


# Data collection / Simulation

**Other Conference Item****Author(s):**

Axhausen, Kay W. 

**Publication date:**

2017-04

**Permanent link:**

<https://doi.org/10.3929/ethz-b-000130422>

**Rights / license:**

In Copyright - Non-Commercial Use Permitted

# Preferred citation style

---

Axhausen, K.W. (2017) Data collection / Simulation, presentation at NSF Workshop “Advancing the science of transportation demand modelling”, Berkeley, April 2017.

.

# Data collection / Simulation

KW Axhausen

IVT

ETH

Zürich

April 2017

 Institut für Verkehrsplanung und Transportsysteme  
Institute for Transport Planning and Systems

**ETH**

Eidgenössische Technische Hochschule Zürich  
Swiss Federal Institute of Technology Zurich

# Acknowledgments

---

B Schmid for the update of the response rate analysis

M Hohenfellner/F Ciari for the MATSim video

A Loder, A Schreiber, T Rutherford for the integrated aggregate model

A Loder and L Ambühl for the 3d MFD

# Example: Surveys

---

# Some survey issues

---

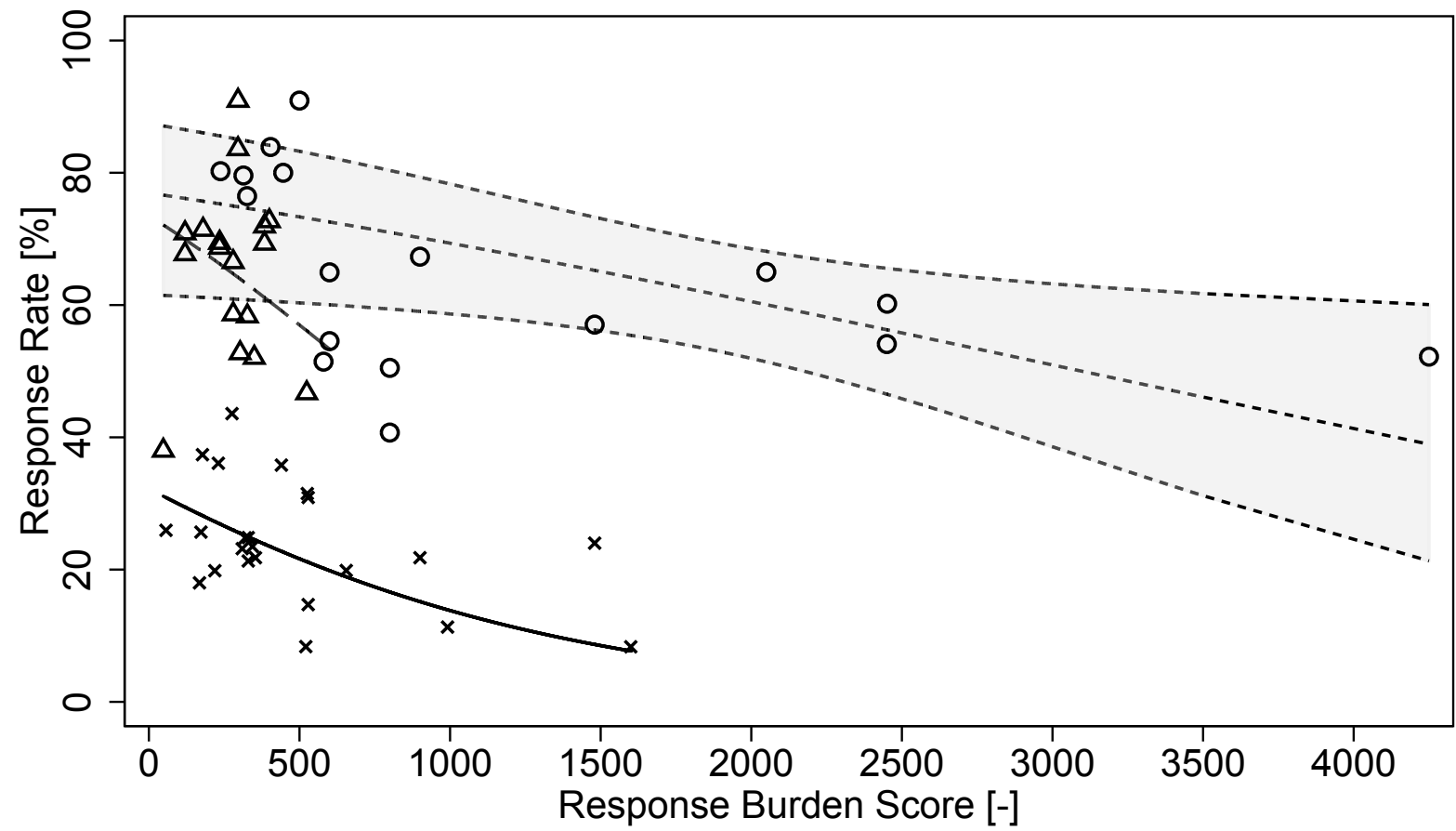
- Selection of the sampling frame
- Self-selection into survey
  - Recruitment rates
  - Response rates
- Soft refusal
  - Being 'immobile'
  - Not reporting whole tours
  - Grouping activities together
  - Missing stages

# Soft refusal: ORANGE versus national travel diary

---

Data	Tracked Interval	Surveyed Persons	Mobile Persons	Mobile Share	Selected for analysis
CDR	30 days	1'388'941	814'381	58.6%	79'874
ENTD	28 days	18'632	4'796	25.7%	4'796
ENTD	91 days	18'632	8'743	46.9%	8'743

# Response rate versus response burden @IVT



○ Prior recruitment and incentive	----- Fit and 95 % CI
△ Prior recruitment, no incentive	- - - - - Fit
× No prior recruitment, no incentive	————— Fit



# Are passive data the solution?

---

# Omissions and issues in passive data

---

- Locations
  - Lack of precision (GPS, GSM, Wifi profiles)
  - Not known (Loop detectors, 'social network data', credit card data)
- Sample composition and bias
  - Imputed socio-demographics only (e.g. via the list of apps for example)
  - Unknown and partial (telco data)
- Precision and completeness
  - Uncalibrated (loop detectors and other counters)
  - Lack of coverage (GSM providers, GPS switched off)
- Lack of (automatic) panel data, unless your operators don't care about privacy

# So

---

- We have no ground truth, and certainly not by the groups of interest
- We have to be more humble (latent construct)
- More focus on change and panel (or SC) data and experiments (e.g. SF Park, AKTA)

# therefore

---

- We should share the enriched data and the data sources and tools for the enrichment (e.g. networks, software)
- We should rerun the models in the forecast year with the den current data
- We should properly archive the data
  - ‘Data paper’ for *Transportation*

# Simulation

---

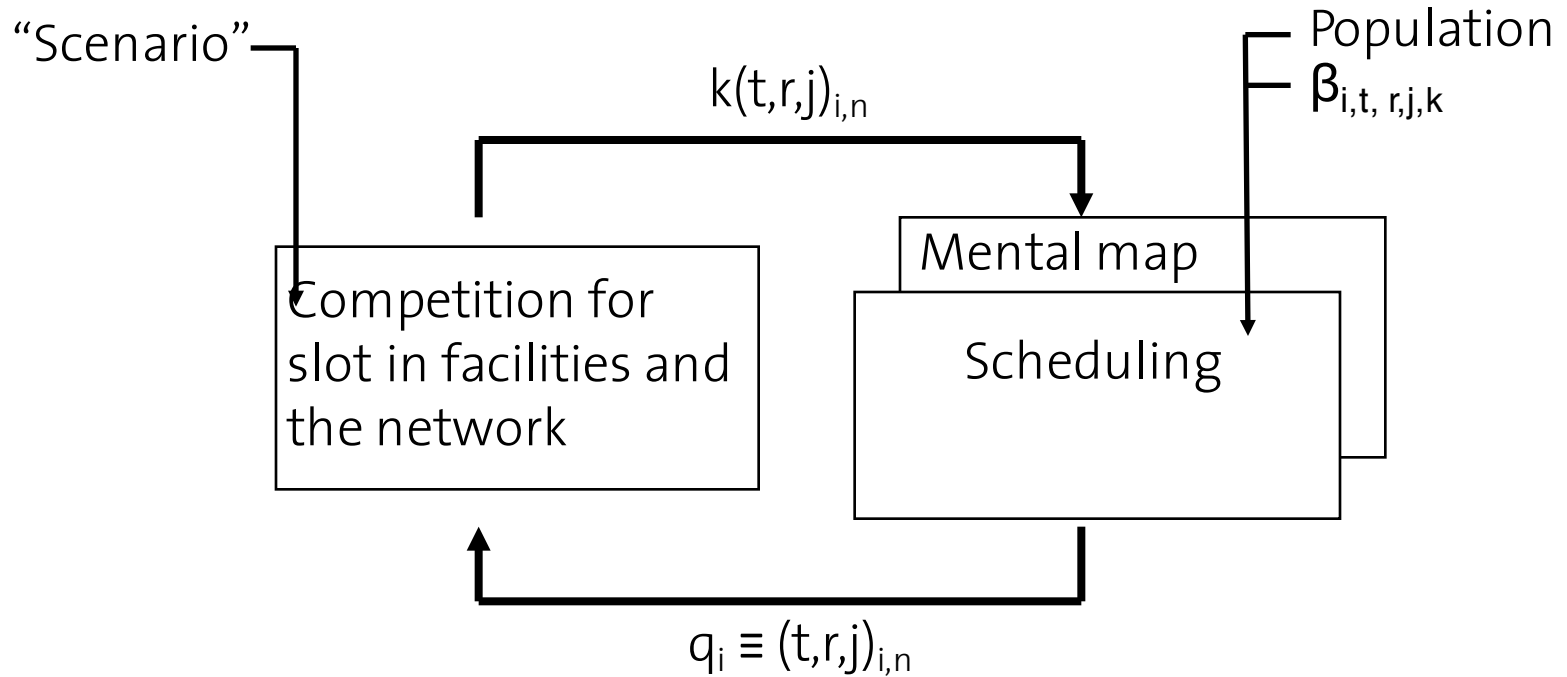
# Time horizons of transport planning

---

	System	Person
Long term	<i>slots</i> Regulation	Home and work locations Mobility tool ownership Social networks
Medium term	Services Prices Awareness	Personal projects
Short term	Operations	<b>Scheduling</b>

# Generic model structure

---



# Current progress: MATSim Switzerland

---





# Activity scheduling dimensions

---

Number and type of activities (and secondary activities)

Sequence of activities

- Start and duration of activity
- Composition of the group undertaking the activity
- Expenditure/income and its 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

# Activity scheduling dimensions – now in MATSim

---

Number and type of activities (and secondary activities)

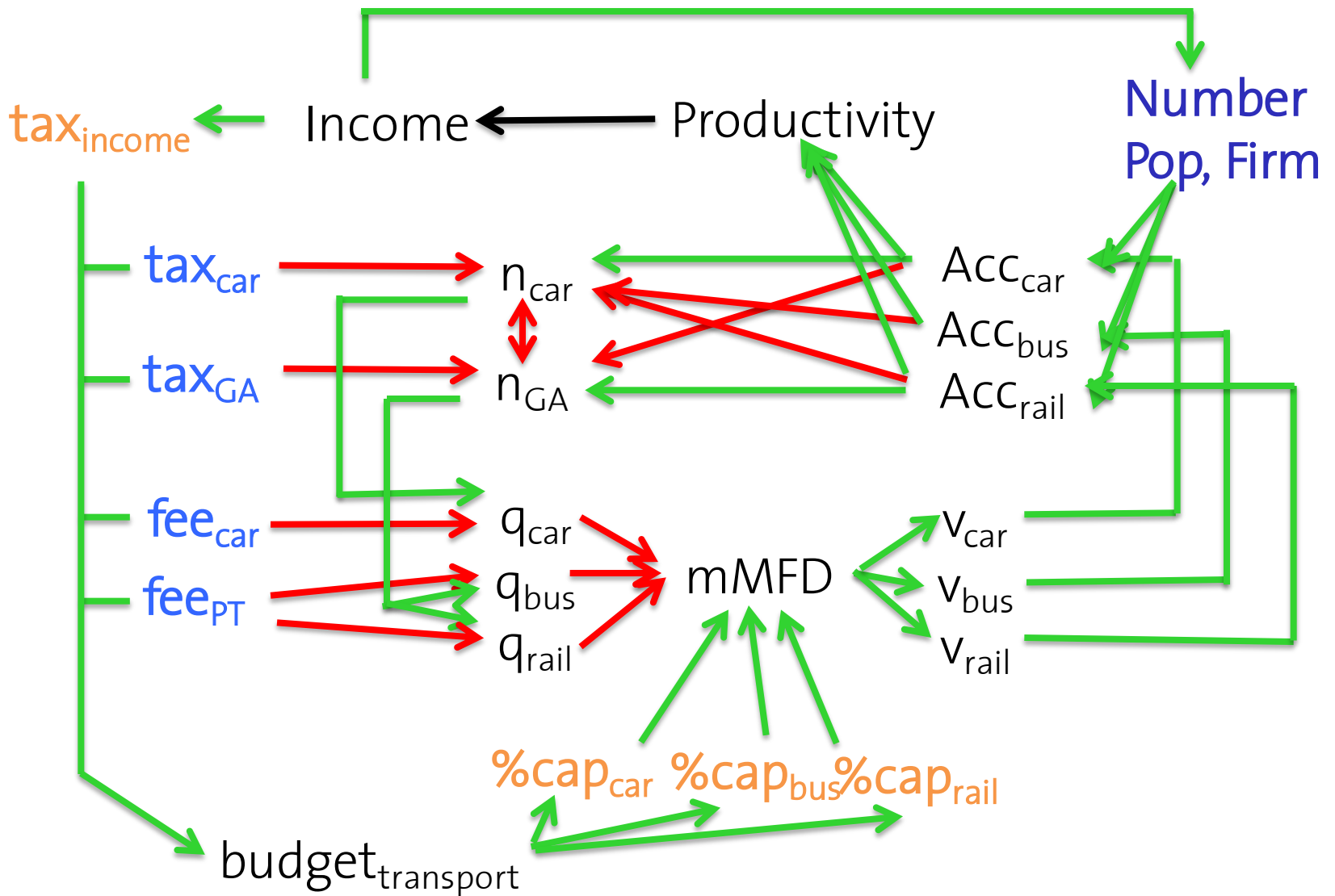
Sequence of activities

- Start and duration of activity
- Composition of the group undertaking the activity
- Expenditure/income and its 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

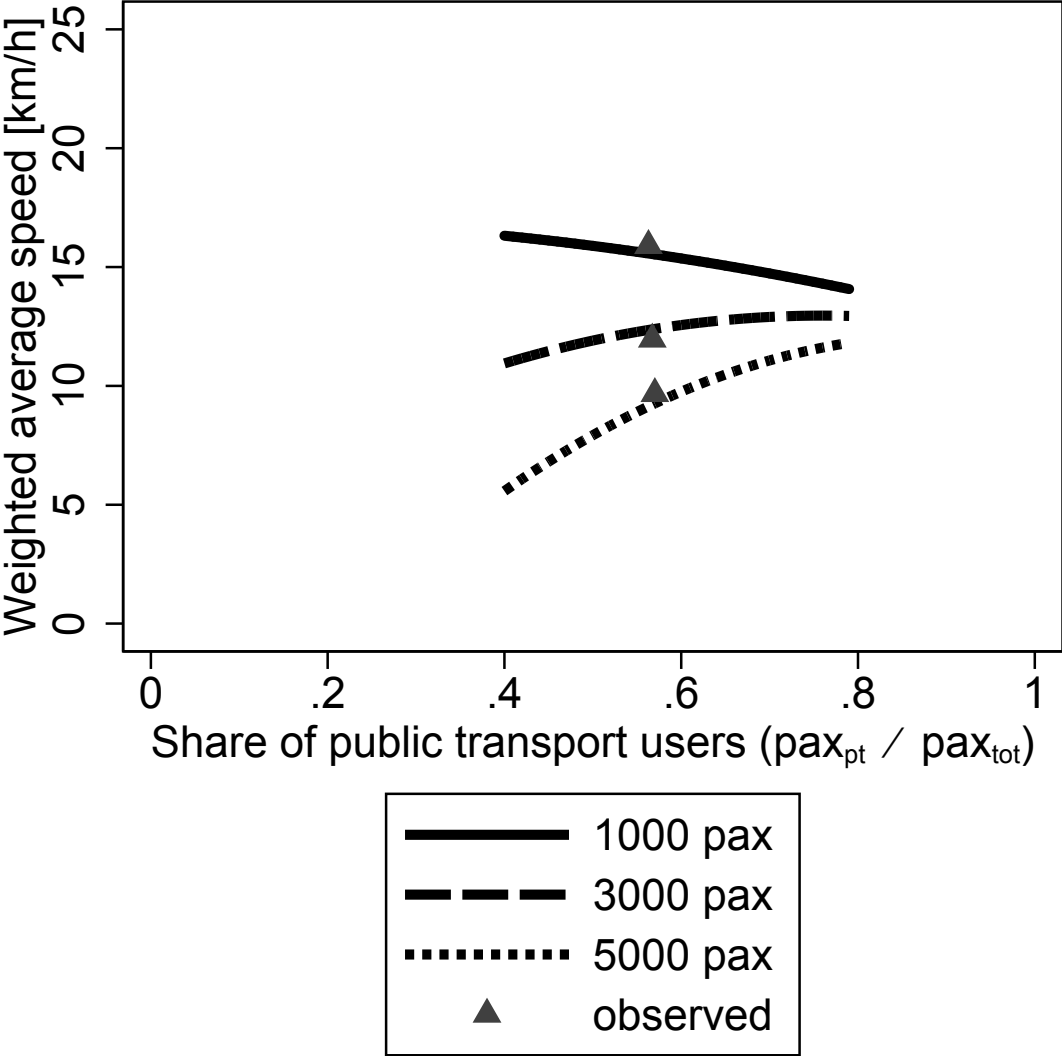
# But do we need this?

---

# Or would this be enough ?



# 3d MFD (Zürich, loops) City centre



Loder et al., 2017

# Questions ?

---

[www.matsim.org](http://www.matsim.org)

[www.ivt.ethz.ch](http://www.ivt.ethz.ch)

[www.futurecities.ethz.ch](http://www.futurecities.ethz.ch)

[www.senozon.com](http://www.senozon.com)

# Questions ?

---

## The Multi-Agent Transport Simulation MATSim

edited by  
Andreas Horni, Kai Nagel, Kay W. Axhausen

```
graph LR; A[initial demand] --> B[mobsim]; B --> C[scoring]; C --> D[analyses]; C --> E[replanning]; E --> B;
```

The diagram illustrates the MATSim workflow. It begins with 'initial demand', which feeds into 'mobsim'. 'mobsim' then leads to 'scoring'. From 'scoring', the process can proceed to 'analyses' or loop back to 'replanning', which then feeds back into 'mobsim'. The background features XML code, a train, a graph, and a bar chart.

**MATSim**  
Multi-Agent Transport Simulation