

---

# A DYNAMIC DISCRETE-CONTINUOUS CHOICE MODEL OF CAR OWNERSHIP, USAGE AND FUEL TYPE

---

Workshop session on 'discrete-continuous models of car ownership and use'  
University of Leeds  
12th November 2013

Aurélie Glerum EPFL

Emma Frejinger Université de Montréal

Anders Karlström KTH

Muriel Beser Hugosson KTH

Michel Bierlaire EPFL



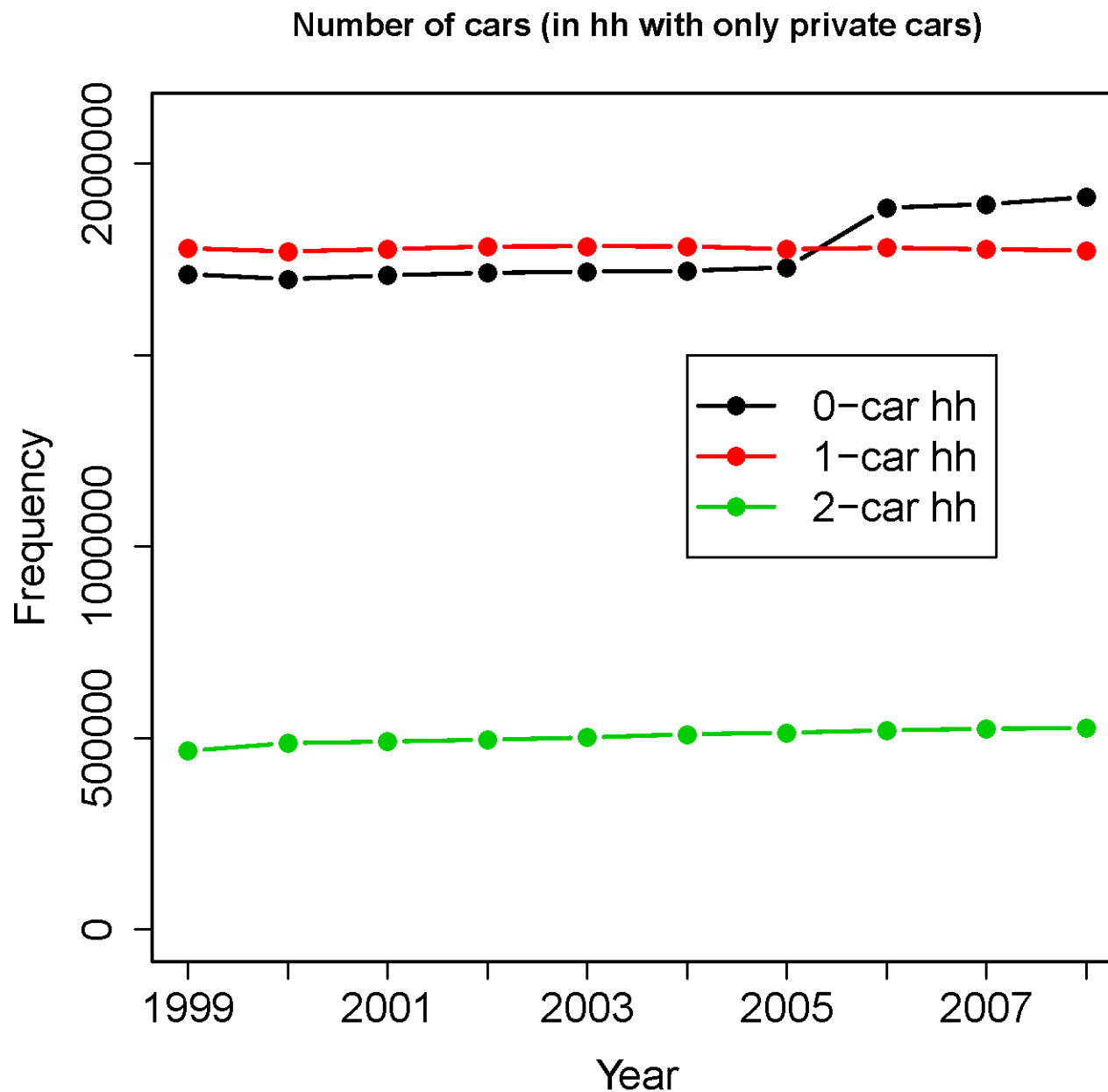
- Introduction
- Background and data
- Dynamic discrete-continuous choice model
- Illustrative example
- Estimation on synthetic data
- Conclusion and future works

## 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

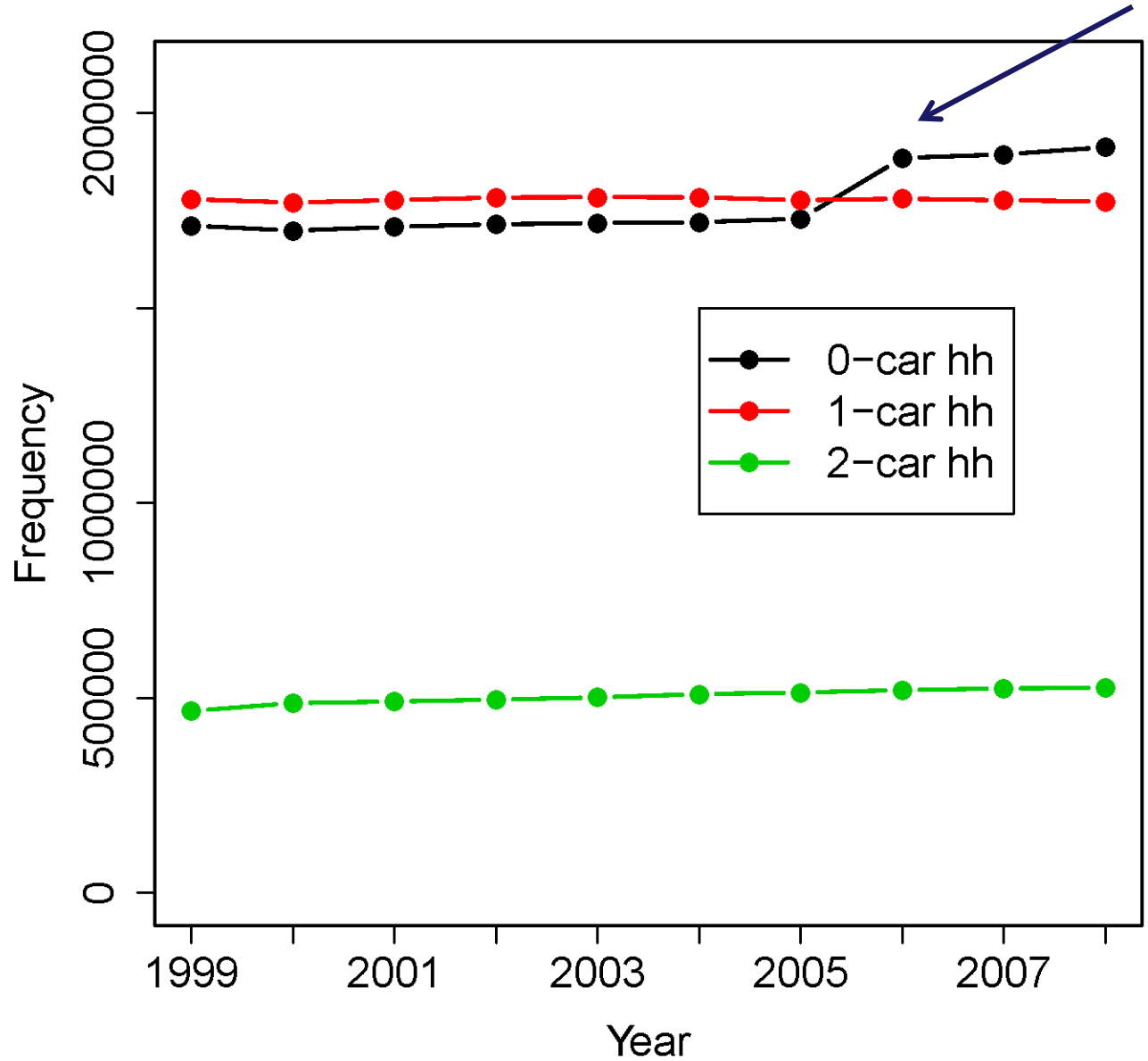
## Register data of Swedish population and car fleet:

- **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**



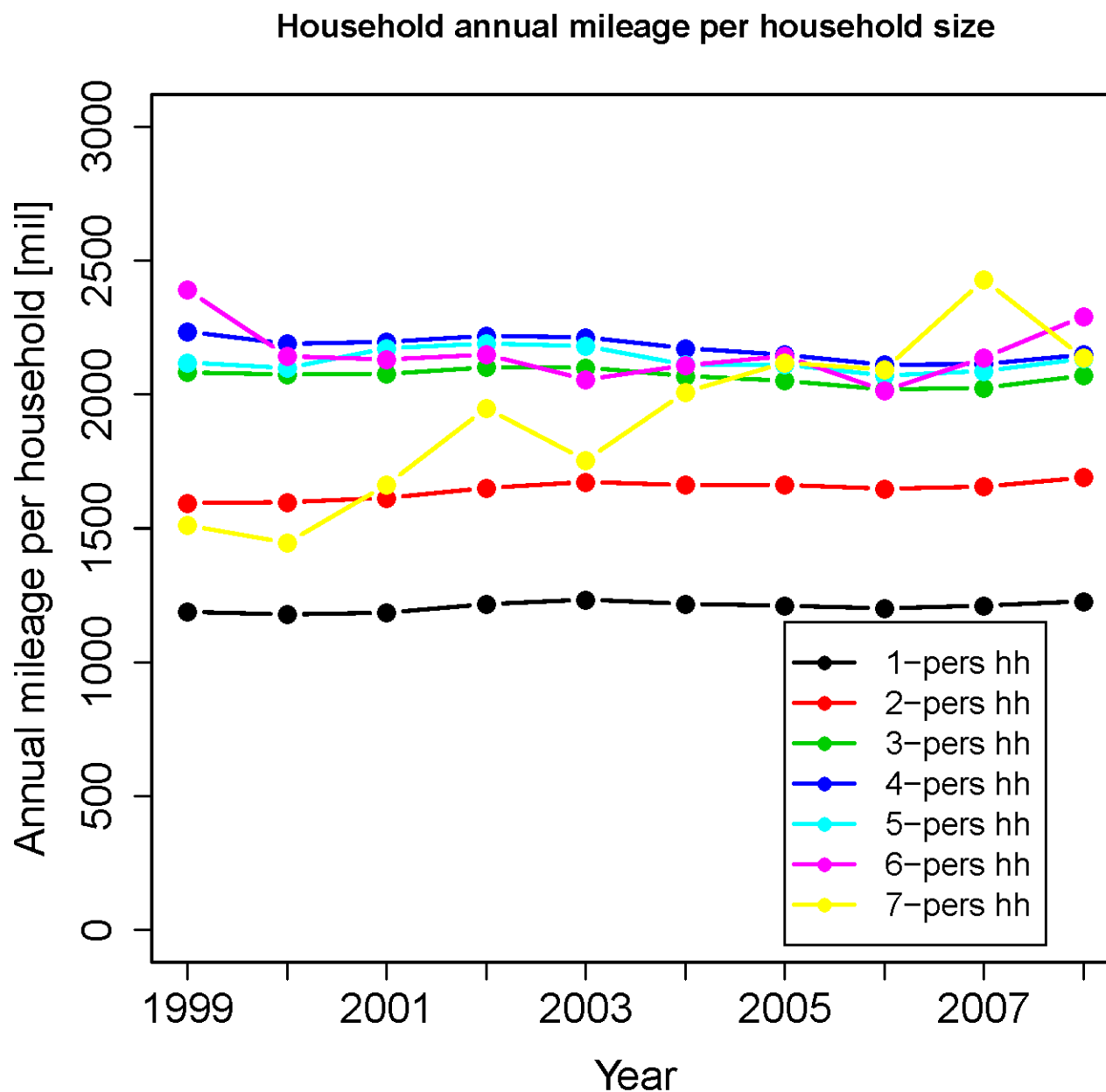
# BACKGROUND AND DATA

Number of cars (in hh with only private cars)



OWNERSHIP  
Congestion tax  
Stockholm





- **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**

- 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**

- 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

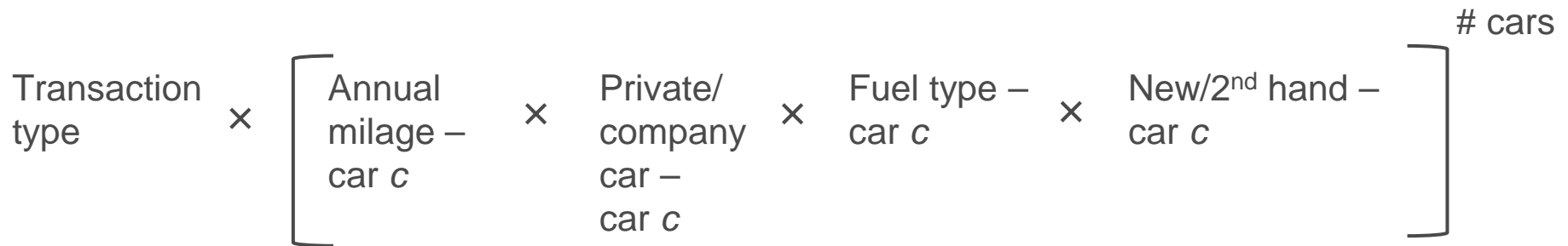
## MAIN FEATURES

- 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.

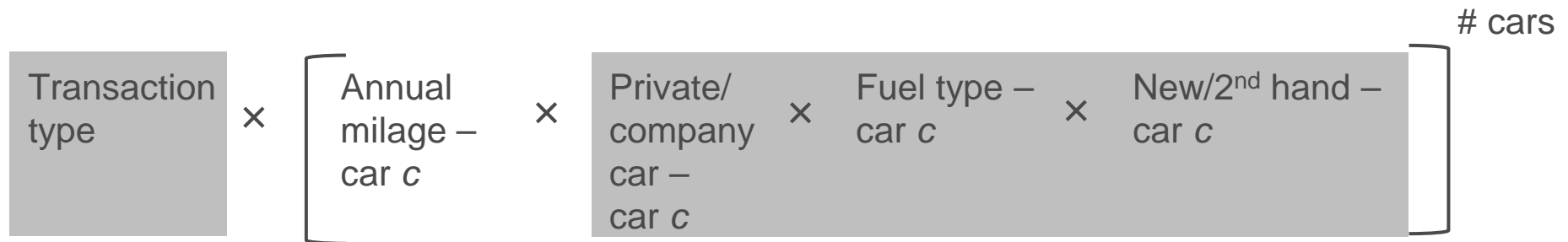
**In details:** model **simultaneous choice** of



### Objective

Model simultaneously car ownership, usage and fuel type.

**In details:** model **simultaneous choice** of

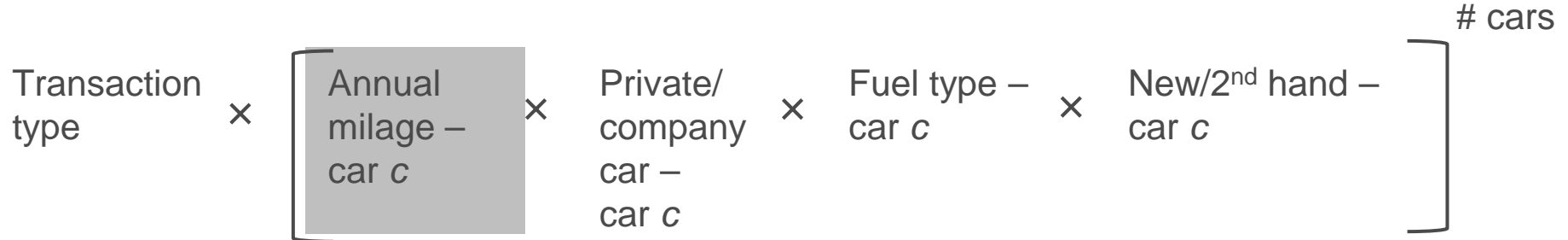


**Discrete variables**

### Objective

Model simultaneously car ownership, usage and fuel type.

**In details:** model **simultaneous choice** of



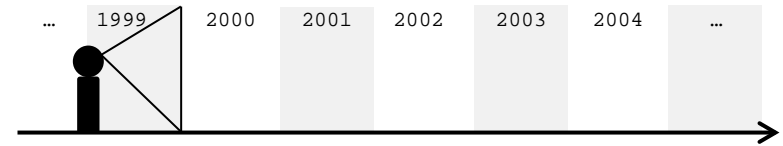
### Continuous variables

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 2<sup>nd</sup>-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)



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

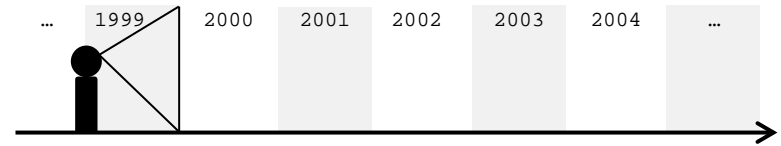
### Strategic choice (**dynamic** case)



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

### 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)

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

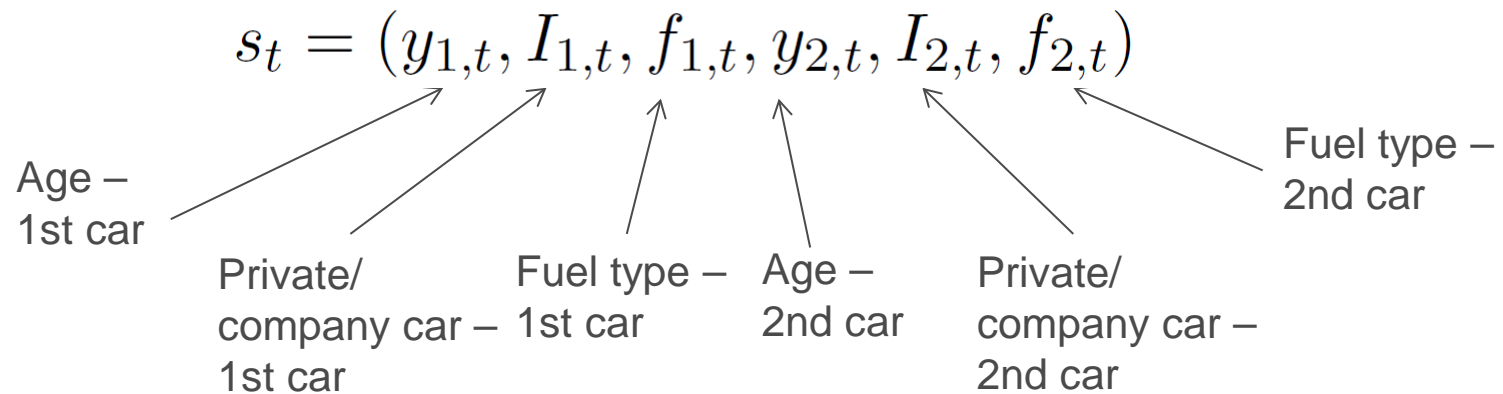


Embeds a **choice model** into a **dynamic programming framework**

### Components of the DDCCM:

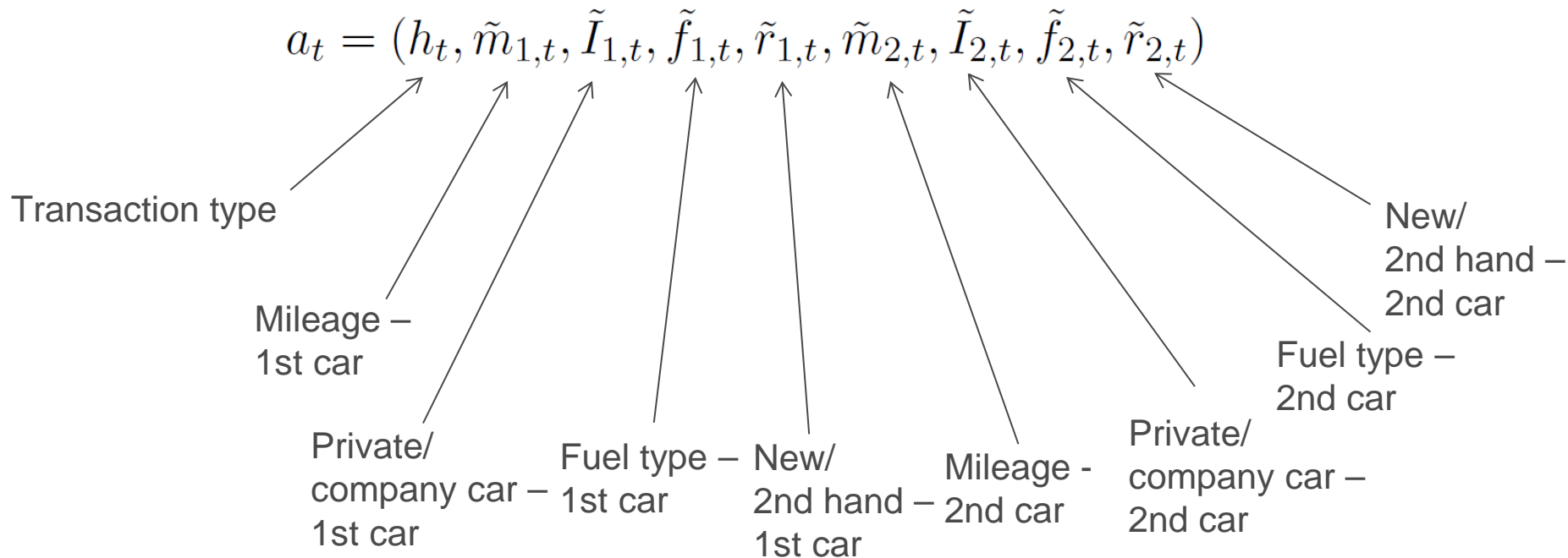
- Agent
- Time step
- State space
- Action space
- Transition rule
- Instantaneous utility function

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



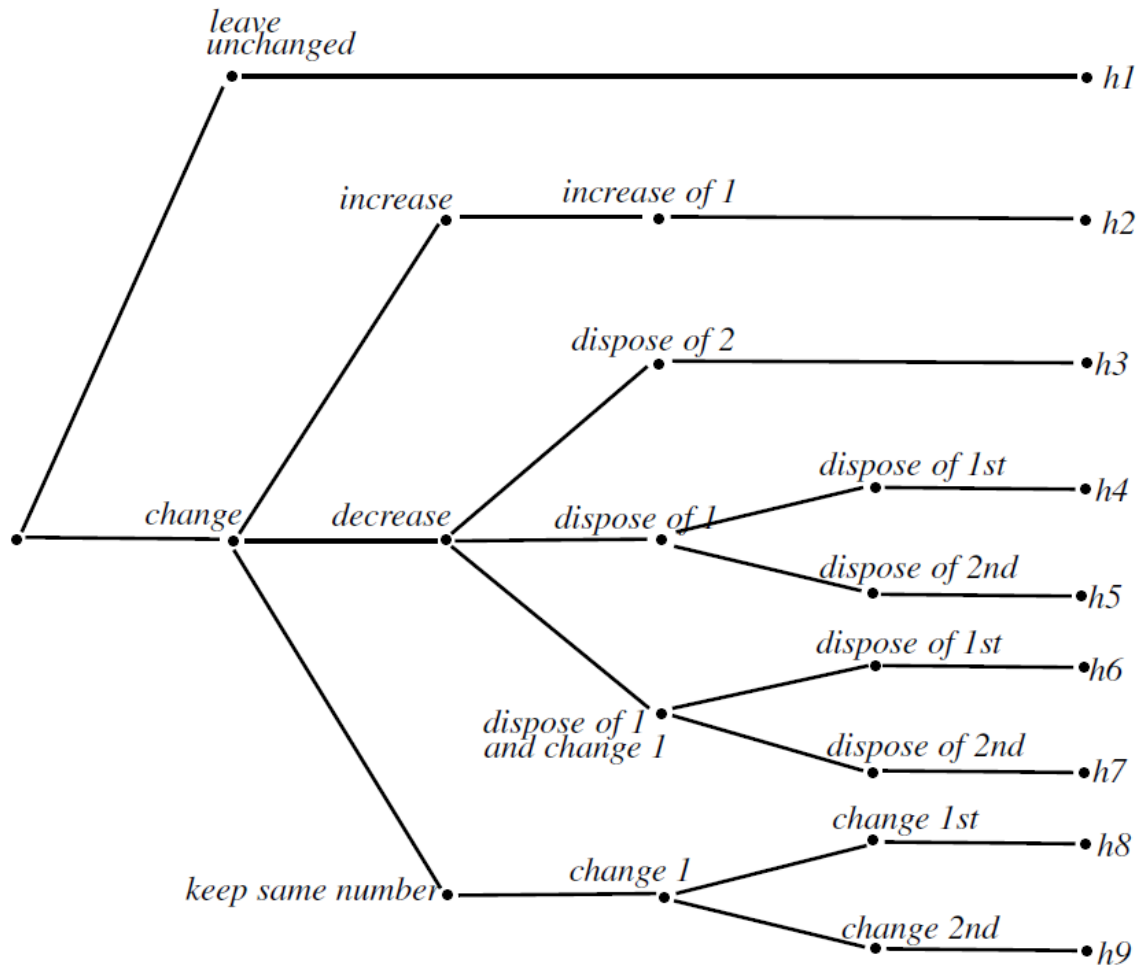
## DEFINITION OF THE COMPONENTS

- Action space  $A$



- Action space  $A$

*Transaction types*



## 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.

## 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.
- **Instantaneous utility function:**

$$u(s_t, a_t^C, a_t^D, x_t, \theta) = \underbrace{v(s_t, a_t^C, a_t^D, x_t, \theta)}_{\text{Deterministic term}} + \underbrace{\varepsilon_D(a_t^D)}_{\text{Random term for discrete choices}}$$

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

$$v(s_t, a_t^C, a_t^D, x_t, \theta) = \underbrace{v_t^D(s_t, a_t^D, x_t, \theta)}_{\text{Utility for discrete actions}} + \underbrace{v_t^C(s_t, a_t^D, a_t^C, x_t, \theta)}_{\text{Utility for continuous actions}}$$



- Instantaneous utility function

- Utility for continuous actions:

**Constant elasticity of substitution (CES)** utility function (e.g. Zabalza, 1983):

$$v_t^C(s_t, a_t^D, a_t^C, x_t, \theta) = (m_{1,t}^{-\rho} + \alpha \cdot m_{2,t}^{-\rho})^{-1/\rho}$$

- 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**:

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

- 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
- 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:  $\max_{m_{1,t}, m_{2,t}} v_t^C$   
 s.t.  $p_{1,t}m_{1,t} + p_{2,t}m_{2,t} = \text{Inc}_t$

- 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

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

Constant for  
households with  
at least one car

Transaction  
costs

Transaction-dependant  
parameters relative to age  
of oldest car

- Chose arbitrary values for parameters

# ILLUSTRATIVE EXAMPLE

Transaction name	Case	$\beta_{Age}$			$\tau$
		0 car	1 car	2 cars	all households
$h_1$ : leave unchanged		0	-1	-1	0
$h_2$ : increase 1		0	0	-	-3
$h_3$ : dispose 2		-	-	1	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

$C = 5$ , for 1- or 2-car households

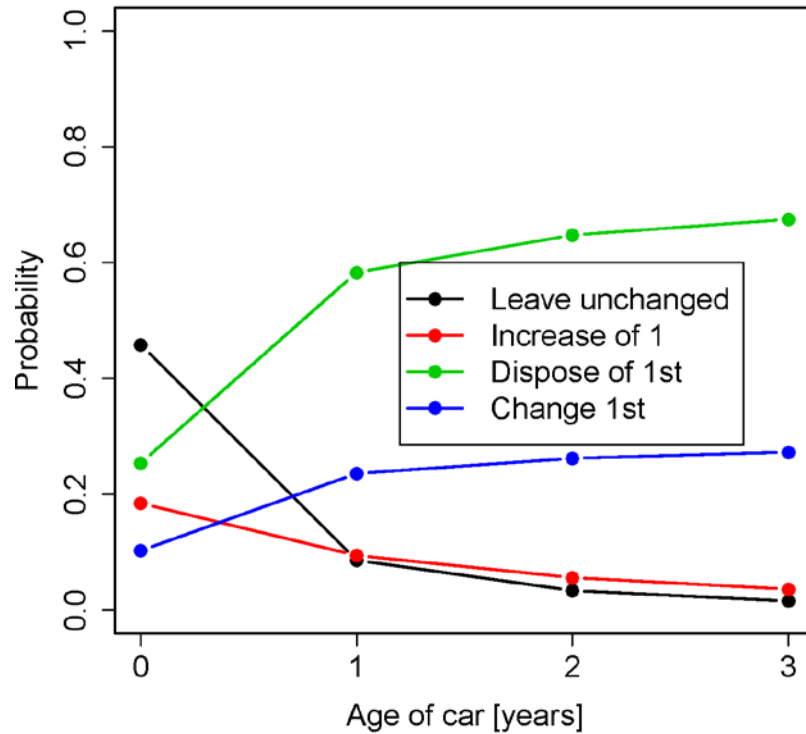
# ILLUSTRATIVE EXAMPLE

## Assumptions for the example:

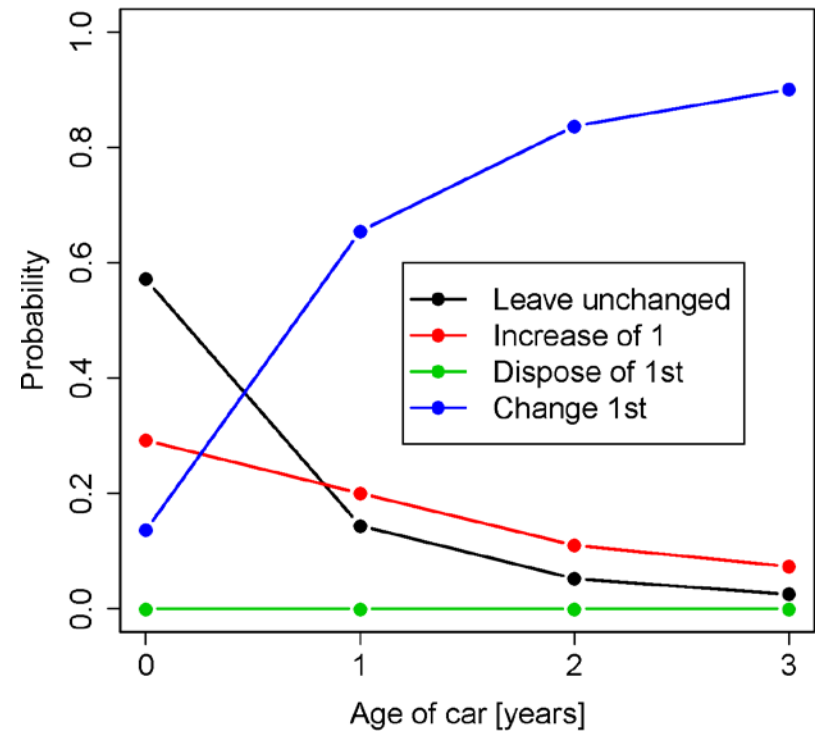
- Visualize choice probabilities for one observation:
  - 1-car household
  - Annual income = 530'000 SEK ( $\approx$  60'200 €, 50'300 GBP)
  - 8% expenses on fuel



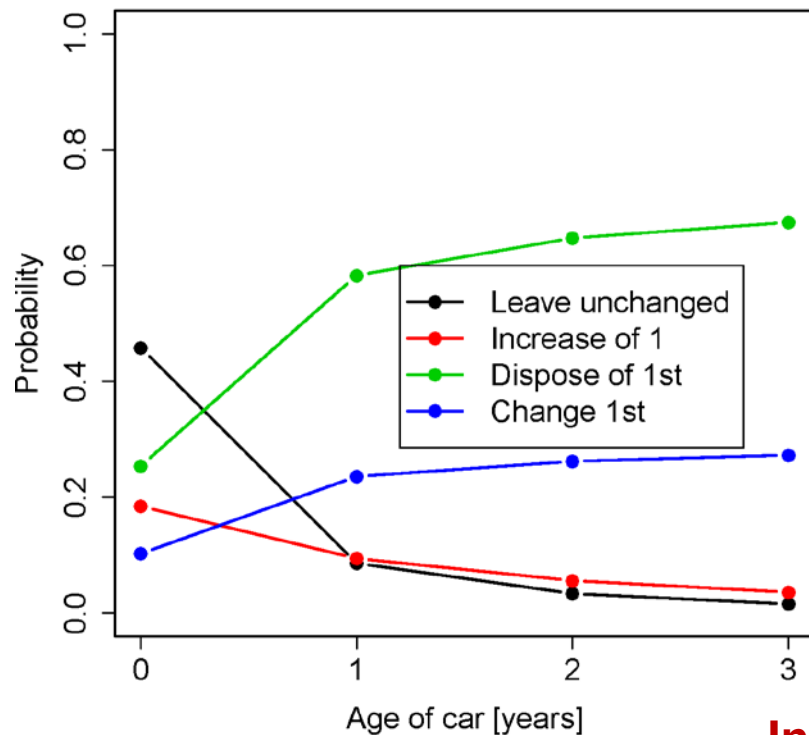
## Static



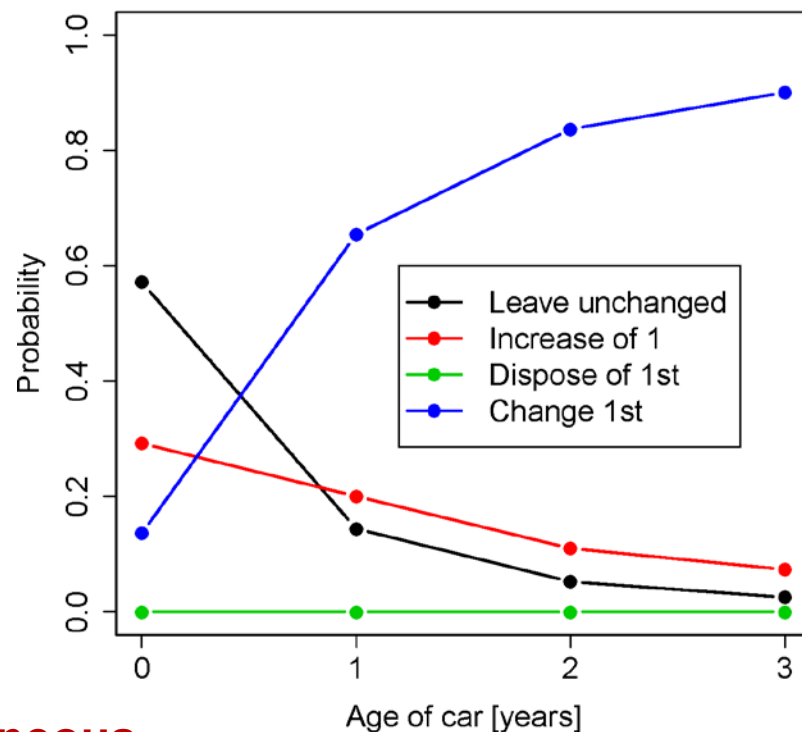
## Dynamic



## Static



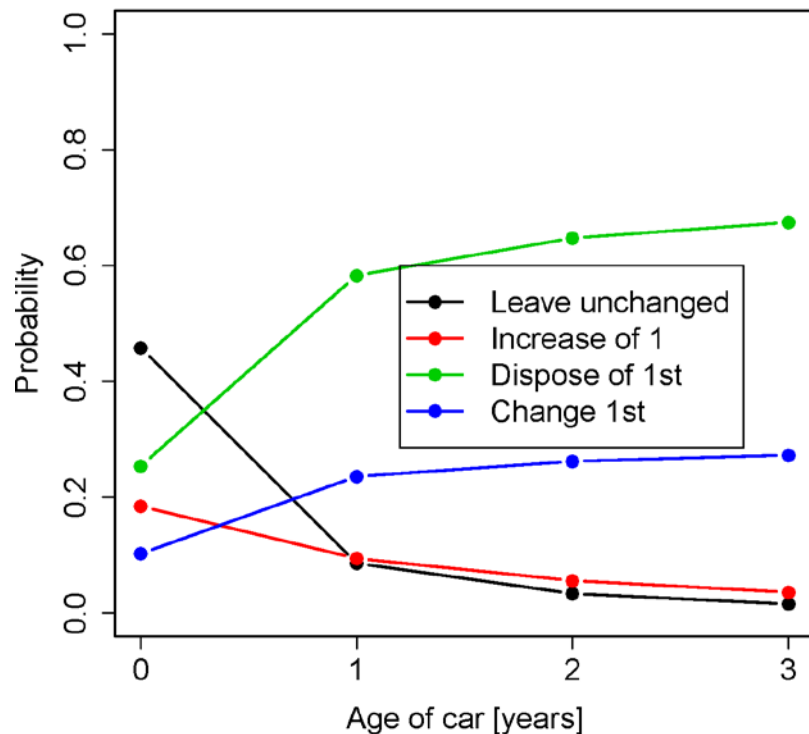
## Dynamic



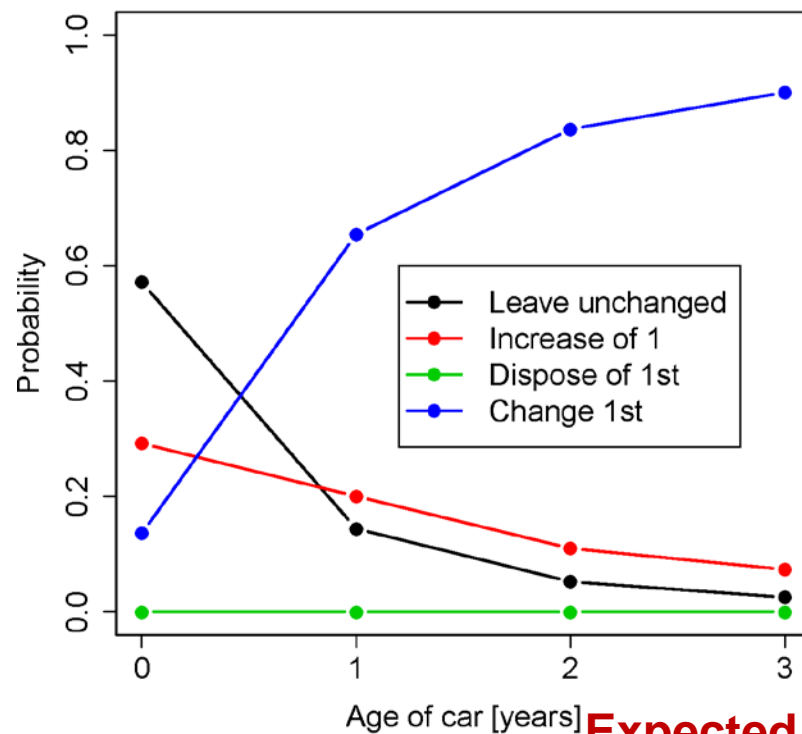
**Instantaneous utility**

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

## Static



## Dynamic



**Expected  
discounted  
utility**

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

## Approach to validate the model framework

- Generate 500 observations 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**

	Rho	Alpha	Beta_Age			
			1 car		2 cars	
			Leave unchanged	Dispose/change	Leave unchanged	Dispose of 2
<b>Fixed parameters</b>	<b>0.5</b>	<b>0.3</b>	<b>-1</b>	<b>1.5</b>	<b>-1</b>	<b>1</b>

# ESTIMATION ON SYNTHETIC DATA

	Rho	Alpha	Beta_Age			
			1 car		2 cars	
			Leave unchanged	Dispose/change	Leave unchanged	Dispose of 2
Fixed parameters	0.5	0.3	-1	1.5	-1	1

Run	Beta_Age 2 cars				Tau								Neg LL
	Dispose/change oldest car				Increase of 1				Dispose of 1 and change the other / change 1				
	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	
	1	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	
2	1.25	0.28	4.40	-0.90	-3.23	0.53	-6.04	-0.43	-4.15	0.36	-11.53	-0.43	1277.08
3	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	-0.24	1254.81
4	1.46	0.28	5.23	-0.14	-3.34	0.54	-6.21	-0.63	-4.06	0.36	-11.38	-0.16	1238.02
5	1.32	0.26	5.05	-0.68	-3.61	0.51	-7.10	-1.21	-4.02	0.34	-11.77	-0.04	1242.86
6	1.33	0.24	5.50	-0.72	-3.40	5.60	-0.61	-0.07	-3.96	2.94	-1.34	0.02	1253.79
7	1.27	0.29	4.42	-0.80	-3.13	0.52	-6.04	-0.25	-4.10	0.35	-11.66	-0.29	1261.26
8	1.31	0.28	4.66	-0.69	-3.30	0.52	-6.30	-0.56	-4.07	0.35	-11.58	-0.20	1256.60
9	1.46	0.27	5.50	-0.15	-3.18	0.83	-3.83	-0.21	-4.17	0.53	-7.88	-0.32	1255.81
10	1.35	0.25	5.33	-0.58	-3.45	0.51	-6.79	-0.89	-4.02	0.34	-11.69	-0.07	1247.57
Average	1.34	0.27			-3.32	1.06			-4.07	0.63			1254.26
Parameters used to generate the data	1.5				-3				-4				

# ESTIMATION ON SYNTHETIC DATA

	Rho	Alpha	Beta_Age			
			1 car		2 cars	
			Leave unchanged	Dispose/change	Leave unchanged	Dispose of 2
Fixed parameters	0.5	0.3	-1	1.5	-1	1

Run	Beta_Age 2 cars				Tau								Neg LL
	Dispose/change oldest car				Increase of 1				Dispose of 1 and change the other / change 1				
	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	
	1	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	
2	1.25	0.28	4.40	-0.90	-3.23	0.53	-6.04	-0.43	-4.15	0.36	-11.53	-0.43	1277.08
3	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	-0.24	1254.81
4	1.46	0.28	5.23	-0.14	-3.34	0.54	-6.21	-0.63	-4.06	0.36	-11.38	-0.16	1238.02
5	1.32	0.26	5.05	-0.68	-3.61	0.51	-7.10	-1.21	-4.02	0.34	-11.77	-0.04	1242.86
6	1.33	0.24	5.50	-0.72	-3.40	0.56	-6.01	-0.07	-3.96	0.29	-11.34	0.02	1253.79
7	1.27	0.29	4.42	-0.80	-3.13	0.52	-6.04	-0.25	-4.10	0.35	-11.66	-0.29	1261.26
8	1.31	0.28	4.66	-0.69	-3.30	0.52	-6.30	-0.56	-4.07	0.35	-11.58	-0.20	1256.60
9	1.46	0.27	5.50	-0.15	-3.18	0.83	-3.83	-0.21	-4.17	0.53	-7.88	-0.32	1255.81
10	1.35	0.25	5.33	-0.58	-3.45	0.51	-6.79	-0.89	-4.02	0.34	-11.69	-0.07	1247.57
Average	1.34	0.27			-3.32	1.06			-4.07	0.63			1254.26
Parameters used to generate the data	1.5				-3				-4				

Parameters retrieved



# ESTIMATION ON SYNTHETIC DATA

	Rho	Alpha	Beta_Age			
			1 car		2 cars	
			Leave unchanged	Dispose/change	Leave unchanged	Dispose of 2
Fixed parameters	0.5	0.3	-1	1.5	-1	1

Run	Beta_Age 2 cars				Tau								Neg LL
	Dispose/change oldest car				Increase of 1				Dispose of 1 and change the other / change 1				
	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	
	1	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	
2	1.25	0.28	4.40	-0.90	-3.23	0.53	-6.04	-0.43	-4.15	0.36	-11.53	-0.43	1277.08
3	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	-0.24	1254.81
4	1.46	0.28	5.23	-0.14	-3.34	0.54	-6.21	-0.63	-4.06	0.36	-11.33	-0.16	1238.02
5	1.32	0.26	5.05	-0.68	-3.61	0.51	-7.10	-1.21	-4.02	0.34	-11.77	-0.04	1242.86
6	1.33	0.24	5.50	-0.72	-3.40	0.60	-0.61	-0.07	-3.96	2.94	-1.34	0.02	1253.79
7	1.27	0.29	4.42	-0.80	-3.13	0.52	-6.04	-0.25	-4.10	0.35	-11.66	-0.29	1261.26
8	1.31	0.28	4.66	-0.69	-3.30	0.52	-6.30	-0.56	-4.07	0.35	-11.53	-0.20	1256.60
9	1.46	0.27	5.50	-0.15	-3.18	0.83	-3.83	-0.21	-4.17	0.53	-7.83	-0.32	1255.81
10	1.35	0.25	5.33	-0.58	-3.45	0.51	-6.79	-0.89	-4.02	0.34	-11.69	-0.07	1247.57
Average	1.34	0.27			-3.32	1.06			-4.07	0.63			1254.26
Parameters used to generate the data	1.5				-3				-4				



Parameters: not significantly different from value used to generate the data





# ESTIMATION ON SYNTHETIC DATA

	Rho	Alpha	Beta_Age			
			1 car		2 cars	
			Leave unchanged	Dispose/change	Leave unchanged	Dispose of 2
Fixed parameters	0.5	0.3	-1	1.5	-1	1

Run	Beta_Age 2 cars				Tau								Neg LL
	Dispose/change oldest car				Increase of 1				Dispose of 1 and change the other / change 1				
	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	
	1	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	
2	1.25	0.28	4.40	-0.90	-3.23	0.53	-6.04	-0.43	-4.15	0.36	-11.53	-0.43	1277.08
3	1.31	0.27	4.90	-0.69	-3.30	0.51	-6.47	-0.58	-4.08	0.35	-11.83	-0.24	1254.81
4	1.46	0.28	5.23	-0.14	-3.34	0.54	-6.21	-0.63	-4.06	0.36	-11.38	-0.16	1238.02
5	1.32	0.26	5.05	-0.68	-3.61	0.51	-7.10	-1.21	-4.02	0.34	-11.77	-0.04	1242.86
6	1.33	0.24	5.50	-0.72	-3.40	0.60	-0.61	-0.07	-3.96	2.94	-1.34	0.02	1253.79
7	1.27	0.29	4.42	-0.80	-3.13	0.52	-6.04	-0.25	-4.10	0.35	-11.66	-0.29	1261.26
8	1.31	0.28	4.66	-0.69	-3.30	0.52	-6.30	-0.56	-4.07	0.35	-11.58	-0.20	1256.60
9	1.46	0.27	5.50	-0.15	-3.18	0.83	-3.83	-0.21	-4.17	0.53	-7.88	-0.32	1255.81
10	1.35	0.25	5.33	-0.58	-3.45	0.51	-6.79	-0.89	-4.02	0.34	-11.69	-0.07	1247.57
Average	1.34	0.27			-3.32	1.06			-4.07	0.63			1254.26
Parameters used to generate the data	1.5				-3				-4				

Parameters: significantly different from 0



## 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:

- Estimation of  $\rho$  and  $\alpha$  on synthetic data
- 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!

