Preferred citation style

Axhausen, K.W. (2015) Activity-based and agent-based modelling: Reflections on choice modelling, simulation and time horizons, presentation at the Nagoya University Transport and Environmental Dynamics Laboratory (NUTREND), June 2015.

.

Activity-based and agent-based modelling: Reflections on choice modelling, simulation and time horizons

KW Axhausen

IVT
ETH
Zürich

June 2015

Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Activity-based versus agent-based approaches

A terminological problem ?

Resolution	Agents, flows
Scheduling model	Trip, tour, daily chain (with breaks)
Choice model	DCM, rules&heuristics
Route choice	Integrated, external (with consistent valuations?)
Choice set construction	Explicit, implicit
Solution method	Whole population (& MSA or similar) Sample enumeration (& MSA or similar), co-evolutionary search
Schedule equilibrium	Yes, no

The typical four-stage model

Resolution	Agents**, flows**
Scheduling model	**Trip, tour**, daily chain (with breaks)
Choice model	**DCM,** rules&heuristics
Route choice	Integrated, **external without consistent valuations**
Choice set construction	**Explicit**, implicit
Solution method	**Whole population** (& MSA or similar) Sample enumeration (& MSA or similar), co-evolutionary search
Schedule equilibrium	**(Yes)**, no

The typical activity-based model (ABM)

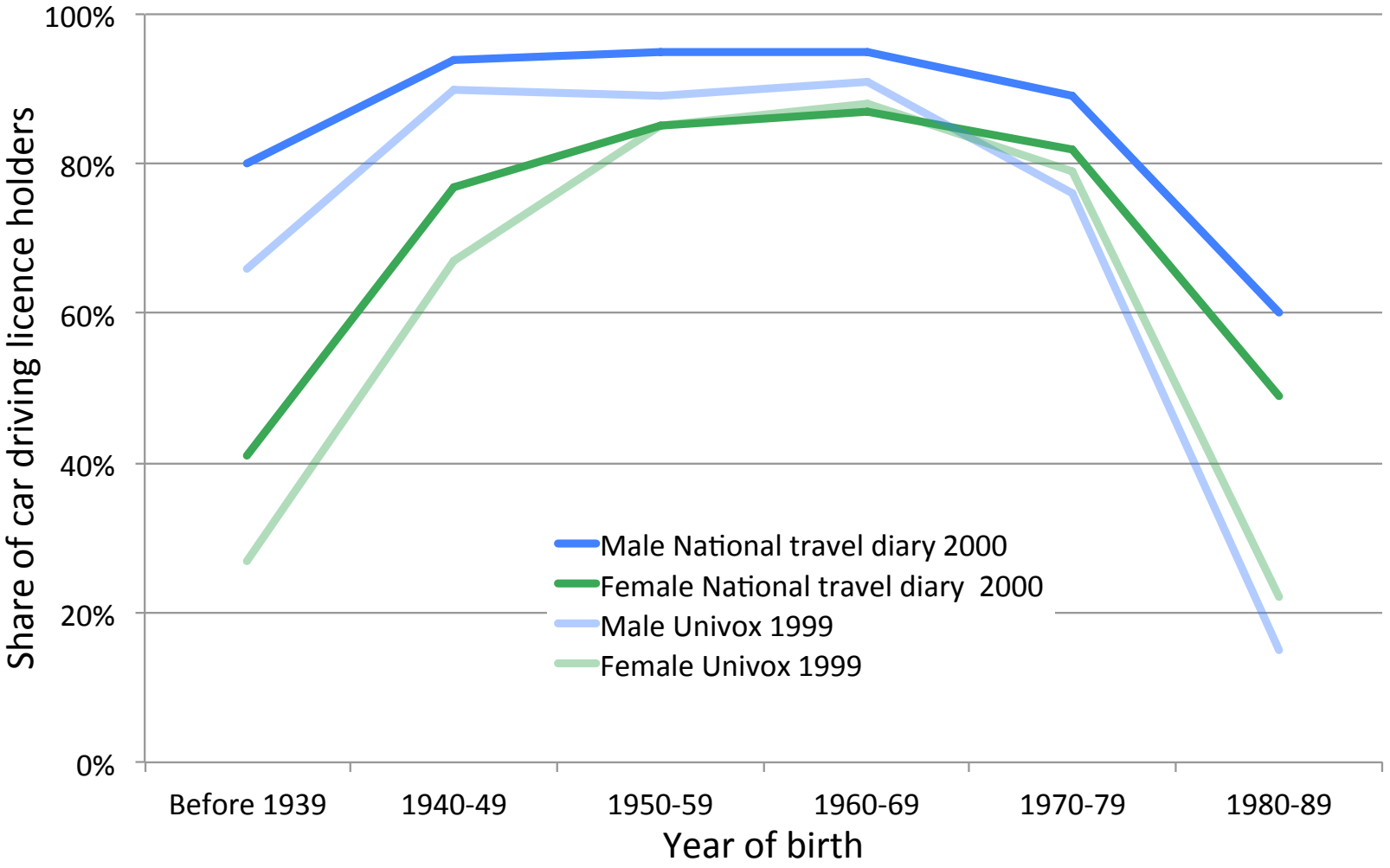
Resolution	**Agents**, flows
Scheduling model	Trip, tour, **daily chain** (with breaks)
Choice model	**DCM** , rules&heuristics
Route choice	Integrated, **external without consistent valuations**
Choice set construction	**Explicit**, implicit
Solution method	Whole population (& MSA or similar) **Sample enumeration** (& MSA or similar), co-evolutionary search
Schedule equilibrium	Yes, **none reported it yet**

MATSim

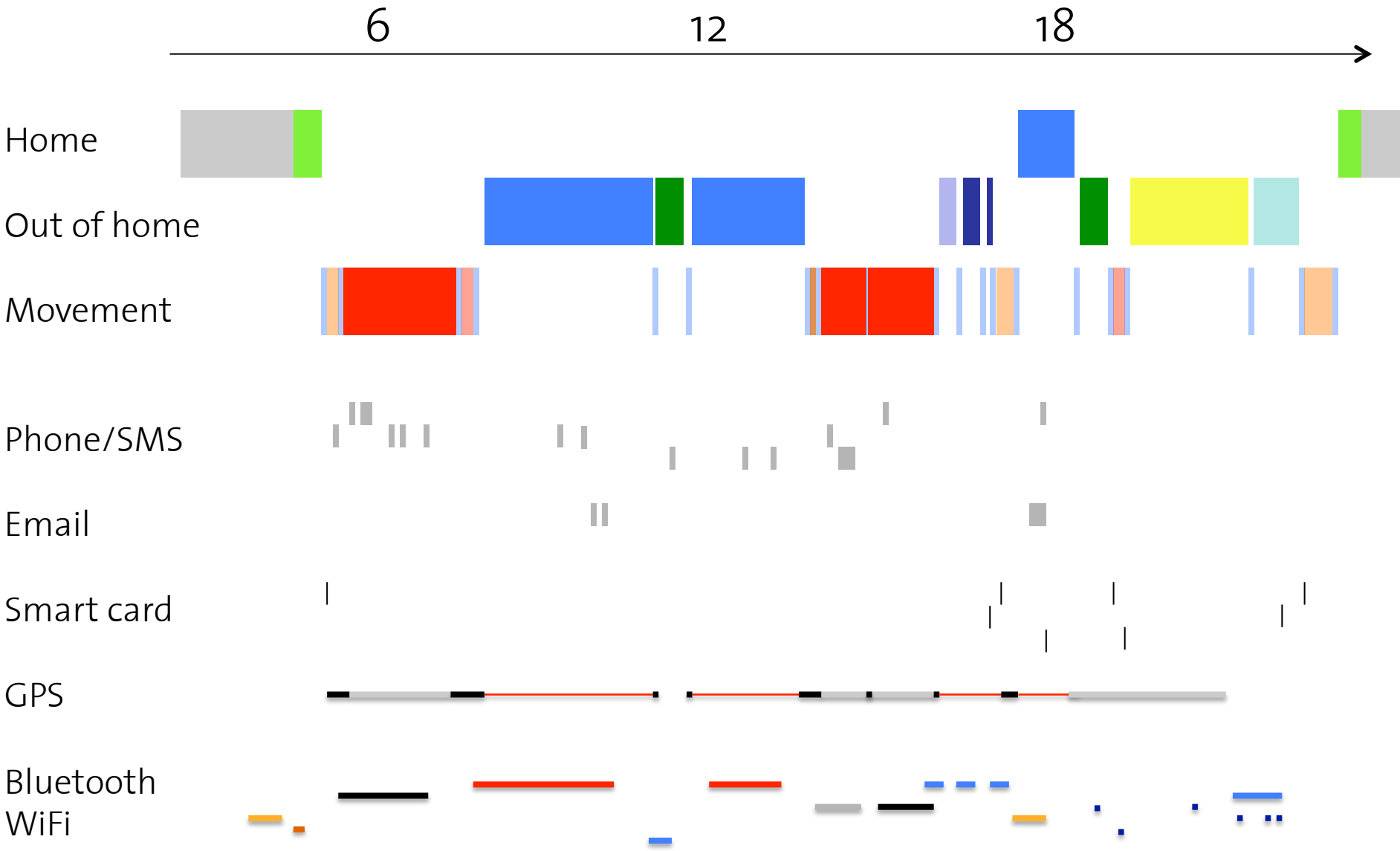
Resolution	**Agents**, flows
Scheduling model	Trip, tour, **daily chain** without breaks
Choice model	DCM, **rules**&heuristics
Route choice	**Integrated with consistent valuations**, external
Choice set construction	Explicit, **implicit**
Solution method	Whole population (& MSA or similar) Sample enumeration (& MSA or similar), **co-evolutionary search**
Schedule equilibrium	**Yes**, no

Data challenges

Do we know the numbers? e.g. drivers licence ownership



Activities, movement and traces: A full example record



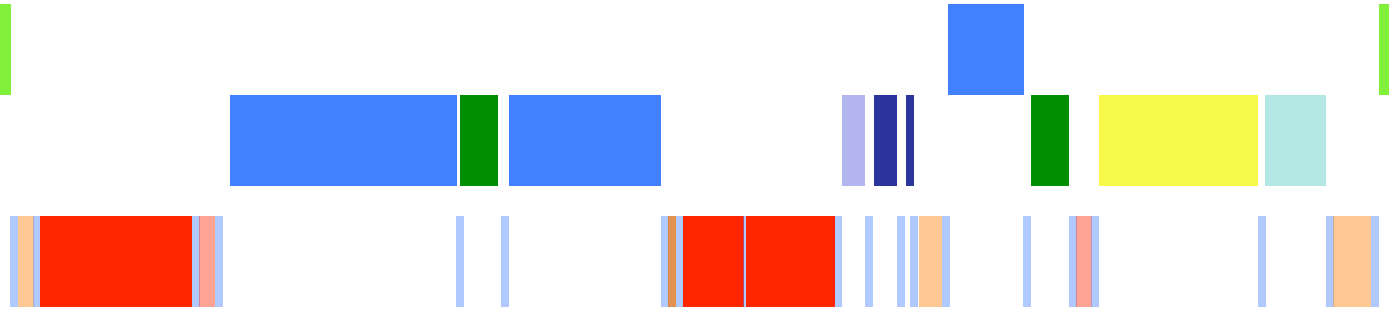
What is left after the known error processes ?

True

5 at home

9 Out of home

26 Stages,
11 trips,
1 subtour,
2 tours

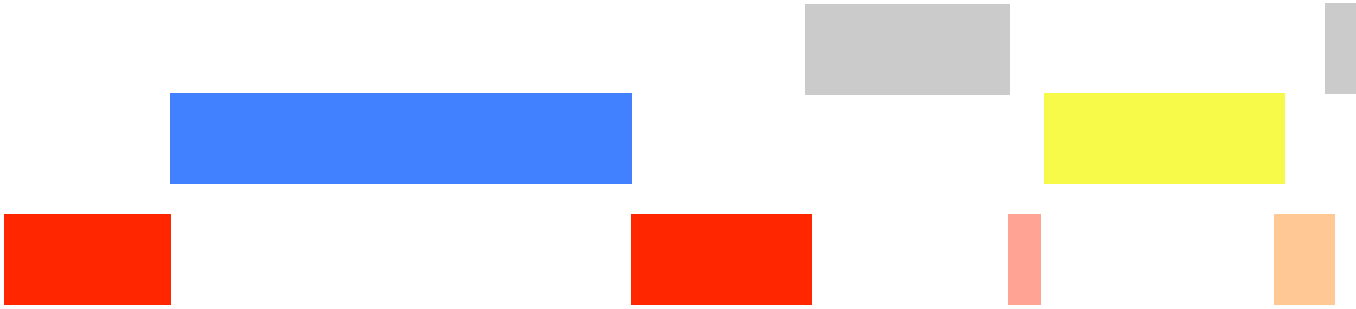


After all processes

3 at home

2 Out of home

4 trips,
2 tours



What happens next ?

Geocoding addresses

Ideal	Street addresses identifying the entry to the network
Best-case	Unambiguous street addresses
State of the art	Street address
State of practice	Street address/mid-street block/street corners; missing conversion of facility names
Still seen in practice	Arbitrary zonal centroid, e,g post offices

Calculating distances & travel time

Ideal	Complete GPS track for distance and times with pedestrian-networks added
Best-case	Minimal gaps, and state-of-the-art imputation of GPS tracks and modes
State of the art	SUE derived travel times and distances (navigation network)
State of practice	DUE derived travel times and distances (planning networks)
Still seen in practice	Shortest path on empty planning networks

Choice set construction

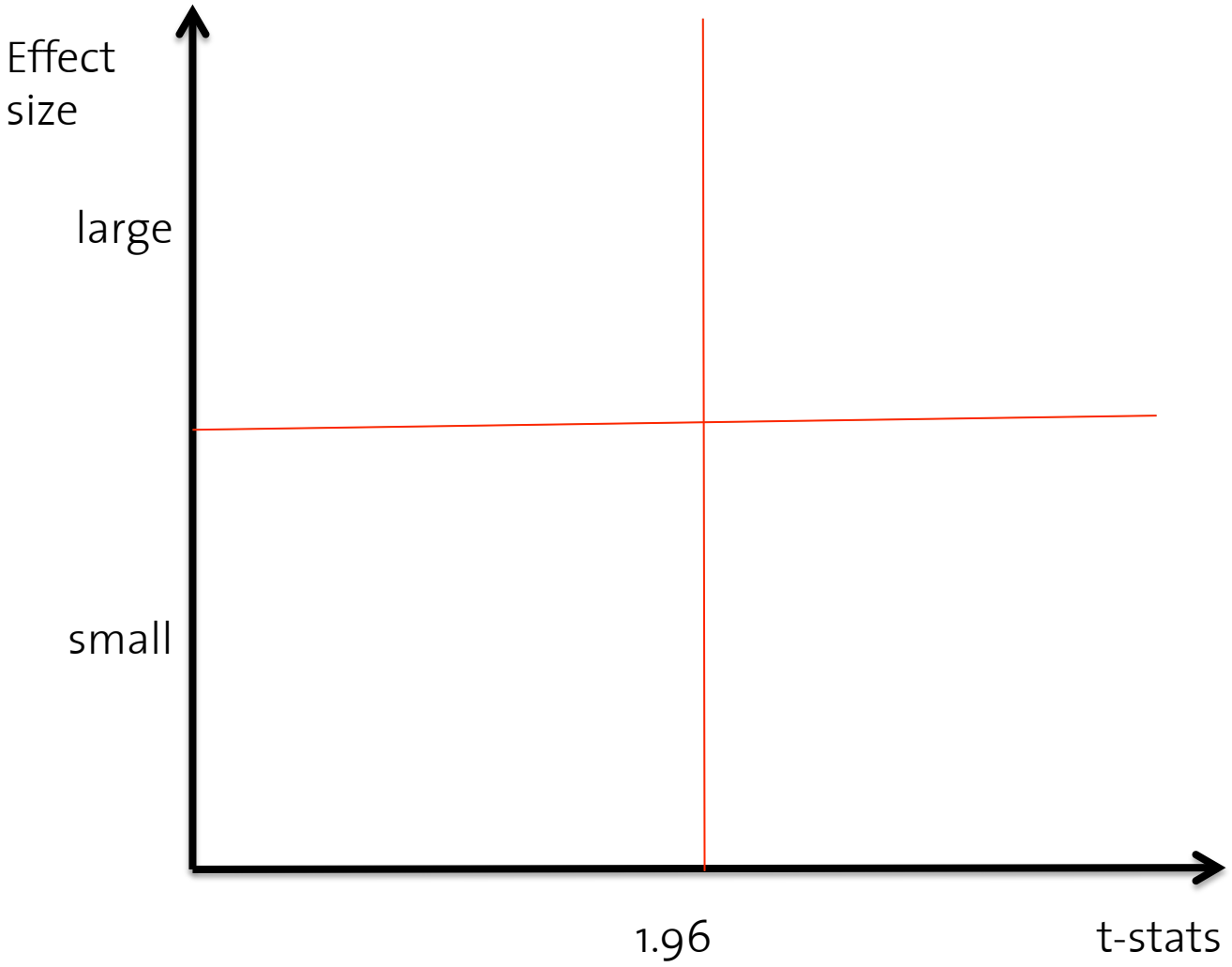
Ideal	identify the awareness set of traveller $P(i) > 0$
Best-case	Identify a set including a) much of the awareness set and b) not too many outside (i.e. $P(i) = 0$)
State of the art	Fast construction algorithms (Route choice); A-priori exclusion based on constraints Latent-class models among the universal set
State of practice	Universal choice sets; sampling; choice of low resolution alternatives

Modelling challenges

Modelling challenges: The usual worries

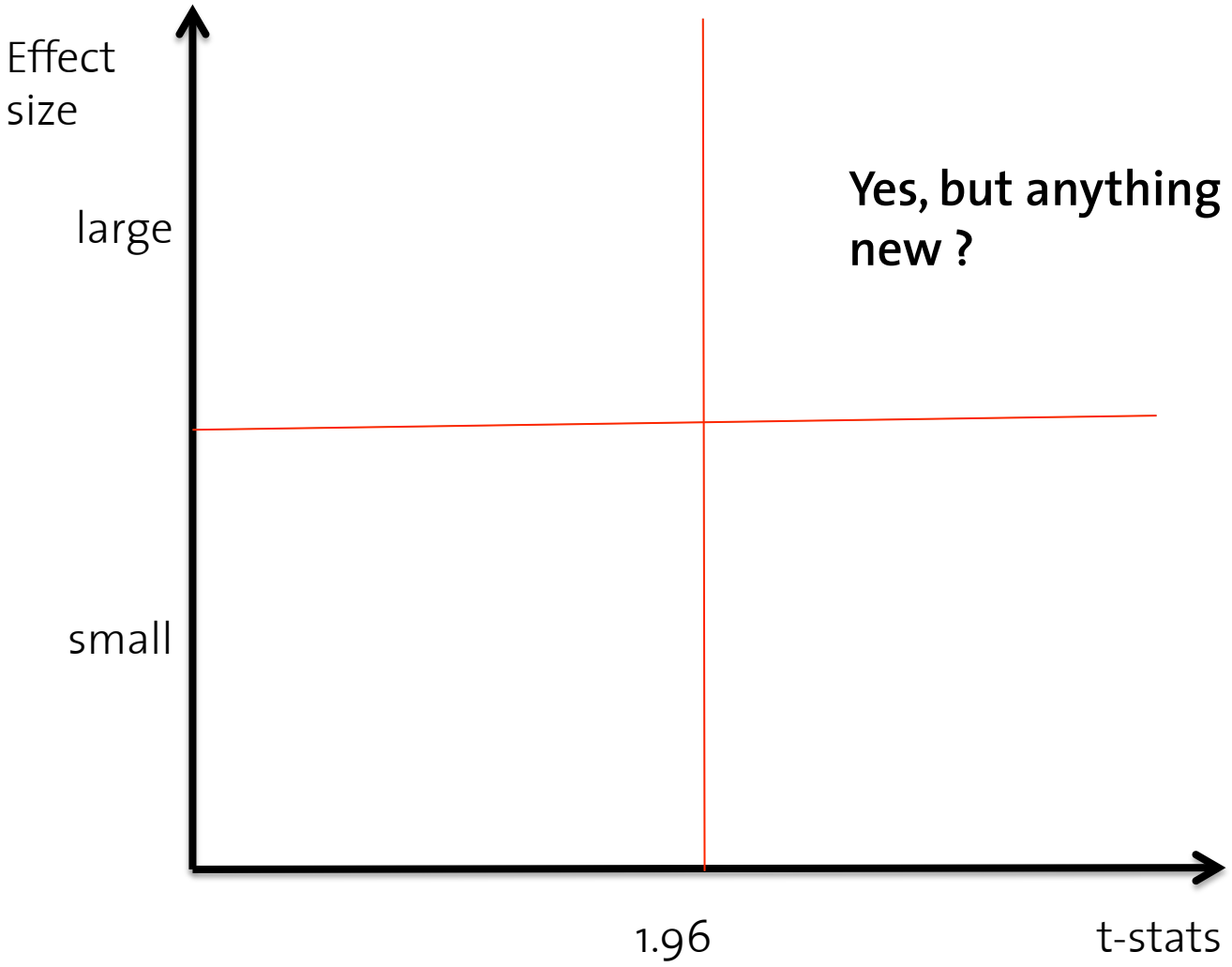
Error heterogenity	Is it always checked ?
Spatial correlations	Are they always checked ?
Temporal correlations	Are they always checked ?
Independence	Do we check the correlations of the independent variables (sample) thoroughly enough?
Endogenity	Do we fully account for it ? (sample selection)
Error of the second kind	Do you calculate it ?
Validation	How often do we ask for out-of-sample tests?
Substance	or do we talk about t-tests ?

Modelling challenges: Substance or t-tests ?



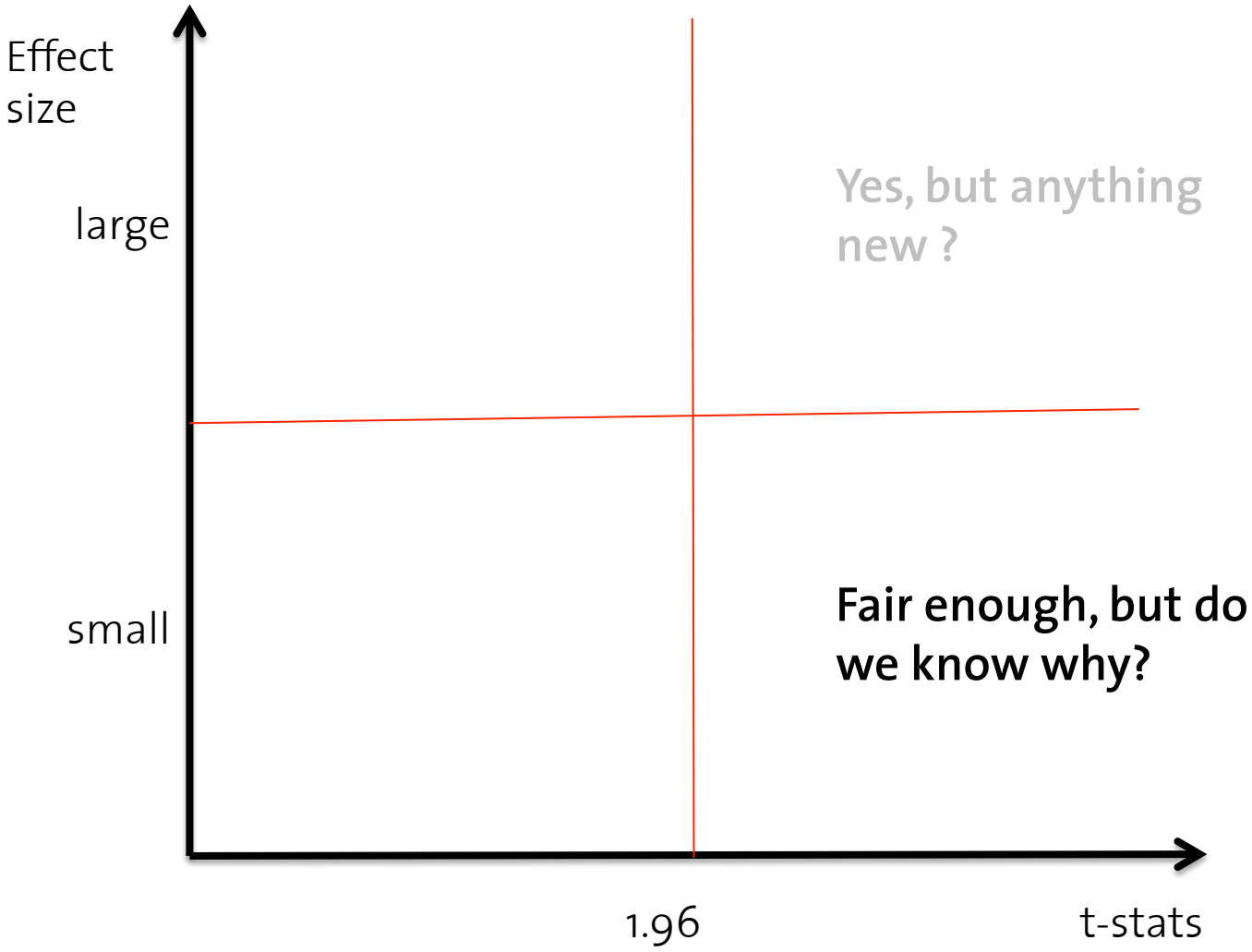
Adapted from Zilliak and McCloskey (2008)

Modelling challenges: Substance or t-tests ?



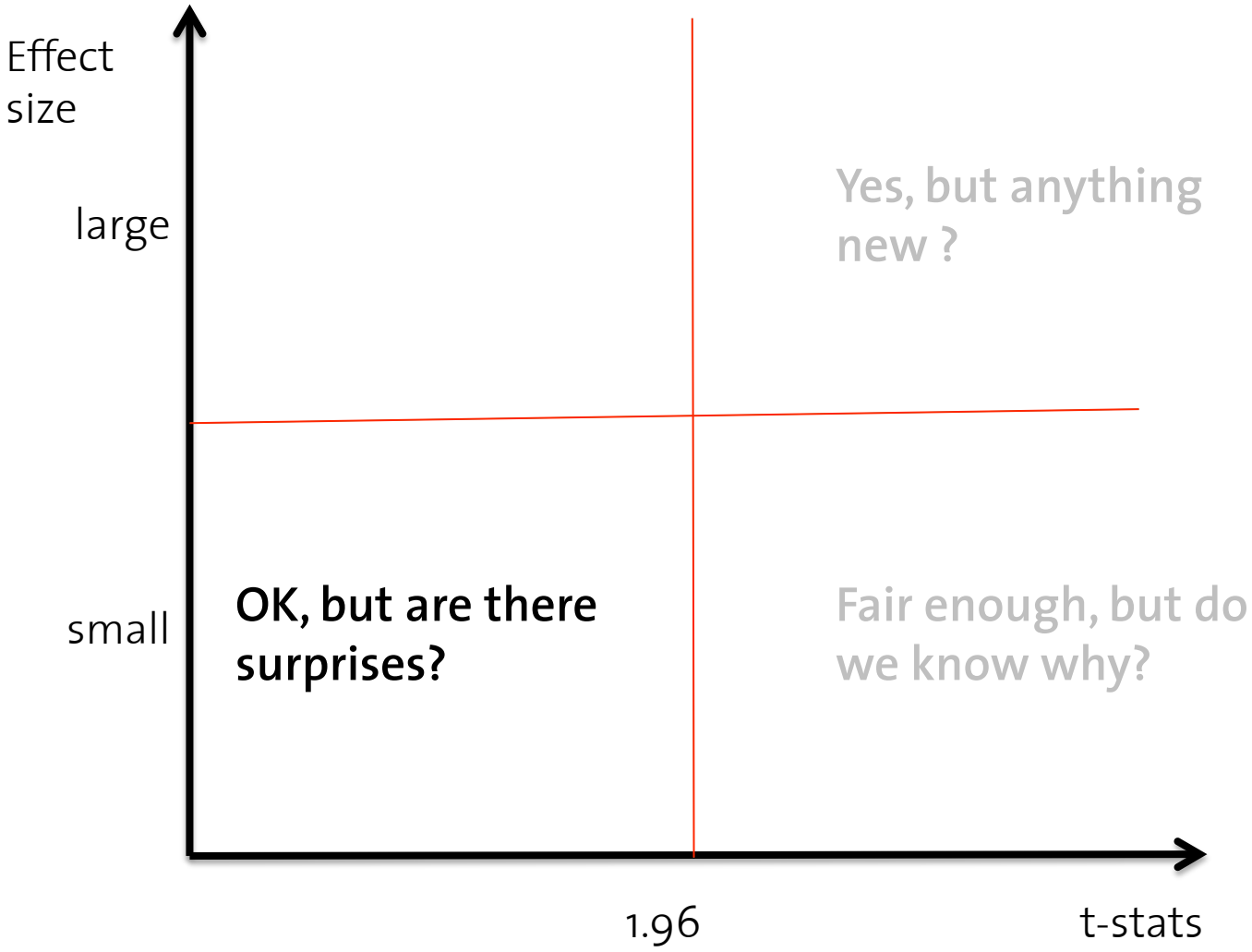
Adapted from Zilliak and McCloskey (2008)

Modelling challenges: Substance or t-tests ?



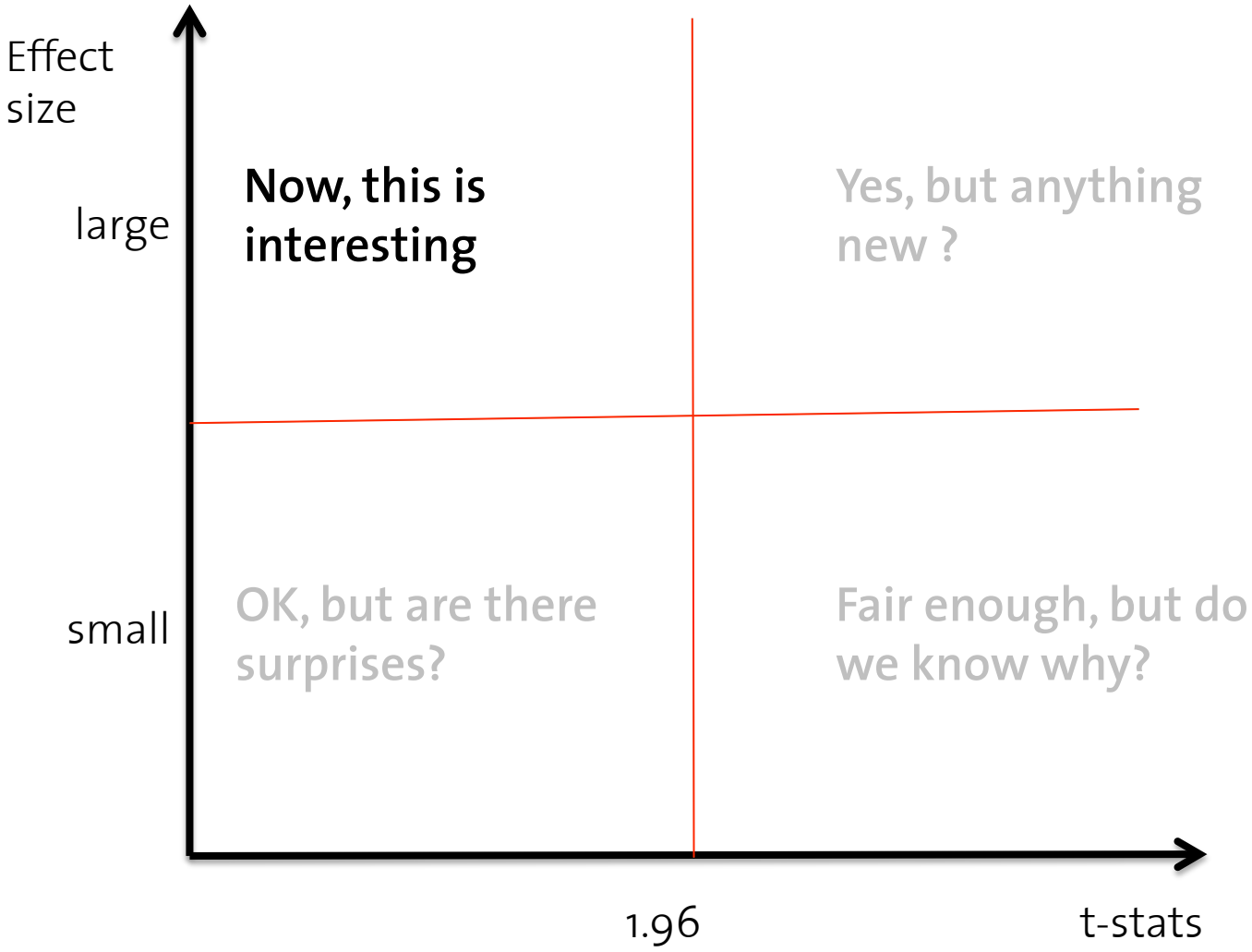
Adapted from Zilliak and McCloskey (2008)

Modelling challenges: Substance or t-tests ?



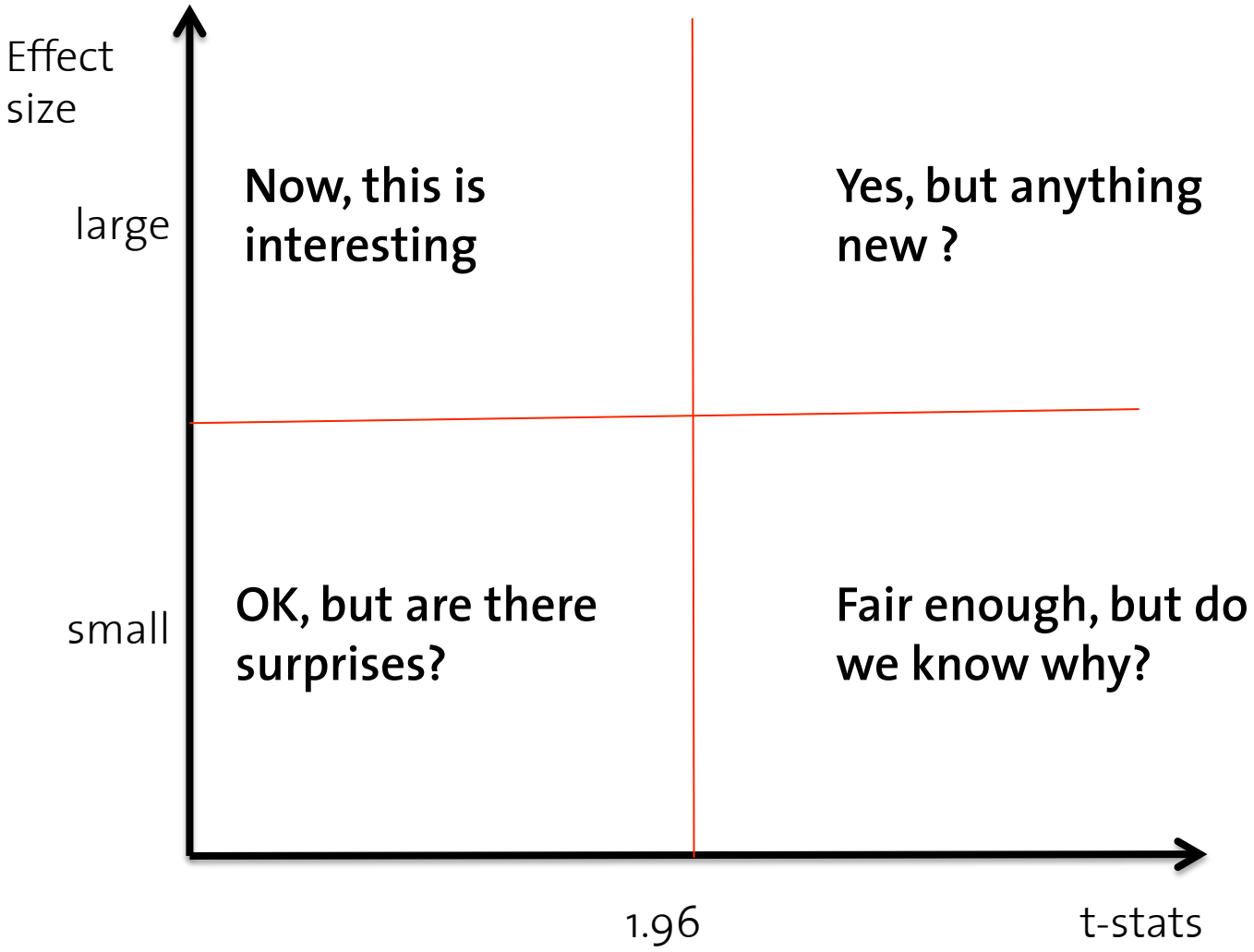
Adapted from Zilliak and McCloskey (2008)

Modelling challenges: Substance or t-tests ?



Adapted from Zilliak and McCloskey (2008)

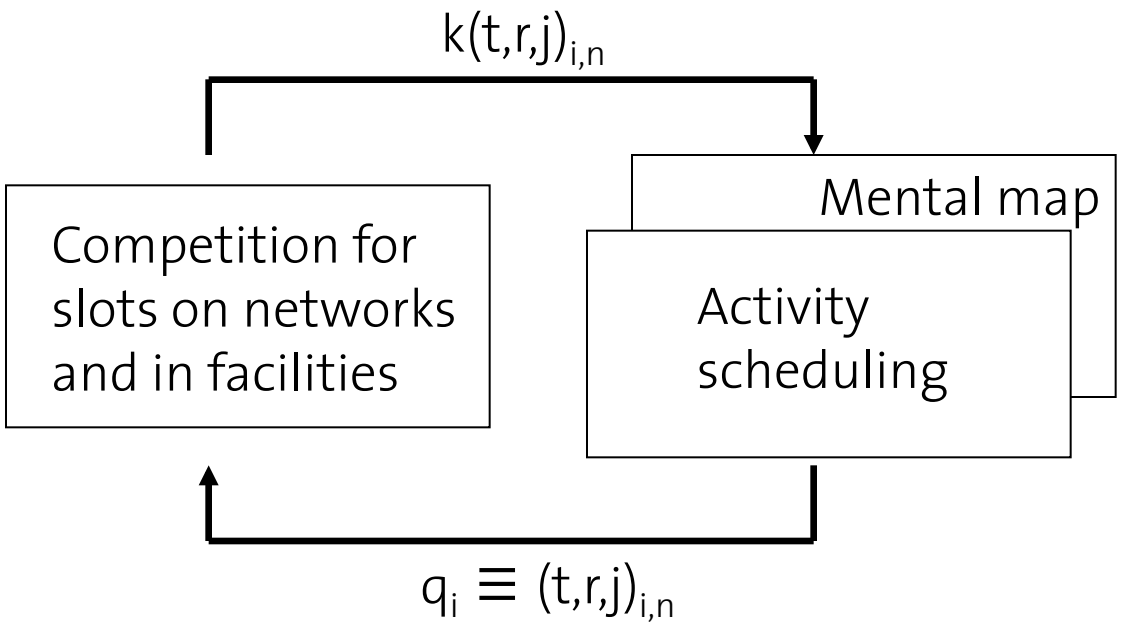
Modelling challenges: Substance or t-tests ?



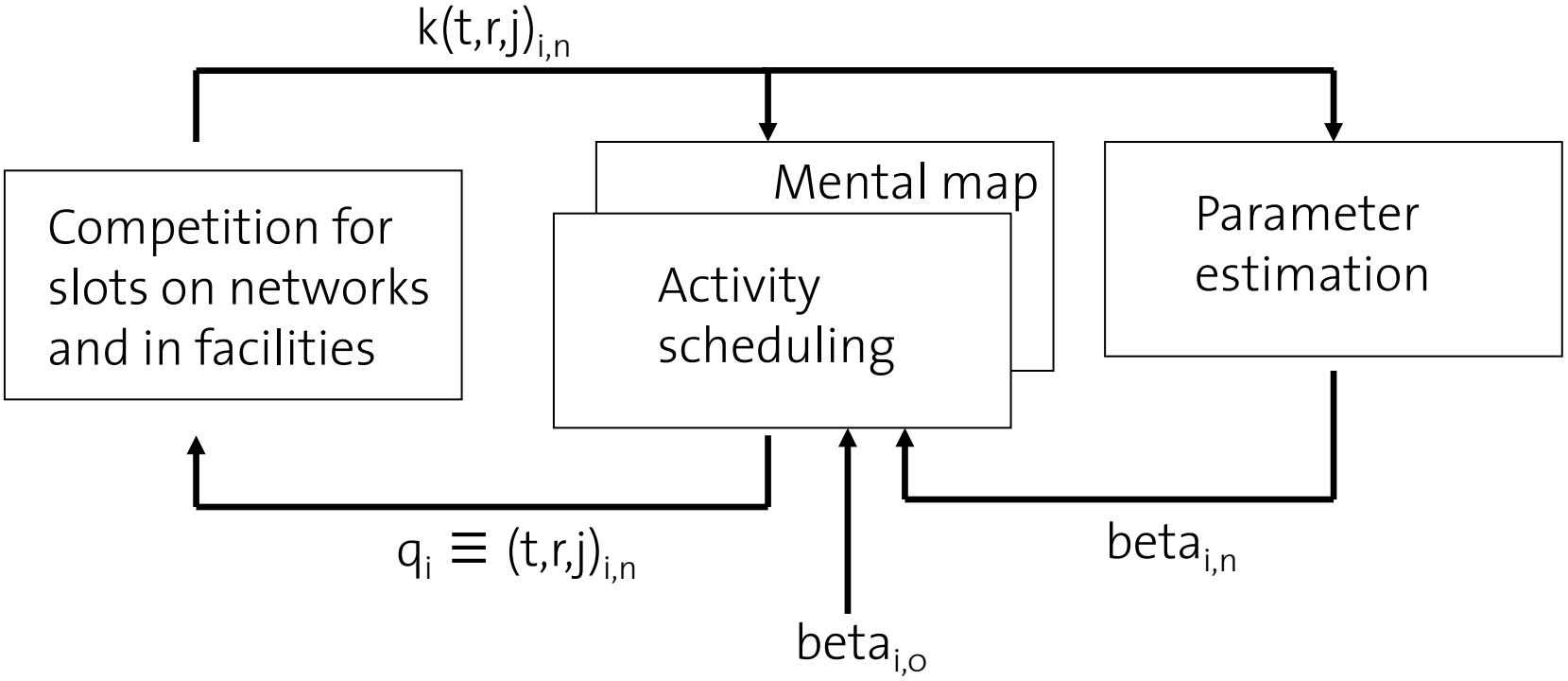
Adapted from Zilliak and McCloskey (2008)

Consistent LOS variables (travel times, rents, crowding, etc.)

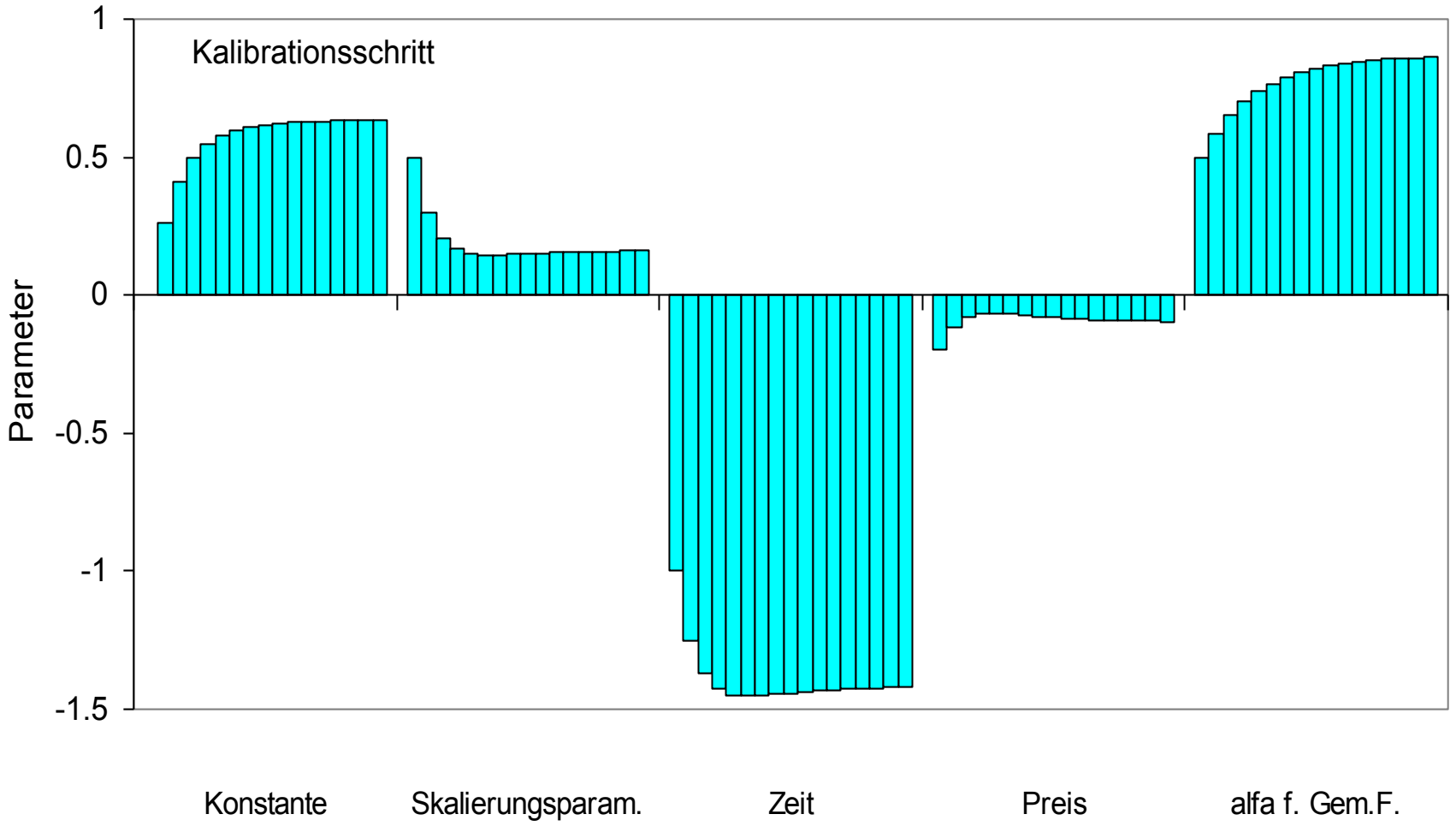
Learning approach of the generic one-day transport model



Model estimation: $beta_{i,o} = beta_{i,n}$? $beta_{i,n-1} = beta_{i,n}$?



Model estimation: $beta_{i,o} = beta_{i,n}$? Route and mode



Source: Vfrtic (2003)

Do we have a MAUP problem ?

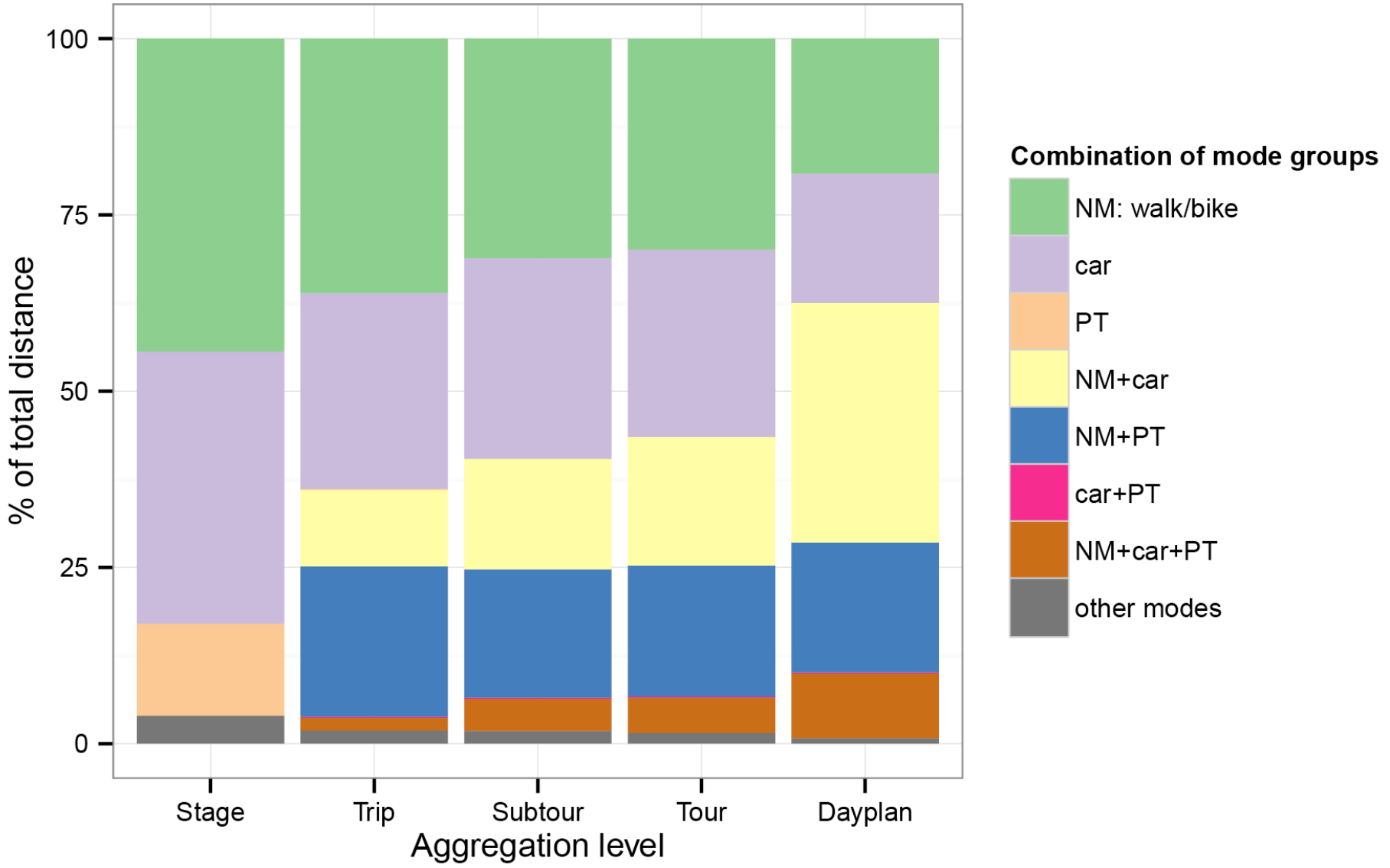
Do we have a MAUP-like problem for DCM?

- Location choice, obviously via zonal resolution
- Route choice, obviously via network resolution
- Time of day choice, obviously via temporal resolution

- But also, mode choice

    - Stage
    - Trip
    - Sub-tour
    - Tour
    - Daily schedule

Swiss national travel diary 2010: Main mode by aggregation



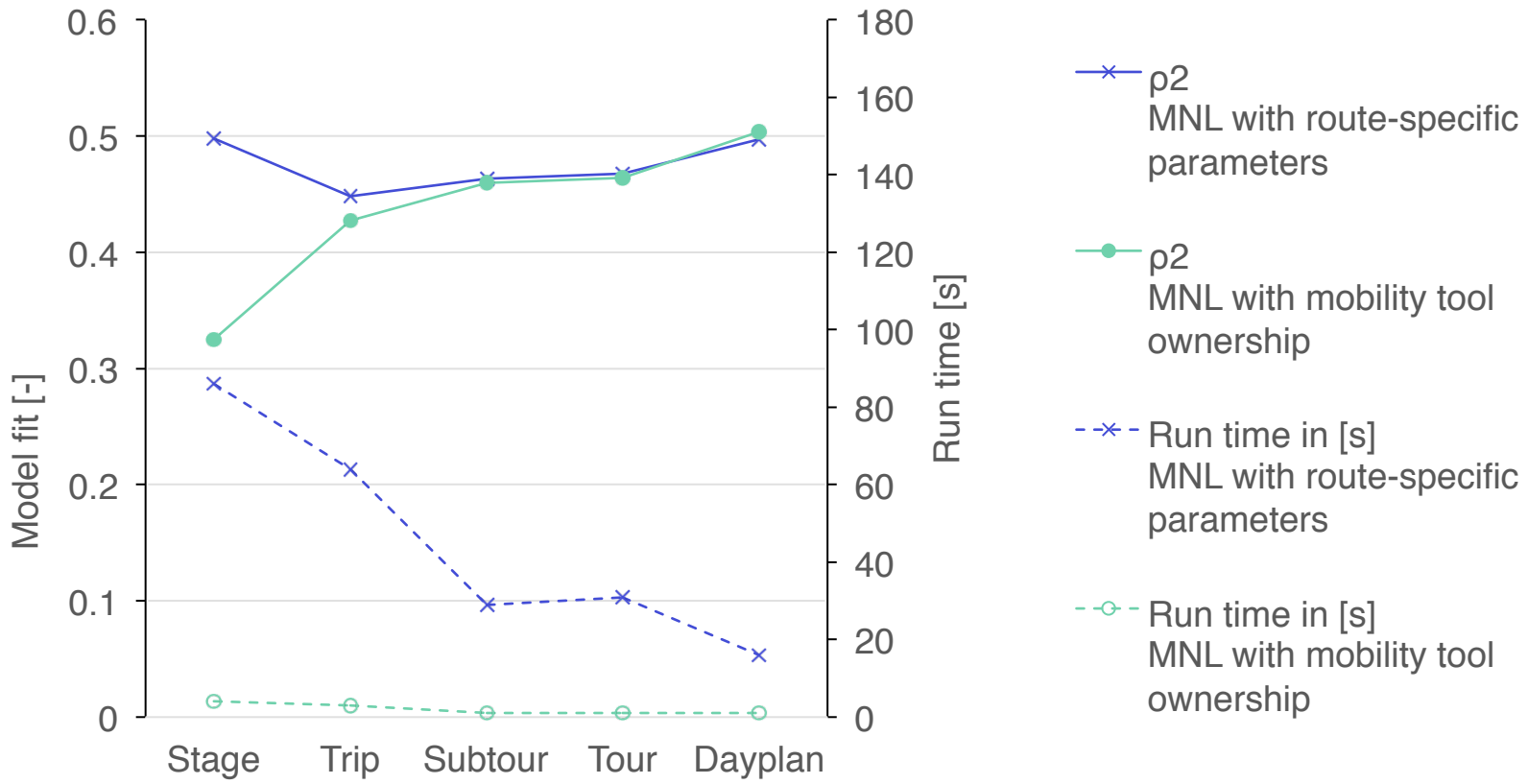
Do we have a MAUP-like problem for DCM?

		Stage	Trip	Subtour	Tour
Value of Time Walking	CHF/h	152	28	26	24
Value of Time Bike	CHF/h	194	39	43	40
Value of Time Car	CHF/h	135	25	30	27
Value of Time PT	CHF/h	-30	2	7	6
Value of Time PT access	CHF/h	819	15	22	22
TT PT / TT Car	-	-4.46	12.33	4.07	4.16
TT Walk / Access time PT	-	0.19	1.83	1.19	1.09
Transfer / TT PT	min	-220.43	107.00	31.28	32.92
Interval / TT PT	-	0.96	7.00	3.47	6.33
Access time / TT PT	-	-27.10	7.67	3.02	3.35

Source: Schmujtz (2015)

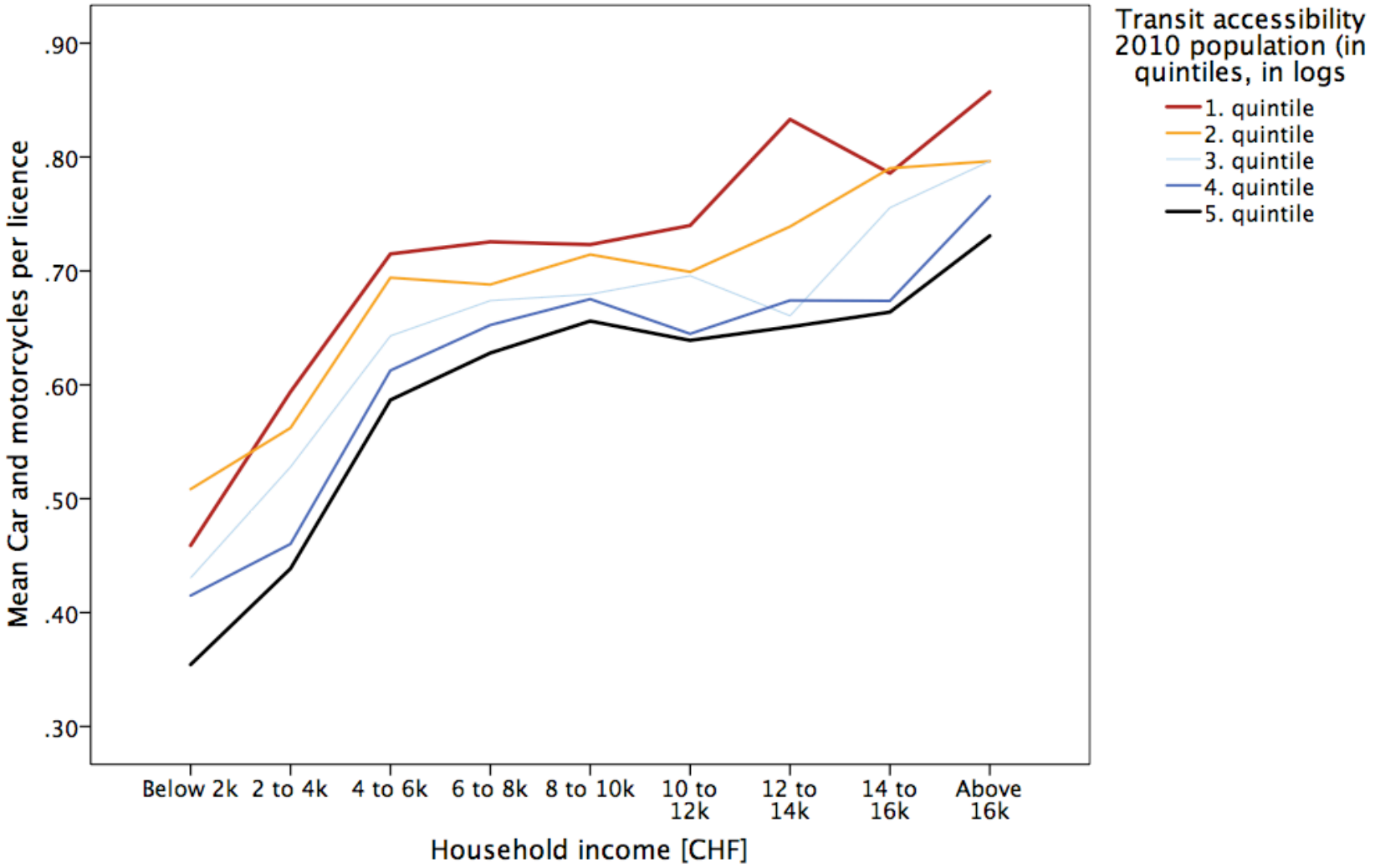
Do we get the time horizon right?

Do we get the time horizon right?



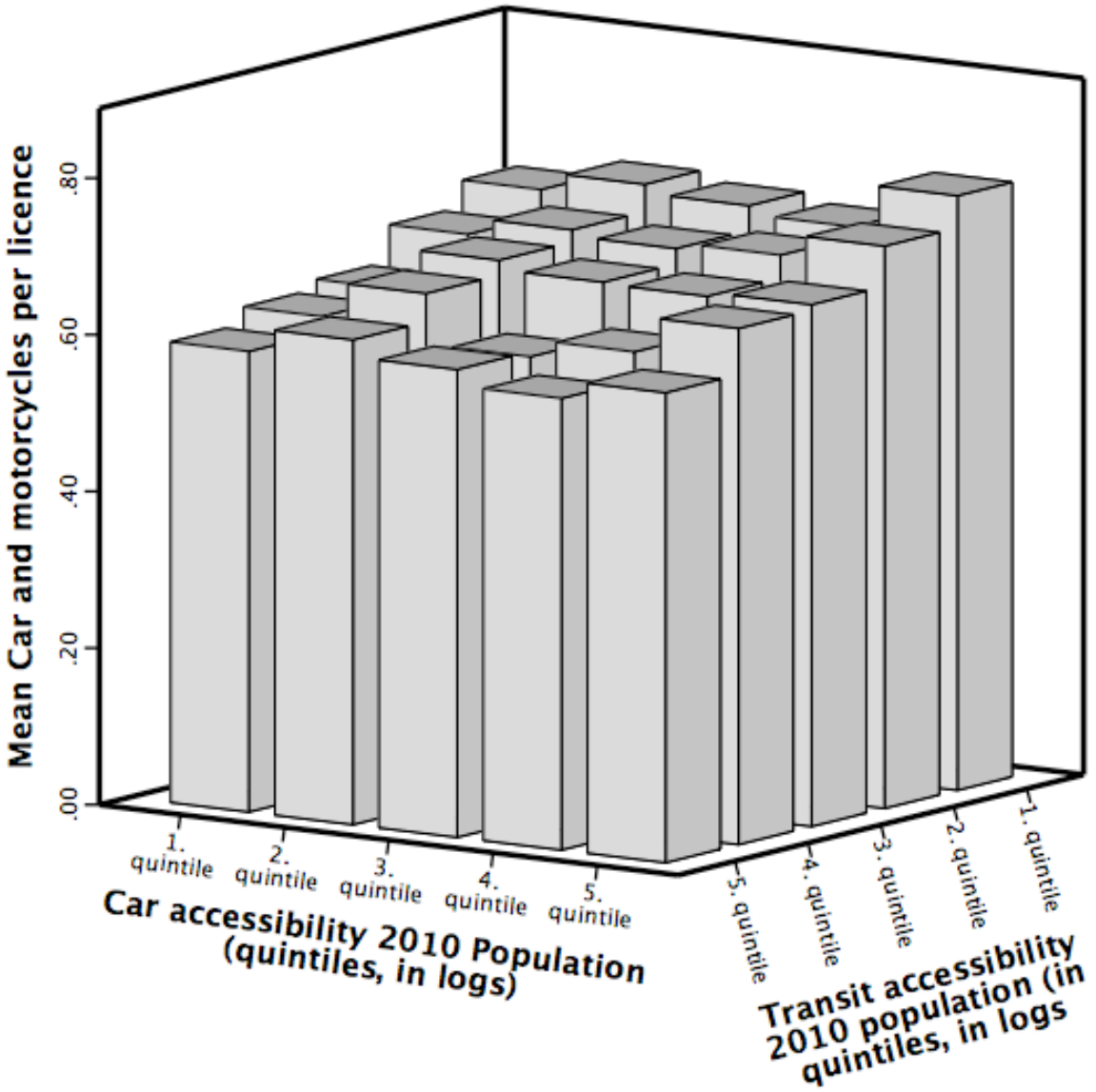
Source: Schmujtz (2015)

Do we get the time horizon right?



Data: Mikrozensus 2010 – Swiss National Travel Diary

Do we get the time horizon right?



Data: Mikrozensus 2010 – Swiss National Travel Diary

Do we get the time horizon right?

- Is daily mode choice the result of trade-off between LOS?
- Is daily mode choice purely the result of longer-term commitments?
- Is mode choice just the result of choosing standard ‘scripts’ or activity – location – mode packages ?
- What drives the commitments ?
  - Accessibility
  - Housing & modal packages
    - Self-selection by lifestyle
    - Self-selection by social commitments

MATSim

MATSim: A GNU public licence software project

MATSim: A GNU public licence software project

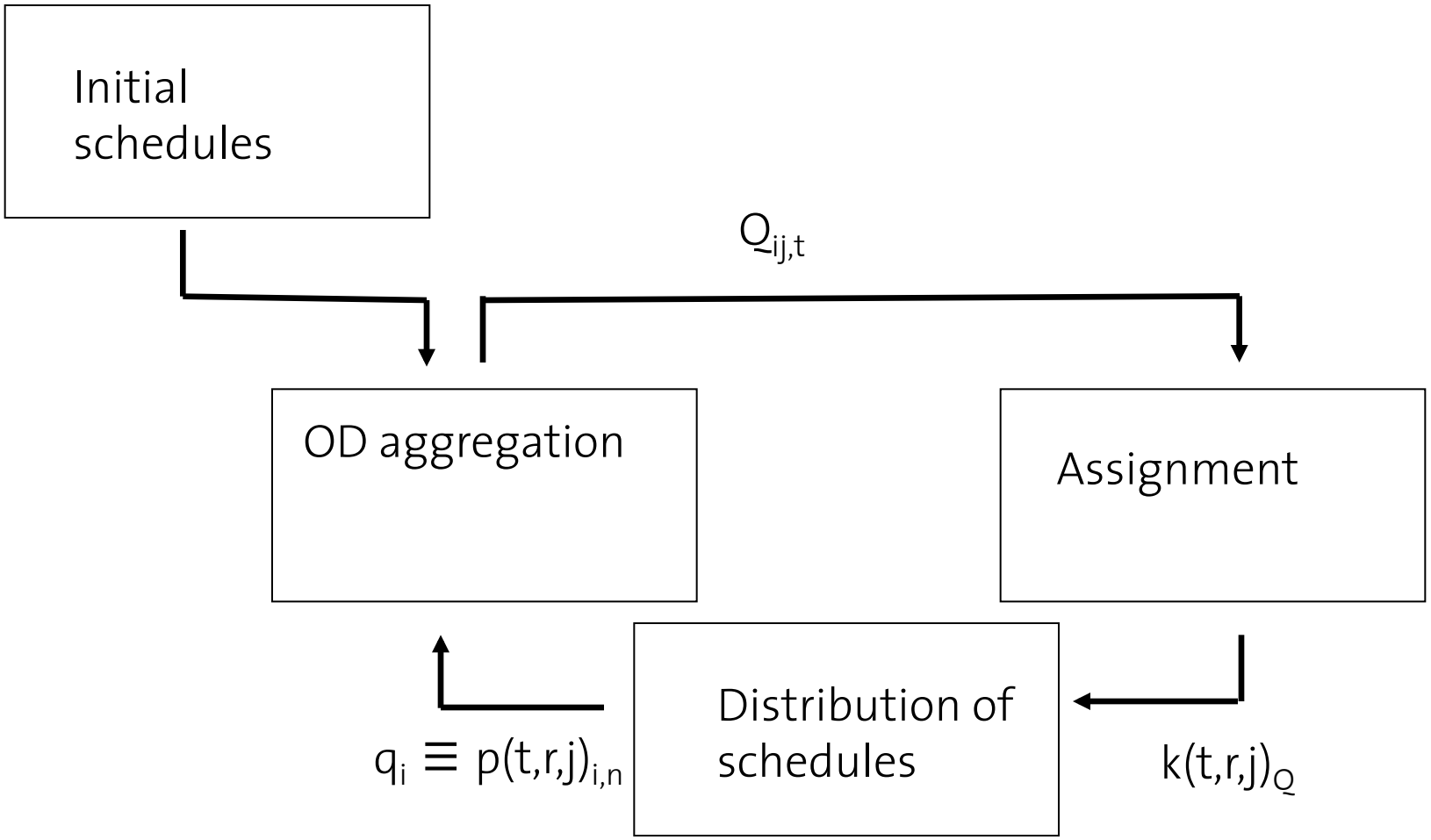
Main partners:

- TU Berlin (Prof. Nagel)
- ETH Zürich & FCL SIngapore
- senozon (Dr. Balmer, Dr. Rieser)

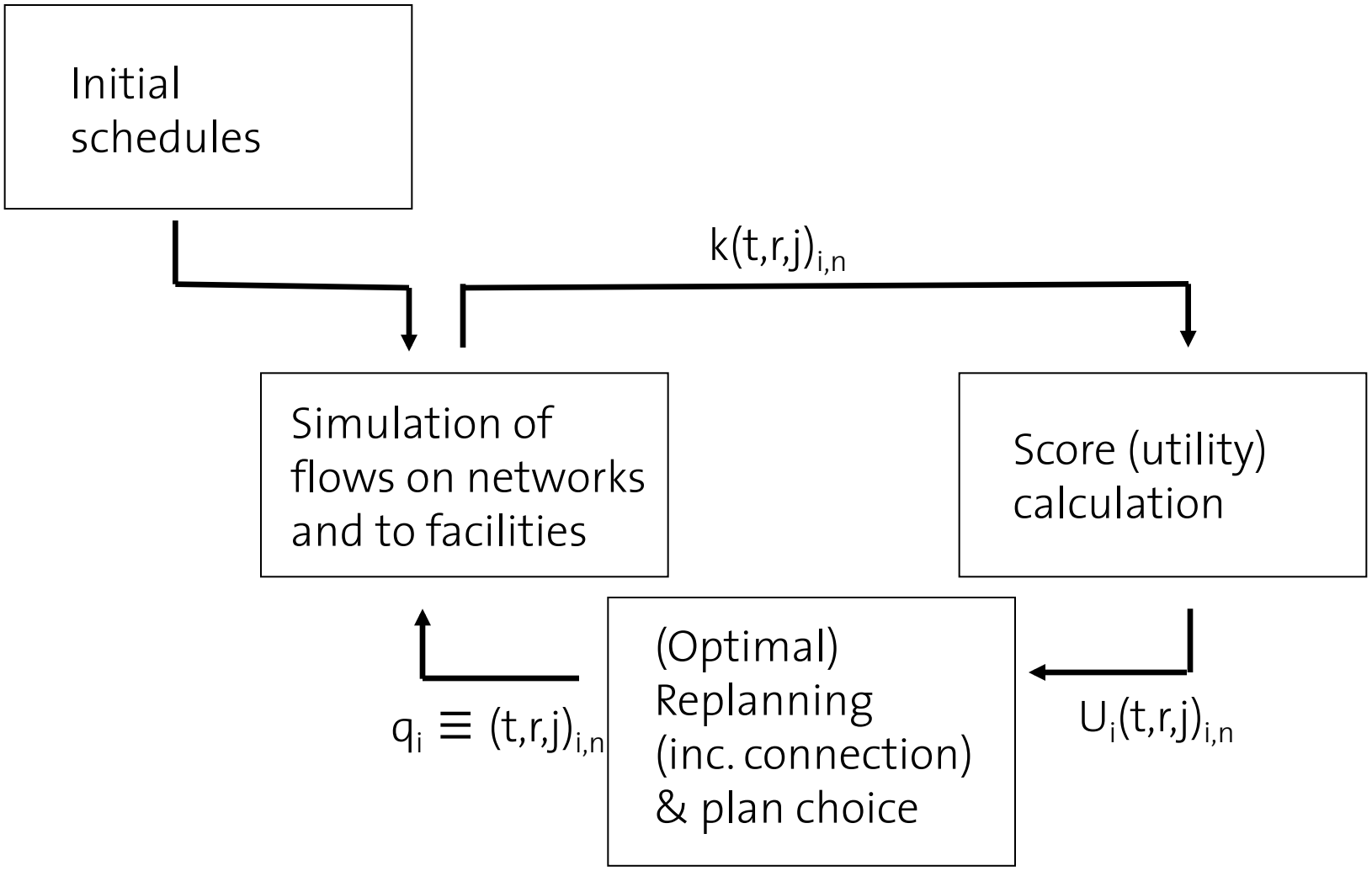
Contributors, users, e.g.:

- TU Poznan
- University of Pretoria
- CASA, UCL, London
- Forschungszentrum Jülich

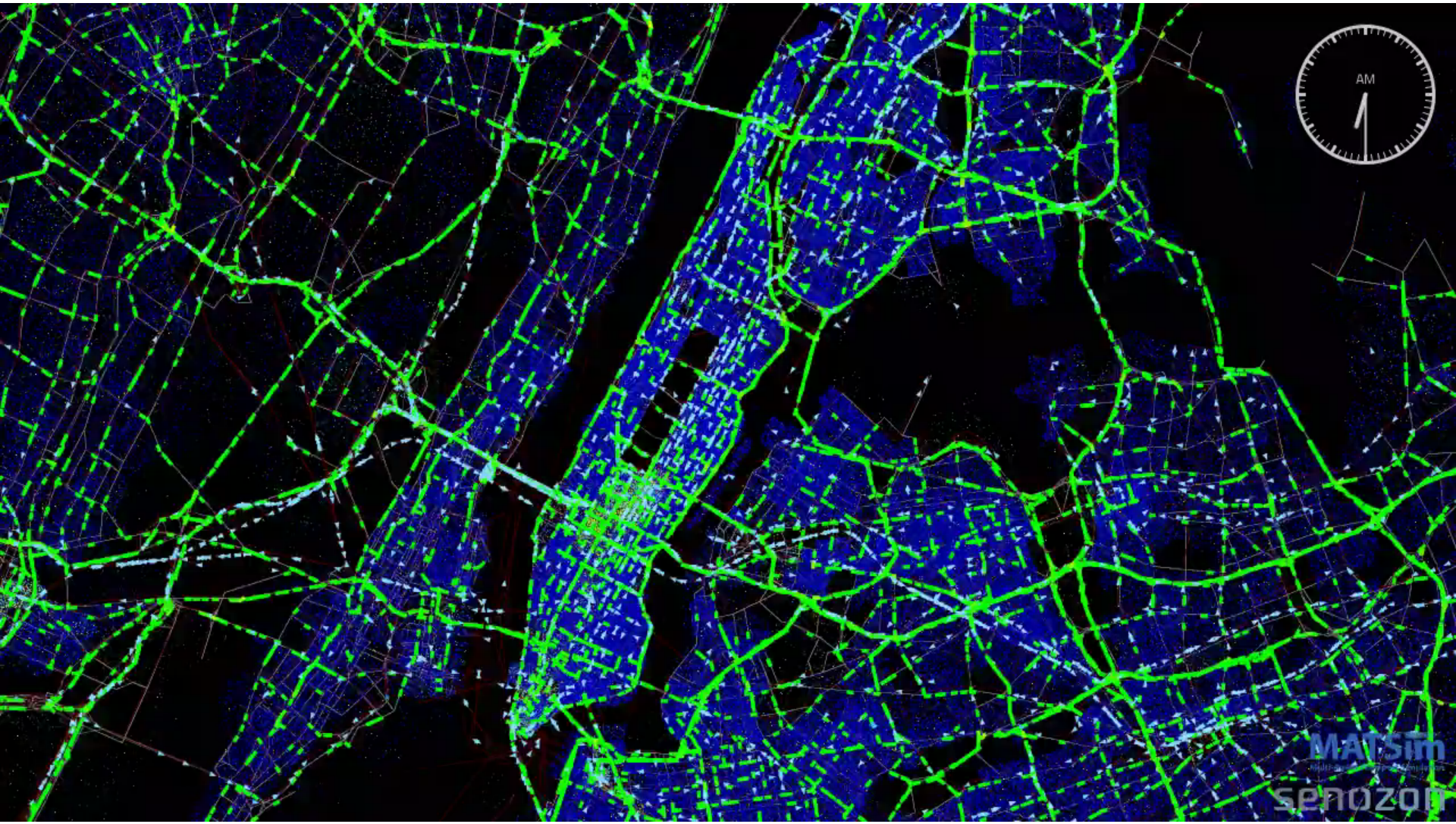
Equilibrium search in „ABM“ & assignment combinations



Equilibrium search in MATSim



MATSim today



Following the agents

MATSim: Logic of the co-evolution – Step 0

Agent 1

Plan 1.1 H-W-H; 8:00, 17:00; C,C;

Agent 2

Plan 2.1 H-W-H; 8:00, 17:00; C,C;

Agent 3

Plan 3.1 H-W-H; 8:00, 17:00; C,C;

Co-evolution – Step 1.1 – Simulation/scoring

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	35
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	35
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	35

Co-evolution – Step 1.2 – After replanning (1/3)

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	35
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	35
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	35
Plan 3.2	**H-W-H; 8:15, 17:30; C,C**	

Co-evolution – Step 1.3 – After plan selection (best/MNL)

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	**100%**
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	**100%**
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	35
Plan 3.2	H-W-H; 8:15, 17:30; C,C;	**New**

Co-evolution – Step 2.1 – Simulation/scoring

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	**45**
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	**45**
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	35
Plan 3.2	H-W-H; 8:15, 17:30; C,C;	**60**

Co-evolution – Step 2.2 – After replanning (1/3)

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	45
Plan 1.2	**H-W-H; 8:00, 17:00; B,B;**	
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	45
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	35
Plan 3.2	H-W-H; 8:15, 17:30; C,C;	60

Co-evolution – Step 2.3 – After plan selection (best/MNL)

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	45
Plan 1.2	H-W-H; 8:00, 17:00; B,B;	New
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	100%
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	38%
Plan 3.2	H-W-H; 8:15, 17:30; C,C;	62%

Co-evolution – Step 3.1 – Simulation/scoring

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	45
Plan 1.2	H-W-H; 8:00, 17:00; B,B;	**70**
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	**45**
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	**45**
Plan 3.2	H-W-H; 8:15, 17:30; C,C;	60

Co-evolution – Step 3.2 – After replanning (1/3)

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	45
Plan 1.2	H-W-H; 8:00, 17:00; B,B;	70
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	45
Agent 3		
Plan 3.1	H-W-H; 8:00, 17:00; C,C;	45
Plan 3.2	H-W-H; 8:15, 17:30; C,C;	60
Plan 3.3	**H-W-H; 7:30, 17:15; B,B**	

Co-evolution – Step 3.3 – After plan selection (best/MNL)

Agent 1		
Plan 1.1	H-W-H; 8:00, 17:00; C,C;	36%
Plan 1.2	H-W-H; 8:00, 17:00; B,B;	**64%**
Agent 2		
Plan 2.1	H-W-H; 8:00, 17:00; C,C;	**100%**
Agent 3		
~~Plan 3.1~~	~~H-W-H; 8:00, 17:00; C,C;~~	~~45~~
Plan 3.2	H-W-H; 8:15, 17:30; C,C;	60
Plan 3.3	H-W-H; 7:30, 17:15; B,B	**New**

(The (worst) plan, more then memory allows, is deleted)

Co-evolution – Summary of best scores

	Iteration 1	Iteration 2	Iteration 3
Agent 1	35	45	80
Agent 2	35	45	45
Agent 3	35	60	60
Mean	**35**	**50**	**62**

Co-evolution – Issues

- Size of search space ~ Behavioural alternatives
- Rate of replanning (~ MSA)
- Size of the choice set ~ RAM
- Similarity of the daily schedules
- Integration into a log-sum term

Activity schedule dimensions

Activity scheduling dimensions

Number and type of activities
Sequence of activities

- Start and duration of activity
- Composition of the group undertaking the activity
- Expenditure division
- Location of the activity
    - Movement between sequential locations
        - Location of access and egress from the mean of transport
            - Parking type
        - Vehicle/means of transport
        - Route/service
        - Group travelling together
        - Expenditure division

Current Vickrey-type utility function

$$U_{plan} = \sum_{i=1}^{n} U_{act,i} + \sum_{i=2}^{n} U_{trav,i-1,i}$$

$$U_{act,i} = U_{dur,i} + U_{late.ar,i}$$

Future whole day utility function?

Time elements	linear
• Travel time	By mode and type of service; by crowding level by comfort level (parking search, stop&go)
• Transfer penalty	
• Late penalty	by activity type
Activity time	log (Vickrey) or S-shape (Joh) (all, individual)
• Minimum duration	by activity type
• Preferred duration	by activity type
• Duration	by time of day (might go away if participation is included)
Destination	Attractiveness, Value for money
Expenditure	by activity
	by mode/type of service

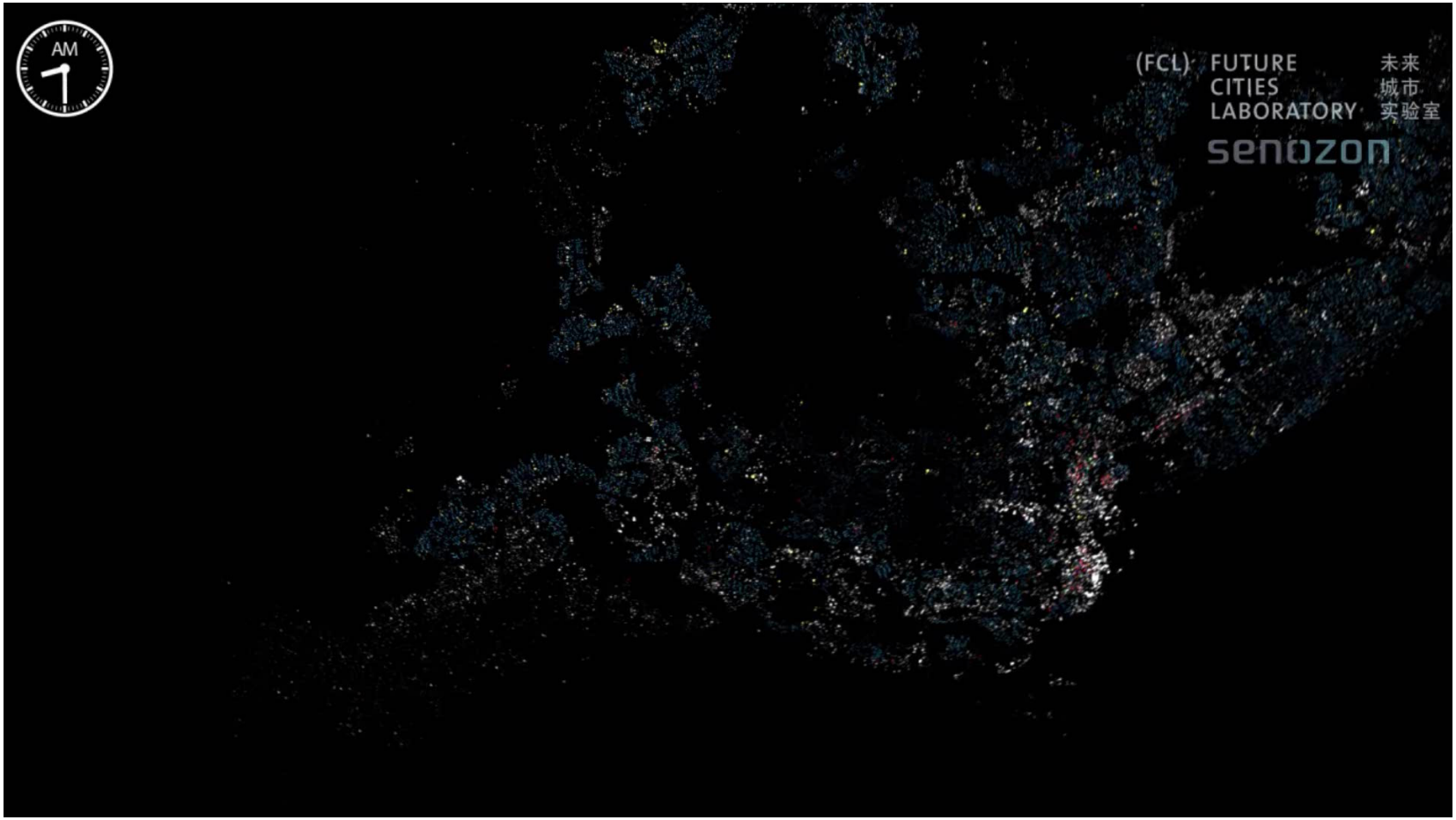
Current status

© Marcel Rieser, senozon

Current status

Known implementations:	About 35 (Europe, Asia, US)
Research groups:	About 25 (including some beyond transport)
Uses:	Research Some initial commercial uses Some policy consulting
Software:	Last reimplementation in 2012/13 Stable API Daily tests JAVA

Current progress: Singapore



Schedule detail possibilities (in current stable MATSim)

Number and type of activities (Feil, Balac)
Sequence of activities (Ordonez)

- **Start and duration of activity**
- Composition of the group undertaking the activity (Kowald, Tan, Fourie)
- Expenditure division
- **Location of the activity** (Horni)
  - Movement between sequential locations
    - **Location of access and egress from the mean of transport**
      - Parking search and type (Waraich)
    - **Vehicle/means of transport** (Ciari, Bösch)
    - **Route/service** (Chakirov)
    - Group travelling together (Dubernet, Fourie)
    - Expenditure division

Recent and current developments and applications at ETH

Integration of walking	Multi-level network resolution
New modes	Escalators and ‘walkways’
Autonomous vehicles	Include a ‘gopher mode’
Car sharing:	Station-based car-sharing Free-float car sharing
Parking search	Specialised within day replanning In conjunction with recent SC experiments
Evacuation	Specialised within-day replanning

Challenges

Challenges for MATSim

- Econometric estimation of the whole day scoring function
- Increase the size and variance of the implicit choice set
- Link to a log-sum formulation
- Accelerating the iterative equilibrium search
- Gridlock modeling (& stability of equilibrium)
- Generation of artificial social networks in the agent-population

Wider challenges

- Consistency of the LOS variables in model estimation
- Integrating the capacity constraints
- ‘MAUP’ at different levels and choice dimensions
- Standards for choice set size
- Daily versus non-daily choices (Overreach of the NL – models ?)
- How many robustness test should we report in choice modelling papers?

MATSim @ ETHZ, TU Berlin, FCL, Senozon (past & present)

Prof. Kay Axhausen
Milos Balac
Dr. Michael Balmer
Henrik Becker
Patrick Bösch
Dr. David Charypar
Dr. Nurhan Cetin
Artem Chakirov
Dr. Yu Chen
Dr. Francesco Ciari
Dr. Christoph Dobler
Thibaut Dubernet
Dr. Alexander Erath
Dr. Matthias Feil
Dr. Gunnar Flötteröd
Pieter Fourie
Dr. Christian Gloor
Dr. Dominik Grether
Dr. Jeremy K. Hackney
Dr. Andreas Horni
Dr. Johannes Illenberger
Dr. Gregor Lämmel
Nicolas Lefebvre
Prof. Kai Nagel
Dr. Konrad Meister
Manuel Moyo
Kirill Müller
Dr. Andreas Neumann
Dr. Thomas Nicolai
Benjamin Kickhöfer
Sergio Ordonez
Dr. Bryan Raney
Dr. Marcel Rieser
Dr. Nadine Rieser
Lijun Sun
Alexander Stahel
Dr. David Strippgen
Michael Van Eggermond
Dr. Rashid Waraich
Michael Zilske

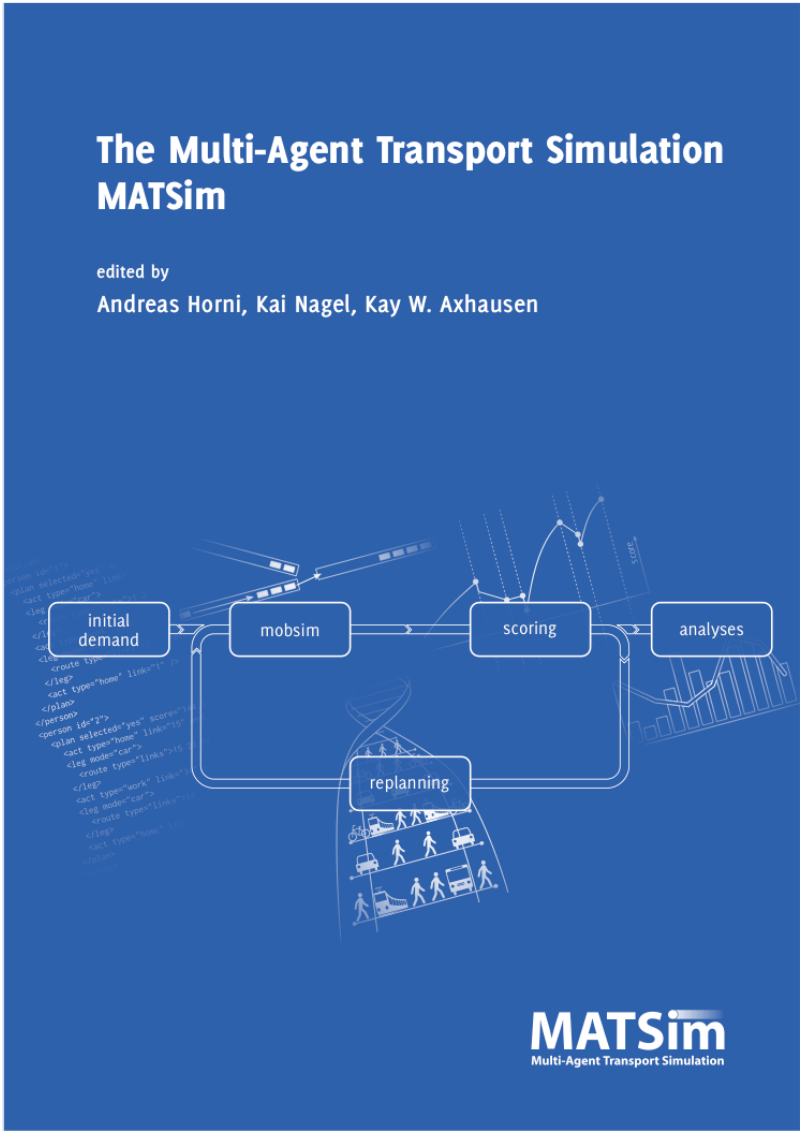
Questions ?

www.matsim.org

www.ivt.ethz.ch
www.futurecities.ethz.ch

www.senozon.com

Questions ?


The Multi-Agent Transport Simulation
MATSim
edited by
Andreas Horni, Kai Nagel, Kay W. Axhausen
initial demand
mobsim
scoring
analyses
replanning
MATSim
Multi-Agent Transport Simulation

References

Chikaraishi, M., A. Fujiwara, F. Nishikawa, H. Seya and J. Zhang (2015) Modeling shopping behavior in a neighborhood with endogenous representation of retail attractiveness, presentation at the The 4th International BinN research seminar "Dynamic modelling based on bounded rationality to understand diversifying travel behavior", University of Tokyo, June 2015.

De Palma, A., N. Picard and P. Waddell (2007) Discrete choice models with capacity constraints: An empirical analysis of the housing market of the greater Paris region, Journal of Urban Economics, 62 (2) 204-230.

Manski, C.F. (1993) Identification of Endogenous Social Effects: The Reflection Problem. The Review of Economic Studies 60, 531-542.

Rust, J. (1987) Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher, Econometrica 55, 999-1033.

Schmutz, Simon (2015) Auswirkung von analytische Einheiten und Aggregationsregeln auf die Verkehrsmittelwahlmodellierung, MSc thesis, Zürich, January 2015.

Schuessler, N. (2010) Accounting for similarities between alternatives in discrete choice models based on high-resolution observations of transport behaviour, ETH Zürich, Zürich.

Vrtic, M. (2003) Simultanes Routen- und Verkehrsmittelwahlmodell, PhD Dissertation, Fakultät für Verkehrswissenschaften, TU Dresden, Dresden.

Walker, J.L., Ehlers, E., Banerjee, I., Dugundji, E.R. (2011) Correcting for endogeneity in behavioral choice models with social influence variables. *Transportation Research Part A: Policy and Practice* **45**, 362-374.

Zilliak, S. and D. McCloskey (2008) *The Cult of Statistical Significance: How the Standard Error Costs Us Jobs, Justice, and Lives*, University of Michigan Press, Ann Arbor.