hEART

Destination choice with longitudinal data

Antonin Danalet, Loïc Tinguely,
Matthieu de Lapparent, Michel Bierlaire

Lyngby,
September 10, 2015

Outline

Motivation: Activity-based model for pedestrian facilities

Correcting endogeneity for dynamic discrete choice models

Case study: Catering location choice on EPFL campus

Outline

Motivation: Activity-based model for pedestrian facilities

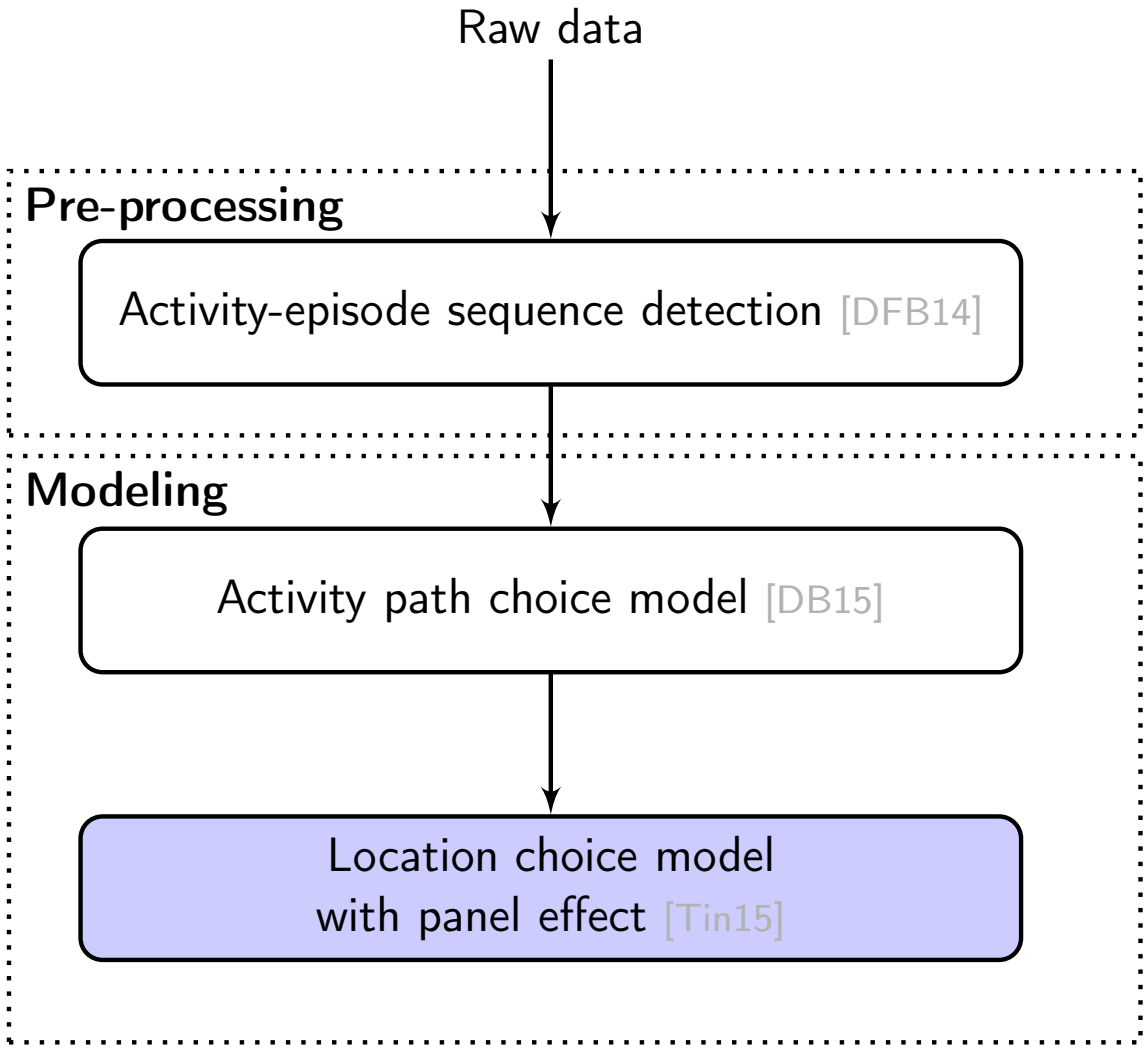
Correcting endogeneity for dynamic discrete choice models

Case study: Catering location choice on EPFL campus

Motivation:
Activity-based model for pedestrian facilities

- **Where, when and for how long** do pedestrians perform activities in public spaces?
- Based on **WiFi traces** from existing access points

Activity path approach



Goal

- Model location choice conditional on an activity type
- Adapted to panel data

Outline

Motivation: Activity-based model for pedestrian facilities

Correcting endogeneity for dynamic discrete choice models

Case study: Catering location choice on EPFL campus

Static model

$$U_{int} = V_{int} + \varepsilon_{int}$$

Ignores two aspects:

- Dynamics
- Serial correlation

Dynamic model without agent effect

$$U_{int} = V_{int} + \rho y_{in(t-1)} + \varepsilon_{int}$$

Assumes

- Dynamic process of order one
- Location-specific dependence
- Previous choice $y_{in(t-1)}$ independent of error term ε_{int}

Relaxing the independence assumption of error terms

- Agent effect α_{in}: time-invariant factor (*"between" individuals variability*)
- Unobserved heterogeneity ε'_{int}: short-term variation of probabilities (*"within" an individual variability*)

$$U_{int} = V_{int} + \rho y_{in(t-1)} + \alpha_{in} + \varepsilon'_{int}$$

Endogeneity issue:

- $y_{in(t-1)}$ and α_{in} are correlated

Wooldridge approach

$$\alpha_{in} = a + by_{in0} + c'\bar{x}_n + \xi_{in}$$

Endogeneity issue solved [Woo05]

Dynamic model with agent effect correction

$$U_{int} = V_{int} + \rho y_{in(t-1)} + a + b y_{in0} + c' \bar{x}_n + \xi_{in} + \varepsilon'_{int}$$

3 different models

Static model	Dynamic model without agent effect	Dynamic model with agent effect
$\rho = 0$	$\rho \neq 0$	$\rho \neq 0$
$a, b, c, \sigma_\alpha^2 = 0$	$a, b, c, \sigma_\alpha^2 = 0$	$a, b, c, \sigma_\alpha^2 \neq 0$

Outline

Motivation: Activity-based model for pedestrian facilities

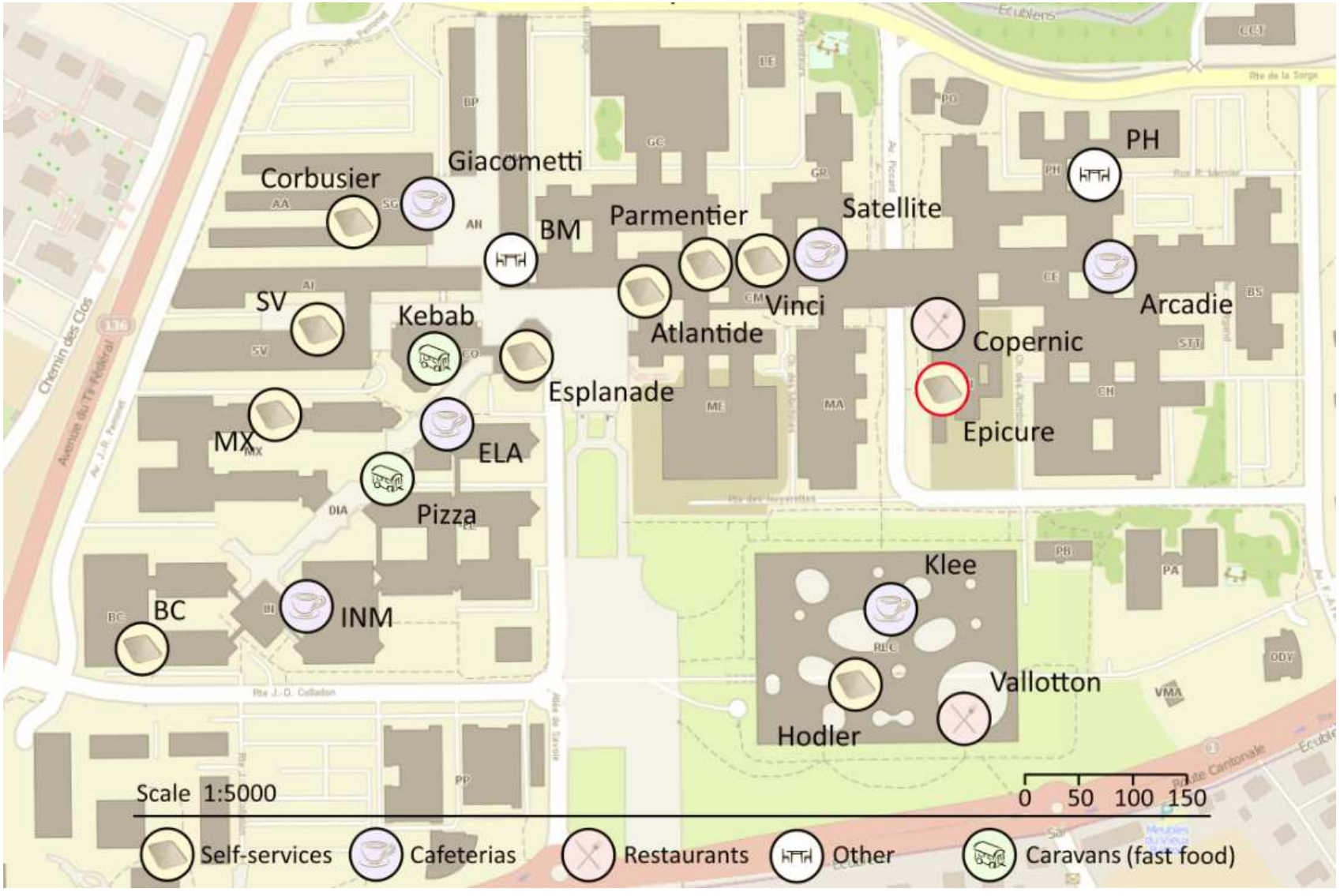
Correcting endogeneity for dynamic discrete choice models

Case study: Catering location choice on EPFL campus

Sequences of activity episodes from WiFi traces

- Bayesian approach [DFB14]
  - detects stops
  - provide semantics
- merging data
  - map information
  - attractivity
  - time constraints

EPFL catering locations



Two specifications for the agent effect

$$\alpha_{in} = a + by_{in0} + \xi_n$$

$$\alpha_{in} = a + by_{in0} + cy_{int}^{\text{count}} + \xi_n$$

4 models estimated

Static model	**Dynamic model without agent effect**	**Dynamic model with agent effect correction**	
		First choice	First choice and frequency
$\rho = 0$	$\rho \neq 0$	$\rho \neq 0$	$\rho \neq 0$
$a = 0$	$a = 0$	$a \neq 0$	$a \neq 0$
$b = 0$	$b = 0$	$b \neq 0$	$b \neq 0$
$c = 0$	$c = 0$	$c = 0$	$c \neq 0$
$\sigma_\alpha^2 = 0$	$\sigma_\alpha^2 = 0$	$\sigma_\alpha^2 \neq 0$	$\sigma_\alpha^2 \neq 0$

Estimation results

- **Distance** has a negative impact
- **Yearly evaluation** has a positive impact
- **Beer after 14:00** has a positive impact
- **Cost** has a negative impact
- **Dinner** has a positive impact
- **Capacity** has a positive impact

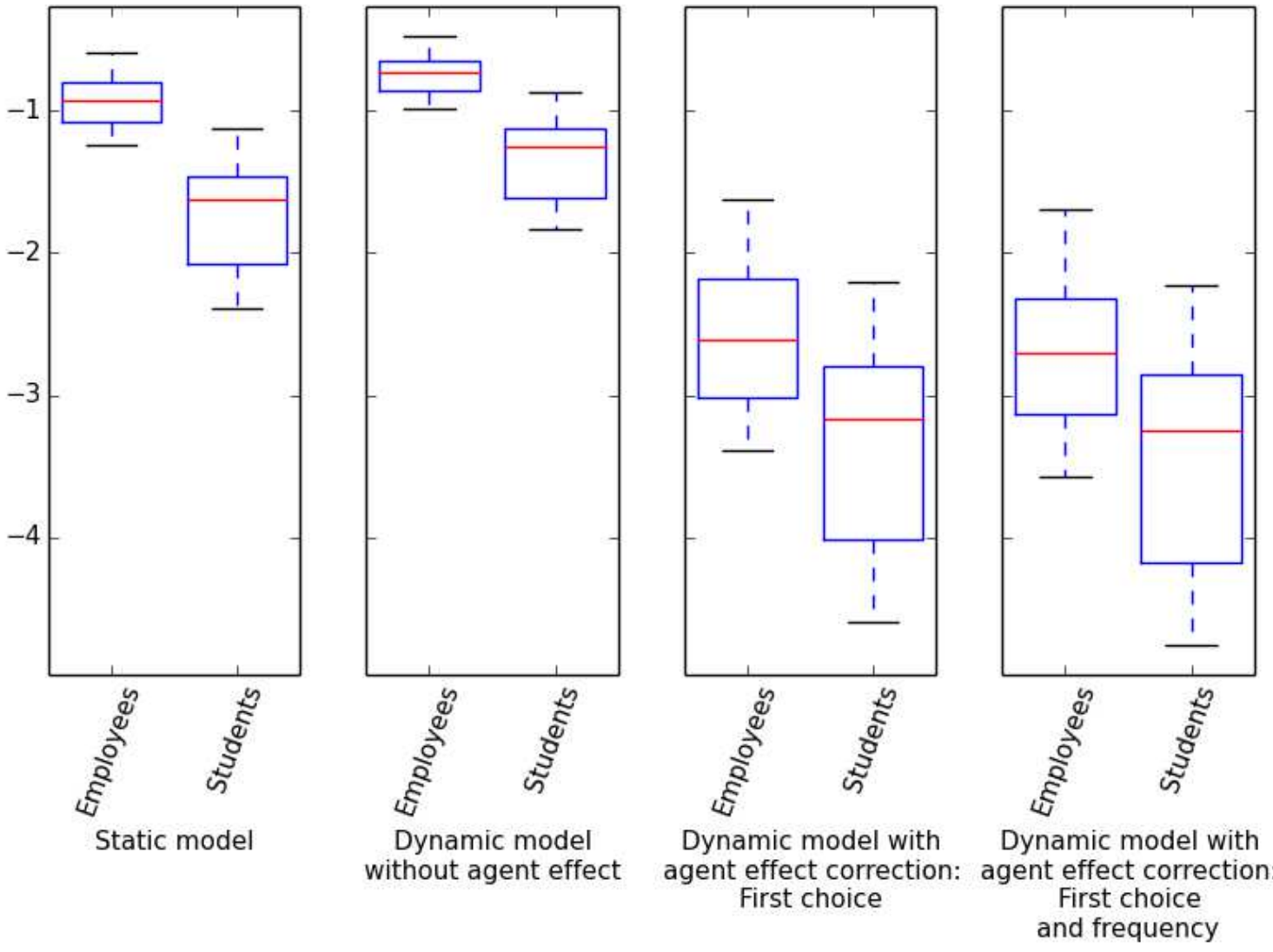
Likelihood ratio tests

Static model	**Dynamic model without agent effect**	**Dynamic model with agent effect correction**	
		First choice	First choice and frequency
354.003 (> 5.99)	920.354 (> 58.12)	16.172 (> 5.99)	

Validation

	Predicting last observations based on past observations			
	Static model	Dynamic model without agent effect	Dynamic model with agent effect correction	
			First choice	First choice and frequency
Sum of the squares of the errors	232.95	204.01	184.16	173.85

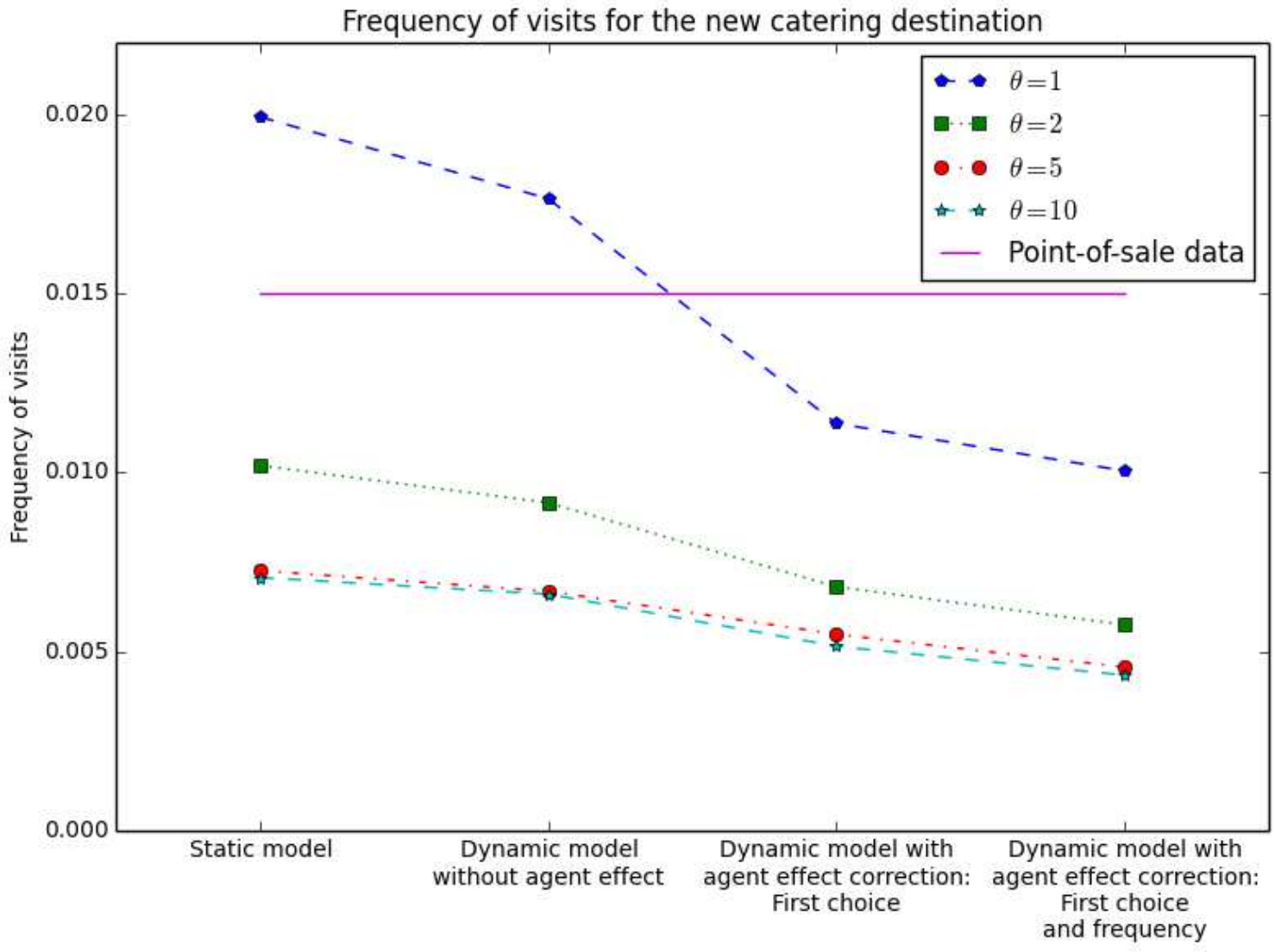
Elasticities to price



Forecasting: opening a new catering location

Nesting structure with the most similar alternative

Forecasting: opening a new catering location



Summary

- Location choice conditional on the choice of activity type
- From WiFi traces
- Including lagged variable
- Correcting for endogeneity
- Applied to case study
  - Cost, capacity, distance, beer, etc. have an impact
  - Validation
  - Elasticity of price
  - Forecasting with new element in the choice set

Conclusion

- WiFi traces allow to understand pedestrian destination choices;
- When having longitudinal data (GSM/GPS/WiFi traces), destination choice models must include Wooldridge correction.

Thank you

hEART:
Destination choice with longitudinal data
Antonin Danalet, Loïc Tinguely,
Matthieu de Lapparent, Michel Bierlaire

– antonin.danalet@epfl.ch

Bibliography I


Antonin Danalet and Michel Bierlaire.
Importance sampling for activity path choice.
In *15th Swiss Transport Research Conference (STRC)*, page 42, Monte Verità, Ascona, Switzerland, 2015.


Antonin Danalet, Bilal Farooq, and Michel Bierlaire.
A Bayesian approach to detect pedestrian destination-sequences from WiFi signatures.
Transportation Research Part C, 44:146–170, 2014.


Loïc Tinguely.
Exploiting pedestrian WiFi traces for destination choice modeling.
Master thesis, EPFL, 2015.

Bibliography II


Jeffrey M. Wooldridge.
Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity.
Journal of Applied Econometrics, 20(1):39–54, 2005.