# Bevorzugter Zitierstil für diesen Vortrag

Axhausen, K.W. (2014) Agent-based or agent based modelling: Reflections on choices, constraints and commitments, keynote at the *14th Conference of the International Association of Travel Behaviour Research (IATBR)*, Windsor, July 2015.

# Agent-based or agent based modelling: Reflections on choices, constraints and commitments

KW Axhausen

IVT
ETH
Zürich

July 2015

Institut für Verkehrsplanung und Transportsysteme
Institute for Transport Planning and Systems

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

# What is the issue?

# Activity-based versus agent-based approaches ?

# A terminological problem ?

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

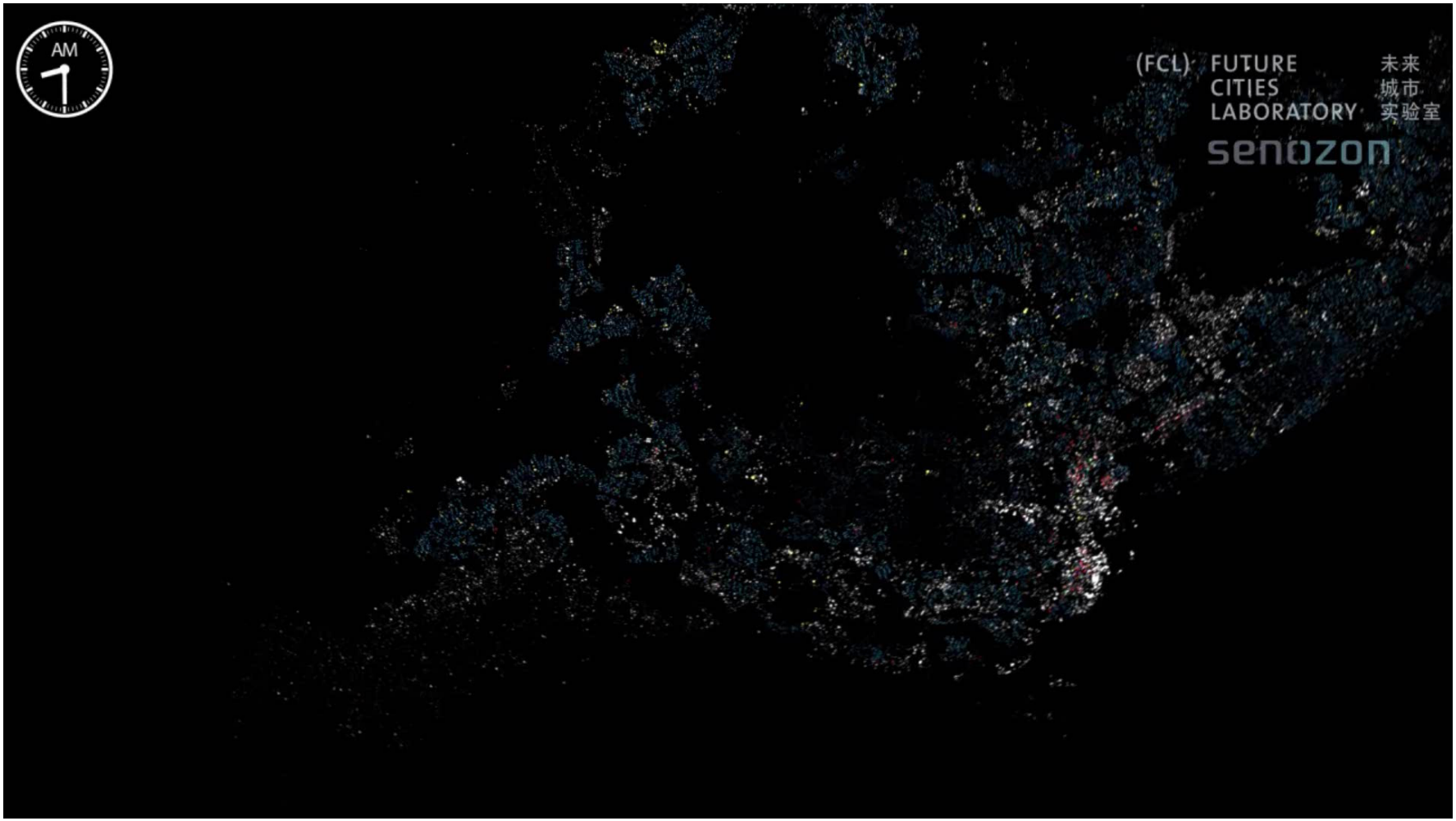
# The typical activity-based model (ABM)

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

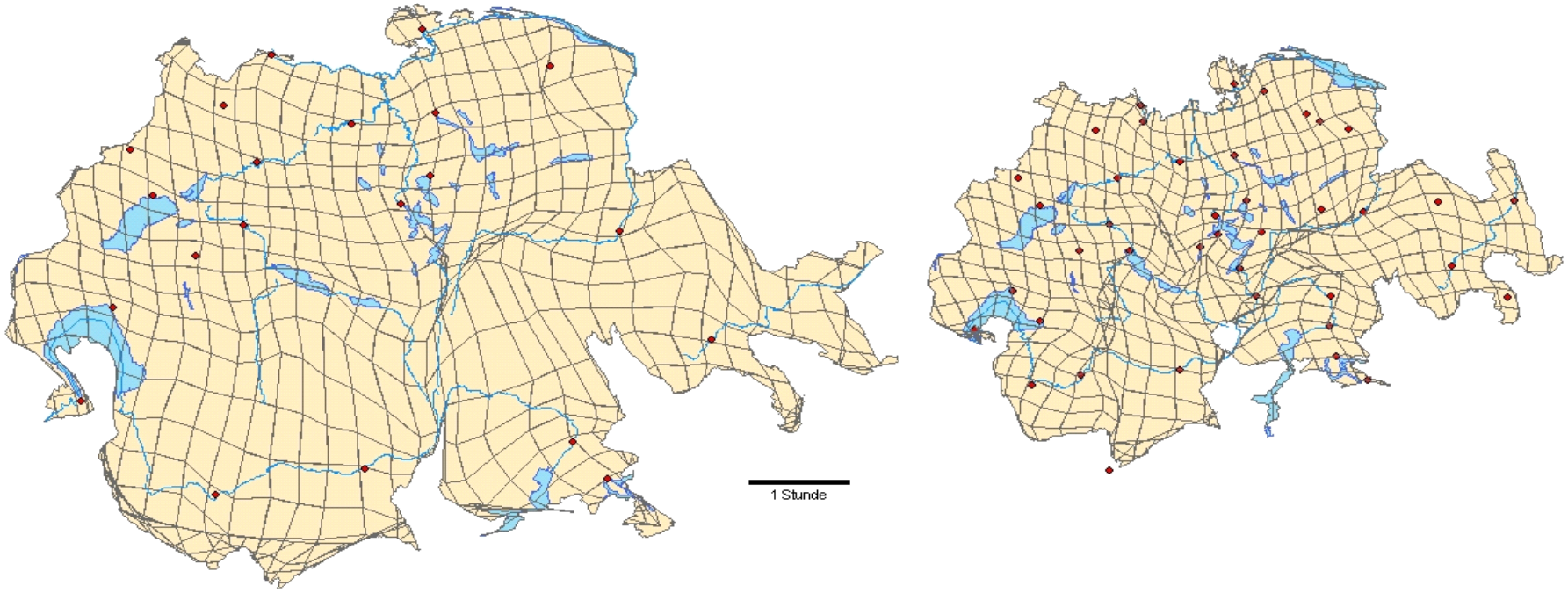
# MATSim

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

# An agent-based model of travel demand: e.g. MATSim Singapore



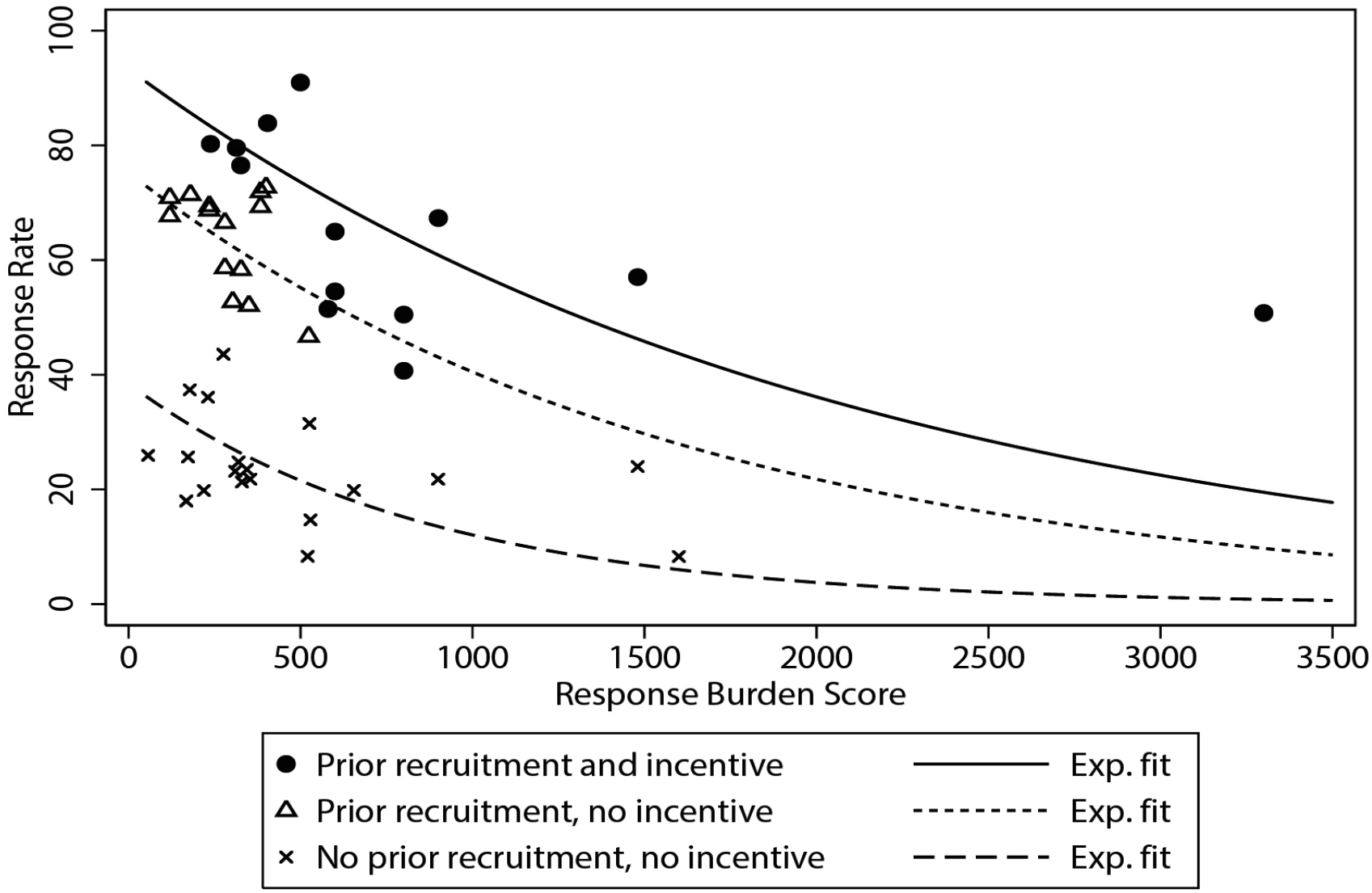
# Long-term response: Road based – Switzerland 1950 and 2000



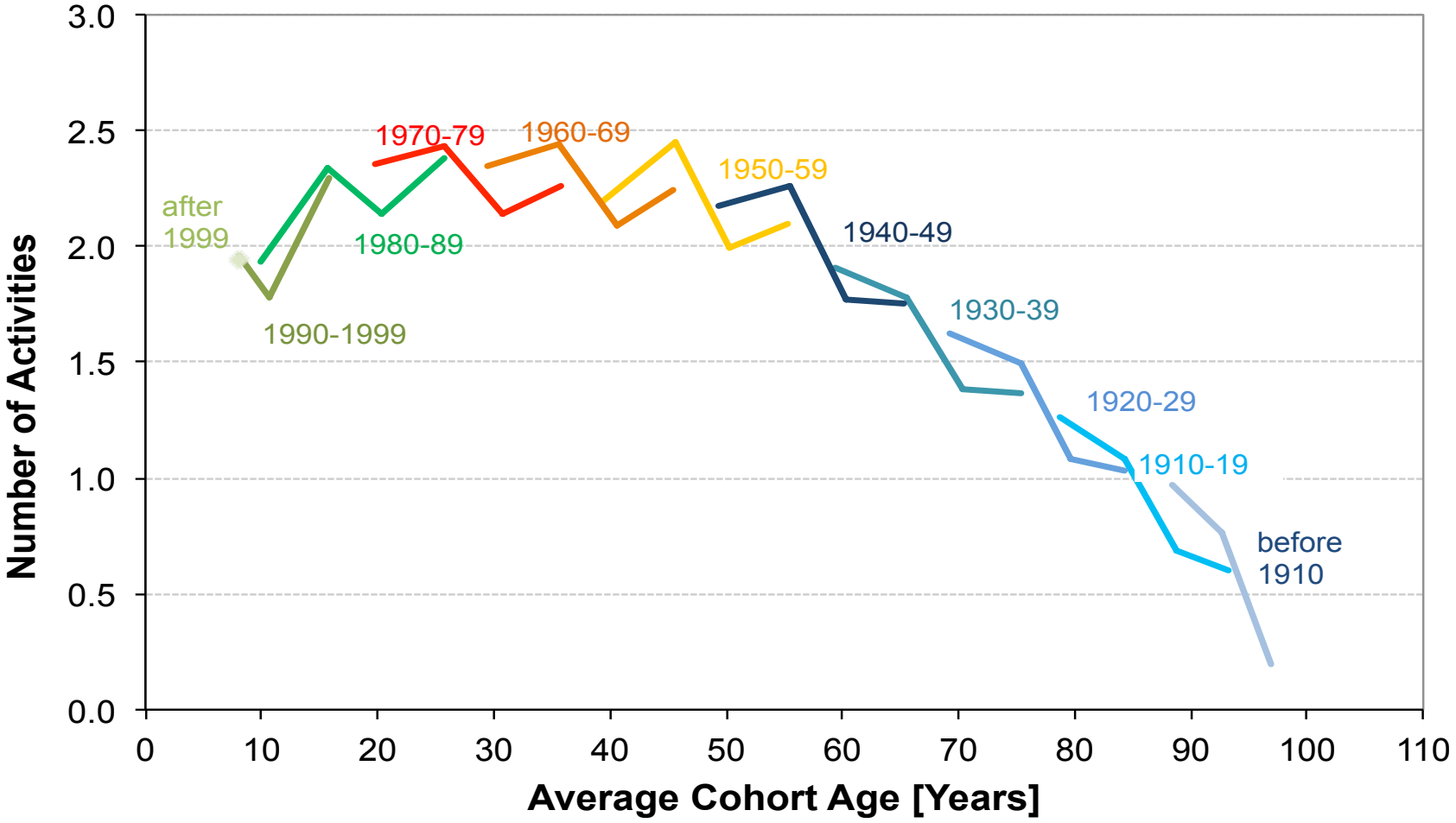

Scherer, 2004

# Data challenges

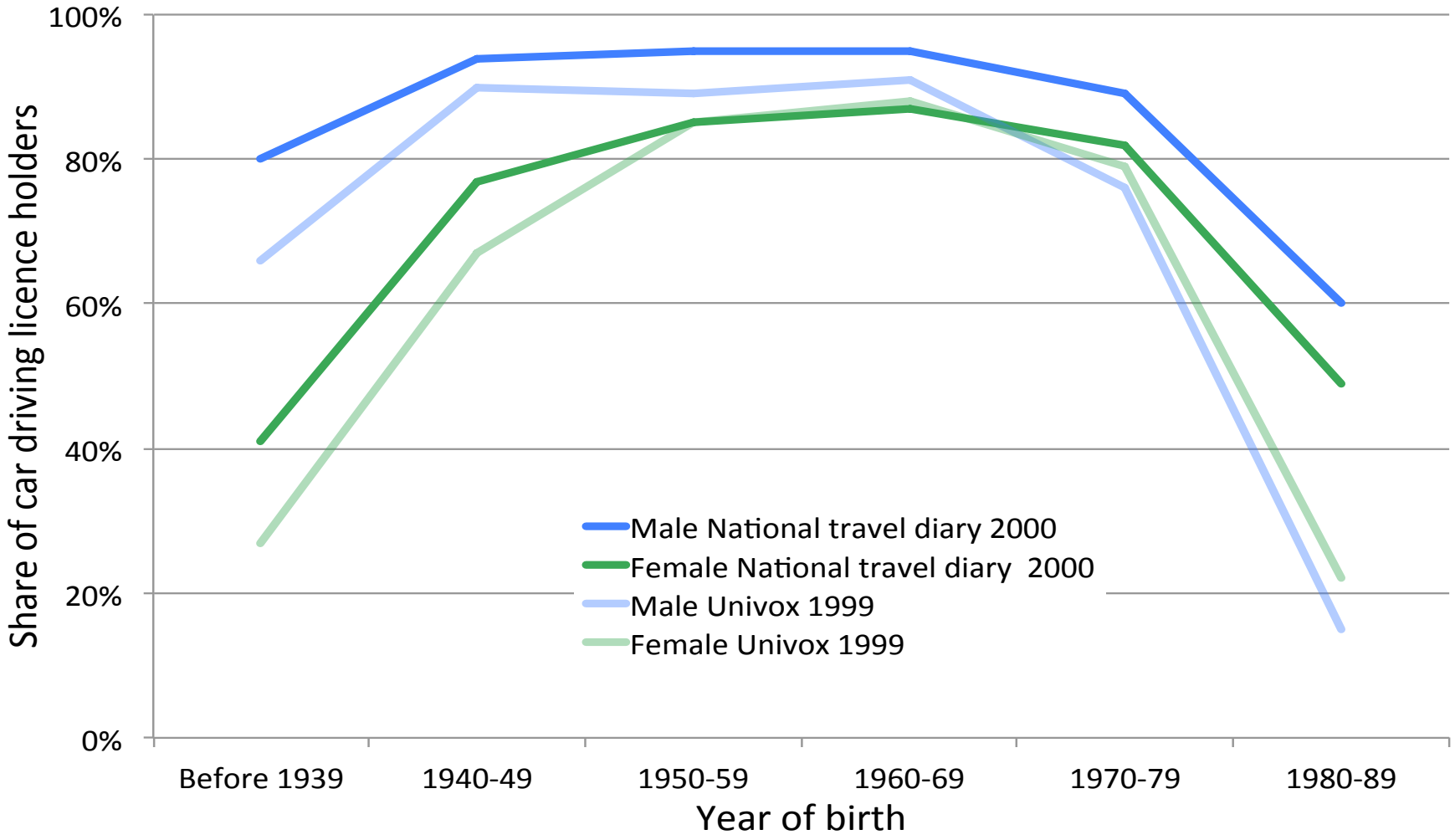
# Response as a function of response burden @IVT, 2015



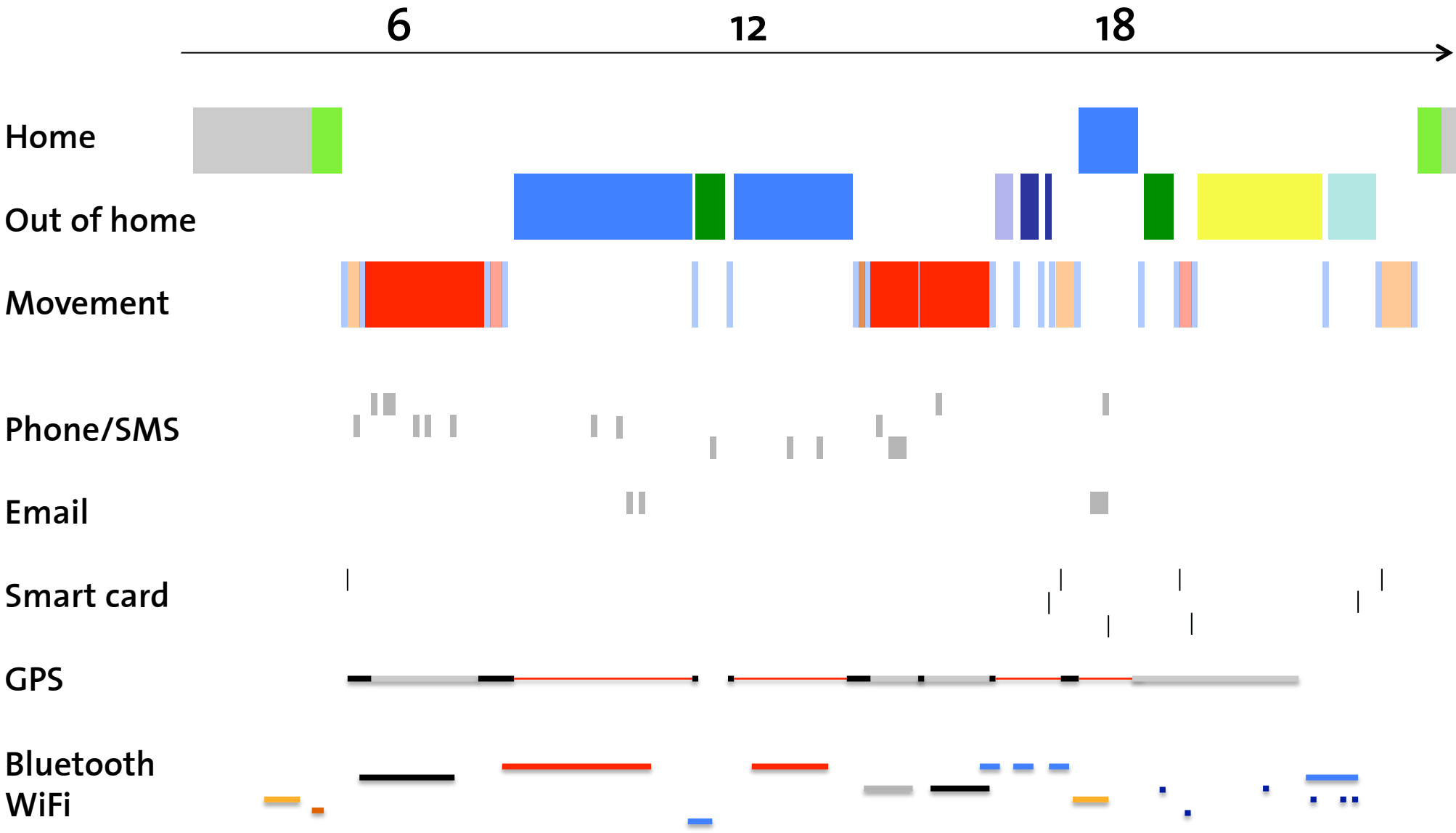
# Do we know the numbers? e.g. daily activities in Switzerland



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



# Activities, movement and traces: A full example record



# What is left ?

True

5 at home

9 Out of home

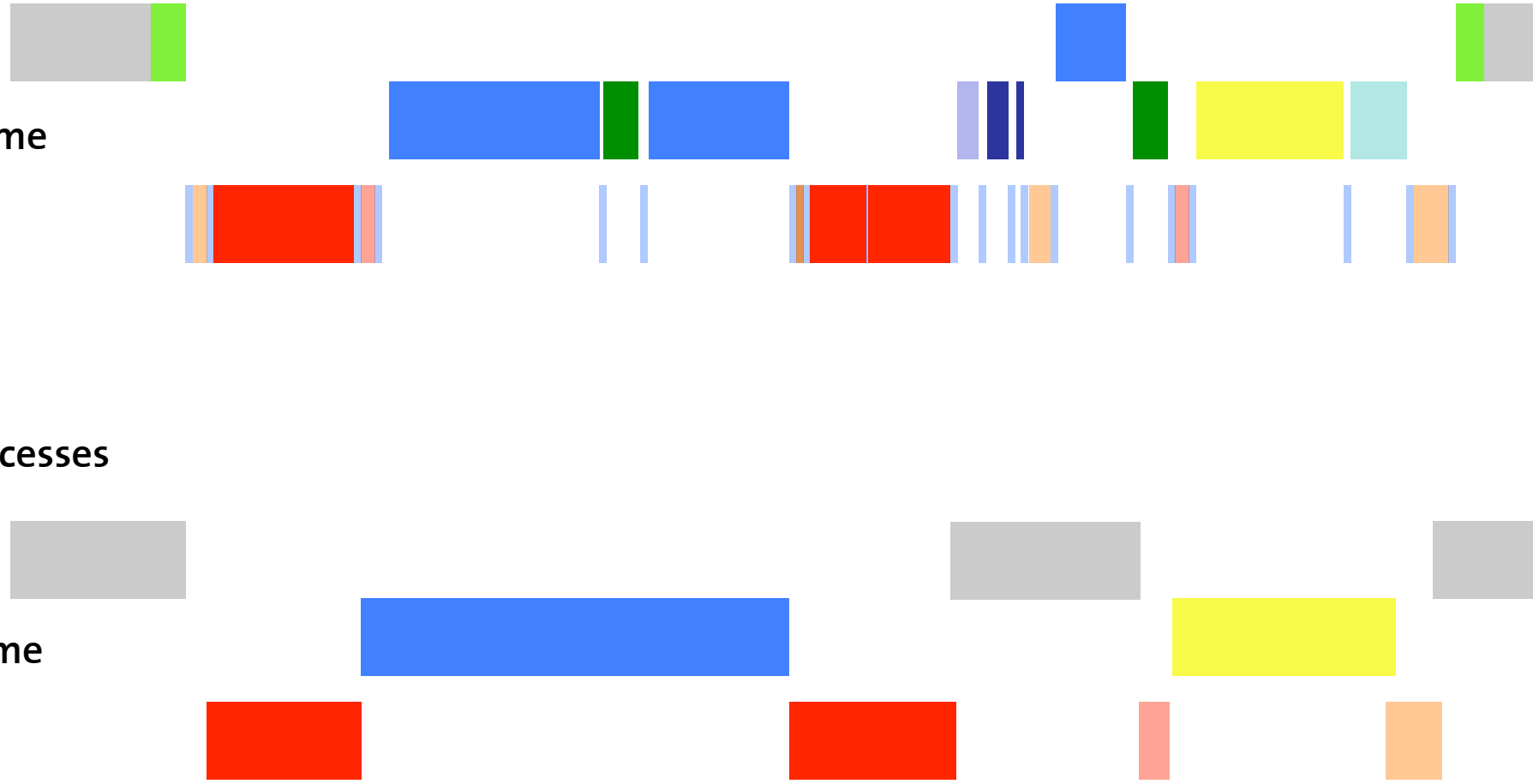
26 Stages,
11 trips,
1 subtour,
2 tours

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) calibrated on counts |
| Still seen in practice | Shortest path on empty planning networks |

# Choice set construction

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

# What should we do ?

# Next steps

- Query what we really need for
  - Cost-benefit analysis
  - Planning of prices and services
  - Planning for the slow modes
  - Social accounting

- High-quality multi-modal surveys to establish the measurement errors (add bluetooth and wifi senders, noise profile)
- Error correction models
- Cross check against third party sources

- Treat survey data as indicators in a measurement model
- Treat traces as indicators in a measurement model

# , but especially

- Remember, that observation/surveys are ‘talk’ and
- Treat respondents as partners in a talk, discussion:

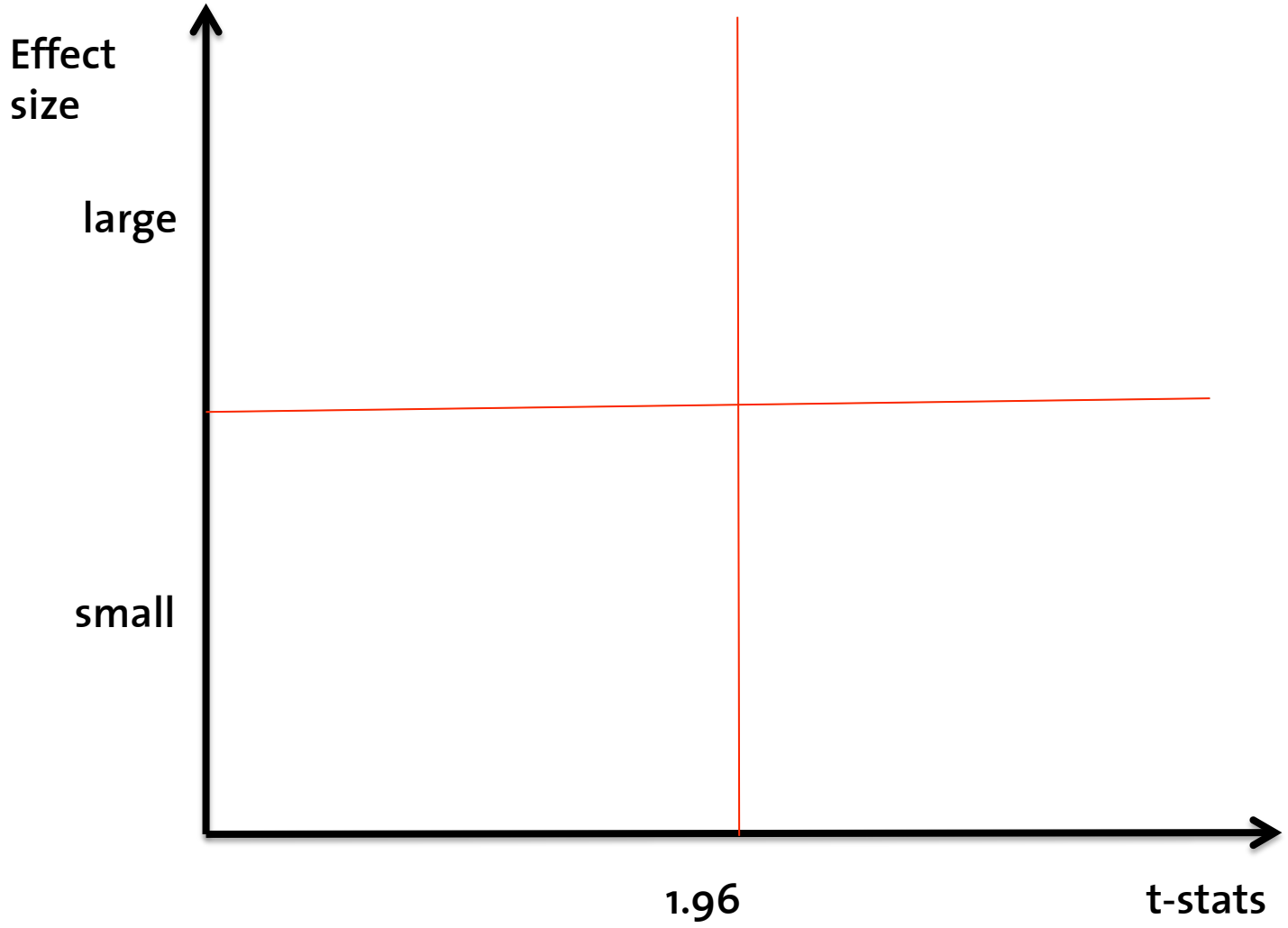
  - Frame your request in a way which addresses them in a clearly defined social role (citizen, driver, customer, etc.)
  - Match your role and the questions
  - Account for their constraints (readability of text, full guidance through the forms, require no calculations – unless necessary, speak their ‘language’)
  - Be as complex, as the topic warrants, requires, but not more so
  - Don’t surprise them with unannounced requests
  - Don’t ask them to do work you can do for them
  - If appropriate, provide an incentive, acknowledgement

# Modelling challenges

# Modelling challenges: The usual worries

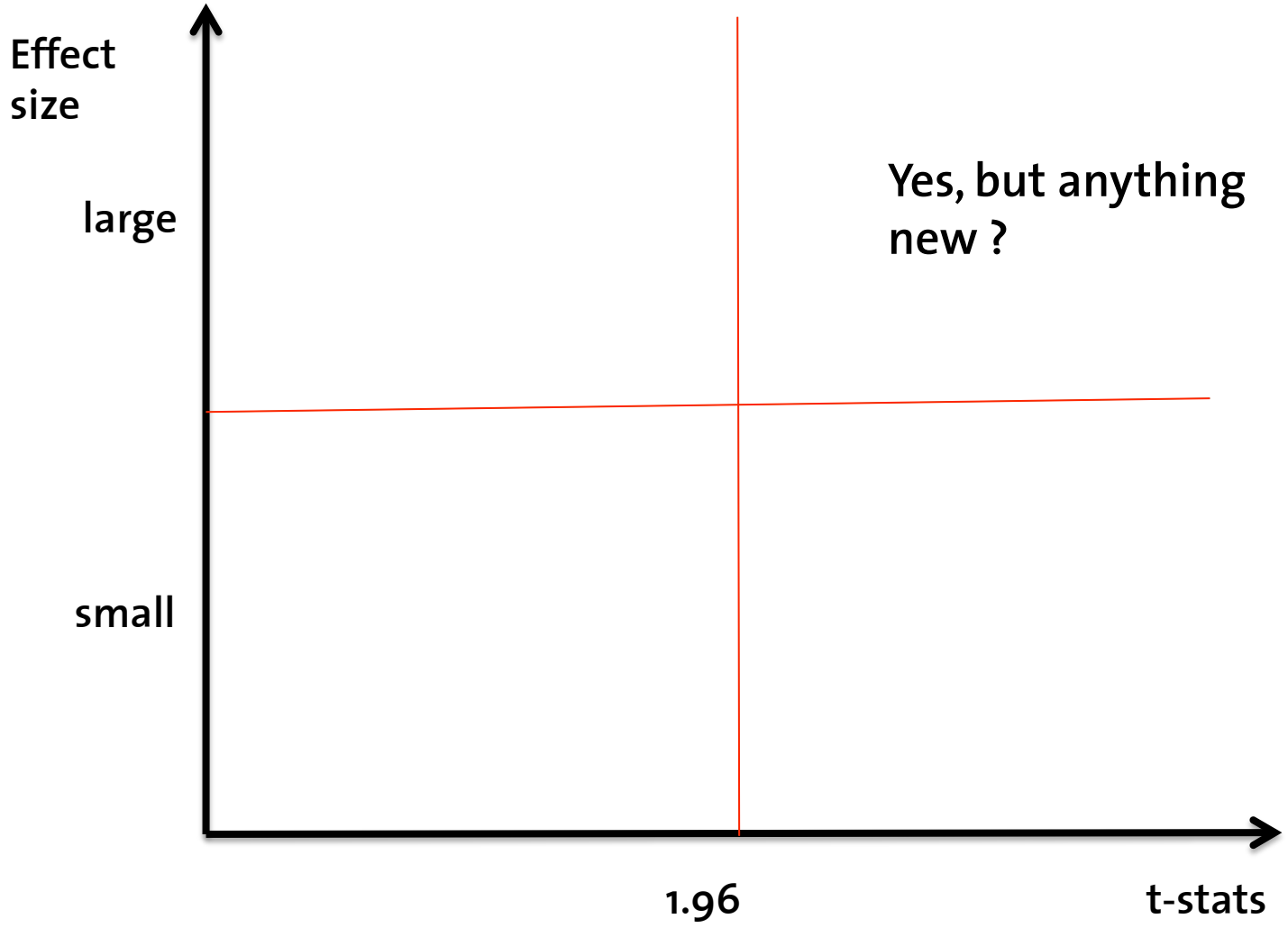
| | |
|---|---|
| Error heterogenity | Is it always checked ? |
| Spatial correlations | Are they always checked ? |
| Temporal correlations | Are they always checked ? |
| Independence | Do we check the correlations of the independent variables (sample) thoroughly enough? |
| Endogenity | Do we fully 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 ?



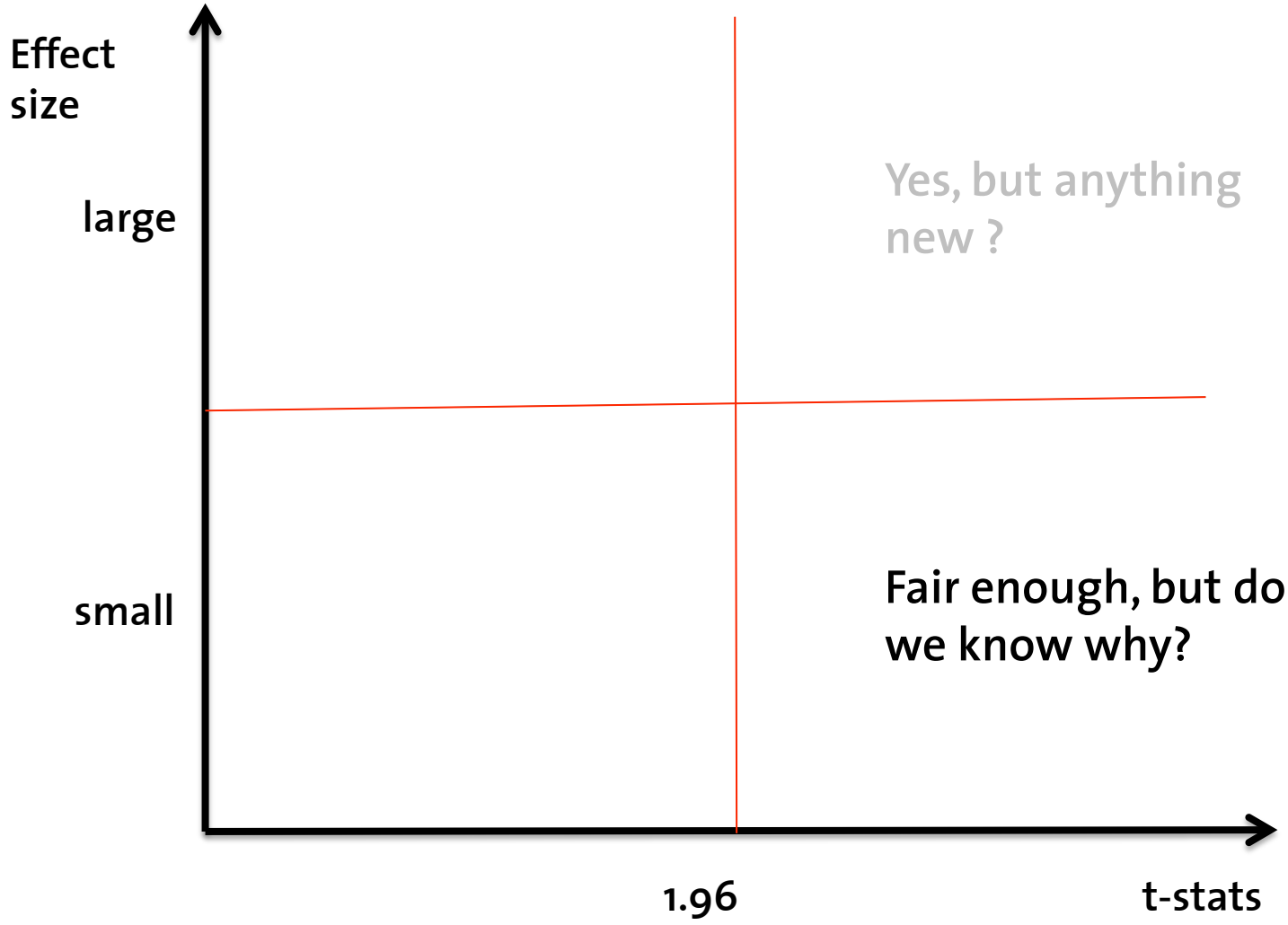

Adapted from Zilliak and McCloskey (2008)

# Modelling challenges: Substance or t-tests ?



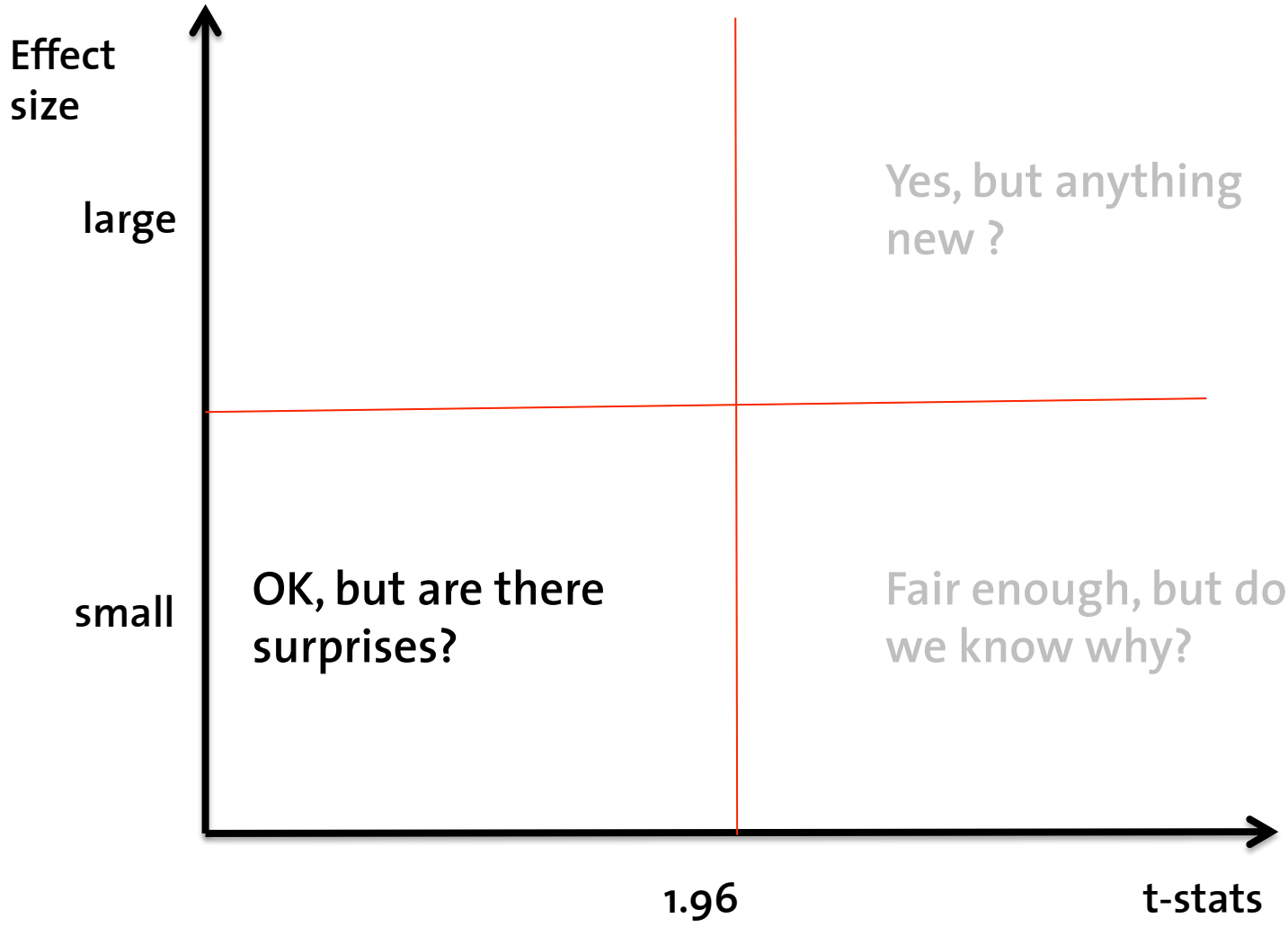

Adapted from Zilliak and McCloskey (2008)

# Modelling challenges: Substance or t-tests ?



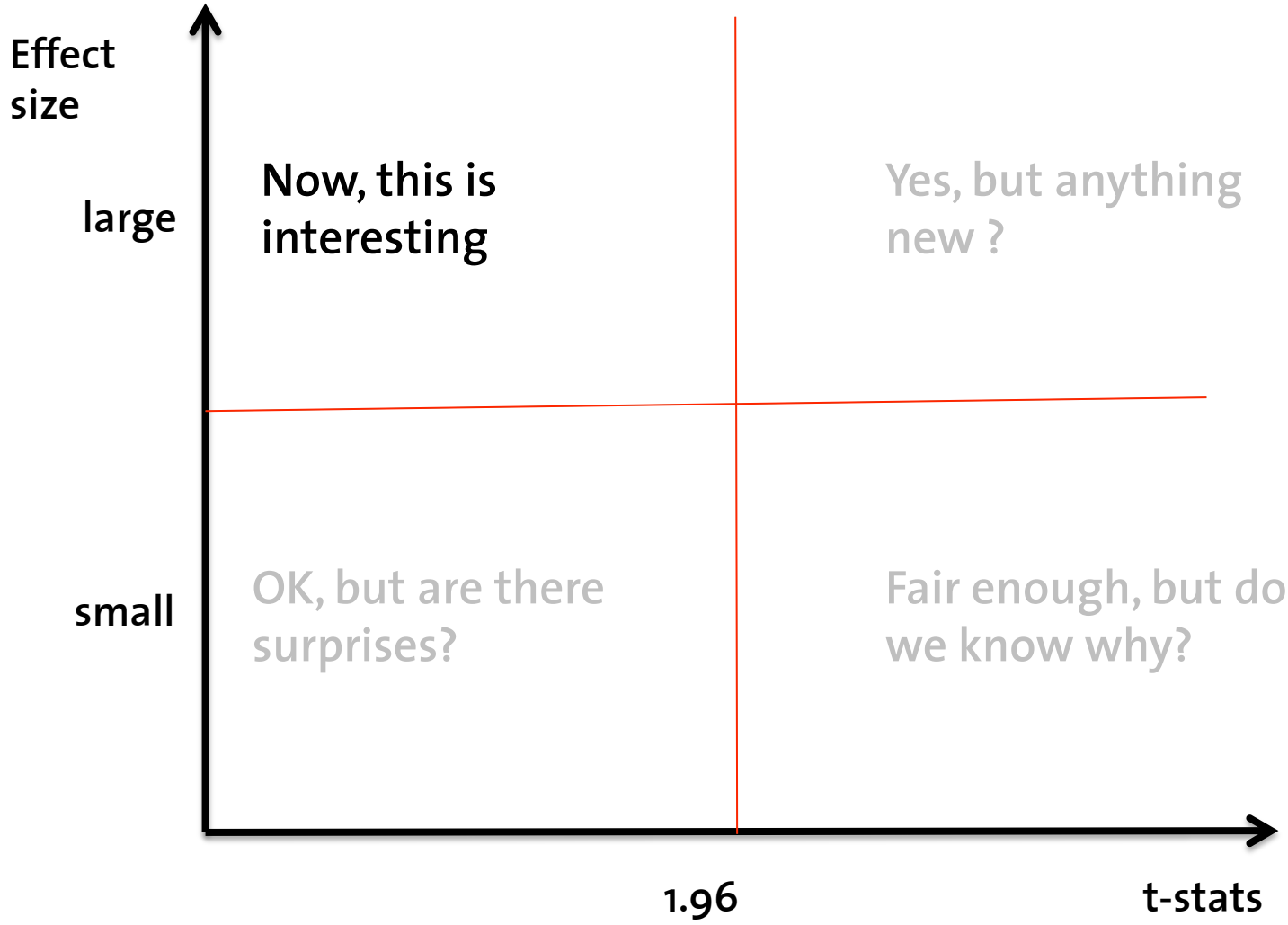

Adapted from Zilliak and McCloskey (2008)

# Modelling challenges: Substance or t-tests ?



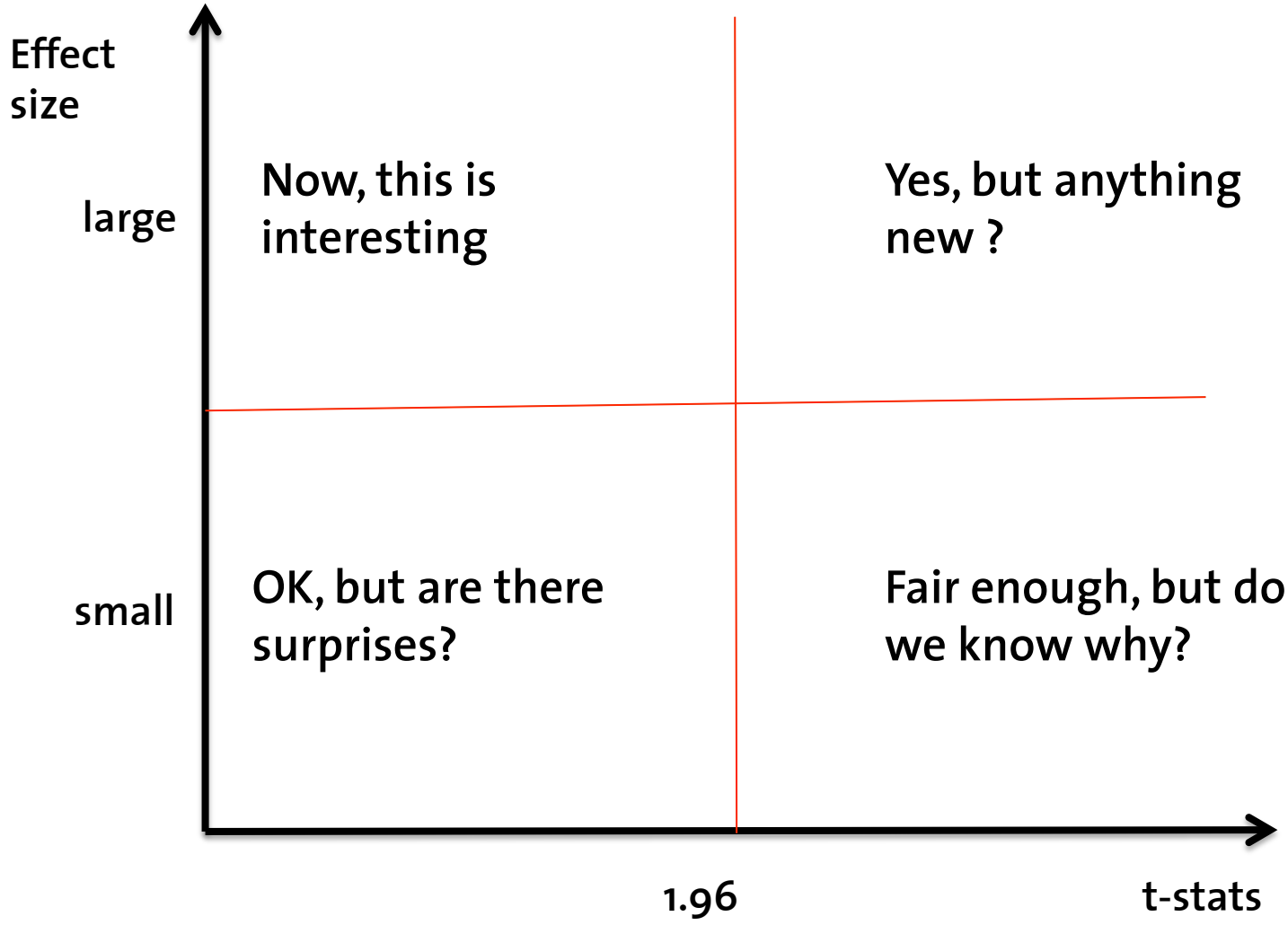

Adapted from Zilliak and McCloskey (2008)

# Modelling challenges: Substance or t-tests ?




Adapted from Zilliak and McCloskey (2008)

# Modelling challenges: Substance or t-tests ?




Adapted from Zilliak and McCloskey (2008)

# Choice modelling challenges

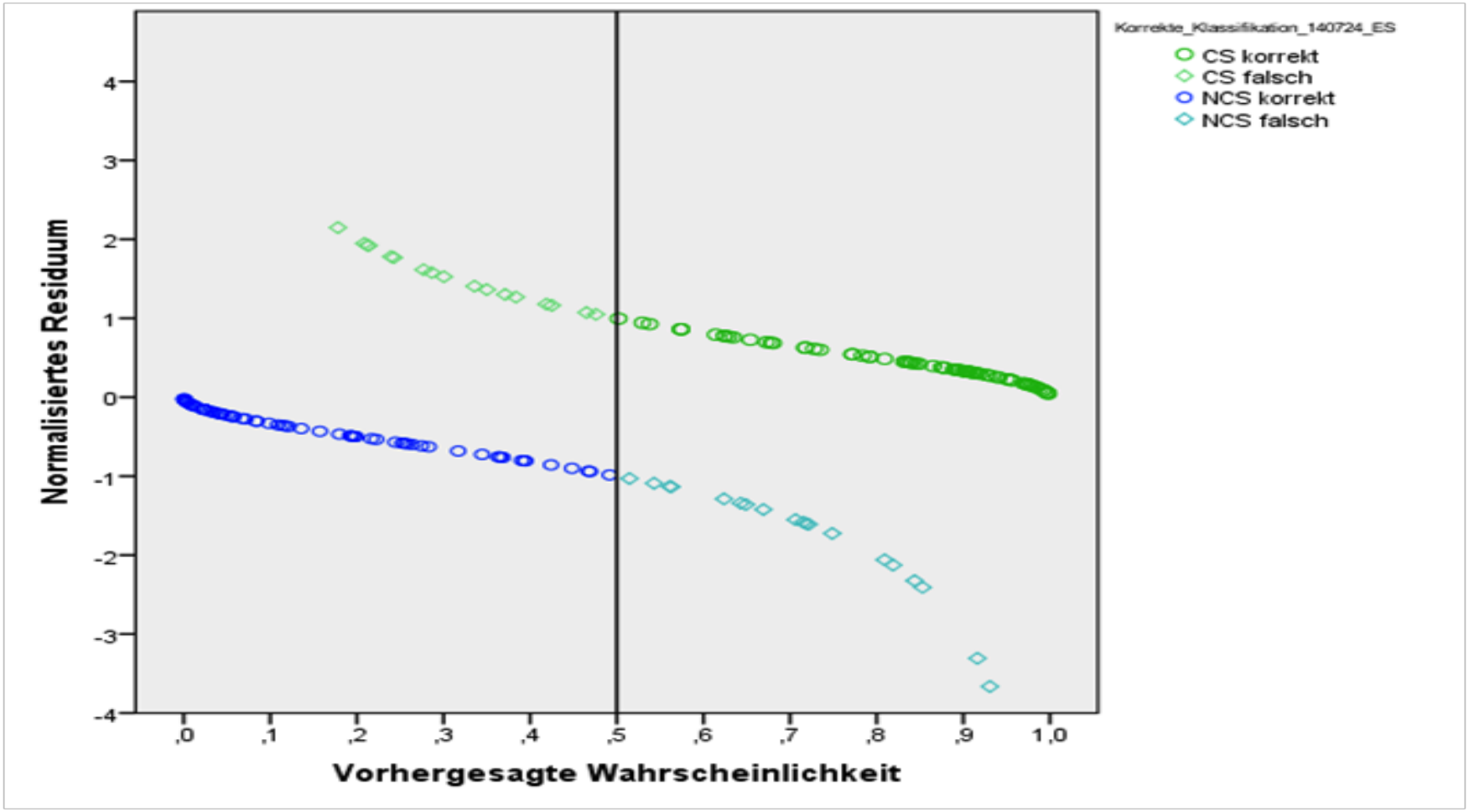
# Choice modelling challenges: The usual worries

| | |
|---|---|
| Error heterogenity | Is it always checked ? |
| Spatial correlations | Are they it 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 ? |

# Choice modelling challenges: less usual concerns

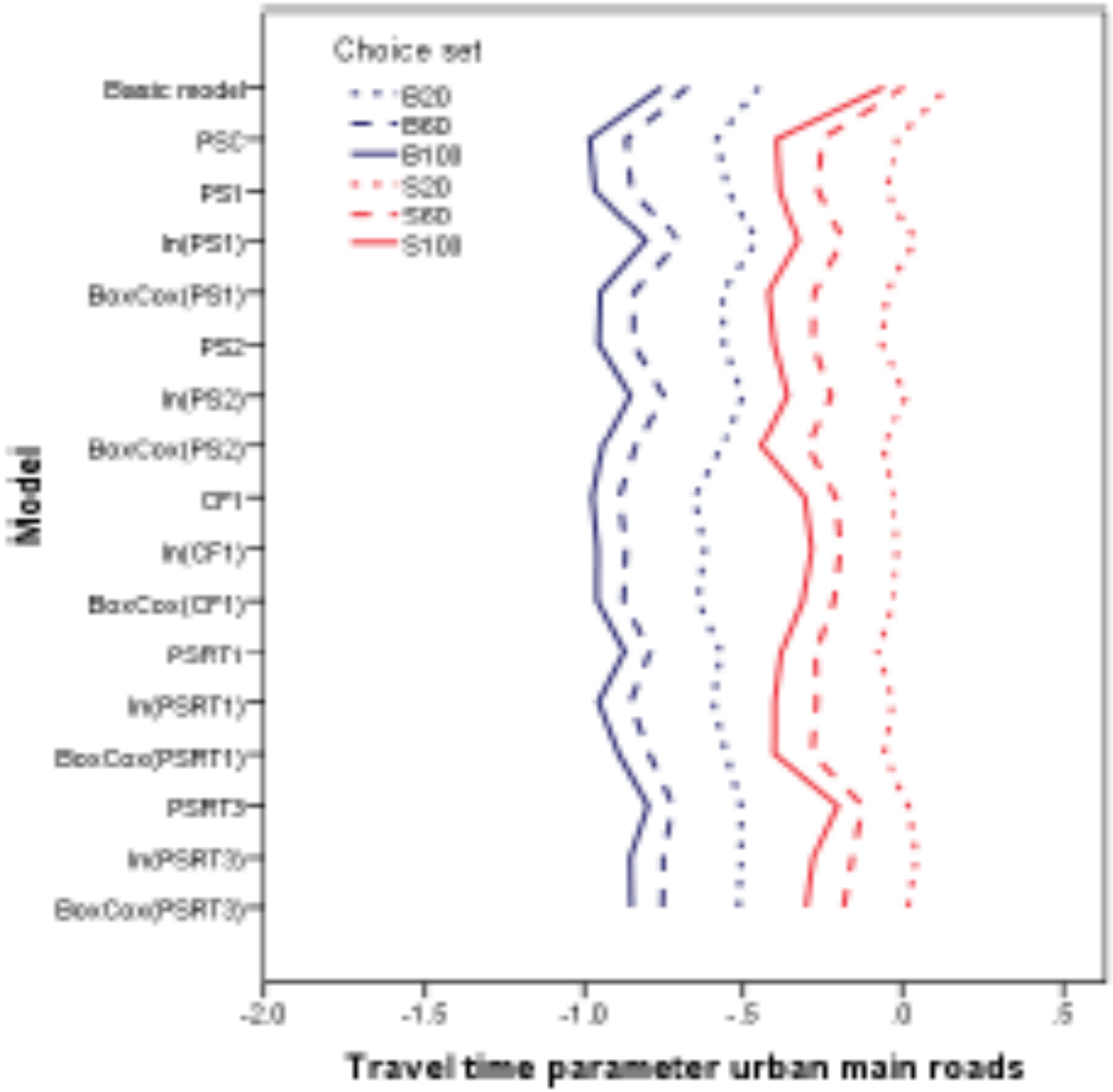
| | |
|---|---|
| Error heterogenity | Why don't we check them ? |
| Number of non-chosen alternatives | How much leverage do they have for your problem? |
| Number of choice sets | How stable are our estimates? |
| Capacity constraints | Do we check for their impact on the parameters? (attribute values of the known (non)chosen alternatives) |
| Unit of analysis | Do we have a MAUP problem? |

# Residuals: False positives of a membership model




Source: Kopp (Forthcomng)

# Number of non-chosen alternatives: routes




Source: Schüssler (2010)

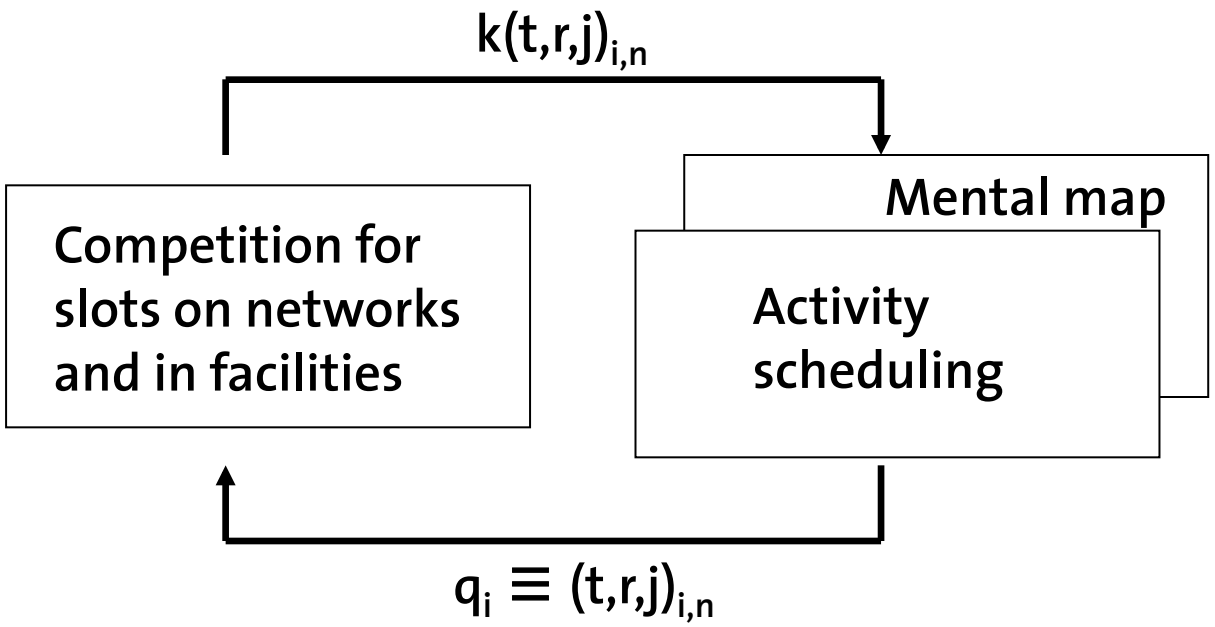
# Number of choice sets: residential choice

| MEASUREMENTS | ESTIMATES | | | | | |
|---|---|---|---|---|---|---|
| | DAT1 | | DAT2 | | DAT3 | |
| **Household** | | | | | | |
| DIST_PREVLOC | -5.440 | ** | -7.070 | ** | -8.740 | ** |
| DIST_WORK | -2.460 | * | -3.220 | * | -3.880 | * |
| ETA_PREVLOC | 0.192 | ** | 0.163 | ** | 0.135 | ** |
| ETA_WORK | 0.218 | ** | 0.203 | ** | 0.166 | ** |
| **Accessibility** | | | | | | |
| MIVACC_CAR | -0.233 | | -0.302 | ** | -0.187 | |
| PTACC_NOCAR | 0.555 | ** | 0.541 | ** | 0.547 | ** |
| **Socioeconomic Environment** | | | | | | |
| SAME_HH_AGE_SHARE | 0.782 | ** | 0.684 | ** | 0.634 | * |
| $R^2$ | 0.508 | | 0.529 | | 0.524 | |
| adj $R^2$ | 0.500 | | 0.522 | | 0.517 | |

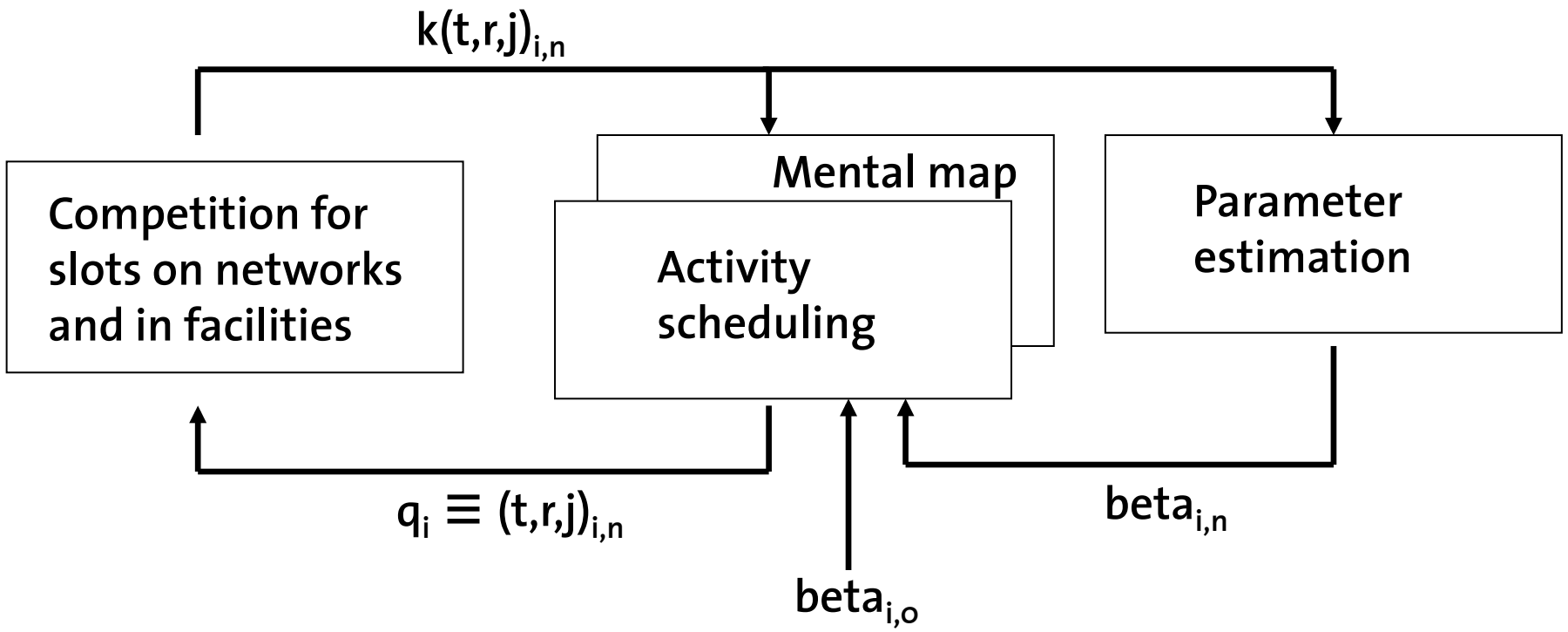
Adapted from: Schirmer (Forthcoming)

# Accounting for consistency

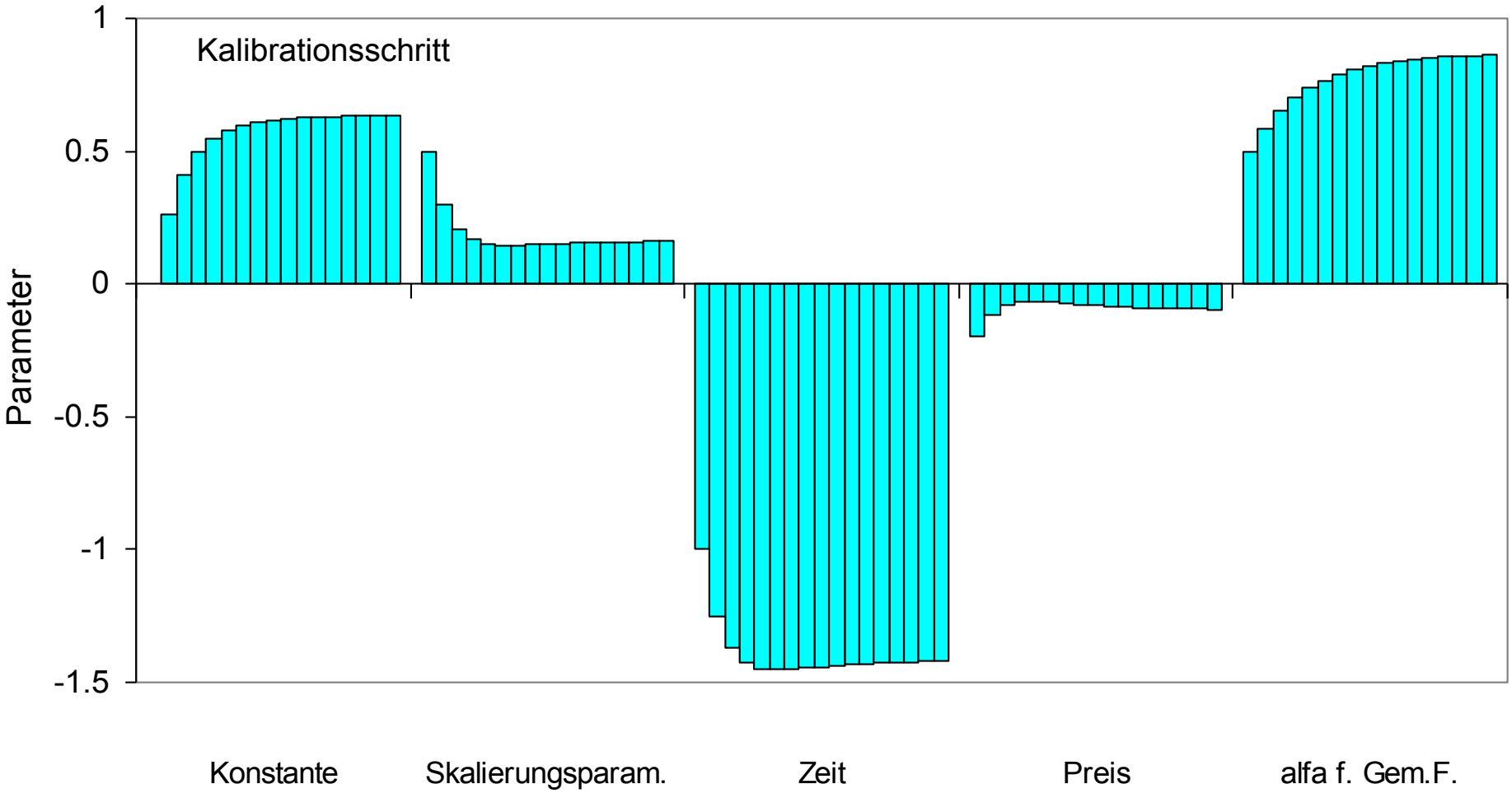
# Learning approach of the generic one-day transport model



# 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 ?

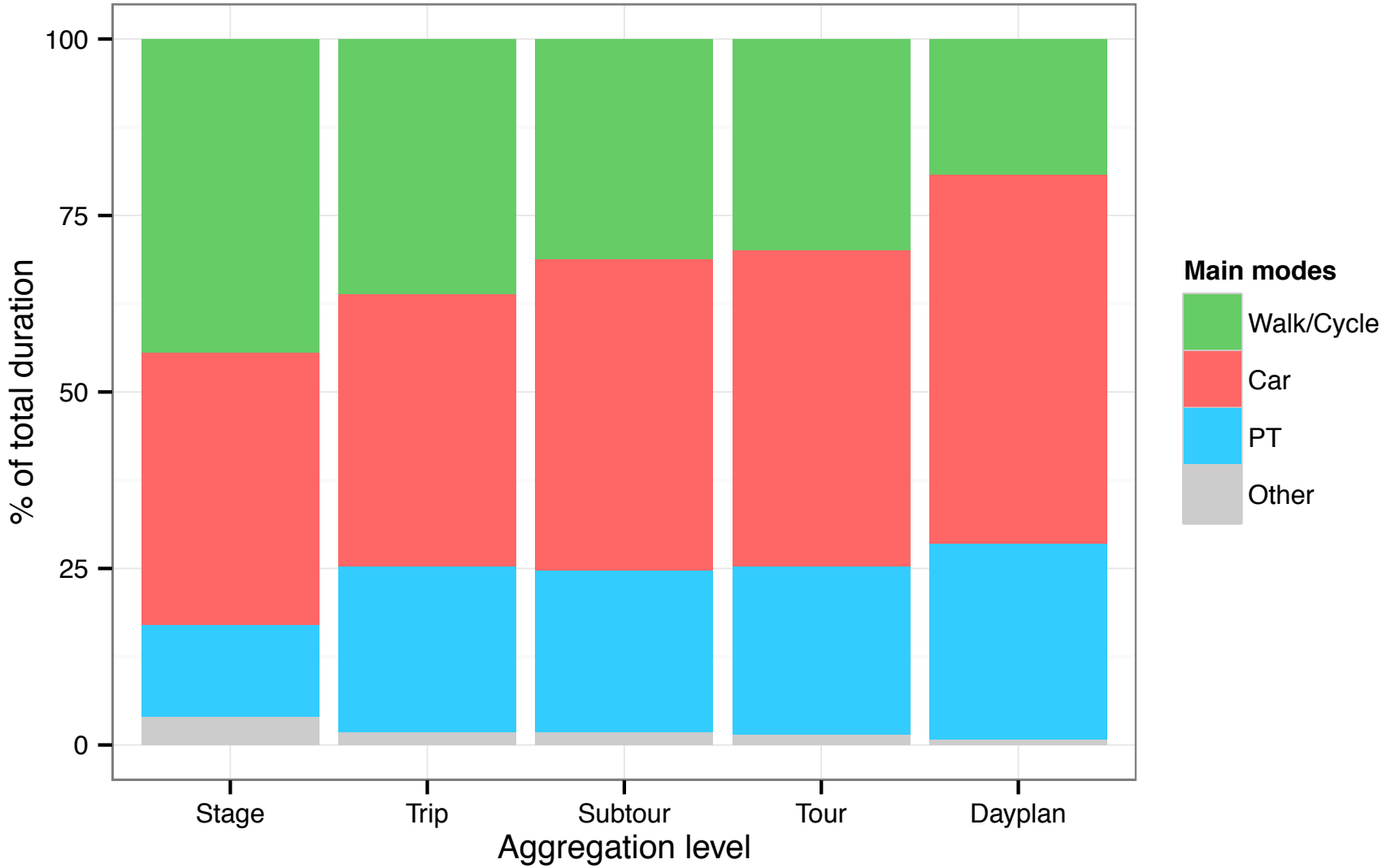
# Do we have a MAUP-like problem for DCM?

- Location choice, obviously
- Route choice, obviously
- Time-of-day choice, obviously

- But also, mode choice

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

# Swiss national travel diary 2010: Main mode by aggregation



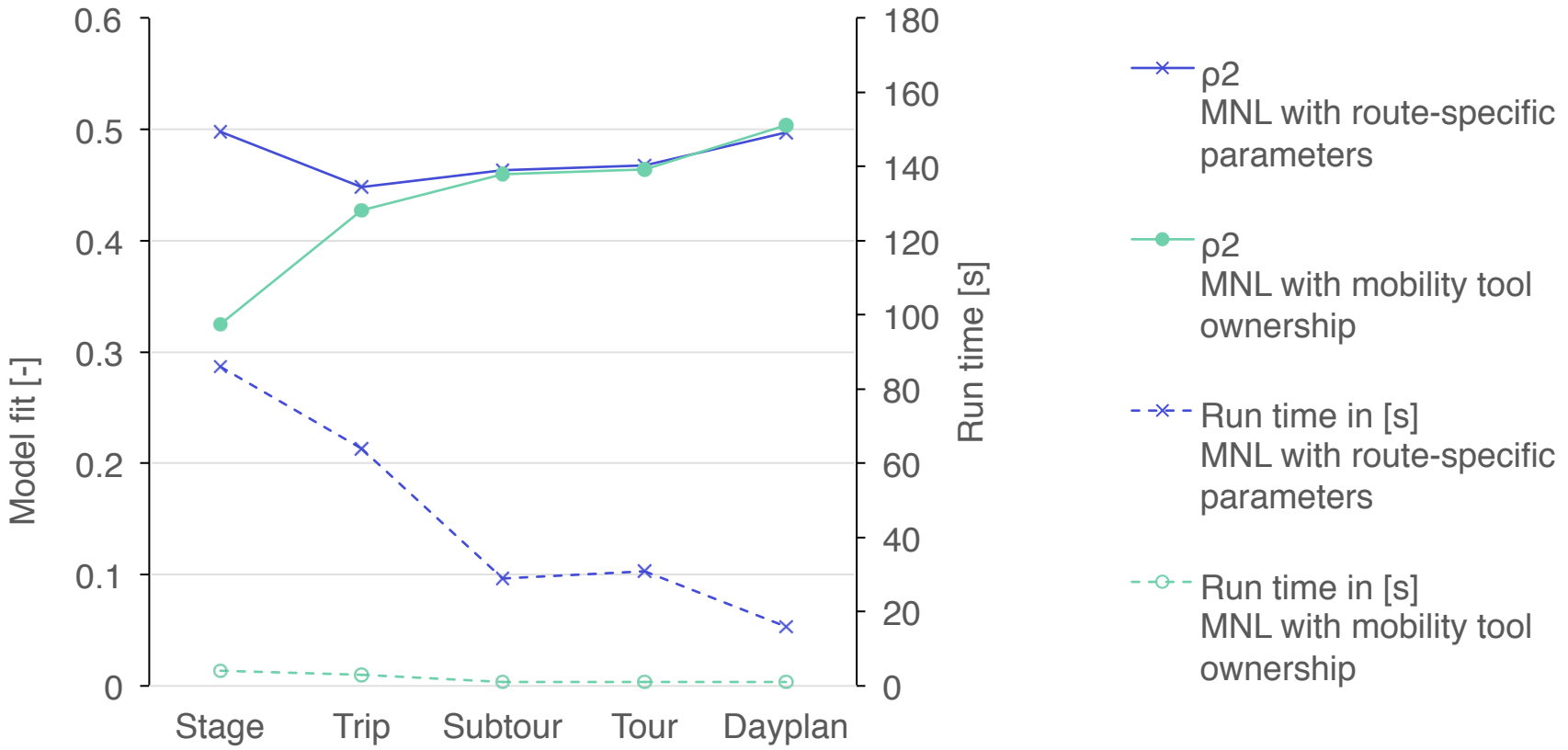
# Do we have a MAUP-like problem for DCM?

| | | 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)

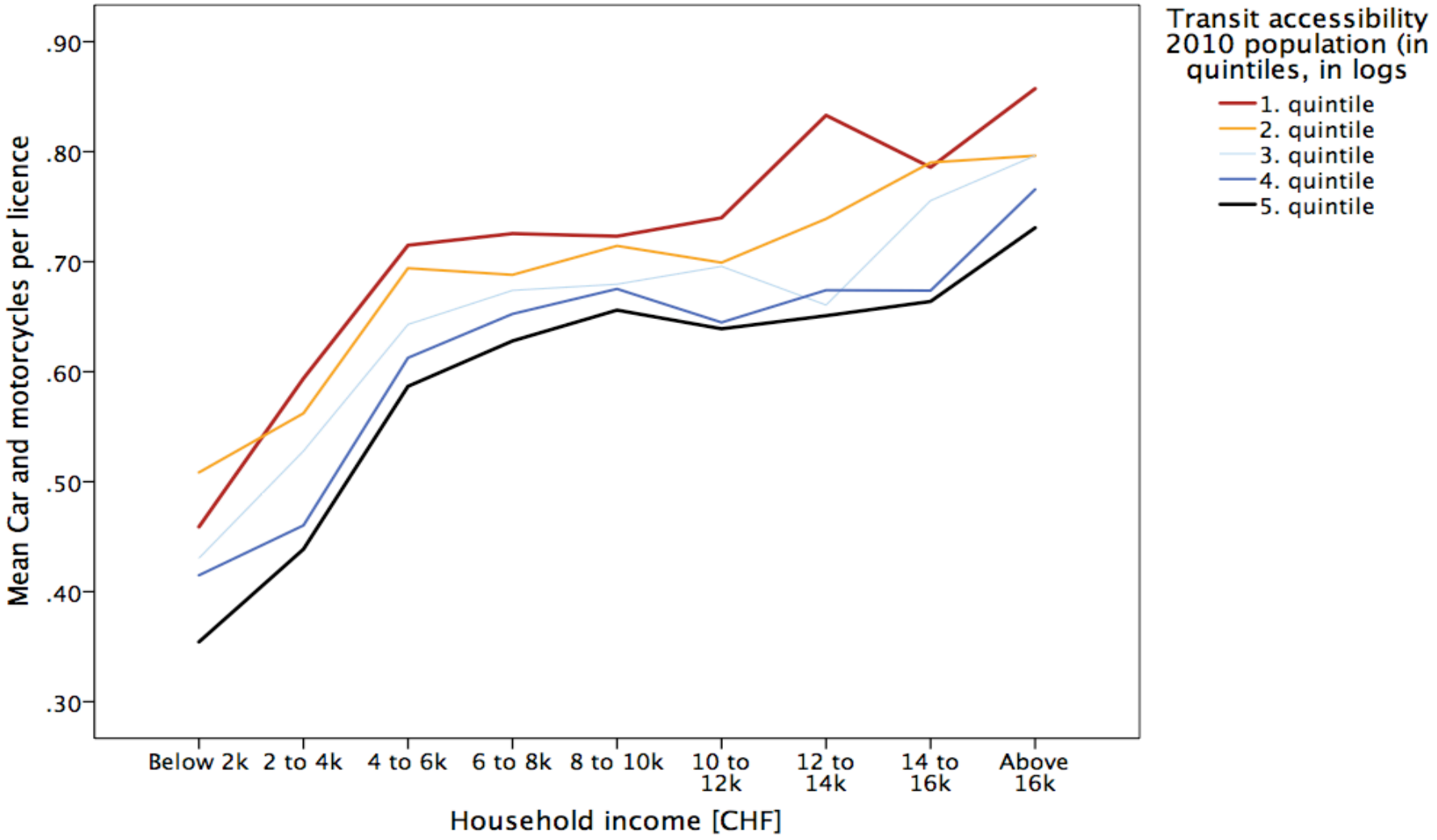
# Do we get the time horizon right?

# Do we get the time horizon right?



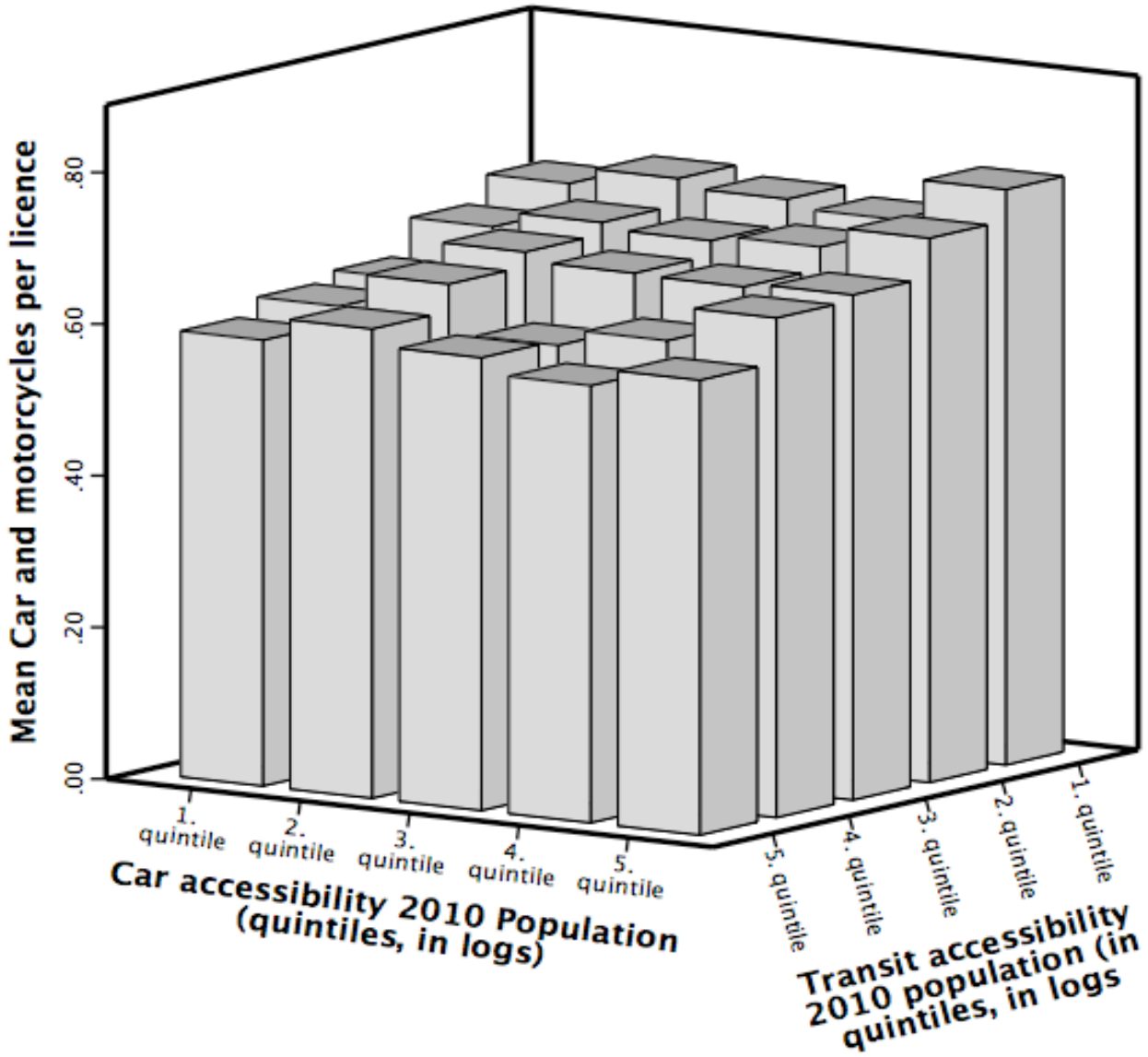

Source: Schmujtz (2015)

# 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 a trade-off between daily 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

# What should we do ?

# Next steps

- Become more systematic
  - Test for choice set size effects
  - Test for the stability of the estimates wrt choice set
  - Test for the stability wrt imputation of the attribute values
- Check for the right unit of analysis
- Check for the right set of explanatory variables
- Check for the appropriate time horizon
- Talk about substance

# MATSim

# MATSim: A GNU public licence software project

# Known implementations

# MATSim: A GNU public licence software project

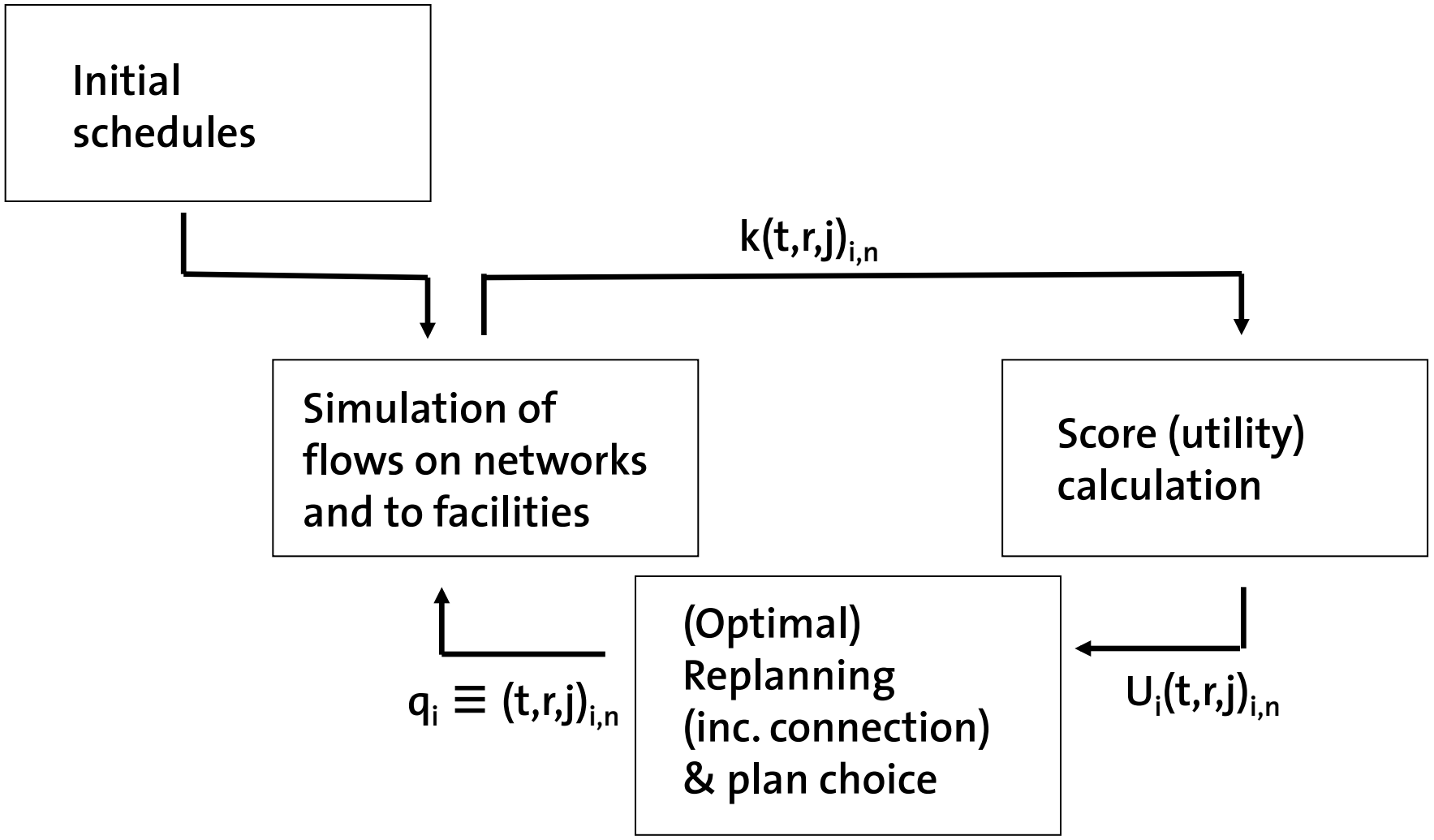
Main partners:

- 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
- University of Seoul

# Equilibrium search in MATSim



# Co-evolution – Issues

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

# Activity schedule dimensions

# 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

# 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 for welfare assessment

- 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

- Daily versus non-daily choices (Overreach of the NL – models ?)

- How many robustness tests should we report in choice modelling papers?

- Talk about substance

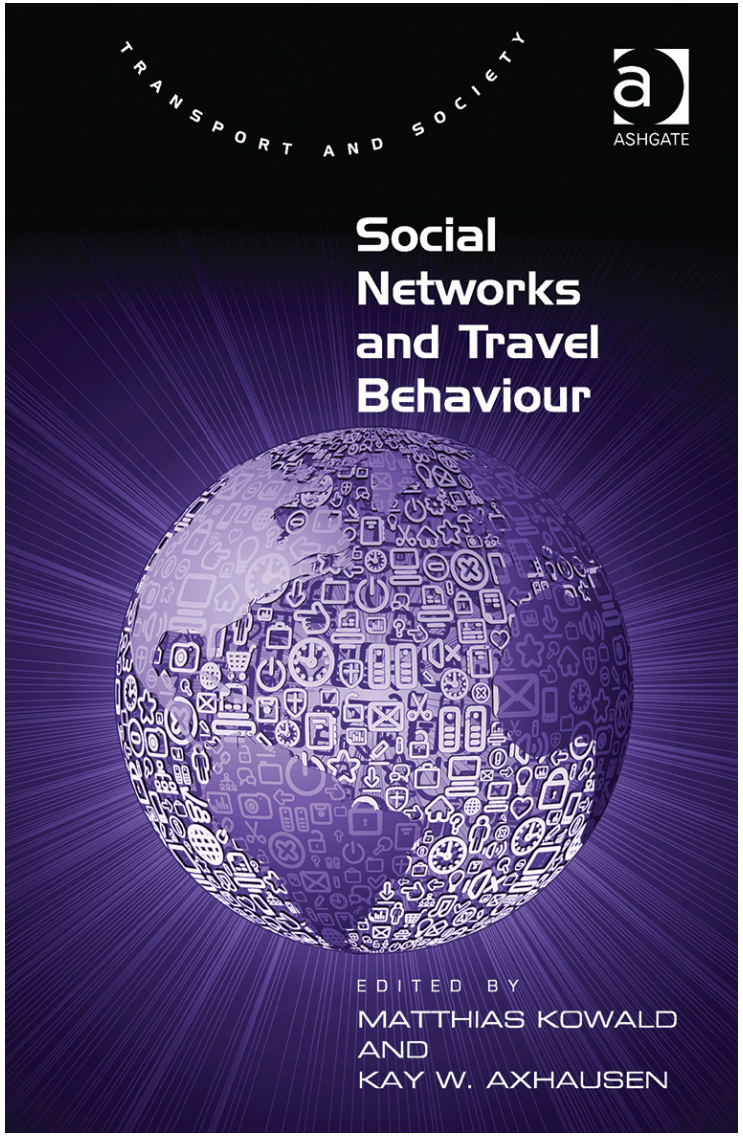
## Questions ?

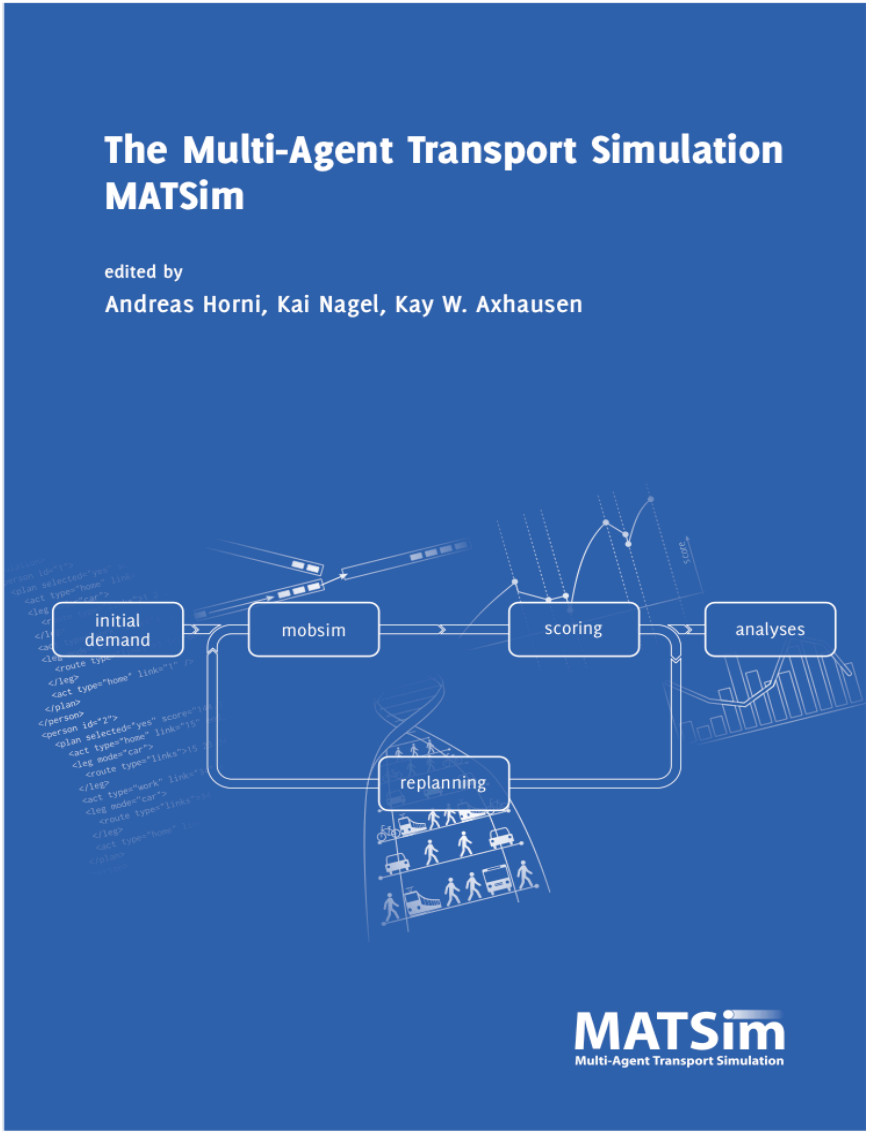
**www.matsim.org**

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

**www.senozon.com**

# Questions ?


TRANSPORT AND SOCIETY
ASHGATE
Social Networks and Travel Behaviour
EDITED BY
MATTHIAS KOWALD
AND
KAY W. AXHAUSEN


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.

Jäggi, B. (Forthcoming) Decision modeling on the household level for energy, fleet choice and expenditure, , Dissertation, ETH Zürich, Zürich.