A DYNAMIC DISCRETE-CONTINUOUS CHOICE MODEL FOR CAR OWNERSHIP AND USAGE

ESTIMATION PROCEDURE

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

14th Swiss Transport Research Conference
15th May 2014

OUTLINE

- Introduction
- Contributions
- Model specification & estimation
- Swedish registers
- Estimation on synthetic & real data
- Conclusion and future works

INTRODUCTION

Objective

Model households' simultaneous choices of car ownership, usage and fuel type, assuming they are forward-looking.

INTRODUCTION

Objective

Model households' simultaneous choices of car ownership, usage and fuel type, assuming they are forward-looking.

Motivations for a dynamic model

- Governmental policies to reduce carbon emissions / car usage (e.g. congestion taxes, independence of fossil fuels,…)
- Technology changes (e.g. increase of alternative-fuel vehicles)
- Variations in economic factors (e.g. fuel price changes)

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

INTRODUCTION

Objective

Model households' simultaneous choices of car ownership, usage and fuel type, assuming they are forward-looking.

Motivations for a dynamic model

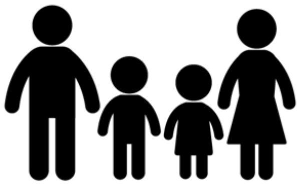
- Governmental policies to reduce carbon emissions / car usage (e.g. congestion taxes, independence of fossil fuels,…)
- Technology changes (e.g. increase of alternative-fuel vehicles)
- Variations in economic factors (e.g. fuel price changes)

Case study: Swedish registers of vehicles and individuals

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

INTRODUCTION



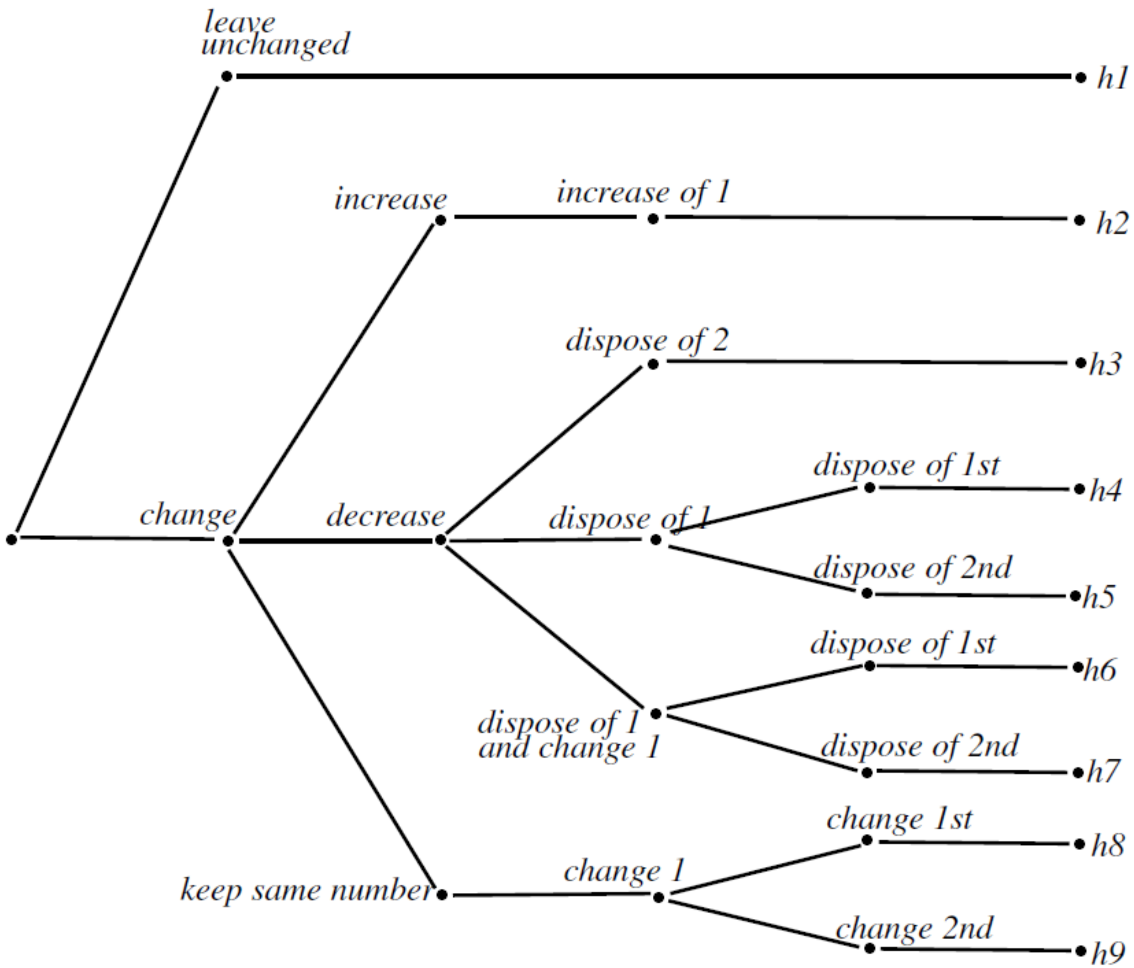
Objective

Model households' simultaneous choices of car ownership, usage and fuel type, assuming they are forward-looking.

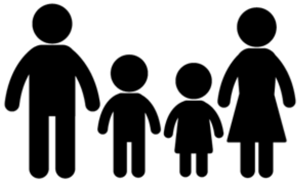
Transaction type × [Annual driving distance × Fuel type] # cars

INTRODUCTION

Transaction type (0-, 1-, 2-car households)



INTRODUCTION

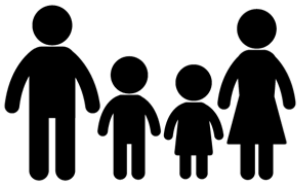


Objective

Model households' simultaneous choices of car ownership, usage and fuel type, assuming they are forward-looking.

Transaction type × [Annual driving distance × Fuel type] # cars

INTRODUCTION



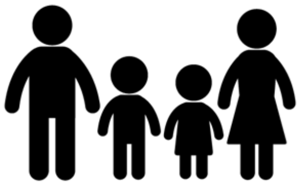
Objective

Model households' simultaneous choices of car ownership, usage and fuel type, assuming they are forward-looking.

Transaction type × [Annual driving distance × Fuel type] # cars

Discrete variables

INTRODUCTION



Objective

Model households' simultaneous choices of car ownership, usage and fuel type, assuming they are forward-looking.

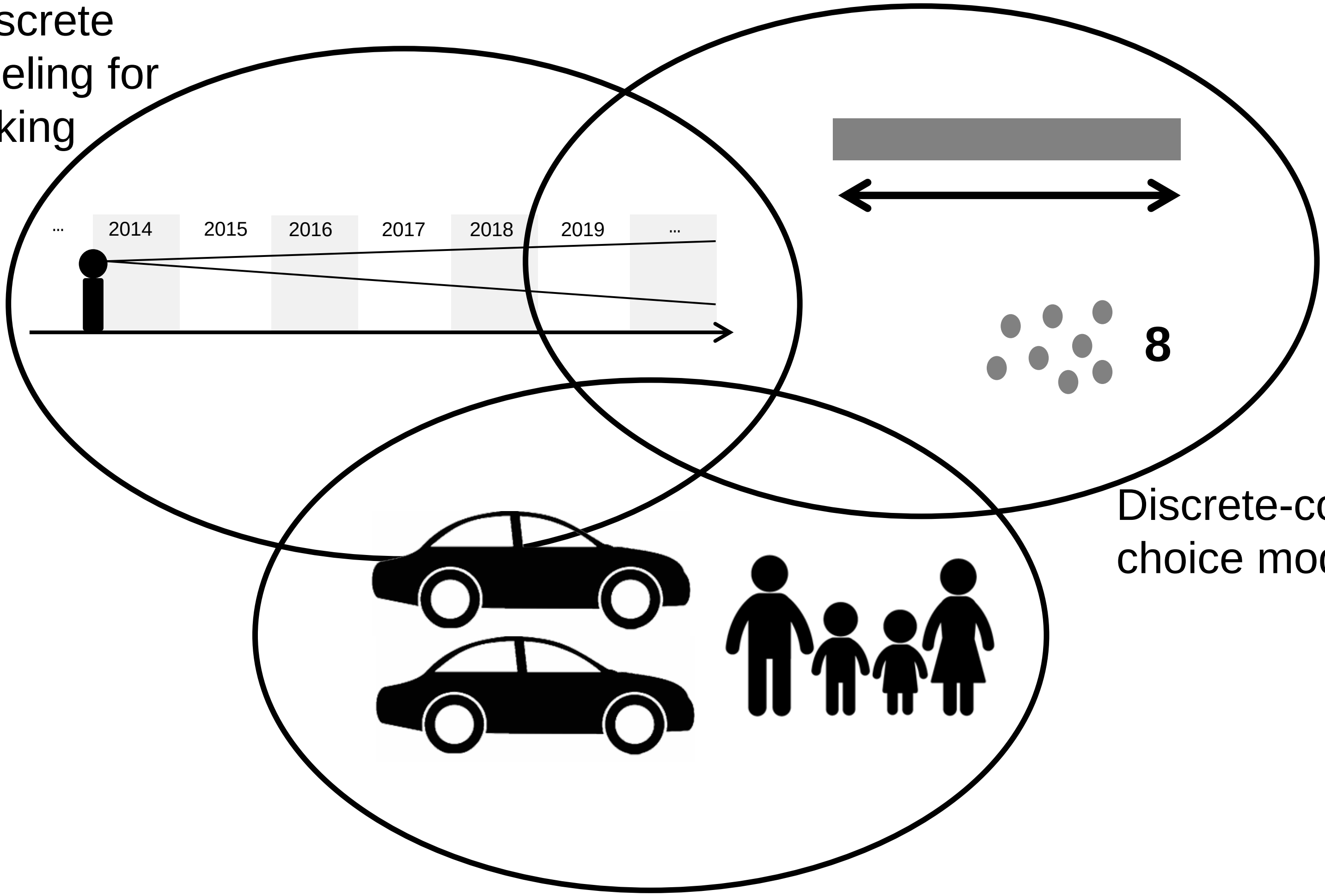
Transaction type × [Annual driving distance × Fuel type] # cars

Continuous variables

CONTRIBUTIONS

The method brings together 3 complex aspects of demand modeling

Dynamic discrete choice modeling for forward-looking agents

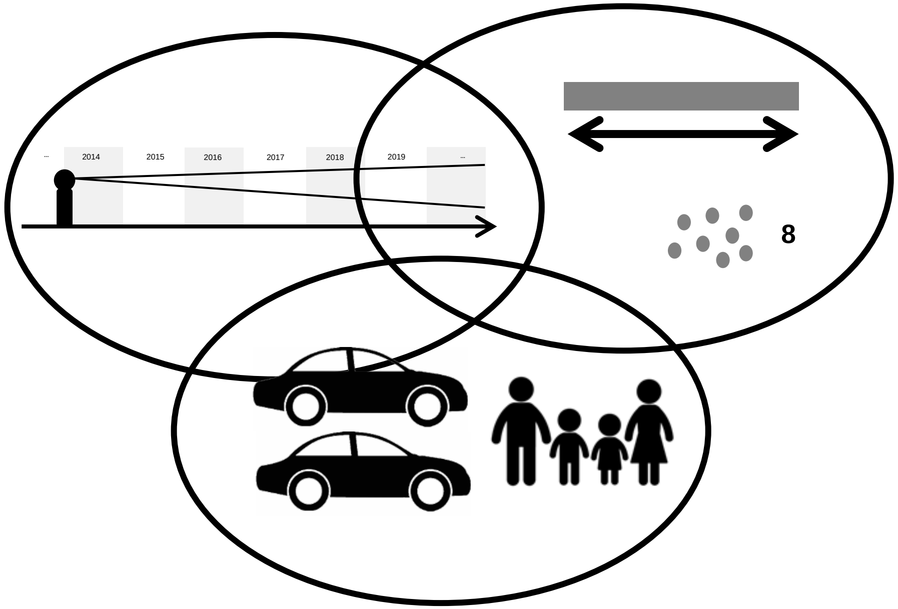


Discrete-continuous choice modeling

Household decisions for multiple (car) holdings

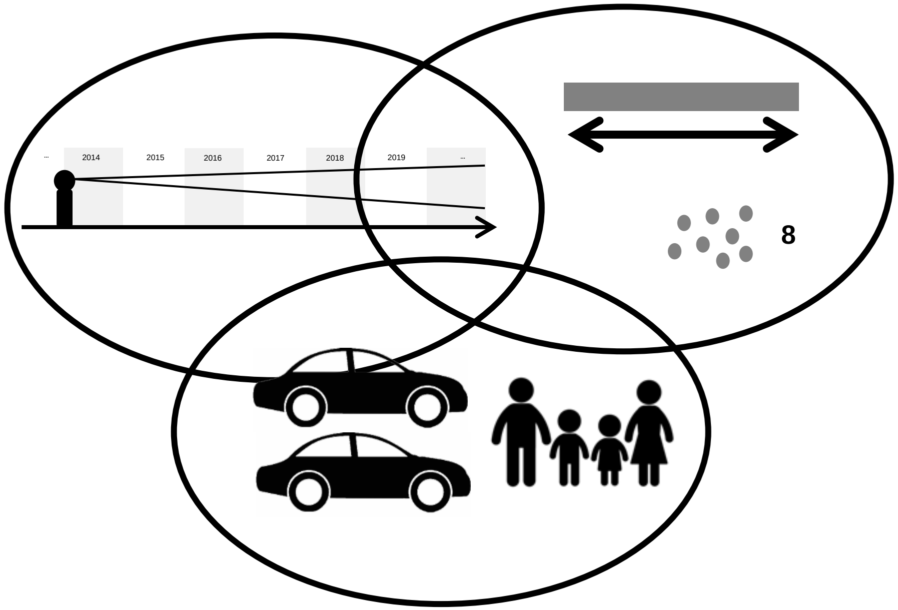
CONTRIBUTIONS

The method brings together 3 complex aspects of demand modeling



CONTRIBUTIONS

The method brings together 3 complex aspects of demand modeling



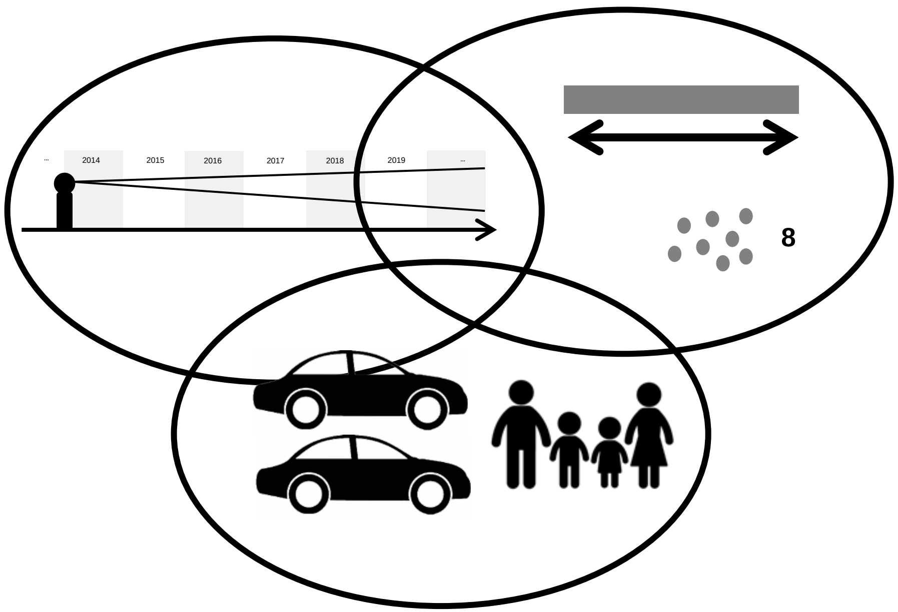
Methodology

Address these issues by applying **dynamic discrete-continuous choice model (DDCCM)**

- Discrete-continuous choice model
- Embedded into a dynamic programming framework

CONTRIBUTIONS

The method brings together 3 complex aspects of demand modeling



Methodology

Address these issues by applying **dynamic discrete-continuous choice model (DDCCM)**

- Discrete-continuous choice model
- Embedded into a dynamic programming framework

Application example

Large **register data** of all **individuals** and **cars** in Sweden

- Approach validated on synthetic data
- Estimation on real data

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL SPECIFICATION

ASSUMPTIONS

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

2. **Strategic choice** of:
    - Transaction
    - Fuel type(s)

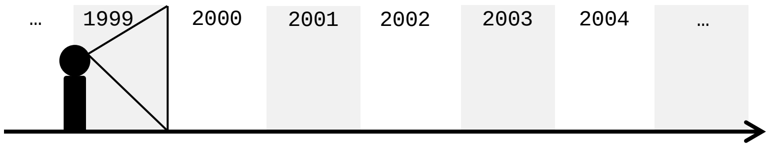
⟹ Account for forward-looking behavior of households

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

MODEL SPECIFICATION

ASSUMPTIONS

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)

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

MODEL SPECIFICATION

DEFINITION OF THE COMPONENTS

Components of the DDCCM

- Agent
- Time step
- State space
- Action space
- Transition rule
- Instantaneous utility function

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL SPECIFICATION

DEFINITION OF THE COMPONENTS

Components of the DDCCM

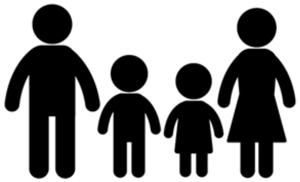
- Agent
- Time step
- State space
- Action space
- Transition rule
- Instantaneous utility function

Fundamental components in a dynamic programming framework

MODEL SPECIFICATION

DEFINITION OF THE COMPONENTS

- Agent: household



- Time step *t*: year

- State space *S*

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

Age – 1st car; Fuel type – 1st car; Age – 2nd car; Fuel type – 2nd car

- Action space *A*

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

Transaction type; Mileage – 1st car; Fuel type – 1st car; Mileage – 2nd car; Fuel type – 2nd car

- Transition rule: deterministic rule: each state s_{t+1} can be inferred exactly once s_t and a_t are known.

MODEL SPECIFICATION

DEFINITION OF THE COMPONENTS

- Agent: household

- Time step t: year

- State space S

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

Age – 1st car; Fuel type – 1st car; Age – 2nd car; Fuel type – 2nd car

- Action space A

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

Transaction type; Mileage – 1st car; Fuel type – 1st car; Mileage – 2nd car; Fuel type – 2nd car

- Transition rule: deterministic rule: each state s_{t+1} can be inferred exactly once s_t and a_t are known.

MODEL SPECIFICATION

DEFINITION OF THE COMPONENTS

- Agent: household
- Time step *t*: year

- State space *S*

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

Age – 1st car; Fuel type – 1st car; Age – 2nd car; Fuel type – 2nd car

- Action space *A*

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

Transaction type; Mileage – 1st car; Fuel type – 1st car; Mileage – 2nd car; Fuel type – 2nd car

- Transition rule: deterministic rule: each state s_{t+1} can be inferred exactly once s_t and a_t are known.

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL SPECIFICATION

DEFINITION OF THE COMPONENTS

- Agent: household

- Time step *t*: year

- State space *S*

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

$y_{1,t}$: Age – 1st car; $f_{1,t}$: Fuel type – 1st car; $y_{2,t}$: Age – 2nd car; $f_{2,t}$: Fuel type – 2nd car

- Action space *A*

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

h_t: Transaction type; $\tilde{m}_{1,t}$: Mileage – 1st car; $\tilde{f}_{1,t}$: Fuel type – 1st car; $\tilde{m}_{2,t}$: Mileage – 2nd car; $\tilde{f}_{2,t}$: Fuel type – 2nd car

- Transition rule: deterministic rule: each state s_{t+1} can be inferred exactly once s_t and a_t are known.

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL SPECIFICATION

DEFINITION OF THE COMPONENTS

- Agent: household

- Time step *t*: year

- State space *S*

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

Age – 1st car; Fuel type – 1st car; Age – 2nd car; Fuel type – 2nd car

- Action space *A*

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

Transaction type; Mileage – 1st car; Fuel type – 1st car; Mileage – 2nd car; Fuel type – 2nd car

- Transition rule: deterministic rule: each state s_{t+1} can be inferred exactly once s_t and a_t are known.

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL SPECIFICATION

- **Instantaneous utility function**

Deterministic utility

$$v(s_t, a_t^C, a_t^D, x_t, \theta) = \underbrace{v_t^D(s_t, a_t^D, x_t, \theta)}_{\text{Utility for the acquisition}} + \underbrace{v_t^C(s_t, a_t^D, a_t^C, x_t, \theta)}_{\text{Utility of driving}}$$

Constant elasticity of substitution (CES) utility function

$$v_t^C(s_t, a_t^D, a_t^C, x_t, \theta) = (m_{g,t}^{\rho} + m_{d,t}^{\rho})^{\frac{1}{\rho}}$$

Choice probability

$$P(a_{n,t}^D | s_{n,t}, x_{n,t}, \theta) = \frac{v_{n,t}^D + v_{n,t}^{C*} + \overbrace{\beta \sum_{s_{n,t+1} \in S} \bar{V} f}^{\text{Expected discounted utility}}}{\sum_{a_{n,t}^{\tilde{D}}} \left\{ v_{n,t}^D + v_{n,t}^{C*} + \beta \sum_{s_{n,t+1} \in S} \bar{V} f \right\}}$$

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL ESTIMATION

- Parameters obtained by maximizing likelihood:

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

- 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

MODEL ESTIMATION

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)

Reasons of step 1.

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL ESTIMATION

1. Finding the optimal value(s) of mileage (e.g. 2-car households with different fuel types)

- Maximization of the continuous utility: $\max_{m_{g,t}, m_{d,t}} v_t^C$

  s.t. $p_{g,t} m_{g,t} + p_{d,t} m_{d,t} = Inc_t$

- Find analytical solutions $m_{g,t}^*$ and $m_{d,t}^*$.
- 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 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 $\bar{V}$

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

MODEL ESTIMATION

1. Finding the optimal value(s) of mileage (e.g. 2-car households with different fuel types)

- Maximization of the continuous utility: $\max_{m_{g,t}, m_{d,t}} v_t^C$

  s.t. $p_{g,t} m_{g,t} + p_{d,t} m_{d,t} = Inc_t$

- Find analytical solutions $m_{g,t}^*$ and $m_{d,t}^*$.
- 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 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 $\bar{\boldsymbol{V}}$

SWEDISH REGISTERS

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
  - Vehicle **characteristics** (make, model, fuel consumption, fuel type, age)
  - **Annual mileage** from odometer readings
  - Privately-owned cars, cars from privately-owned company and **company** cars
  - Car bought **new or second-hand**

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

ESTIMATION ON SYNTHETIC DATA

Approach to validate the model framework

- Generate 5000 observations (households) based on distributions of variables in the Swedish register data
- Generate choice (for each observation) based on postulated parameters (10 different samples generated)
- Estimation of model on 10 samples
- Approach validated once postulated parameters are retrieved

ESTIMATION ON SYNTHETIC DATA

Statistics on the Swedish register

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]	0 year	12.77	12.57	9.02	8.64	8.57	8.21	7.72	7.16	7.42	5.45
		1 year	11.55	13.53	13.45	9.72	9.18	9.25	9.03	8.64	8.24	8.73
		2 years	11.14	11.80	13.81	13.54	9.89	9.46	9.56	9.45	9.22	9.06
		3 years	9.70	11.70	12.42	14.11	13.71	10.40	9.99	10.21	10.16	10.28
		4 years	8.90	9.73	11.99	12.46	13.82	13.66	10.62	10.35	10.63	11.01
		5 years	8.42	8.71	9.52	11.52	11.73	13.22	13.21	10.42	10.34	10.90
		6 years	6.70	8.08	8.43	9.06	10.72	11.10	12.64	12.76	10.25	10.43
		7 years	8.42	6.38	7.76	7.98	8.38	10.03	10.49	12.13	12.41	10.23
		8 years	10.18	7.95	6.08	7.30	7.34	7.82	9.42	9.98	11.74	12.30
		9 years	12.20	9.55	7.53	5.66	6.67	6.84	7.31	8.90	9.58	11.61
	Fuel type [% households]	Gasoline	97.23	97.04	97.10	97.05	97.01	96.96	96.91	96.47	95.37	94.44
		Diesel	2.77	2.96	2.90	2.95	2.99	3.04	3.09	3.53	4.63	5.56
Household	Number of cars [% households]	0 car	43.25	42.93	42.97	42.93	42.91	42.86	43.01	45.04	45.15	45.41
		1 car	44.96	44.76	44.69	44.65	44.57	44.44	44.20	42.54	42.35	42.11
		2 cars	11.79	12.30	12.35	12.41	12.52	12.70	12.79	12.42	12.50	12.48
	Income [SEK]	Mean	185'508	197'706	201'695	210'277	214'197	218'315	226'946	232'715	254'452	259'523
		SD	321'885	667'570	631'202	429'462	298'663	237'607	224'982	465'895	338'340	981'006

KTH VETENSKAP OCH KONST

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

ESTIMATION ON SYNTHETIC DATA

Assumptions for the example

- Deterministic utility function

$$v_t^D(s_t, a_t^D, x_t, \theta) = \tau(a_t^D) + \beta_{\text{Age}}(a_t^D, s_t) \cdot \max(\text{Age1}_t, \text{Age2}_t)$$

Transaction costs

Transaction-dependant parameters relative to age of oldest car

- Chose arbitrary values for parameters

ESTIMATION ON SYNTHETIC DATA

		β_{Age}			τ
Transaction name	Case	0 car	1 car	2 cars	all households
h_1: leave unchanged		0	0	0	0
h_2: increase 1		0	0	-	-3
h_3: dispose 2		-	-	0	0
h_4: dispose 1st	1st car is oldest	-	1.5	1.5	0
	2nd car is oldest	-	-	0	0
h_5: dispose 2nd	1st car is oldest	-	-	0	0
	2nd car is oldest	-	-	1.5	0
h_6: dispose 1st and change 2nd		-	-	0	-4
h_7: dispose 2nd and change 1st		-	-	0	-4
h_8: change 1st	1st car is oldest	-	1.5	1.5	-4
	2nd car is oldest	-	-	0	-4
h_9: change 2nd	1st car is oldest	-	-	0	-4
	2nd car is oldest	-	-	1.5	-4

ESTIMATION ON SYNTHETIC DATA

Outcomes from synthetic data

	Rho				Age							
					1 car Dispose/change				2 cars Dispose/change oldest car			
Run	Value	SD	t-test 0	t-test true value	Value	SD	t-test 0	t-test true value	Value	SD	t-test 0	t-test true value
Synthetic data	0.49	0.03	14.10	-0.29	1.50	0.04	34.07	0.03	1.46	0.04	38.47	-0.93
True value	0.5				1.5				1.5			
Initial value	0.5				1.5				1.5			
Real data	0.56	0.35	1.61	-	-0.52	0.04	-12.81	-	-0.17	0.05	-3.78	-
Initial value	0.5				-0.55				-0.19			

	Transaction cost								Neg L
	Increase of 1				Dispose of 1 and change the other / change 1				
Run	Value	SD	t-test 0	t-test true value	Value	SD	t-test 0	t-test true value	
Synthetic data	-3.01	0.06	-50.74	-0.13	-3.99	0.05	-79.56	0.29	0.92
True value	-3				-4				
Initial value	-3				-4				
Real data	-4.44	0.15	-30.13	-	-1.37	0.07	-18.49	-	7.17
Initial value	-4.3				-1.33				

Tolerance synthetic data: 0.01

Tolerance real data: 0.8

KTH VETENSKAP OCH KONST

EPFL ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

ESTIMATION ON REAL DATA

Outcomes from real data

- Subsample of 446 households from merged registers of individuals and cars in Sweden
- 3431 observations

ESTIMATION ON REAL DATA

Outcomes from real data

Run	Rho				Age: 1 car Dispose/change				Age: 2 cars Dispose/change oldest car			
	Value	SD	t-test 0	t-test true value	Value	SD	t-test 0	t-test true value	Value	SD	t-test 0	t-test true value
Synthetic data	0.49	0.03	14.10	-0.29	1.50	0.04	34.07	0.03	1.46	0.04	38.47	-0.93
True value	0.5				1.5				1.5			
Initial value	0.5				1.5				1.5			
Real data	0.56	0.35	1.61	-	-0.52	0.04	-12.81	-	-0.17	0.05	-3.78	-
Initial value	0.5				-0.55				-0.19			

Run	Transaction cost: Increase of 1				Transaction cost: Dispose of 1 and change the other / change 1				Neg L
	Value	SD	t-test 0	t-test true value	Value	SD	t-test 0	t-test true value	
Synthetic data	-3.01	0.06	-50.74	-0.13	-3.99	0.05	-79.56	0.29	0.92
True value	-3				-4				
Initial value	-3				-4				
Real data	-4.44	0.15	-30.13	-	-1.37	0.07	-18.49	-	7.17
Initial value	-4.3				-1.33				

Tolerance synthetic data: 0.01
Tolerance real data: 0.8

CONCLUSION AND FUTURE WORKS

Contributions

Integrate three complex aspects of demand

- Forward-looking decision-makers
- Discrete-continuous choice: both fixed and operational costs are accounted for.
- Household decisions for multiple-car fleets

Next steps

- Further specification testing on the subsample of real data from Swedish registers
- Improvement of the optimization 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

Thank you!

MODEL ESTIMATION

Reasons of step 1.

Likelihood for the full model

$$L = \int P(D|C)f(C)dC$$

D = discrete variables
C = continuous variables