Parameter estimation for activity-based models

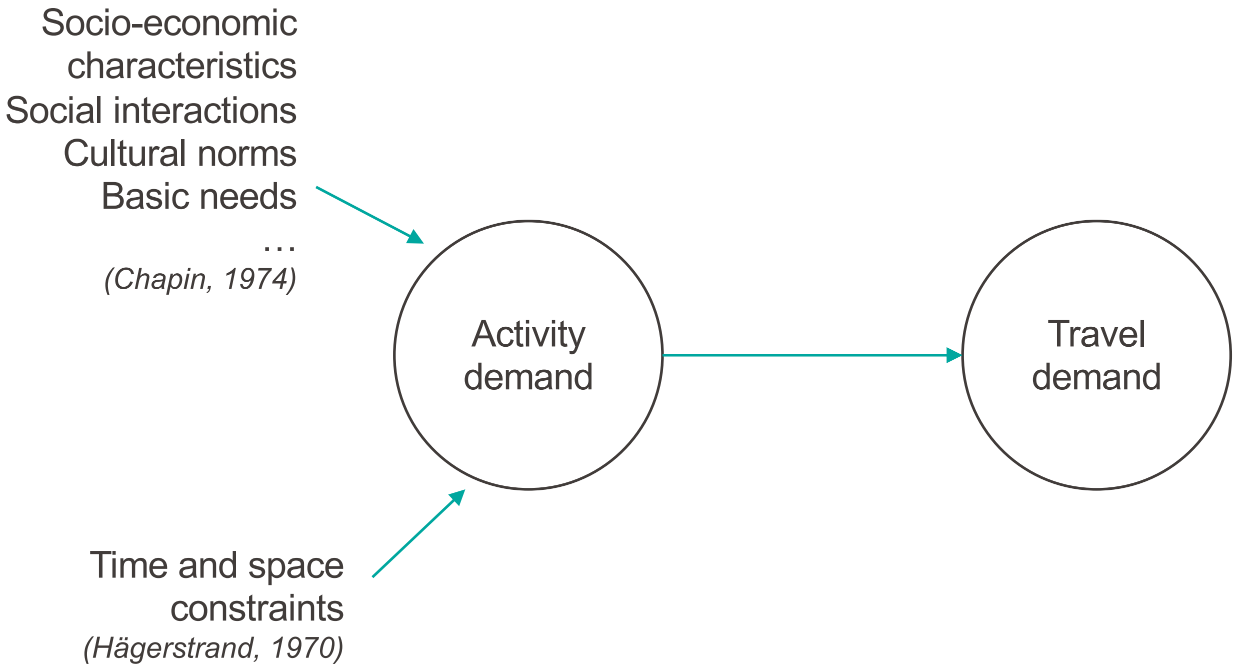
Janody Pougala · Tim Hillel · Michel Bierlaire

EPFL

STRC 2022

TRANSP-OR

Introduction



Introduction

Utility-based models

Decision is made by maximizing utility derived from activities

e.g.

Bowman & Ben-Akiva, 2001
Bhat et al, 2004
Pougala et al, 2021

Rule-based models

Decision is made by considering context-dependent rules

e.g.

Gollegde et al., 1994
Arentze & Timmermans 2000

Introduction

Utility-based models

Decision is made by maximizing utility derived from activities

e.g.

Bowman & Ben-Akiva, 2001
Bhat et al, 2004
Pougala et al, 2021

Rule-based models

Decision is made by considering context-dependent rules

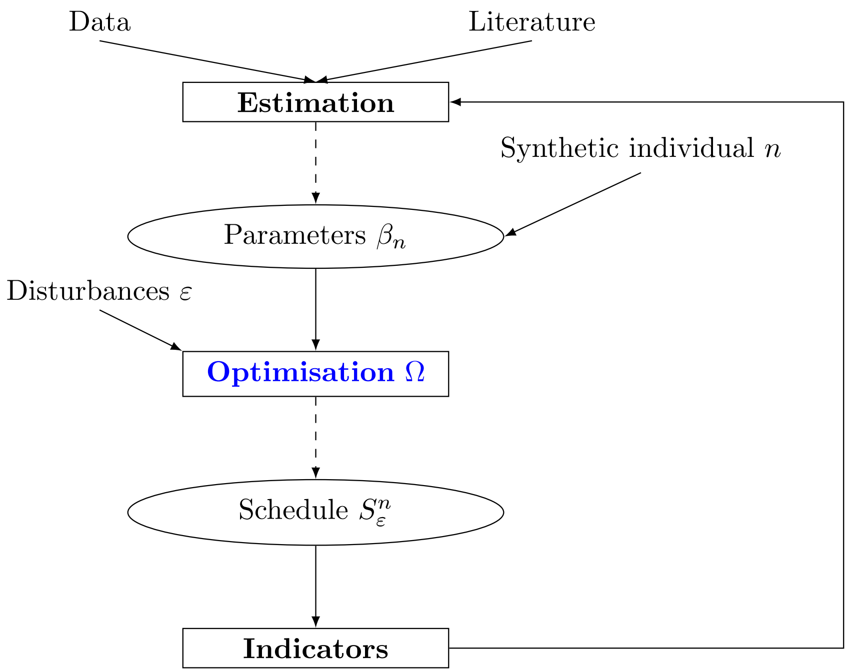
e.g.

Gollegde et al., 1994
Arentze & Timmermans 2000

Background

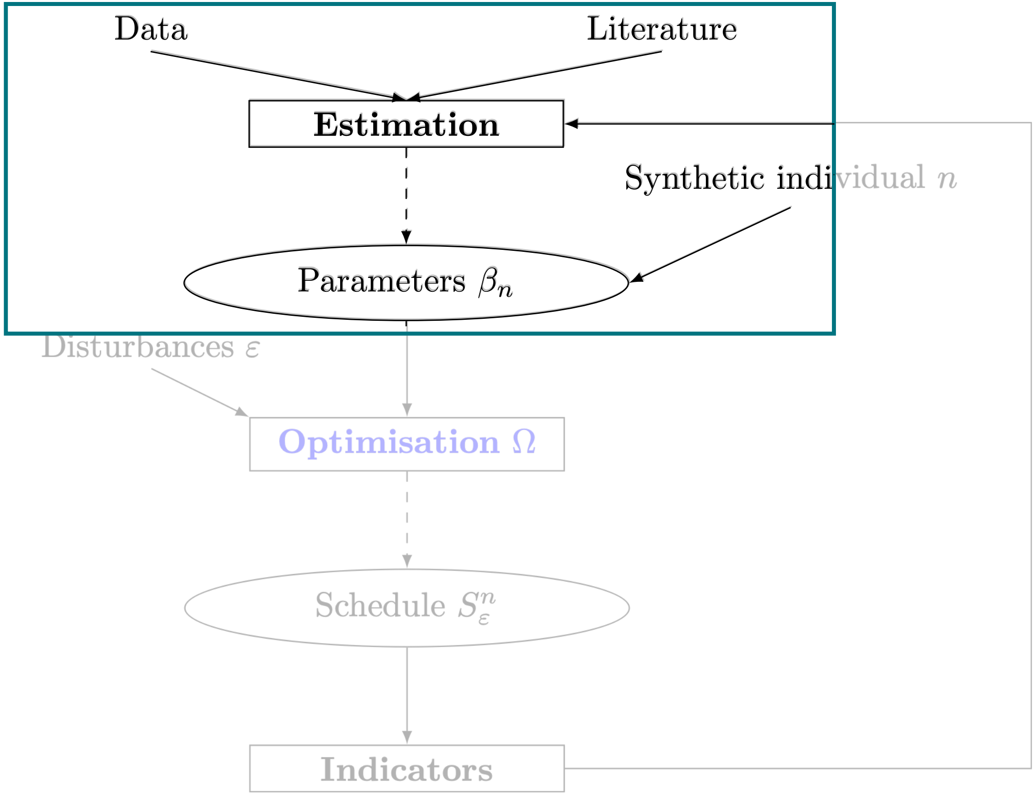
- Optimisation-based simulation framework for activity-based models
  - Pougala et al, 2022

- Joint estimation
  - Activity participation
  - Activity scheduling
  - Mode choice
  - Location choice



Parameter estimation

Parameter estimation



Parameter estimation

- Maximum likelihood estimation (MLE) of parameters in DCM:

$$\hat{\theta} = \arg\max L_n(\theta)$$

$$L_n = \prod_{n=1}^{N} \prod_{i \in C_n} P_n(i)^{y_{in}}$$

Parameter estimation

- Maximum likelihood estimation (MLE) of parameters in DCM:

$$\hat{\theta} = \arg\max L_n(\theta)$$

$$L_n = \prod_{n=1}^{N} \prod_{i \in C_n} P_n(i)^{y_{in}}$$

Enumeration over choice set C_n

- Common assumptions on choice set:
  - Universal across population
  - Fully observed or observable

Parameter estimation

- Maximum likelihood estimation (MLE) of parameters in DCM:

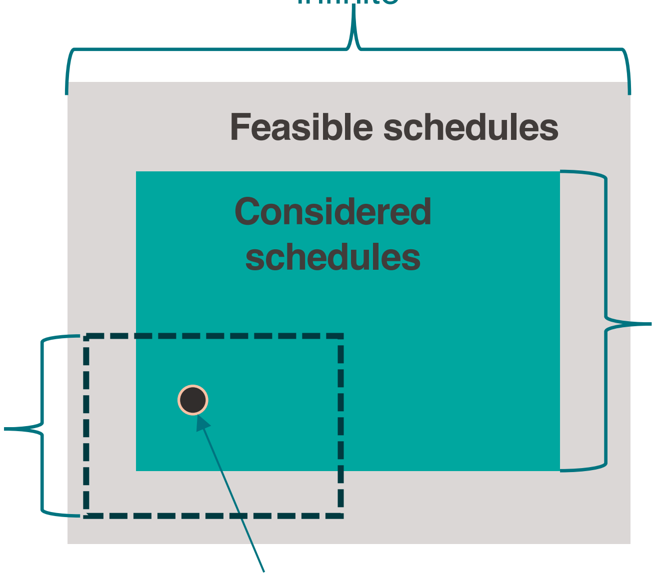
$$\hat{\theta} = \arg\max L_n(\theta)$$

$$L_n = \prod_{n=1}^{N} \prod_{i \in \boldsymbol{C_n^*}} \boldsymbol{P_n(i|C_n^*)^{y_{in}}}$$

- **Sample of alternatives $C_n^* \subset C_n$:**
  - Correction of the choice probabilities (Ben-Akiva & Lerman, 1985)

Choice set

Unobserved and possibly infinite



Actual choice set:
Unobserved

Estimation choice set:
sample of feasible schedules generated for estimation purposes

Realized schedule

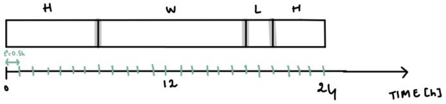
Based on Shocker (1991)

TRANSP-OR

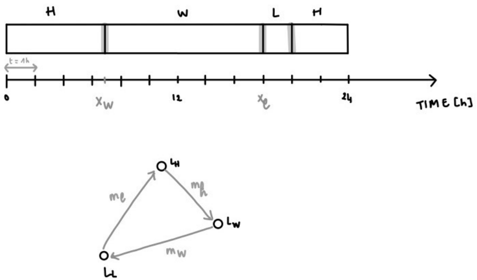
Choice set

- **Choice set generation**
  - Metropolis-Hastings sampling of feasible schedules
  - STRC 2021

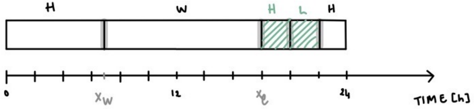
Block



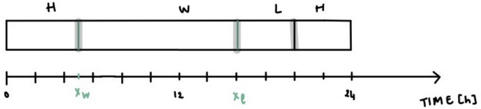
Initial state



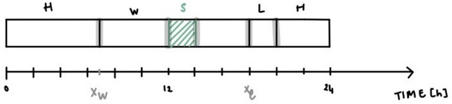
Swap



Inflate/Deflate



Assign



Parameters

- Utility of a schedule: $\mathrm{U_n} = \sum_a U_{an}$
  - For an individual n considering an activity a with a flexibility k:

$$U_{an} = U_{const} + \boxed{U_{early} + U_{late}} + \boxed{U_{long} + U_{short}} + U_{travel} + \varepsilon_{an}$$

Start time deviations:

$$U_{early} = \boldsymbol{\theta_{ek}} \max(0, \boldsymbol{x_a^*} - x_a)$$
$$U_{late} = \boldsymbol{\theta_{lk}} \max(0, x_a - \boldsymbol{x_a^*})$$

Duration deviations:

$$U_{short} = \boldsymbol{\theta_{dsk}} \max(0, \boldsymbol{\tau_a^*} - \tau_a)$$
$$U_{long} = \boldsymbol{\theta_{dlk}} \max(0, \tau_a - \boldsymbol{\tau_a^*})$$

Case study

- **Lausanne population, MTMC 2015** (BFS & ARE, 2017)
  - **3 samples:**
    - S1: Students (236 individuals)
    - S2: Workers (618 individuals)
    - S3: All occupations (1118 individuals)
  - **Choice set sizes:**
    - N = 10 alternatives for S1, S2
    - N = 100 for S3

→

Model 1 (14 parameters):

- Activity-specific constants
- Aggregated penalties (flexible vs. Non flexible)

Model 2 (30 parameters):

- Activity-specific constants
- Activity specific penalties

Results

- **Model 1, Workers: constants**
  - Reference: ASC Home = 0
  - $\overline{\rho}^2 = 0.77$
  - Runtime: 1.36 s

Parameter	Param. estimate	Rob. Std error	Statistical significance ($p < 0.05$)
ASC Education	10.8	2.50	Significant
ASC Errands	7.63	1.28	
ASC Escort	9.79	1.45	
ASC Leisure	15.3	1.38	
ASC Shopping	12.5	1.38	
ASC Work	18.5	2.00	

TRANSP-OR

Results

○ **Model 1, Workers: penalties**

- Flexible (F): errands, leisure, shopping
- Non-Flexible (NF): work, education, escort

Parameter	Param. estimate	Rob. Std error	Statistical significance $(p < 0.05)$
F: early	-0.813	0.160	Significant
F: late	-1.12	0.138	
*F: short**	*0**	- *	*Not significant*
F: long	-0.569	0.165	Significant
NF: early	-0.827	0.160	
NF: late	-1.26	0.236	
NF: short	-3.24	0.555	
NF: long	-0.789	0.229	

TRANSP-OR

Summary

- Estimated ABM parameters using sampled choice sets
- Expected signs, ratios and significance

- Further work:
  - Complex model specifications (logit mixtures, non-linear parameters…)
  - Investigate stability (increasing N)
  - Application on synthetic population for validation

Thank you!

janody.pougala@epfl.ch

tim.hillel@ucl.ac.uk

michel.bierlaire@epfl.ch

TRANSP-OR

References

- Manser P., Haering T., Hillel T., Pougala J., Krueger R., Bierlaire M. (2021). Resolving temporal scheduling conflicts in activity-based modelling. Report TRANSP-OR 211209
- Pougala J., Hillel T., Bierlaire M. (2022). Capturing trade-offs between daily scheduling choices. Journal of Choice Modelling 43 (100354)
- Pougala J., Hillel T., Bierlaire M. (2021) *Choice set generation for activity-based models*. Proceedings of the 21st Swiss Transport Research Conference (STRC), 12-14 September, Ascona, Switzerland
- Pougala J., Hillel T., Bierlaire M. (2022) *Parameter estimation for activity-based models*. Proceedings of the 22nd Swiss Transport Research Conference (STRC), 18-20 May, Ascona, Switzerland (forthcoming).
- Salvadé N., Hillel T., Pougala J., Haering T., Bierlaire M. (2022) *Representing location choice within activity-based models*. Proceedings of the 22nd Swiss Transport Research Conference (STRC), 18-20 May, Ascona, Switzerland (forthcoming).
- Schultheiss M., Spatial familiarity and mobility motifs, Bridging Transportation Researchers, August 2021