DYNAMIC DISCRETE CHOICE MODELING

FOR CAR USE, OWNERSHIP AND FUEL TYPE BASED ON REGISTER DATA

Aurélie Glerum, EPFL
Emma Frejinger, KTH
Anders Karlström, KTH
Muriel Beser Hugosson, KTH
Michel Bierlaire, EPFL

IRUC seminar
Monday 17th December 2012

OUTLINE

- Introduction
- The data
- Possible approaches
- The dynamic discrete choice model
- Value iteration results
- Conclusion and future works

INTRODUCTION

Aim of the research project:

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

- Motivations
  - Governmental policies (Source: Swedish parliament):
    - Goal of fossil-independent vehicle fleet by 2030
    - No emissions by 2050
  - Technology changes:
    - Increase of alternative-fuel vehicles
    - Changes in the supply
  - Company cars: represent important share of new car sales

- Difficulties
  - Car are durable goods ⟹ modeling transactions not straightforward
  - Need to account for forward-looking behavior of individuals

KTH VETENSKAP OCH KONST

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

INTRODUCTION

Current literature on car ownership and usage modeling:

- Car ownership models in transportation literature:
  - 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)** models
  - Recently, **dynamic discrete choice models (DDCM)** starting to be applied in transportation field (Cirillo and Xu, 2011; Schiraldi, 2011)

- Other types of joint models of car ownership and usage:
  - **Discrete-continuous model** of vehicle choice and usage based on register data (Gillingham, 2012)
  - **Duration models** and regression techniques for car holding duration and usage (De Jong, 1996)

THE DATA

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**

THE DATA

Advantage of such detailed data:

- Can observe and analyze **demand shifts** that occurred as a response to changes in policies:
  - Changes in vehicle circulation taxes
  - Changes in fuel prices

    } at a national/regional level
  - Introduction of congestion pricing } at a local/regional level

- Can test the impact of **planned policies** on the demand for vehicles and car usage (e.g. fossil-independent fleet by 2030)

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MOTIVATIONS

Aim of the project:

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

$$\text{Transaction type} \times \left[\text{Annual milage – car } c \times \text{Private/company car – car } c \times \text{Fuel type – car } c \times \text{New/2}^{\text{nd}}\text{ hand – car } c \right]^{\#\text{ cars}}$$

- No more than 2 cars in household

- Account for forward-looking behavior of households

POSSIBLE APPROACHES

How can car ownership and usage be modeled?

Studied 2 approaches:

1. Discrete-continuous dynamic programming:
   mileage(s) considered as continuous and other variables as discrete.

2. Dynamic discrete choice modeling:
   all components of a choice variable are discretized.

POSSIBLE APPROACHES

How can car ownership and usage be modeled?

1. Discrete-continuous DP

   - **Endogeneous grid method (EGM)** to solve continuous choice problems (Carroll, 2006) in a fast way: avoids a numerical rootfinding procedure when the Euler equation is solved

   - EGM **generalized for discrete-continuous** choices (Iskhakov et al, 2012). Valid when one continuous variable is considered.

POSSIBLE APPROACHES

How can car ownership and usage be modeled?

2. DDCM

- All variables considered as discrete ⟹ mileage(s) discretized
- In transportation research, DP methods using random utility theory developed for discrete actions
  - Account for random term in utility function
  - Assumption that choices are affected by unobserved attributes, taste variations, etc.
- These DP methods lead to simple closed-form formula for the Bellman equation (e.g. Aguirregabiria and Mira, 2010)

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POSSIBLE APPROACHES

How can car ownership and usage be modeled?

2. DDCM

- All variables considered as discrete ⇒ mileage(s) discretized
- In transportation research, DP methods using random utility theory developed for discrete actions
  - Account for random term in utility function
  - Assumption that choices are affected by unobserved attributes, taste variations, etc.
- These DP methods lead to simple closed-form formula for the Bellman equation (e.g. Aguirregabiria and Mira, 2010)

Approach selected

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POSSIBLE APPROACHES

Motivations for DDCM

- In transportation engineering, usually work with random term in utility
- Though treating mileage as continuous is more adapted, no work on discrete-continuous DP models accommodating random utility theory
- Current DDCM were developed for a different setting:
  - Rust (1987) (and other applications) uses discrete actions and discretizes a continuous state space.
  - Here: large and discrete-continuous action space (while the state space is discrete and small)

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- 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

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

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- 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

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

2-car households

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- 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

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

1-car households

KTH VETENSKAP OCH KONST

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- Agent: household
- Time step *t*: year
- State space *S*

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

- $y_{1,t}$: Age – 1st car
- $I_{1,t}$: Private/ company car – 1st car
- $f_{1,t}$: Fuel type – 1st car
- $y_{2,t}$: Age – 2nd car
- $I_{2,t}$: Private/ company car – 2nd car
- $f_{2,t}$: Fuel type – 2nd car

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

0-car households

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- 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

= 6 = 4 = 4

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

= 1407 ➔ relatively small state space

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- Action space *A*

$$a_t = (h_t, \tilde{m}_{1,t}, \tilde{I}_{1,t}, \tilde{f}_{1,t}, \tilde{r}_{1,t}, \tilde{m}_{2,t}, \tilde{I}_{2,t}, \tilde{f}_{2,t}, \tilde{r}_{2,t})$$

Transaction type

Mileage – 1st car

Private/ company car – 1st car

Fuel type – 1st car

New/ 2nd hand – 1st car

Mileage - 2nd car

Private/ company car – 2nd car

Fuel type – 2nd car

New/ 2nd hand – 2nd car

KTH VETENSKAP OCH KONST

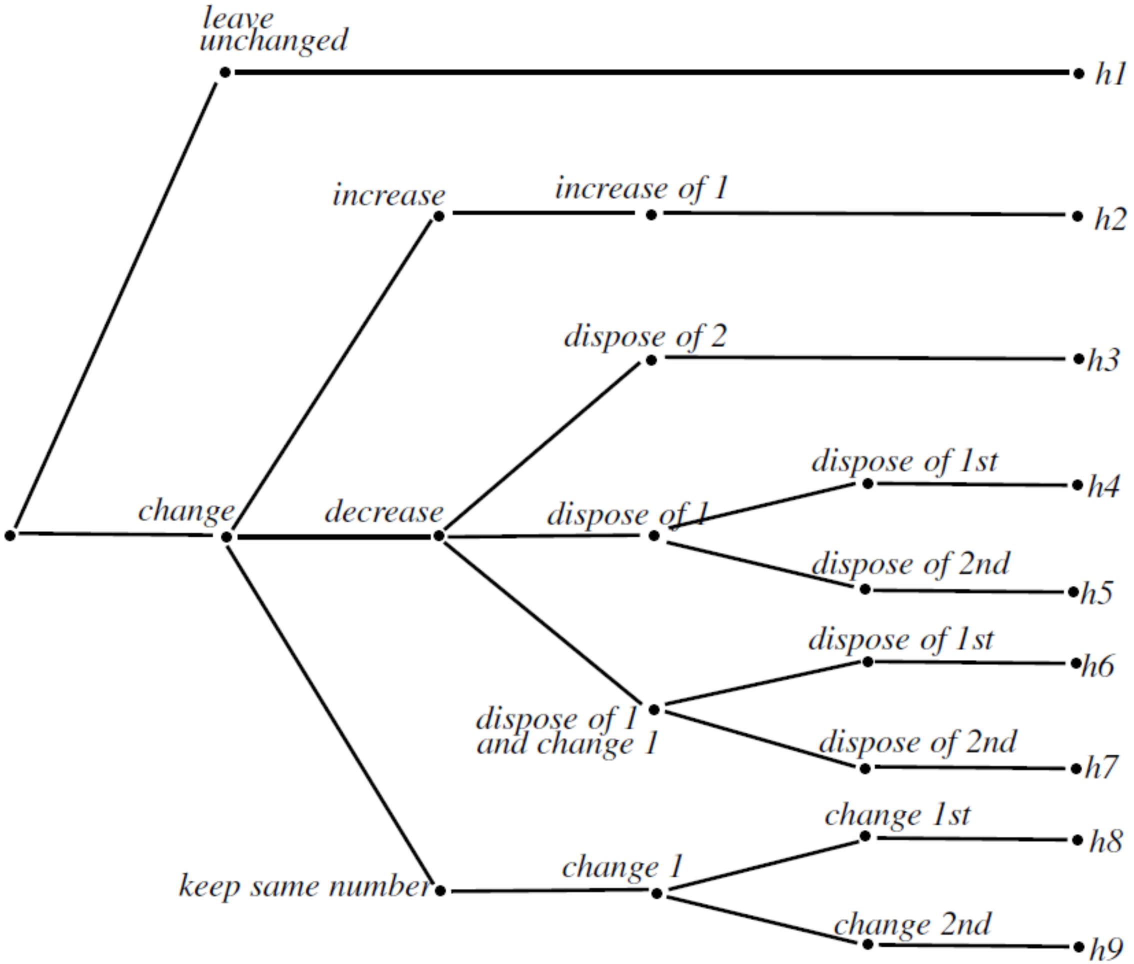
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- Action space *A*

Transaction type: details



THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- Action space *A*

Transaction name	0 car	1 car	2 cars
$h1$: leave unchanged	1	4	16
$h2$: increase 1	72	288	-
$h3$: dispose 2	-	-	1
$h4$: dispose 1st	-	1	4
$h5$: dispose 2nd	-	-	4
$h6$: dispose 1st and change 2nd	-	-	72
$h7$: dispose 2nd and change 1st	-	-	72
$h8$: change 1st	-	72	288
$h9$: change 2nd	-	-	288
Sum	73	365	745

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- Action space *A*

Size of the action space assuming:

- 4 levels mileage
- 3 levels company car
- 3 levels fuel type
- 2 levels new/2nd hand

Transaction name	0 car	1 car	2 cars
$h1$: leave unchanged	1	4	16
$h2$: increase 1	72	288	-
$h3$: dispose 2	-	-	1
$h4$: dispose 1st	-	1	4
$h5$: dispose 2nd	-	-	4
$h6$: dispose 1st and change 2nd	-	-	72
$h7$: dispose 2nd and change 1st	-	-	72
$h8$: change 1st	-	72	288
$h9$: change 2nd	-	-	288
Sum	73	365	745

KTH VETENSKAP OCH KONST

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- Action space *A*

Size of the action space assuming:

- 4 levels mileage
- 3 levels company car
- 3 levels fuel type
- 2 levels new/2nd hand

Transaction name	0 car	1 car	2 cars
$h1$: leave unchanged	1	4	16
$h2$: increase 1	72	288	-
$h3$: dispose 2	-	-	1
$h4$: dispose 1st	-	1	4
$h5$: dispose 2nd	-	-	4
$h6$: dispose 1st and change 2nd	-	-	72
$h7$: dispose 2nd and change 1st	-	-	72
$h8$: change 1st	-	72	288
$h9$: change 2nd	-	-	288
Sum	73	365	745

Maximum size of action space

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Definition of the components of the dynamic programming model:

- Instantaneous utility: $u(s_t, a_t, x_t, \theta) = v(s_t, a_t, x_t, \theta) + \varepsilon(a_t)$
- Transition rule: deterministic rule: each state s_{t+1} can be inferred exactly once s_t and a_t are known.

Example:

2 ans | private car | diesel | increase 1 | mileage level 1

If $s_t = [2,1,2,0,0,0]$ and $a_t = [1,2,0,0,0,1,3,0,0]$

mileage level 2 | company car

then $s_{t+1} = [3,1,2,0,3,0]$.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Resolution of the dynamic programming problem:

Value iteration

- Value function:

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

- Using the integrated value function $\bar{V}(s_t, x_t, \theta) = \int V(s_t, x_t, \varepsilon_t) dG_\varepsilon(\varepsilon_t)$
  - Another update rule that has a closed-form expression can be defined
  - Obtained from the expected maximum utility

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

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

THE DYNAMIC DISCRETE CHOICE MODEL

Resolution of the DDCM problem:

- Parameters obtained by maximizing likelihood:

$$\mathcal{L} = \prod_{n=1}^{N} \prod_{t=1}^{T} P(a_{n,t} | s_{n,t}, x_{n,t}, \theta) f(s_{n,t} | s_{n,t-1}, a_{n,t-1})$$

- Optimization algorithm is Rust's **nested fixed point algorithm** (Rust, 1987):
  - **Outer optimization algorithm:** search algorithm to obtain parameters maximizing likelihood
  - **Inner value iteration algorithm:** solves the DP problem for each parameter trial

VALUE ITERATION RESULTS

Preliminary results for the value iteration algorithm:

Inputs:

- Size state space = 1407
  - Max age = 5
  - Company car levels = 3
  - Number of fuel types = 3
- Size action space = max 745
  - Number of transaction types = 9
  - Number of mileage levels = 4
  - Number of levels of new/old = 2
- Utility function contains:
  - Transaction-dependent parameters for fuel type
  - Transaction-dependent parameters for age of oldest car

  } arbitrarily defined
- Parameters of DP problem:
  - Discount factor $\beta = 0.7$
  - Stopping criterion $\varepsilon = 0.01$

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

VALUE ITERATION RESULTS

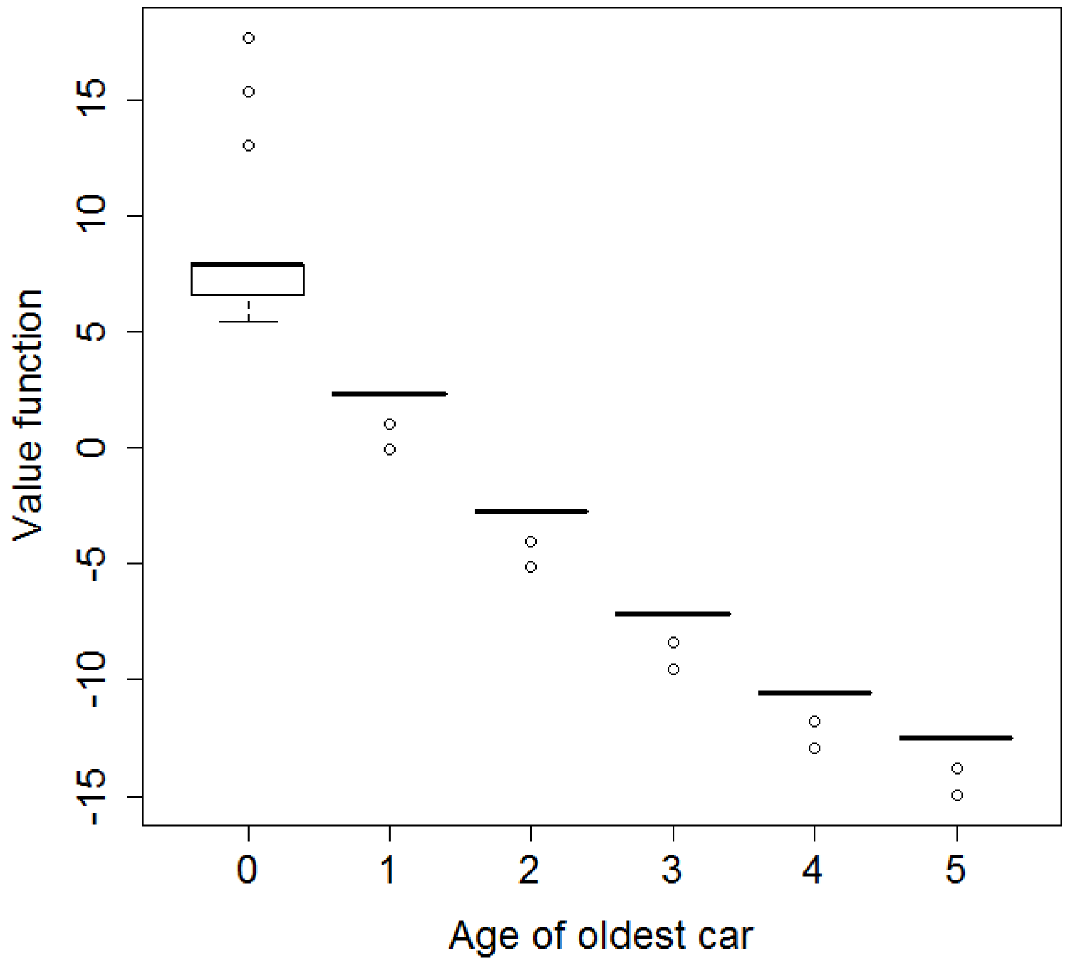
Preliminary results for the value iteration algorithm:

Program:

- Prototype in Matlab
- 15 hours on 8-core server

Graph:

- $\bar{V}$ vs age of oldest of 2 cars



VALUE ITERATION RESULTS

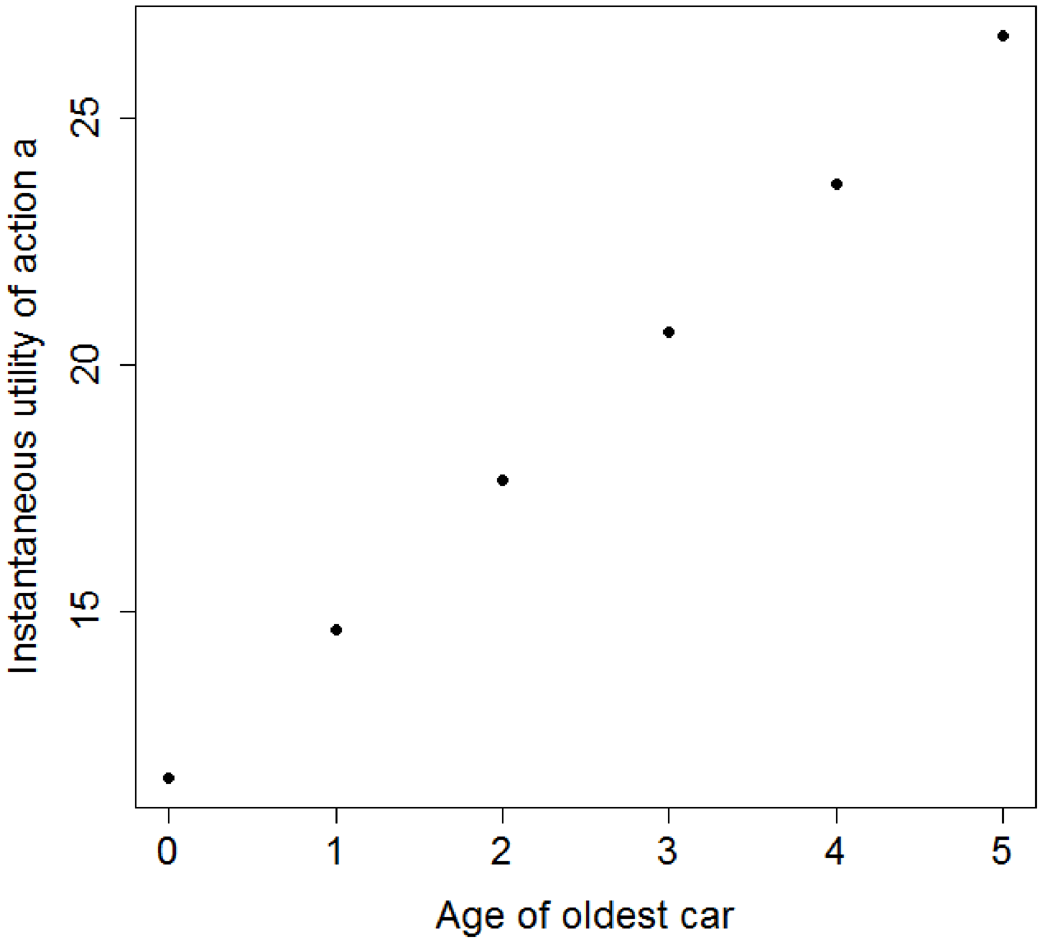
Preliminary results for the value iteration algorithm:

Program:

- Prototype in Matlab
- 15 hours on 8-core server

Graph:

- Assume an initial action $a = [3, 1, 0, 0, 0, 0, 0, 0, 0]$ $u(a)$ vs age of oldest car



VALUE ITERATION RESULTS

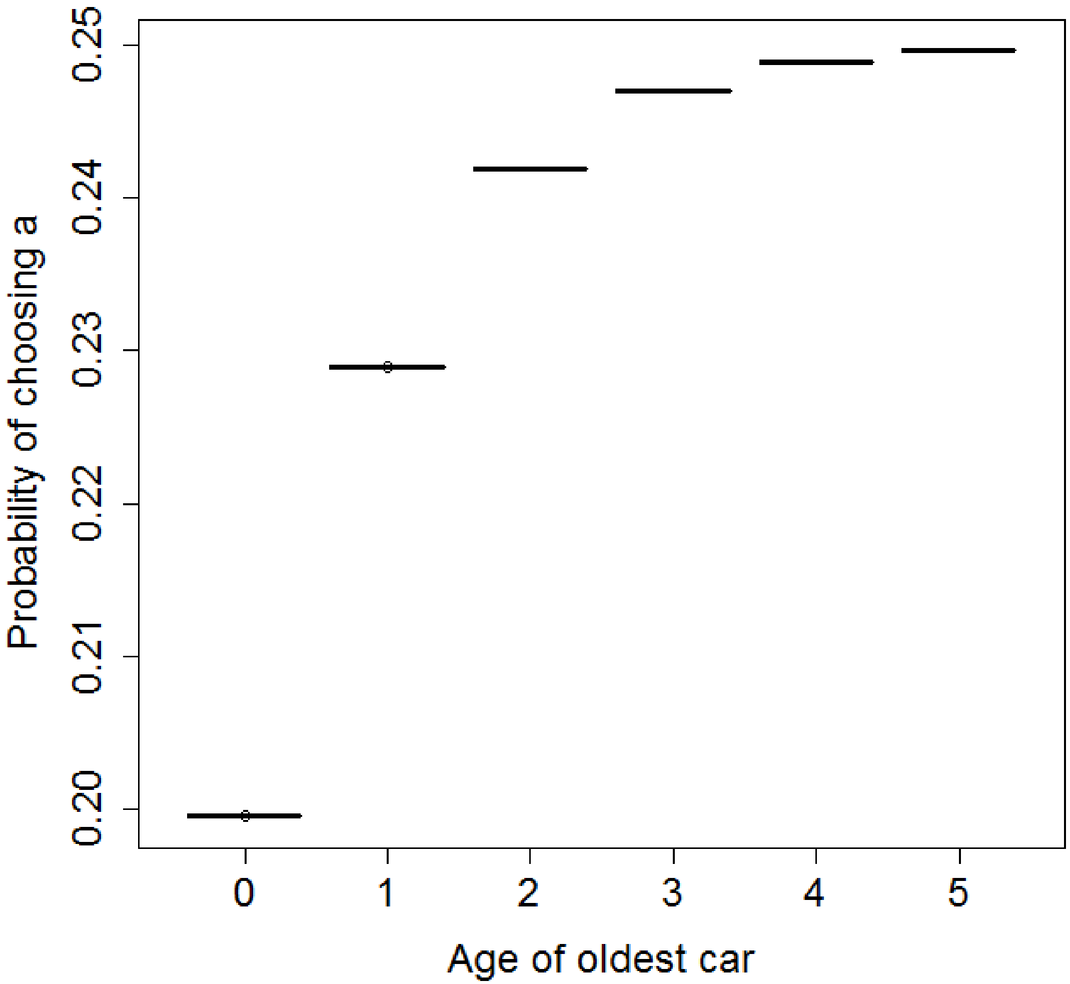
Preliminary results for the value iteration algorithm:

Program:

- Prototype in Matlab
- 15 hours on 8-core server

Graph:

- Assume an initial action $a = [3, 1, 0, 0, 0, 0, 0, 0, 0]$ $P(a)$ vs age of oldest car



CONCLUSION AND FUTURE WORKS

Conclusion:

- First results of value function
- Shows feasibility of problem

Future works:

- Speed up computation of integrated value function ⟹ C++
- Sequential choice of two vehicles if computational time still too large
- Exploratory analysis to specify instantaneous utility
- Implement outer loop (maximization algorithm)
- 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

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Thanks!