A dynamic discrete-continuous choice model of car ownership, usage and fuel type

Aurélie Glerum [*] Emma Frejinger [†] Anders Karlström [‡]
Muriel Beser Hugosson [‡] Michel Bierlaire [*]

2nd Symposium of the European Association for Research in Transportation
4th-6th September 2013, Stockholm

1 Introduction

Car purchase behavior and usage have evolved importantly during the last decades. Over time, the vehicle market has been characterized by technology changes in terms of the availability of different fuel types and the improvement of fuel efficiency. At the same time, policies have been implemented in order to reduce carbon dioxide emissions and favor more environmentally friendly cars. In that context, it is crucial to identify the factors underlying households' vehicle purchase decisions in order to understand and predict changes in the demand.

Many vehicle ownership and usage models have been developed in the transportation literature, but most of them are static models that do not account for the fact that a car is a durable good. Indeed individuals keep their vehicles for several years and their expectations about the future affect their current decisions. A few recent studies have started to account for that feature. Among them, Xu (2011) models car replacement decisions and vehicle type choice in a stated preferences context using a dynamic discrete choice model (DDCM). Schiraldi (2011) specifies such a model on car register data in Italy to evaluate the impact of of scrappage policies on the demand. Heterogeneity of preferences

---

[*] ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory (TRANSP-OR), {aurelie.glerum, michel.bierlaire}@epfl.ch

[†] UNIVERSTITÉ DE MONTRÉAL, Facult des arts et des sciences, Department of Computer Science and Operations Research, frejinge@iro.umontreal.ca

[‡] ROYAL INSTITUTE OF TECHNOLOGY (KTH), School of Architecture and the Built Environment, Department of Transport Science, {anders.karlstrom, muriel.hugosson}@abe.kth.se

is captured by introducing aggregate socio-economic information into the model. Other modeling approaches have been considered in order to jointly model car ownership and usage. For instance, Gillingham (2012) models cars monthly mileage conditional on vehicle type. Other researches focus on applying a dynamic programming mixed logit (DPMXL) approach (Schjerning, 2008) to model vehicle type choice, usage and replacement decisions (according to discussions and unpublished work by Munk-Nielsen). The particularity of this model is that it can handle both discrete and continuous decision variables.

In this study, we are using a data set that combines the registers of all individuals and vehicles in Sweden from 1998 to 2008. Detailed socio-economic information is available for each individual (e.g. net income, house and work location, type of employment) and in addition, we have information about household composition, i.e. married individuals and unmarried couples living together with children can be identified from the data. Extensive information about the vehicles is also available (e.g. make, model, fuel consumption, fuel type, age and annual mileage from odometer readings).

The objective of this paper is to specify and estimate a *dynamic discrete-continuous choice model (DDCCM)* on the register of individuals and cars in Sweden in order to jointly explain car replacement decisions and usage. This research contributes to the following aspects. First, we model the above decisions at the household level and account for the fact that each household can have at most two cars, hence allowing for a realistic representation of the decision maker. Second, we are considering an extensive choice variable, consisting of the number of cars owned in the household, the annual distance that each car will be driven, the fuel type of each car, the decision to choose a company car and to take a new or second-hand car. This enables us to analyze decisions that are closely related. Third, the availability of the whole register data allow us to analyze and predict variations in car holding behavior and usage for the whole Swedish population, from the local level to the national level.

2 Model specification

The DDCCM is formulated as a discrete-continuous choice model that is embedded in a *dynamic programming (DP)* framework. We make the following assumptions:

- Decisions are taken at a household level.
- We consider an infinite-horizon problem to account for the fact that households make long-term decisions in the context of car acquisition.
- The choice of mileage(s) is conditional on the choice of the discrete decision variables.
- 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.

The *state space* S is constructed based on the following car attributes: the discretized total mileage $m_{c,t}$ for car c at year t, a discrete variable $I_{c,t}$ indicating whether car c is owned privately, by sole proprietorship or by another type of company, the fuel type $f_{c,t}$ of car c. Therefore, each state $s_t \in S$ can be represented as $s_t = (m_{1,t}, I_{1,t}, f_{1,t}, m_{2,t}, I_{2,t}, f_{2,t})$.

The *action space* A is constructed based on the following variables: the transactions h_t in the household composition of the car fleet at year t (e.g. increase, decrease and replacement of vehicles), the annual mileage $\tilde{m}_{c,t}$ for car c, the choice $I_{c,t}$ of taking company car, the fuel type $f_{c,t}$, the choice $r_{c,t}$ to buy a new or second-hand car. Each action $a_t \in A$ can hence be represented as $a_t = (h_t, \tilde{m}_{1,t}, I_{1,t}, f_{1,t}, r_{1,t}, \tilde{m}_{2,t}, I_{2,t}, f_{2,t}, r_{2,t})$.

Given that a household is in a state s_t and has chosen an action a_t, the transition function $f(s_{t+1}|s_t, a_t)$ is defined as the rule mapping s_t and a_t to the next state s_{t+1}. In our case, s_{t+1} can be inferred deterministically from s_t and a_t.

At each year t, the household chooses an action $a_t \in A$ that maximizes its current utility and expected discounted utility for the future years. The value function is therefore:

$$V(s_t, x_t, \theta) = \max_{a_t \in A}\{u(s_t, a_t, x_t, \theta) + \beta \sum_{s_{t+1} \in S} \bar{V}(s_{t+1}, x_{t+1}, \theta) f(s_{t+1}|s_t, a_t)\}, \tag{1}$$

where $\beta \in (0, 1)$ is a discount factor, θ is a set of parameters to be estimated, x_t is a vector of explanatory variables of households' choices which are not part of the state or action spaces. Moreover $u(s_t, a_t, x_t, \theta)$ is the single-period utility.

The integrated value function is defined as $\bar{V}(s_t, x_t, \theta) = \int V(s_t, x_t, \theta) dG_\varepsilon(\varepsilon_t)$, where G_ε is the CDF of ε_t. In the case where all actions are discrete and the random terms are i.i.d. extreme value, it corresponds to the logsum (see e.g. Aguirregabiria and Mira, 2010). In our case, the choice of mileage(s) is a continuous variable. Assuming that $a_t^D = (h_t, I_{1,t}, f_{1,t}, r_{1,t}, I_{2,t}, f_{2,t}, r_{2,t})$ gathers the discrete components of an action a_t and $a_t^C = (\tilde{m}_{1,t}, \tilde{m}_{2,t})$ gathers the continuous components, the single-period utility can be defined as:

$$u(s_t, a_t, x_t, \theta) = u(s_t, a_t^C, a_t^D, x_t, \theta) = v(s_t, a_t^C, a_t^D, x_t, \varepsilon_C(a_t^C), \theta) + \varepsilon_D(a_t^D), \tag{2}$$

where $v(s_t, a_t^C, a_t^D, x_t, \varepsilon_C(a_t^C), \theta)$ is the deterministic part, $\varepsilon_D(a_t^D)$ is a random error term for the discrete actions and $\varepsilon_C(a_t^C)$ captures the randomness inherent to the continuous decision(s). Under special conditions determined by Munk-Nielsen, we can find a closed-form formula for the Bellman equation in the discrete-continuous case too:

$$\bar{V}(s_t, x_t, \theta) = \log \sum_{a_t^D} \exp\{\max_{a_t^C}\{v(s_t, a_t^C, a_t^D, x_t, \varepsilon_C(a_t^C), \theta)\} + \beta \sum_{s_{t+1} \in S} \bar{V}(s_{t+1}, x_{t+1}, \theta) f(s_{t+1}|s_t, a_t)\}. \tag{3}$$

Technical details on the specification of the instantaneous utility will be described in the paper. At a first stage, the model parameters are estimated based on Rust's nested fixed point algorithm (Rust, 1987).

3 Perspectives

In this research we are developing a dynamic model of car ownership, fuel type and usage that can accommodate discrete and continuous decision variables. This is work in progress and we are currently estimating the model on synthetic data to validate our approach. In the presentation, we will also present results of the estimation of the model on the whole Swedish register.

This work is part of an extensive research project and intermediate results have been accepted to two other conferences. The present work is hence related, but in the presentation we will show more complete results.

References

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

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

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

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

B. Schjerning. *Dynamic aspects of entrepreneurial behavior*. PhD thesis, Department of Economics, University of Copenhagen, 2008.

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