A DYNAMIC DISCRETE-CONTINUOUS CHOICE MODEL TO EXPLAIN CAR OWNERSHIP, USAGE AND FUEL TYPE DECISIONS

Aurélie Glerum EPFL
Emma Frejinger Université de Montréal
Anders Karlström KTH
Muriel Beser Hugosson KTH
Michel Bierlaire EPFL

IRE seminar, USI
11th October 2013
• Introduction
• Background and data
• Dynamic discrete-continuous choice model
• Illustrative example
• Estimation on synthetic data
• Conclusion and future works
INTRODUCTION

Aim of the research:

• Model dynamics of car transactions, usage and choice of fuel type in the Swedish car fleet

• Motivations
  • Governmental policies to reduce carbon emissions / car usage:
    • Stockholm congestion tax
    • Independence of fossil fuels
  • Technology changes:
    • Increase of alternative-fuel vehicles
  • Economical features:
    • Financial crisis
    • Fuel price changes

Car ownership and usage vary importantly over time.

Model needed to analyze and predict impact of policies on ownership and usage
Register data of Swedish population and car fleet:

- Data from 1998 to 2008

- **All individuals**
  - **Individual information**: socio-economic information on car holder (age, gender, income, home/work location, employment status/sector, etc.)
  - **Household information**: composition (families with children and married couples)

- **All vehicles**
  - Privately-owned cars, cars from privately-owned company and **company** cars
  - Vehicle **characteristics** (make, model, fuel consumption, fuel type, age)
  - **Annual mileage** from odometer readings
  - Car bought **new or second-hand**
Number of cars (in hh with only private cars)

- 0-car hh
- 1-car hh
- 2-car hh

Year:
- 1999
- 2001
- 2003
- 2005
- 2007
BACKGROUND AND DATA

Number of cars (in hh with only private cars)

Ownership
Congestion tax
Stockholm

- 0-car hh
- 1-car hh
- 2-car hh

Year

1999 2001 2003 2005 2007
• Car ownership models in transportation literature:

  • Discrete choice models (DCM) widely used, but mostly static models.
    • Main drawback: do not account for forward-looking behavior
    • Important aspect to account for since car is a durable good

  • Econometric literature: dynamic programming (DP) + DCM

  • Recently, dynamic discrete choice models (DDCM) starting to be applied in transportation field (Cirillo and Xu, 2011; Schiraldi, 2011)
• Joint models of car ownership and usage:

  • **Duration models** and regression techniques for car holding duration and usage (De Jong, 1996)

  • Vehicle type, usage and replacement decisions using **dynamic programming, discrete-continuous, mixed logit** (Schjerning, 2008, and Munk-Nielsen, 2012)

  • **Discrete-continuous model** of vehicle choice and usage based on register data: includes expectation of fuel prices & car future resale price (Gillingham, 2012)

  • Wide literature on car ownership and usage models
Models for discrete-continuous choices:

• One of the base references: Joint choice of energy portfolio and energy consumption (Dubin and McFadden, 1984)

• Joint choice of vehicle type and usage (Munk-Nielsen, 2012; Gillingham, 2012)

• Other references (Hanemann, 1984; Bhat, 2005, 2013)
BACKGROUND AND DATA

RESEARCH ISSUES

• Car are durable goods \overset{\implies}{\rightarrow} Need to account for \textit{forward-looking behavior} of agents

• Difficulty of modeling a \textit{discrete-continuous choice} when jointly modeling car ownership and usage

• Many models focus on individual decisions, but choices regarding car ownership and usage made at \textit{household level}
Car are durable goods \(\Rightarrow\) Need to account for **forward-looking behavior** of agents

Difficulty of modeling a **discrete-continuous choice** when jointly modeling car ownership and usage

Many models focus on individual decisions, but choices regarding car ownership and usage made at **household level**

**Proposed methodology:**

- Attempt to address these issues by applying **dynamic discrete-continuous choice model (DDCCM)**

Large **register data** of all **individuals** and **cars** in Sweden
• In the area of **dynamic choice modeling**
  • Choices modeled at **household level**
  • **Up to two cars** allowed (only 4% households with > 2 cars in 2007)

• **Constant elasticity of substitution (CES) utility** to model annual driving distance for 2-car households

• Several **choices modeled simultaneously**
Objective

Model simultaneously car ownership, usage and fuel type.
In details: model simultaneous choice of

- Transaction type
- Annual milage – car $c$
- Private/company – car $c$
- Fuel type – car $c$
- New/2nd hand – car $c$
- # cars
Objective

Model simultaneously car ownership, usage and fuel type.

In details: model simultaneous choice of

- Transaction type
- Annual milage – car $c$
- Private/ company car – car $c$
- Fuel type – car $c$
- New/2$^{nd}$ hand – car $c$
- # cars

Discrete variables
Objective

Model simultaneously car ownership, usage and fuel type.

In details: model simultaneous choice of

- Transaction type
- Annual milage – car $c$
- Private/company car – car $c$
- Fuel type – car $c$
- New/2nd hand – car $c$
- # cars

Continuous variables
1. **Choice** at **household level**: up to 2 cars in household

2. **Strategic choice** of:
   - Transaction
   - Type(s) of ownership (company vs private car)
   - Fuel type(s)
   - Car state(s) (new vs 2nd-hand)

   Account for forward-looking behavior of households

3. **Myopic choice** of:
   - Annual mileage(s)

4. **Choice** of **mileage conditional** on choice of discrete variables
Myopic choice (static case)

\[ P(\text{action}) = \frac{\exp\{\text{instantaneous utility}\}}{\sum_{\text{all poss. actions}} \exp\{\text{instantaneous utilities}\}} \]

Strategic choice (dynamic case)

\[ P(\text{action}) = \frac{\exp\{\text{instantaneous utility} + \text{expected discounted utility of future choices}\}}{\sum_{\text{all poss. actions}} \exp\{\text{instantaneous utilities} + \text{expected discounted utilities of future choices}\}} \]
Myopic choice (static case)

\[ P(\text{action}) = \frac{\exp\{\text{instantaneous utility}\}}{\sum_{\text{all poss. actions}} \exp\{\text{instantaneous utilities}\}} \]

Strategic choice (dynamic case)

\[ P(\text{action}) = \frac{\exp\{\text{instantaneous utility + expected discounted utility of future choices}\}}{\sum_{\text{all poss. actions}} \exp\{\text{instantaneous utilities + expected discounted utilities of future choices}\}} \]

Embeds a choice model into a dynamic programming framework
Components of the DDCCM:

- Agent
- Time step
- State space
- Action space
- Transition rule
- Instantaneous utility function
• Agent: household
• Time step $t$: year
• State space $S$

$$s_t = (y_{1,t}, I_{1,t}, f_{1,t}, y_{2,t}, I_{2,t}, f_{2,t})$$

- Age – 1st car
- Private/company car – 1st car
- Fuel type – 1st car
- Age – 2nd car
- Private/company car – 2nd car
- Fuel type – 2nd car
DEFINITION OF THE COMPONENTS

- Action space $A$

$$a_t = (h_t, \tilde{m}_{1,t}, \tilde{I}_{1,t}, f_{1,t}, \tilde{r}_{1,t}, \tilde{m}_{2,t}, \tilde{I}_{2,t}, f_{2,t}, \tilde{r}_{2,t})$$
• Action space $A$

Transaction types

DYNAMIC DISCRETE-CONTINUOUS CHOICE MODEL
DEFINITION OF THE COMPONENTS
• **Transition rule:** deterministic rule: each state $s_{t+1}$ can be inferred exactly once $s_t$ and $a_t$ are known.
• Transition rule: deterministic rule: each state $s_{t+1}$ can be inferred exactly once $s_t$ and $a_t$ are known.

• Instantaneous utility function:

$$u(s_t, a_t^C, a_t^D, x_t, \theta) = v(s_t, a_t^C, a_t^D, x_t, \theta) + \varepsilon_D(a_t^D)$$

Assume additive **deterministic utility** for simplicity (see also Munk-Nielsen, 2012):

$$v(s_t, a_t^C, a_t^D, x_t, \theta) = v_t^D(s_t, a_t^D, x_t, \theta) + v_t^C(s_t, a_t^D, a_t^C, x_t, \theta)$$
• Instantaneous utility function

• Utility for continuous actions: **Constant elasticity of substitution (CES)** utility function (e.g. Zabalza, 1983):

\[ v_t^C (s_t, a^D_t, a^C_t, x_t, \theta) = (m_{1,t}^{-\rho} + \alpha \cdot m_{2,t}^{-\rho})^{-1/\rho} \]

• Captures substitution patterns between the choice of both annual driving distances
• \( \rho \) = elasticity of substitution
• \( \alpha \) = share parameter

• Formulation could be extended to introduce randomness in \( \alpha \).
Parameters obtained by maximizing likelihood:

\[ \mathcal{L} = \prod_{n=1}^{N} \prod_{t=1}^{T_n} P(a^D_{n,t} | s_{n,t}, x_{n,t}, \theta) \]

Optimization algorithm: Rust’s nested fixed point algorithm (NFXP) (Rust, 1987):

- **Outer optimization algorithm**: search algorithm to obtain parameters maximizing likelihood
- **Inner value iteration algorithm**: solves the dynamic programming problem for each parameter trial
- Plan to investigate variants of NFXP to speed up computational time (e.g. swapped algorithm from Aguirregabiria and Mira, 2002)
Outer algorithm:

- Standard estimation procedure (as for static discrete choice models)
- Here: BHHH algorithm

Inner algorithm:

Two steps
1. Finding the optimal value(s) of annual mileage conditional on the discrete choices
2. Finding the expected discounted utility of future choices
   (= value function)
1. Finding the optimal value(s) of mileage

- Maximization of the continuous utility:
  \[
  \max_{m_{1,t},m_{2,t}} \, v_t^C
  \]
  \[\text{s.t. } \quad p_{1,t}m_{1,t} + p_{2,t}m_{2,t} = \text{Inc}_t \]

- Find analytical solutions: \( m_{1,t}^* \) and \( m_{2,t}^* \)
  \[
  m_{2,t}^* = \frac{\text{Inc}_t \cdot p_{2,t}^{(-1/(\rho+1))}}{p_{2,t}^{(\rho/(\rho+1))} + p_{1,t}^{(\rho/(1+\rho))} \alpha^{(-1/(\rho+1))}}
  \]
  \[
  m_{1,t}^* = \frac{\text{Inc}_t}{p_{1,t}} - \frac{p_{2,t}}{p_{1,t}} m_{2,t}^*
  \]

- Optimal continuous utility
  \[
  v_t^C(s_t, a_t^D, a_t^C^*, x_t, \theta)
  \]
2. Finding the expected discounted utility of future choices (= value function)

- **Logsum** formula used in the completely discrete case (DDCM) (Aguirregabiria and Mira, 2010; Cirillo and Xu, 2011)

- Logsum can be applied here given the key assumptions:
  - Choice of mileage(s) is conditional on discrete actions
  - Choice of mileage(s) is myopic

\[
\bar{V}(s_t, x_t, \theta) = \log \sum_{a_t^D} \exp\{v_t^D(s_t, a_t^D, x_t, \theta) + v_t^{C^*}(s_t, a_t^D, a_t^{C^*}, x_t, \theta) + \beta \sum_{s_{t+1} \in S} \bar{V}(s_{t+1}, x_{t+1}, \theta)f(s_{t+1}|s_t, a_t)\}
\]

- Iterate on **Bellman equation** to find integrated value function \(\bar{V}\)
Assumptions for the example:

- Deterministic utility function

\[ v^D_t(s_t, a^D_t, x_t, \theta) = C(s_t) + \tau(a^D_t) + \beta_{\text{Age}}(a^D_t, s_t) \cdot \max(\text{Age1}_t, \text{Age2}_t) \]

- Chose arbitrary values for parameters

Constant for households with at least one car

Transaction costs

Transaction-dependant parameters relative to age of oldest car
### ILLUSTRATIVE EXAMPLE

<table>
<thead>
<tr>
<th>Transaction name</th>
<th>Case</th>
<th>$\beta_{\text{Age}}$</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0 car</td>
<td>1 car</td>
</tr>
<tr>
<td>$h_1$: leave unchanged</td>
<td></td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>$h_2$: increase 1</td>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$h_3$: dispose 2</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$h_4$: dispose 1st</td>
<td>1st car is oldest</td>
<td>-</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>2nd car is oldest</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$h_5$: dispose 2nd</td>
<td>1st car is oldest</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2nd car is oldest</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$h_6$: dispose 1st and change 2nd</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$h_7$: dispose 2nd and change 1st</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$h_8$: change 1st</td>
<td>1st car is oldest</td>
<td>-</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>2nd car is oldest</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>$h_9$: change 2nd</td>
<td>1st car is oldest</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>2nd car is oldest</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

C = 5, for 1- or 2-car households
Assumptions for the example:

- Visualize choice probabilities for one observation:
  - 1-car household
  - Annual income = 530’000 SEK (≈ 60’200 €, 74’200 CHF)
  - 8% expenses on fuel
ILLUSTRATIVE EXAMPLE

Static

Dynamic

Instantaneous utility

\[
P(d_{n,t}^D | s_{n,t}, x_{n,t}, \theta) = \frac{v_{n,t}^D + v_{n,t}^{C*} + \beta \sum_{s_{n,t+1} \in S} \tilde{V}f}{\sum_{d_{n,t}^D} \left( v_{n,t}^D + v_{n,t}^{C*} + \beta \sum_{s_{n,t+1} \in S} \tilde{V}f \right)}
\]
ILLUSTRATIVE EXAMPLE

Static

Dynamic

\[ P(d_{n,t}^D | s_{n,t}, x_{n,t}, \theta) = \frac{v_{n,t}^D + v_{n,t}^{C*} + \beta \sum_{s_{n,t+1} \in S} \tilde{V} f}{\sum_{a_{n,t}} \left\{ v_{n,t}^D + v_{n,t}^{C*} + \beta \sum_{s_{n,t+1} \in S} \tilde{V} f \right\} } \]
Approach to validate the model framework

• Generate 100 observations based on distributions of variables in the Swedish register data

• Generate choice (for each observation) based on postulated parameters

• Estimation of model

• Approach validated once postulated parameters are retrieved
Conclusion:

- Methodology to model choice of car ownership and usage dynamically
- Example of application shows how static & dynamic cases can differ
- Currently validating the approach (using synthetic data)

Next steps:

- Model estimation on register data
- Scenario testing:
  - Validation of policy measures taken during the years available in the data
  - Test policy measures that are planned to be applied in future years
Thank you!