study on the data (4 447 households). More precisely we do repeated random subsampling selecting

Variable type	Variable name	Level	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
Market	Average fuel price [SEK]	Gasoline	8.4	9.56	9.52	9.37	9.46	10.05	11.13	11.55	11.65	12.54
		Diesel	6.66	8.44	8.69	8.36	7.92	8.61	10.48	11.27	10.88	13.12
Car fleet	Car age [% households]	1 years	6.61	6.88	6.04	5.05	5.50	5.62	5.47	6.19	4.82	5.88
		2 years	5.97	6.91	7.55	7.01	5.87	6.11	6.55	6.49	7.03	5.81
		3 years	5.63	6.30	7.16	8.05	7.88	6.60	6.61	6.84	6.64	7.21
		4 years	4.77	6.16	6.85	7.37	8.68	8.89	7.02	7.41	7.57	7.15
		5 years	4.63	4.59	6.24	7.01	7.39	8.78	8.82	7.41	7.63	7.49
		6 years	3.04	4.84	4.32	6.15	6.79	7.42	8.39	8.71	7.39	7.43
		7 years	4.60	2.87	4.76	4.37	5.70	6.29	7.43	7.88	8.53	7.25
		8 years	5.38	4.48	2.98	4.74	4.58	5.27	6.03	6.78	7.72	7.95
		9 years	59.37	56.97	54.11	50.26	47.61	45.02	43.68	42.30	42.65	43.84
	Fuel type [% households]	Gasoline	98.38	98.09	98.16	98.16	98.14	98.23	98.34	98.16	97.76	96.53
		Diesel	1.62	1.91	1.84	1.84	1.86	1.77	1.66	1.84	2.24	3.47
Household	Number of cars [% households]	0 car	13.83	12.69	12.06	12.34	11.86	12.51	11.88	12.71	12.27	12.21
		1 car	71.78	70.48	69.59	69.60	69.06	66.98	65.00	63.21	61.03	59.94
		2 cars	14.11	14.76	15.20	15.20	15.47	16.51	17.50	18.39	19.49	20.03
	Income [SEK]	Mean	262'426	279'744	288'578	303'098	311'509	320'611	336'153	357'921	429'227	403'757
		SD	134'702	304'521	154'571	159'050	153'863	142'893	164'311	210'944	1'620'366	261'754

Table 2: Descriptive statistics of the subpopulation per year.

Table 3: Estimation results, the translations of transaction and ownership costs into SEK are for a single car household owning a gasoline car in 2004.

	Model 1		Model 2	
Param.	Estimate (Rob. t-test)	[SEK]	Estimate (Rob. t-test)	[SEK]
β	0.92 (65.13)		-	
ρ	0.75 (33.40)		0.98 (116.63)	
θ_v	0.90 (13.58)		2.58 (13.23)	
$\theta_{CES diesel}$	-7.23 (-4.52)		-9.91 (-4.27)	
θ_0	1.12 (12.49)		2.73 (12.78)	
θ_1	-0.44 (-3.25)	-39592	-0.09 (-0.67)	-2917
θ_2	-6.31 (-36.69)	-563138	-3.70 (-35.63)	-115499
θ_3	-1.05 (-10.15)	-93621	-3.27 (-40.32)	-102052
θ_4	0.77 (3.60)	68452	0.49 (2.36)	15162
θ_5	0.57 (6.63)	50996	0.16 (1.91)	5002
θ_6	4.08 (56.23)	363911	3.83 (56.68)	119621
θ_7	-0.12 (-2.69)	-11099	-0.50 (-4.40)	-15669
θ_8	-0.49 (-7.83)	-44065	-0.91 (-12.00)	-28275
θ_9	0.42 (5.34)	37345	0.49 (6.00)	15262
θ_{10}	-2.91 (-20.75)	-260010	-2.94 (-21.37)	-91641
LL	23173.70	23508.70		

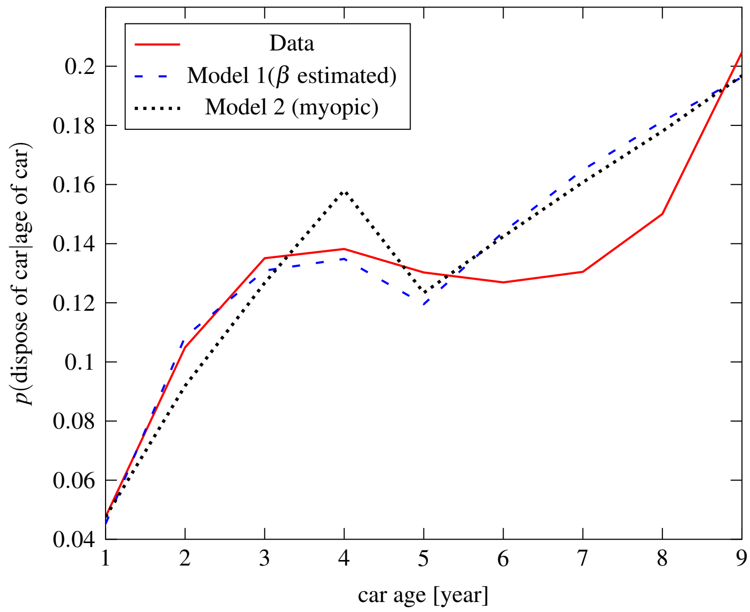

Figure 2: Probability to dispose of a car conditional on the age of the car

80 % of the households into a training set that is used for estimation, and validating the model on the remaining 20 %. 31 such sub-sets were sampled and used in the results reported below.

First let us consider the probability of car disposal. In Figure 2, the probability to dispose of a car conditional on its age is as a transaction frequency according to the data, Model 1 (β is free) and Model 2 ($\beta = 0$). The overall shapes of all three distributions are the same, and both models are better explaining the data for newer cars. In addition, we note that Model 1 better fit the data than Model 2 for cars less than 5 years old.

Tables 4 and 5 show the transaction frequencies in the data and according to models 1 and 2. The error is quite small for all transactions, except for disposing of two cars (transaction 3), but we note that in terms of absolute numbers, there are very few observations of this transaction in the data.

Finally, as a measure of goodness of fit we have compared the log-likelihood of the model on the training sets and the validation sets to see if the dynamic model performs better than the static model. The difference in log-likelihood between Model 1 and Model 2 is denoted ΔLL and is scaled by the log-likelihood of Model

Table 4: Comparison when estimating and validating on the full sample.

Transaction	Observed share	Model 1 error	Model 2 error
$h1$	79.942	0.111	0.500
$h2$	3.759	-0.039	0.017
$h3$	0.052	0.030	0.020
$h4$	2.422	-0.127	-0.582
$h5$	1.086	0.051	0.081
$h6$	0.518	0.041	0.459
$h7$	0.000	0.000	0.000
$h8$	10.577	0.031	-0.232
$h9$	1.645	-0.098	-0.262

1. The mean (and standard deviation of the mean) for this difference

$$100 \cdot \frac{\Delta LL_{training}}{LL_{1,training}} = 1.43805 \pm 0.01 \tag{39}$$

$$100 \cdot \frac{\Delta LL_{validation}}{LL_{1,validation}} = 1.49822 \pm 0.05. \tag{40}$$

The difference between the two models remains as large on the validation sets as on the training sets and in both cases the dynamic model performs better.

4.4 Policy scenario

In this section we report the analysis of a policy scenario, using the above estimated model. We analyze a hypothetical policy that change the cost structure for diesel cars. The choice of policy is motivated by the fact that the market share for diesel cars has increased considerably, from 10 per cent in 2005 to 45 per cent in 2009 and 60 per cent in 2011 (Kageson, 2013). There are several reasons for this increase. Firstly, the technological development of the diesel engine has brought diesel cars to the market that are less noisy, more comfortable and more powerful than before. This has made diesel cars more accepted and more models have been launched on the Swedish market. Secondly, this technical development together with environmental regulations have led to substantial improvements in emissions of CO_2, PM and NOx. Diesel cars are hence perceived as less polluting than before. And thirdly, the taxation of diesel cars has been changed in a favorable way.

In 2006, the definition of a “clean cars” was changed so that it included diesel and gasoline cars emitting less than 120 g CO_2 per km (in addition to flexifuel, CNG, and electric cars that were already included). A notable effect of this change

Table 5: Comparison on validation set consisting of 20% of the households when estimating on training set consisting of the remaining 80%. Report mean root mean square error (rmsq), mean error, and standard deviation of the error.

Transaction	Observed share	(std)	Model 1 rmsq.	mean error	(std)	Model 2 rmsq.	mean	(std)
$h1$	79.862	(0.54)	0.609	0.086	(0.61)	0.782	0.479	(0.63)
$h2$	3.804	(0.23)	0.305	0.043	(0.31)	0.317	0.094	(0.31)
$h3$	0.058	(0.03)	0.044	0.037	(0.02)	0.036	0.027	(0.02)
$h4$	2.466	(0.16)	0.209	-0.088	(0.19)	0.580	-0.546	(0.20)
$h5$	1.102	(0.12)	0.143	0.062	(0.13)	0.162	0.096	(0.13)
$h6$	0.520	(0.09)	0.124	0.042	(0.12)	0.469	0.461	(0.09)
$h7$	0.000	(0.00)	0.000	0.000	(0.00)	0.000	0.000	(0.00)
$h8$	10.560	(0.31)	0.338	-0.048	(0.34)	0.457	-0.313	(0.34)
$h9$	1.628	(0.15)	0.175	-0.133	(0.12)	0.319	-0.299	(0.11)

was that popular car models such as Volvo V70 and Volkswagen Passat equipped with the smallest diesel engine suddenly fulfilled the definition of a clean car. These clean cars enjoyed a purchase subsidy of 10 000 SEK between 2006 and 2009, an incentive that was replaced in 2009 with a five year exempt from the annual vehicle tax. The latter incentive favoured diesel cars in particular since the annual vehicle tax is higher for diesel cars than for gasoline and alternative fuel cars. The annual vehicle tax depends on fuel type and CO_2 emissions. For a diesel car emitting 119 g CO_2 per km, the annual vehicle tax is 1520 SEK.

With this background in mind, the purpose of this subsection is to analyze how much a yearly subsidy of diesel cars with 0 to 15 000 SEK/year would influence the predicted share of diesel cars in 2009 according to the model.

To implement such a subsidy in our model, we first note that when calculating the utility of using a car, given by the solution to the optimization problem in Equation (17), the share of the income a household spends on fuel each year is assumed to be fixed. To calculate the change in utility of owning a diesel car after the change in vehicle tax we will rather assume that it is the yearly budget on car related expenses that is fixed. This means that, according to the model, a decreased vehicle ownership cost implies larger spendings on fuel and the utility of owning a car hence increases.

Figure 3 shows the policy effects of a fictitious subsidy for owning a diesel car. The base level (for no tax change) is 3.6 diesel cars and 103.3 petrol cars per 100 households. We note that the number petrol cars are decreasing, but the increase in diesel cars is greater than the decrease in petrol cars. An increase in 0.9 diesel cars/100 households would mean a 25% increase in a single year due to this subsidy. It should however be noted that given the initial low absolute number of diesel cars, the total effect as measured by the absolute number of diesel cars in the *fleet* is small.

5 Conclusion

This paper presents a methodology to model jointly car ownership, usage and choice of fuel type. The main feature of the model is that we account for the forward-looking behavior of decision-makers within a dynamic programming framework. This is crucial in the case of demand for durable goods such as cars, since the purchase of a car is affected by the utility gained from that car for the present and future years of ownership. We present empirical results that supports the model and we have an estimated discount factor of 0.92. We present both estimation and validation results as well as prediction study analyzing a fictitious policy scenario. In addition to the dynamic aspects of the model, the continuous choice variables impose a major challenge. In particular since we consider that households can own two cars. We deal with this using a CES function that allows us to model the substitution between the usage of the two cars.

In order to obtain an operational model we have been forced to make a number

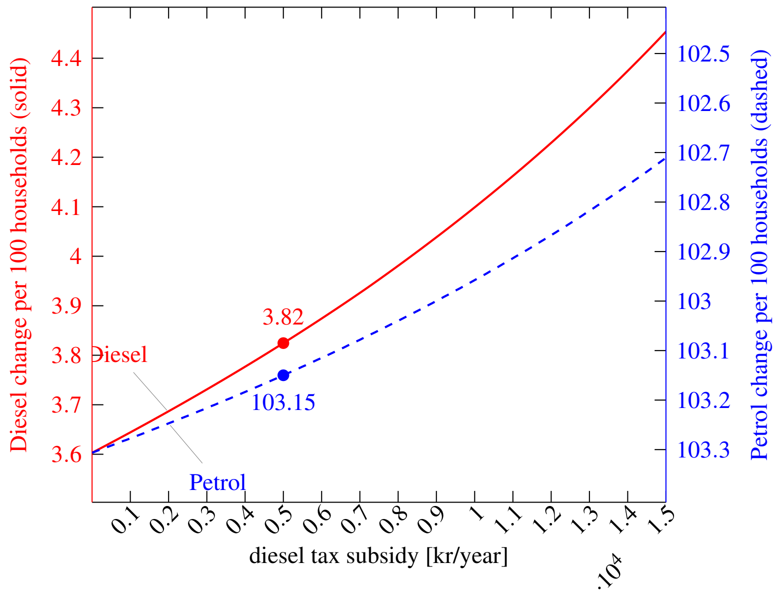

Figure 3: The change in number of diesel cars and petrol cars in the fleet under policy scenario

of simplifying assumptions. The strongest assumptions are associated with the continuous choice variable. First, the annual milage is assumed to be a myopic and deterministic choice. Second, we assume a budget that is exogenously defined. While these assumptions are restrictive, they allows us to model the substitution between cars.

Future research will be dedicated to applying the model to the register data as it becomes available. In recent years a number of policies have been implemented to increase the share of clean cars in the fleet. We plan to apply this model to study these policies as well as predicting the effect of policies scenarios defined by the Swedish government for the upcoming years.

Acknowledgements

This research was partially funded by the Center for Transport Studies, Stockholm. The authors would like to thank Roger Pyddoke for sharing part of the data and Per Olsson for his help on IT issues. Previous results of this research have been presented at the 13th Swiss Transport Research Conference, TRISTAN VIII, the third International Choice Modelling Conference and the Second Symposium of the European Association for Research in Transportation. The work presented in this paper has greatly benefited from comments received from the participants of these conferences.

References

Aguirregabiria, V. and Mira, P. (2010). Dynamic discrete choice structural models: A survey, *Journal of Econometrics* **156**(1): 38–67.

Cernicchiaro, G. and de Lapparent, M. (2015). A dynamic discrete/continuous choice model for forward-looking agents owning one or more vehicles, *Computational Economics* **46**(1): 15–34.
URL: *http://dx.doi.org/10.1007/s10614-014-9449-4*

de Jong, G. (1996). A disaggregate model system of vehicle holding duration, type choice and use, *Transportation Research Part B: Methodological* **30**(4): 263–276.

de Jong, G. C. and Kitamura, R. (2009). A review of household dynamic vehicle ownership models: holdings models versus transactions models, *Transportation* **36**(6): 733–743.

de Jong, G., Fox, J., Daly, A., Pieters, M. and Smit, R. (2004). Comparison of car ownership models, *Transport Reviews* **24**(4): 379–408.

de Lapparent, M. and Cernicchiaro, G. (2012). How long to own and how much to use a car? a dynamic discrete choice model to explain holding duration and driven mileage, *Economic Modelling* **29**(5): 1737–1744.

Gillingham, K. (2012). Selection on anticipated driving and the consumer response to changing gasoline prices, *Technical report*, Yale University.

Glerum, A., Frejinger, E., Karlstrom, A., Hugosson, M. B. and Bierlaire, M. (2014). A dynamic discrete-continuous choice model for car ownership and usage: Estimation procedure, *Proceedings of the 14th Swiss Transport Research Conference*, Ascona, Switzerland.

Hugosson, M. B. and Algers, S. (2012). Accelerated introduction of 'clean' cars in Sweden, *in* T. I. Zachariadis (ed.), *Cars and Carbon*, Springer Netherlands, pp. 247–268.

Kageson, P. (2013). Dieselization in Sweden, *Energy Policy* **54**(0): 42 – 46.

Rust, J. (1987). Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher, *Econometrica* **55**(5): 999–1033.

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

Xu, R. (2011). *Dynamic discrete choice models for car ownership modeling*, PhD thesis, University of Maryland.

Zabalza, A. (1983). The CES utility function, non-linear budget constraints and labour supply. results on female participation and hours, *The Economic Journal* **93**(370): 312–330.