Impact of Route Choice Set on Route Choice Probabilities

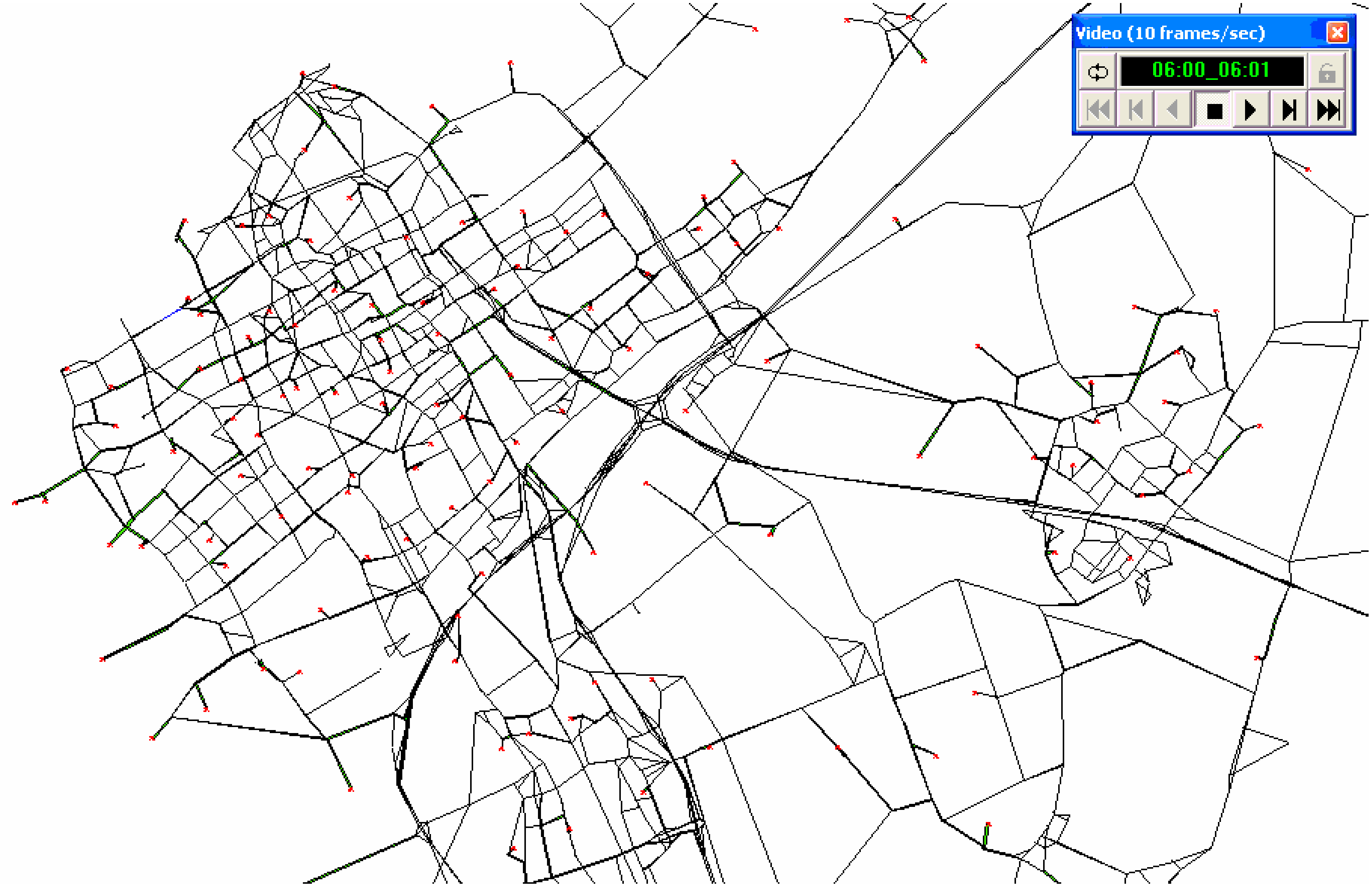
Michiel Bliemer

Delft University of Technology, The Netherlands
& The University of Sydney, Australia

Seminar ITLS, Sydney, April 1st 2008

The University of Sydney

TUDelft
Delft University of Technology


Video (10 frames/sec)
06:00_06:01

Contents

- General route choice model

- Route utility function
- Route choice set
- Route choice probabilities

- Case study with MNL, C-logit, PSL, PSCL, PCL, CNL

- Further research and conclusions

General route choice model

$$\psi_i^{rs} = f(V_i^{rs} \mid C^{rs})$$

For each origin-destination pair (r, s),

ψ_i^{rs} = probability of choosing route i

V_i^{rs} = utility of route i

C^{rs} = route choice set (set of alternative routes)

The route choice probabilities depend on the route choice set.

Route utility function

$$\psi_i^{rs} = f(V_i^{rs} \mid C^{rs})$$

Route utility function

$$V_i^{rs} = \sum_k \beta_k X_{ki}^{rs}$$

Different travelers from the same origin-destination pair (r,s) may choose different routes i because of:

- different perceptions of the attributes X (e.g., travel time, toll costs)
- different preferences β (e.g., depending on income)

The same traveler from the same origin-destination pair (r,s) may choose different routes i at different times because of:

- changes in the attributes X (e.g., due to congestion)
- changes in preferences β (e.g., due to different trip purpose)

Route choice set

$$\psi_i^{rs} = f(V_i^{rs} \mid C^{rs})$$

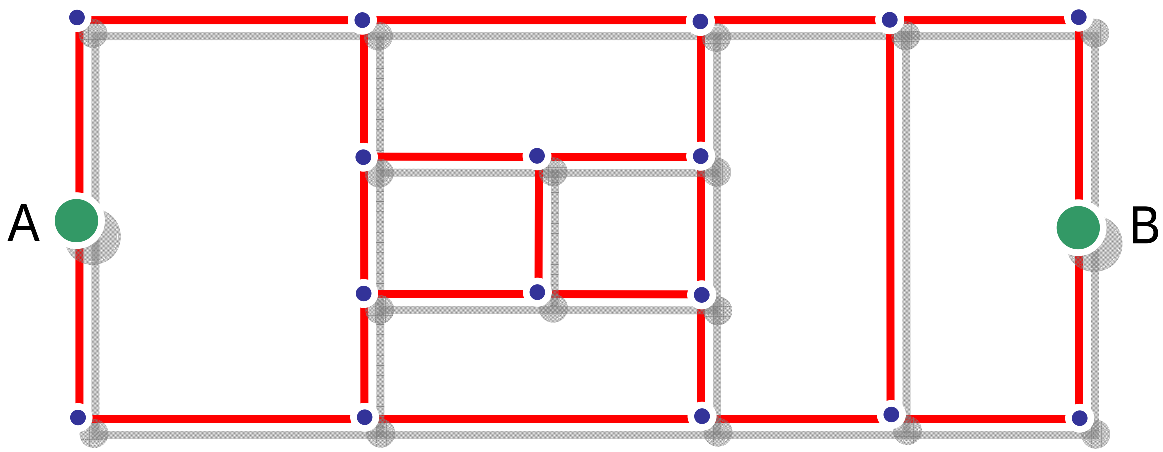
Route choice set

Which routes to include in route choice set C^{rs}?

- at least all relevant routes
- irrelevant routes may be included
  (but should not bias the route choice probabilities!)

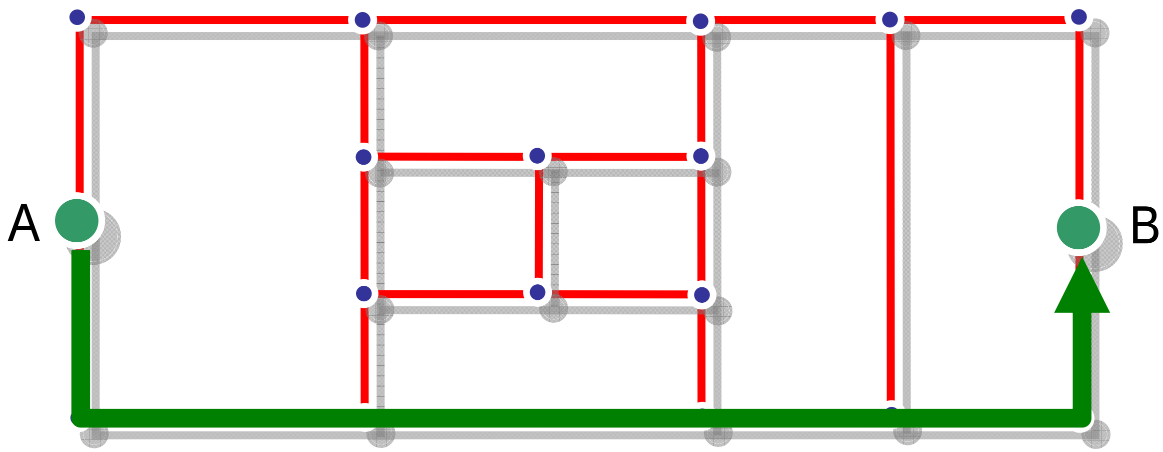
irrelevant route =
route that has a significantly lower utility than the best route alternative
(and as such will be chosen by only a very small percentage of travelers)

Route choice set

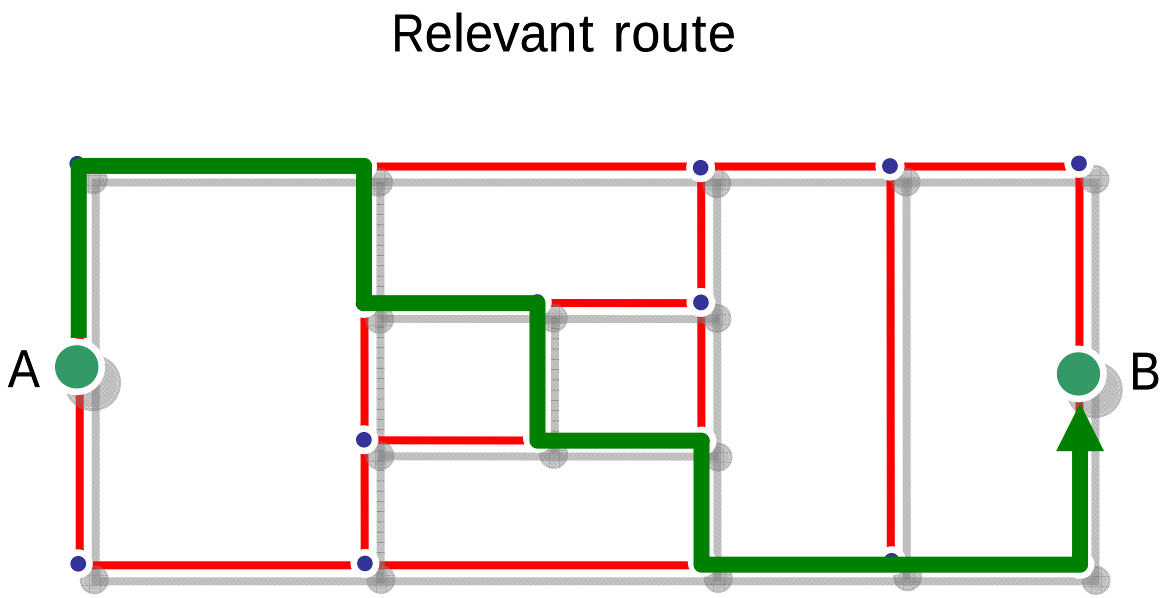


Route choice set

Relevant route

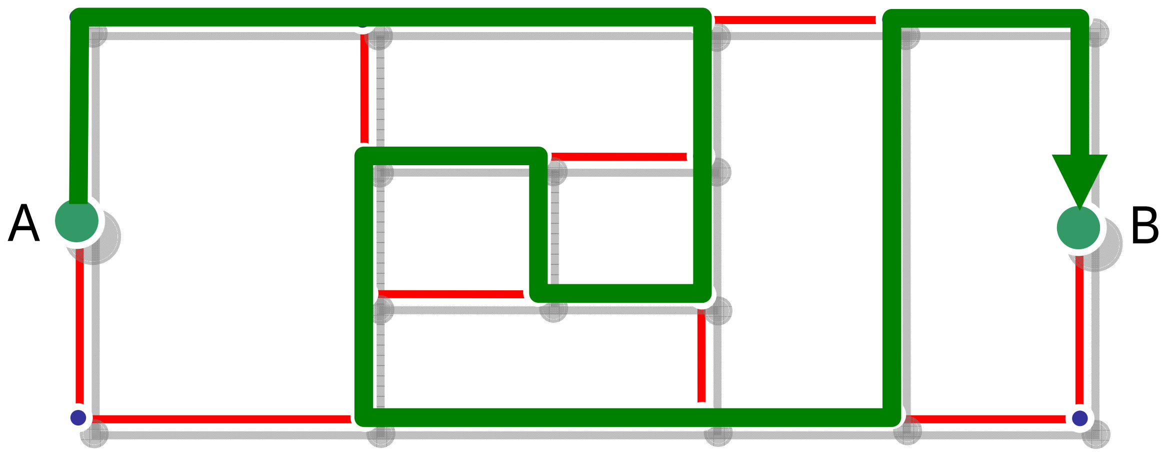


Route choice set



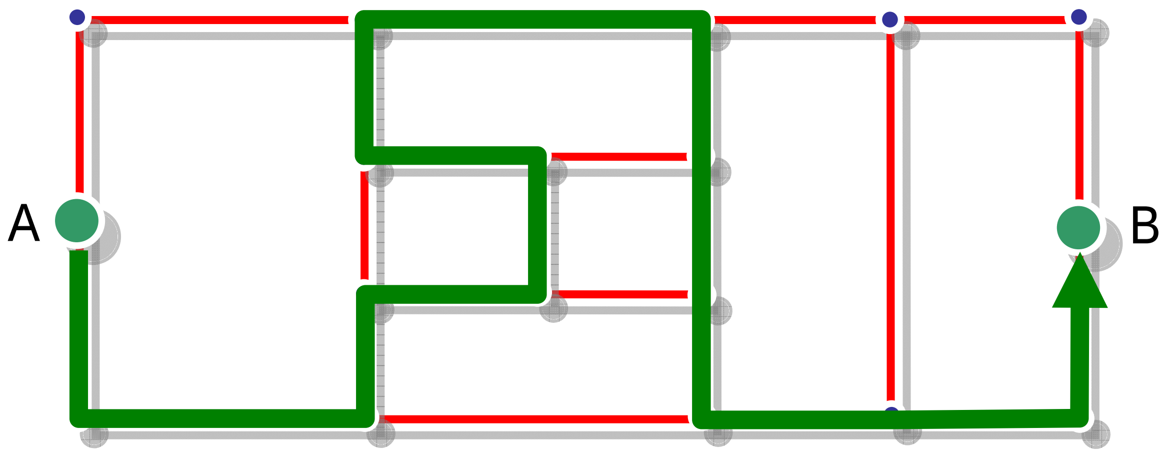
Route choice set

Irrelevant route



Route choice set

Irrelevant route



Route choice set

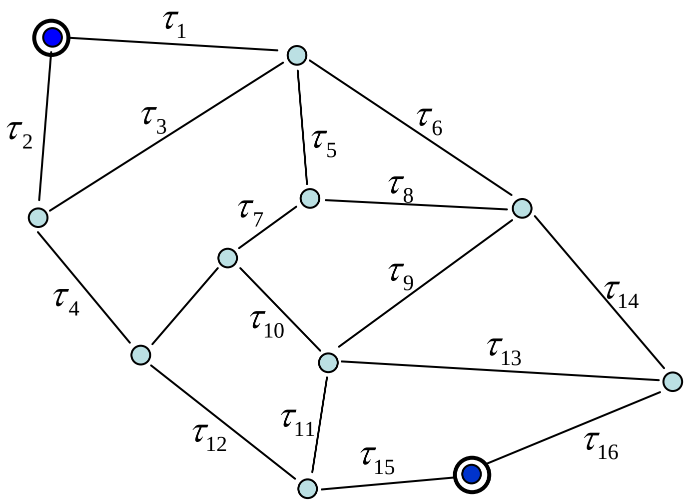
How to generate a route choice set?

1. Deterministic route choice set generation
   - Enumerate all routes
   - Generate *k*-shortest routes
   - Generate essentially least-cost routes
   - Generate constrained routes
     e.g. Hoogendoorn-Lanser, 2005; Prato and Behkor, 2006
     (branch-and-bound algorithms)
2. Stochastic route choice set generation
   - Generate likely routes
     e.g. Bovy and Fiorenzo-Catalano, 2006
     (repeated shortest path algorithms)

Route choice set

Stochastic route choice set generation:

for $j = 1, \ldots, N$,



Step 1: Draw values of link attributes from stochastic distributions

e.g., $\tau_a = \tau_a^0 + |\varepsilon_a|, \qquad \varepsilon_a \sim N(0,\sigma^2)$

Step 2: Determine the route with the highest utility ("shortest" path computation)

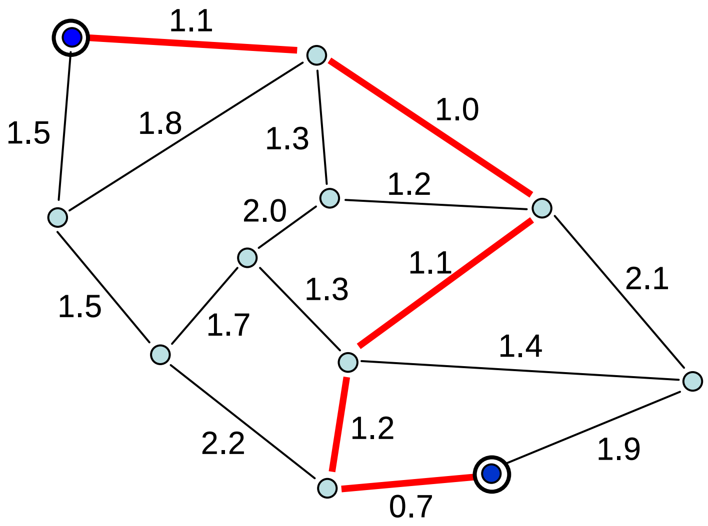
Step 3: Add this route to the route set C^{rs}

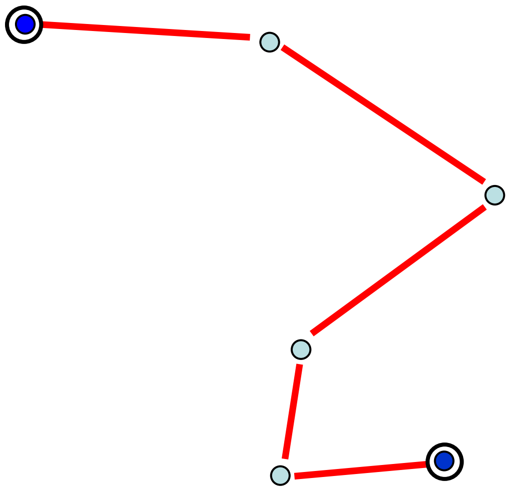
The University of Sydney

TU Delft

Route choice set

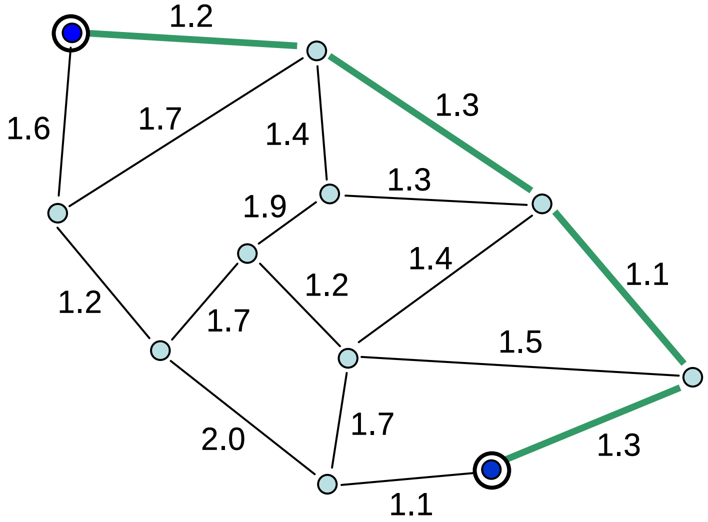
Iteration 1 ...

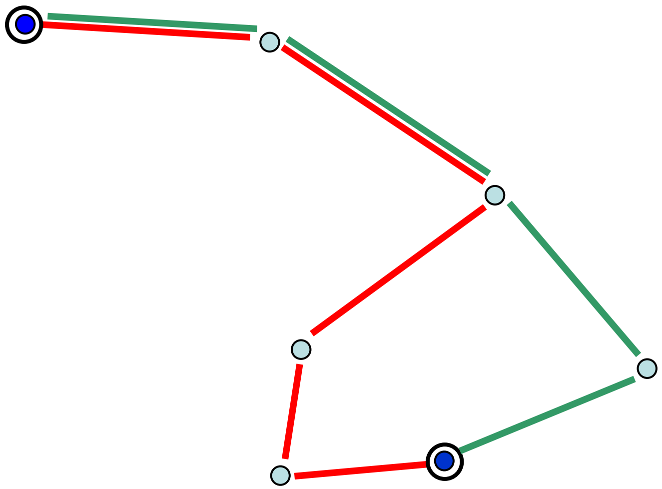




Route choice set

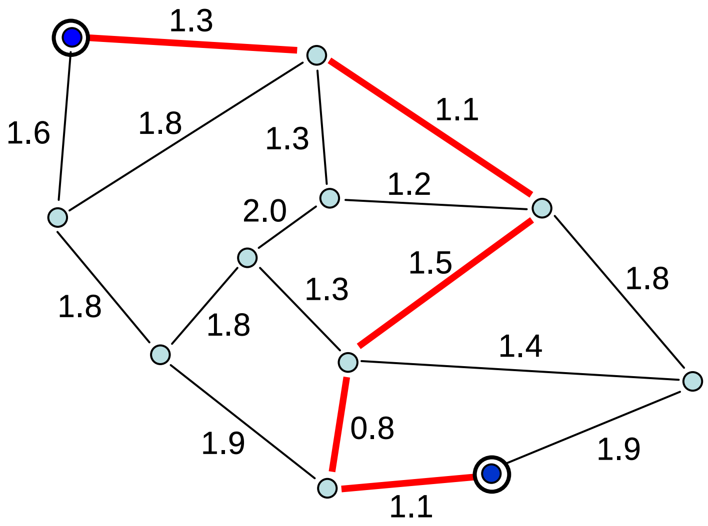
Iteration 2 ...

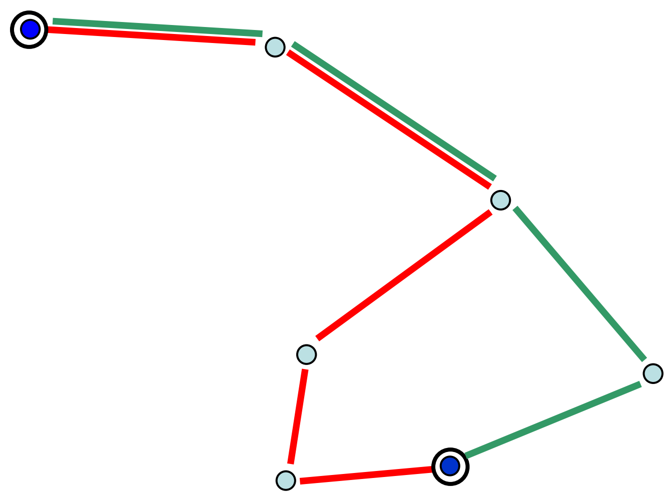




Route choice set

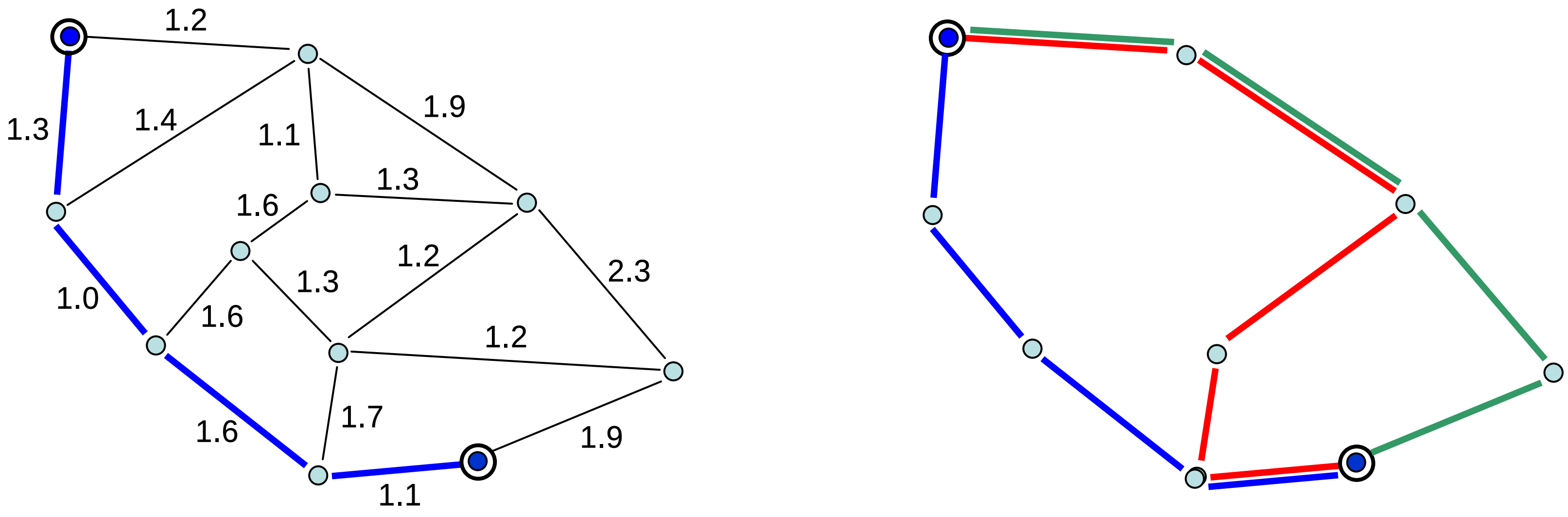
Iteration 3 ...





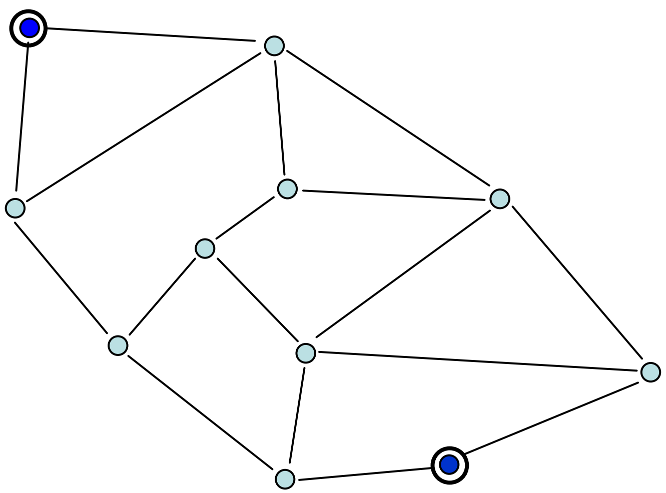
Route choice set

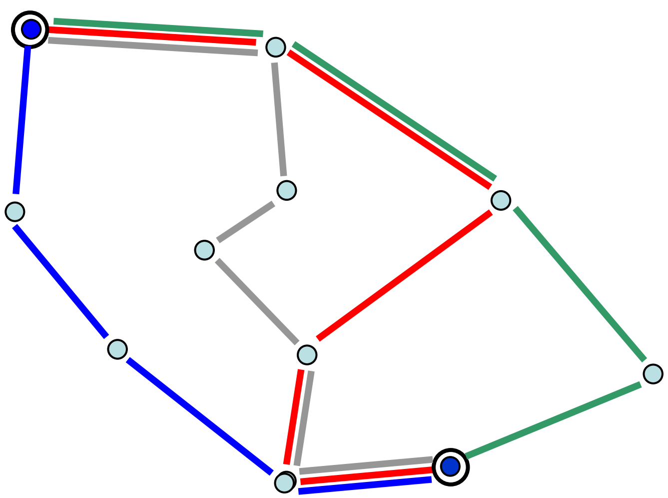
Iteration 4 ...



Route choice set

Iteration *N* ...





Route choice probabilities

$$\psi_i^{rs} = f(V_i^{rs} \mid C^{rs})$$

Route choice probabilities

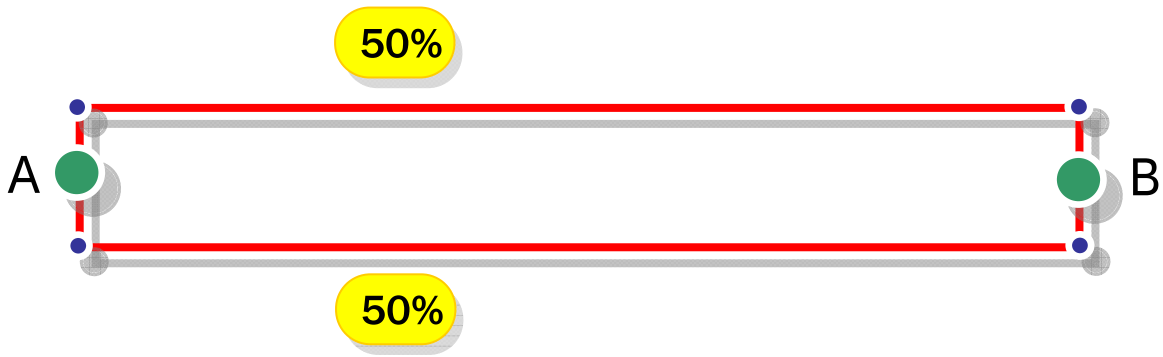
Desired properties of the route choice model:

- Route overlap should be taken into account
- Route choice probabilities should be robust
  (irrelevant routes in the route choice set should not bias the route choice probabilities)

How does the basic multinomial logit (MNL) model perform?

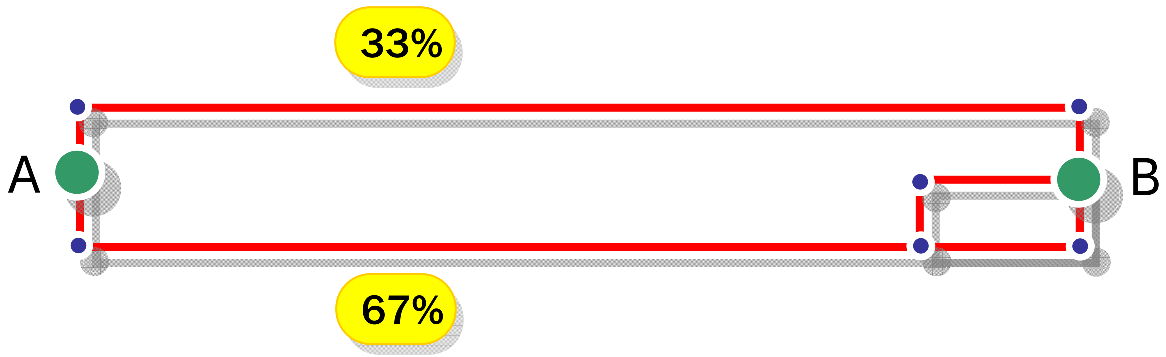
Route choice probabilities

Route choice probabilities in the MNL model:



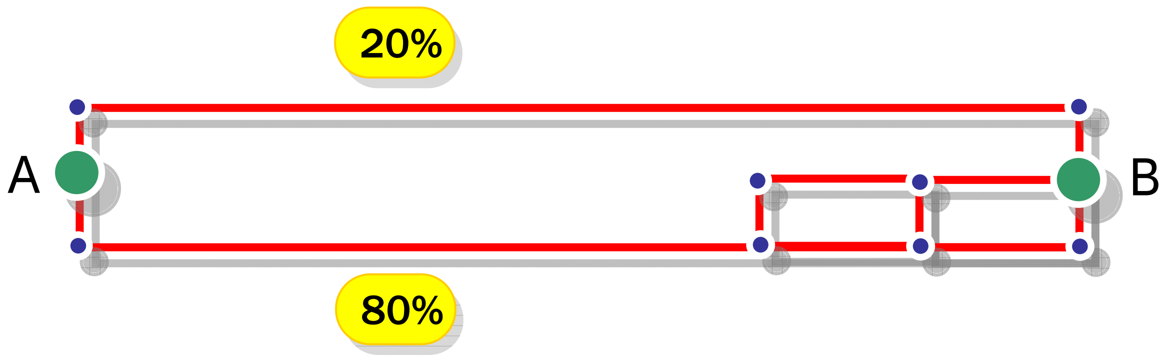
Route choice probabilities

Route choice probabilities in the MNL model:



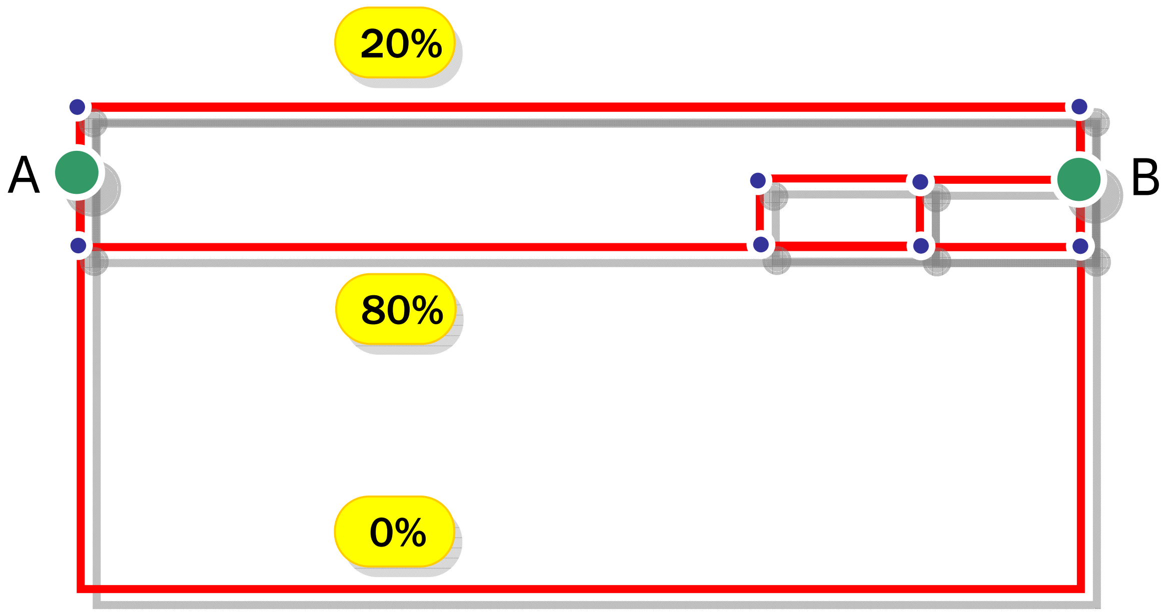
Route choice probabilities

Route choice probabilities in the MNL model:



Route choice probabilities

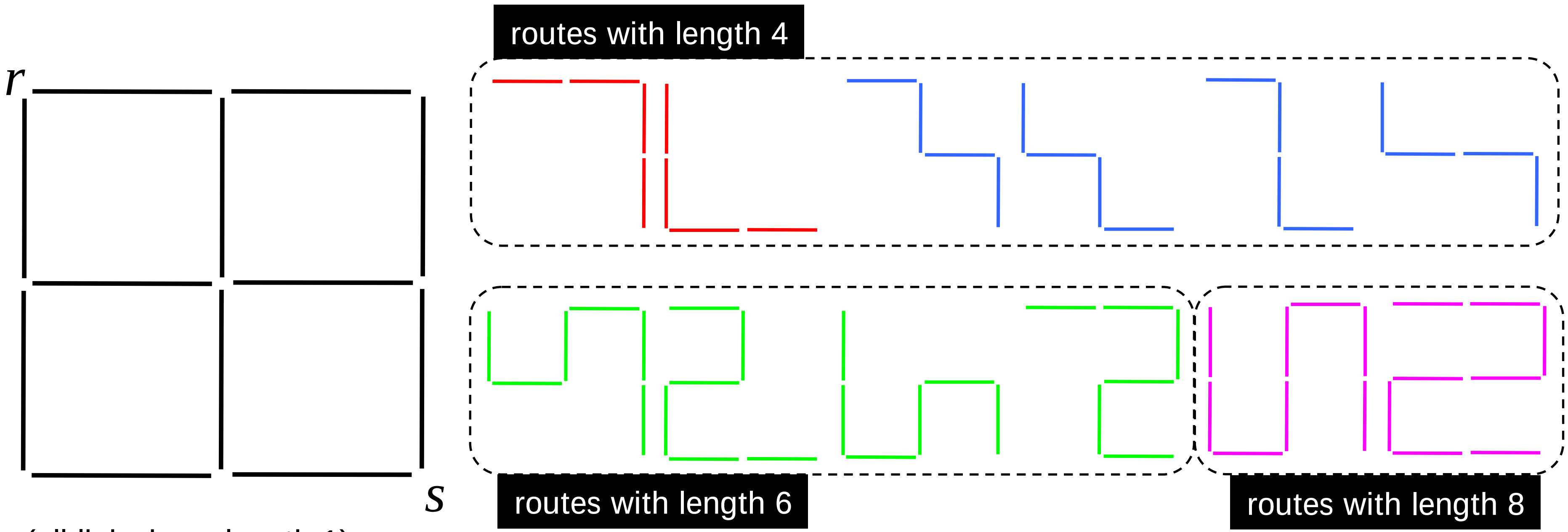
Route choice probabilities in the MNL model:



Route choice probabilities

- The multinomial logit model does not take **route overlap** into account, but is **robust**
- The probit model and the mixed logit model are able to deal with **route overlap** and are **robust**, however require simulation
- Other simple models have been proposed to correct for **route overlap**, such as:
  - ❑ C-logit
  - ❑ path-size (correction) logit
  - ❑ paired combinatorial logit
  - ❑ cross-nested logit
- These models are simple extensions of the MNL model and require no simulation. However, are these models **robust**?

Case study



(all links have length 1)

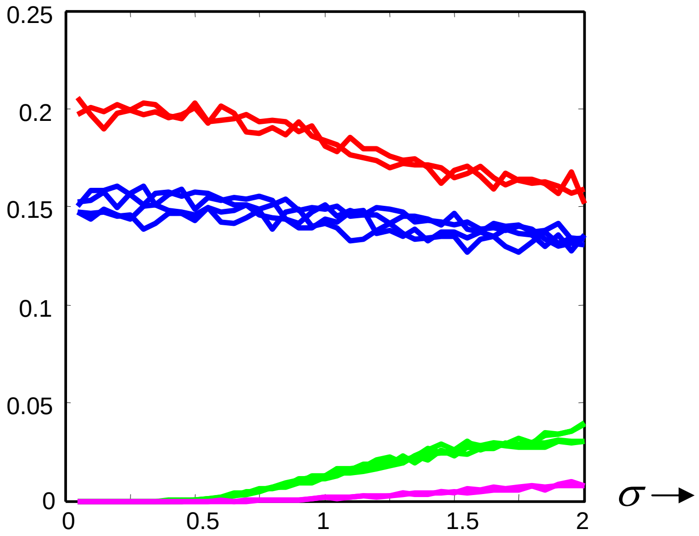
12 distinct routes

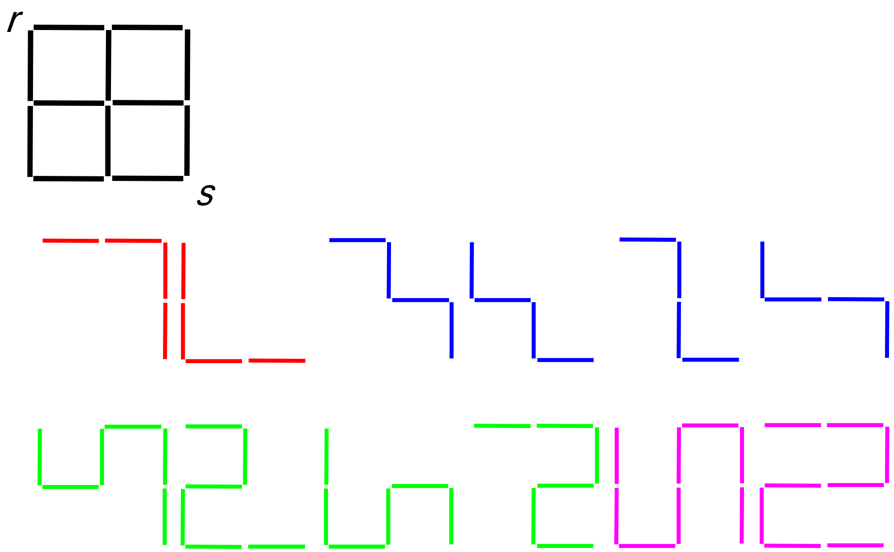
Case study

Route choice set generation

Using $\tau_a = 1 + \varepsilon_a, \quad \varepsilon_a \sim N(0, \sigma^2)$,
mainly the red and blue routes (of length 4) enter route choice set C^{rs}.
The irrelevant routes are only generated with high σ.

Route selection probabilities
(10,000 Monte Carlo simulations)



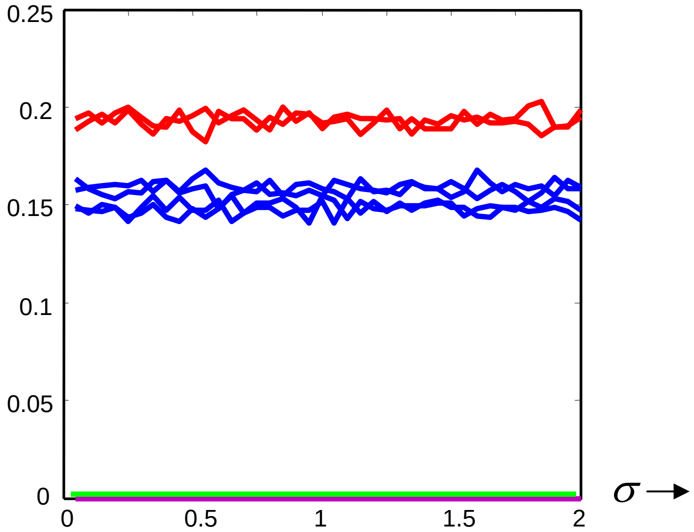


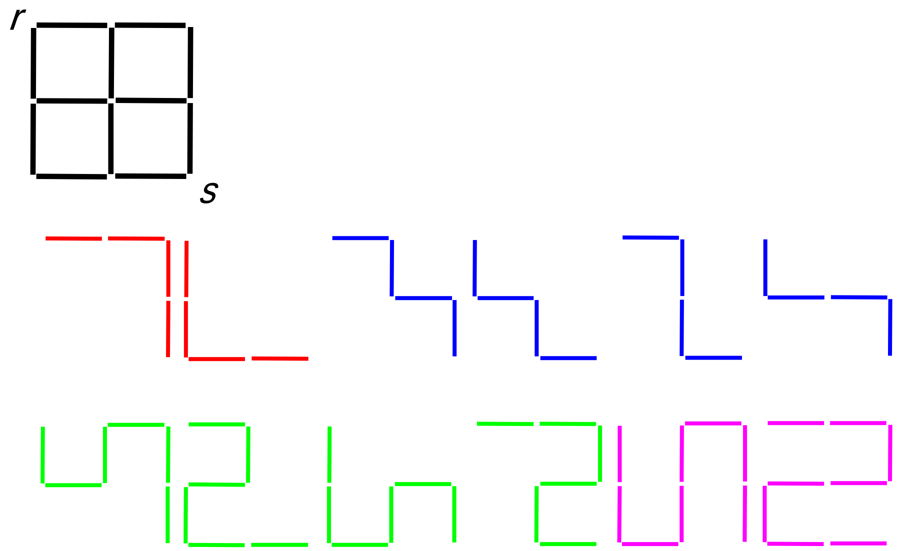
Case study

Route choice set generation

Using $\tau_a = 1 + |\varepsilon_a|, \qquad \varepsilon_a \sim N(0,\sigma^2)$,
only the red and blue routes (of length 4) enter route choice set C^{rs}.
It is impossible to generated the irrelevant routes, independent of σ.

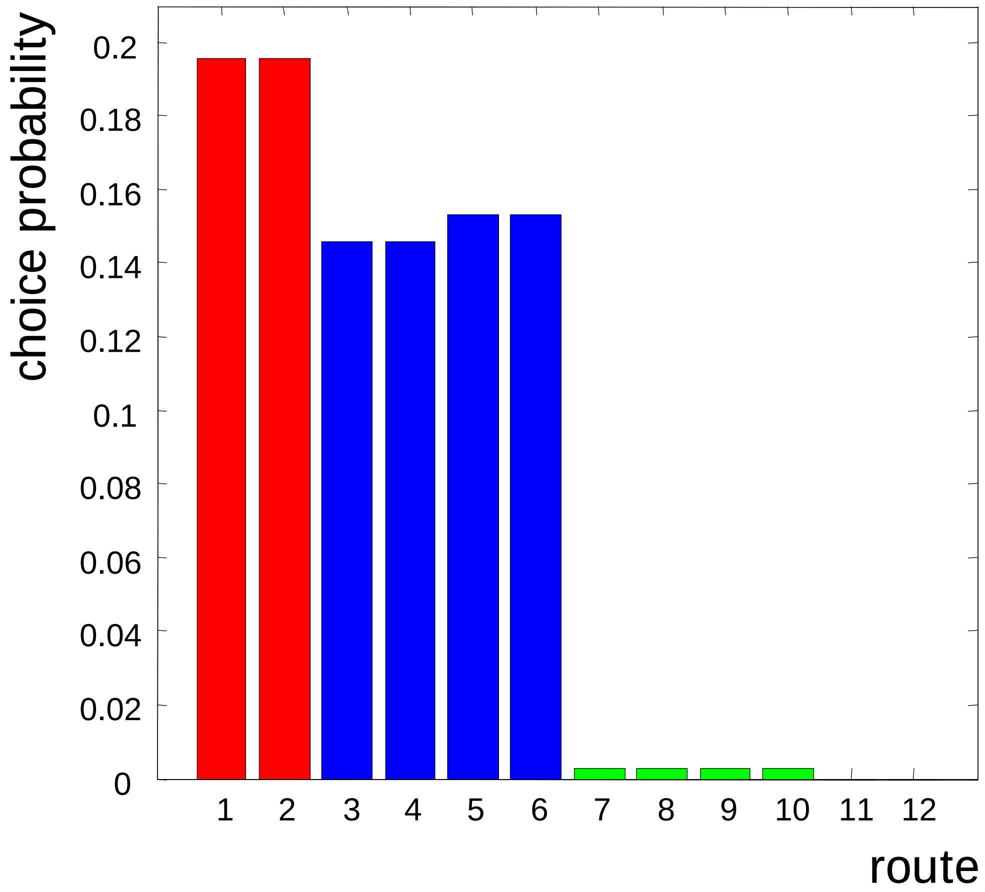
Route selection probabilities
(10,000 Monte Carlo simulations)





Case study

Probit route choice probabilities



generated using
10,000 Monte Carlo
simulations

Relevant routes

Routes 1 – 6 (length 4)

Irrelevant routes

Routes 7 – 12 (length 6 or 8)

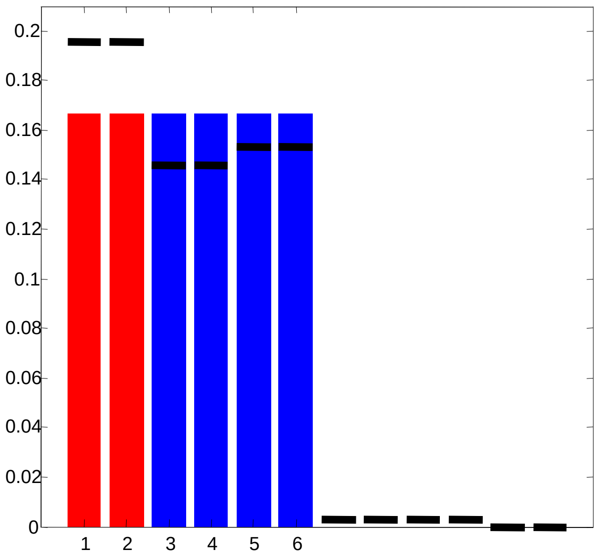
Case I

Calibration on the 6 relevant routes

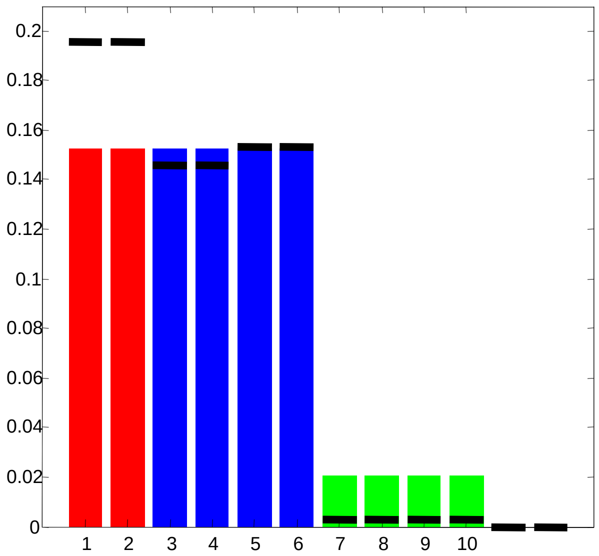
Multinomial logit model

calibrated on 6 routes

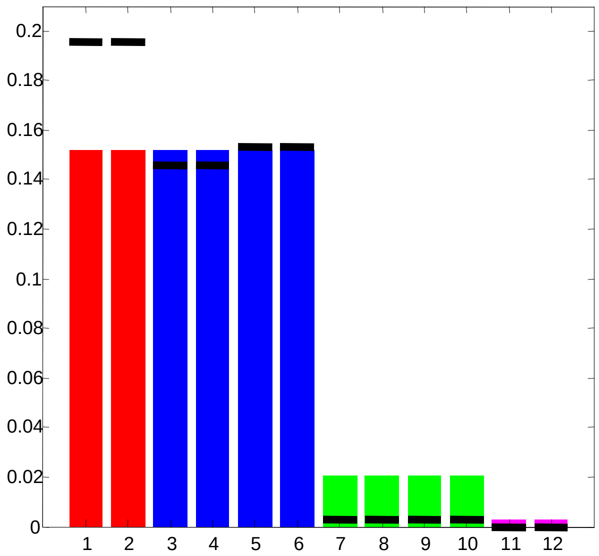
$$\psi_i^{rs} = \frac{\exp\left(\mu V_i^{rs}\right)}{\sum\limits_{j\in C^{rs}} \exp\left(\mu V_j^{rs}\right)}$$



6 routes



10 routes

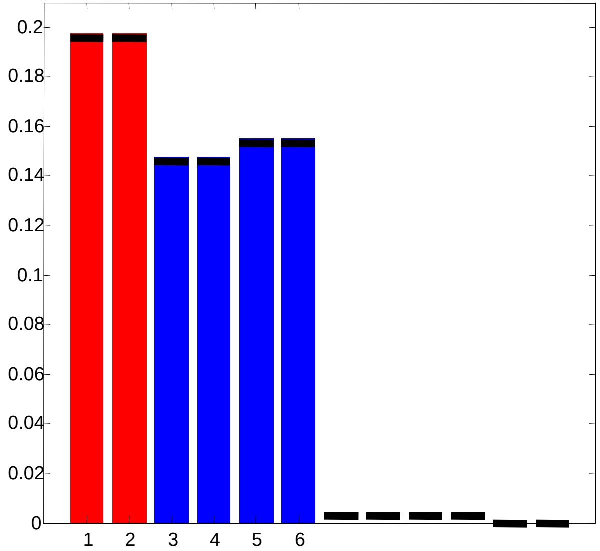


12 routes

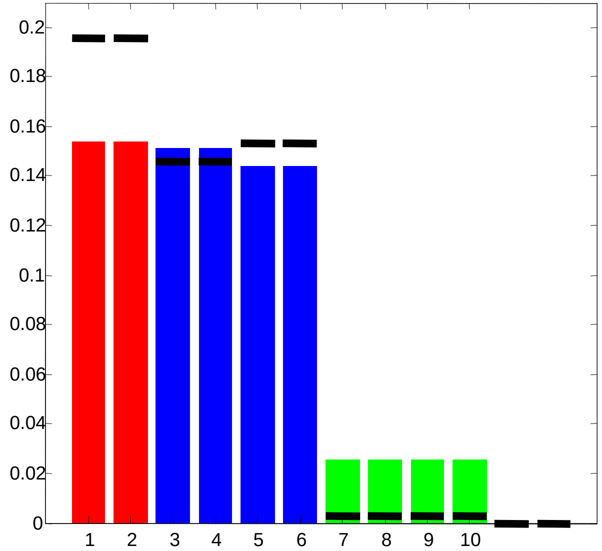
C-logit model

calibrated on 6 routes

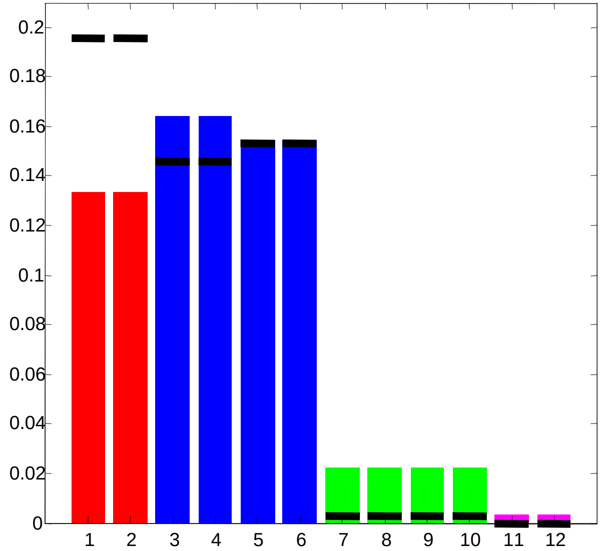
$$\psi_i^{rs} = \frac{\exp\left(\mu\left(V_i^{rs} - \beta \ln \mathrm{CF}_i^{rs}\right)\right)}{\sum_{j \in C^{rs}} \exp\left(\mu\left(V_j^{rs} - \beta \ln \mathrm{CF}_j^{rs}\right)\right)}, \qquad \mathrm{CF}_i^{rs} = \sum_{j \in C^{rs}} \left(\frac{L_{ij}^{rs}}{\sqrt{L_i^{rs} L_j^{rs}}} \right)^{\gamma}$$



6 routes



10 routes

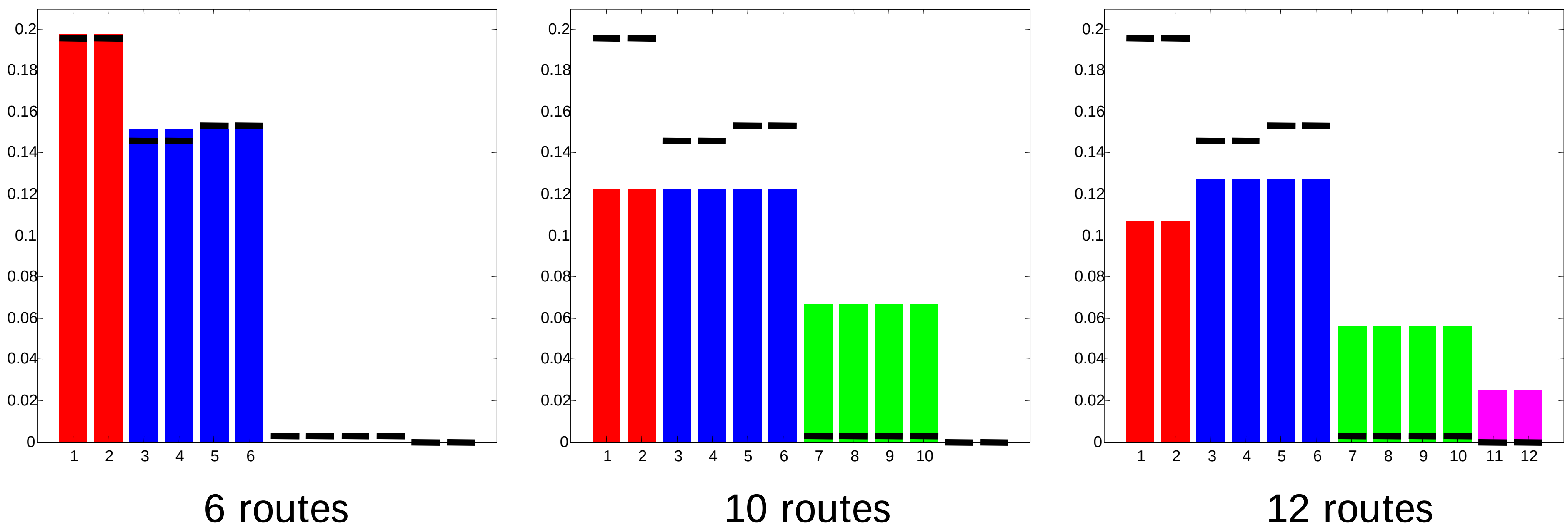


12 routes

Path-size logit model

calibrated on 6 routes

$$\psi_i^{rs} = \frac{\exp\left(\mu\left(V_i^{rs} + \beta \ln \mathrm{PS}_i^{rs}\right)\right)}{\sum_{j \in C^{rs}} \exp\left(\mu\left(V_j^{rs} + \beta \ln \mathrm{PS}_j^{rs}\right)\right)}, \quad \mathrm{PS}_i^{rs} = \sum_{a \in \Gamma_i^{rs}} \left(\frac{l_a}{L_i^{rs}}\right) \frac{1}{\sum_{j \in C^{rs}} \delta_{aj}^{rs} \left(\frac{\overline{L}^{rs}}{L_j^{rs}}\right)^{\gamma}}.$$



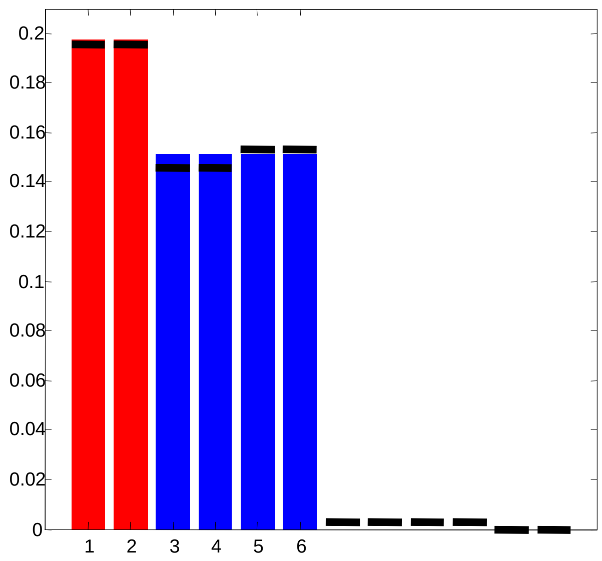
The University of Sydney

TU Delft

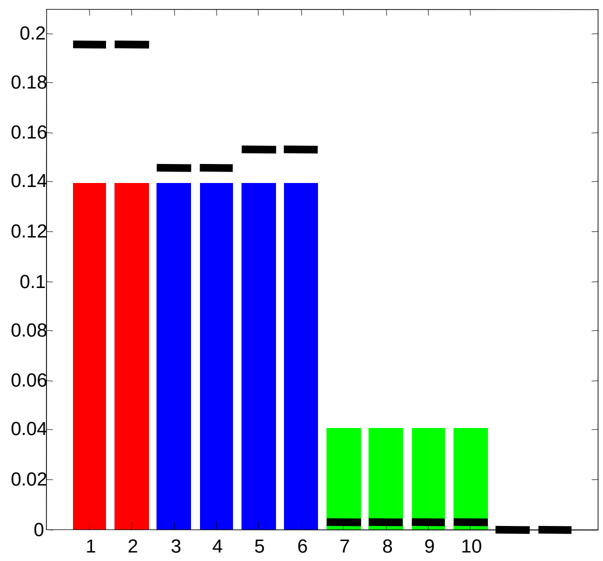
Path-size correction logit model

calibrated on 6 routes

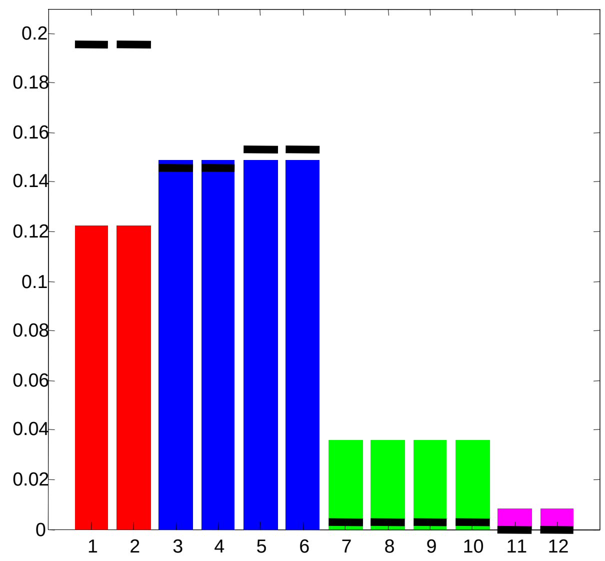
$$\psi_i^{rs} = \frac{\exp\left(\mu\left(V_i^{rs} + \beta \cdot \mathrm{PSC}_i^{rs}\right)\right)}{\sum_{j\in C^{rs}} \exp\left(\mu\left(V_j^{rs} + \beta \cdot \mathrm{PSC}_j^{rs}\right)\right)}, \quad \mathrm{PSC}_i^{rs} = -\sum_{a\in\Gamma_i^{rs}} \frac{l_a}{L_i^{rs}} \ln \sum_{j\in C^{rs}} \delta_{aj}^{rs}.$$



6 routes



10 routes



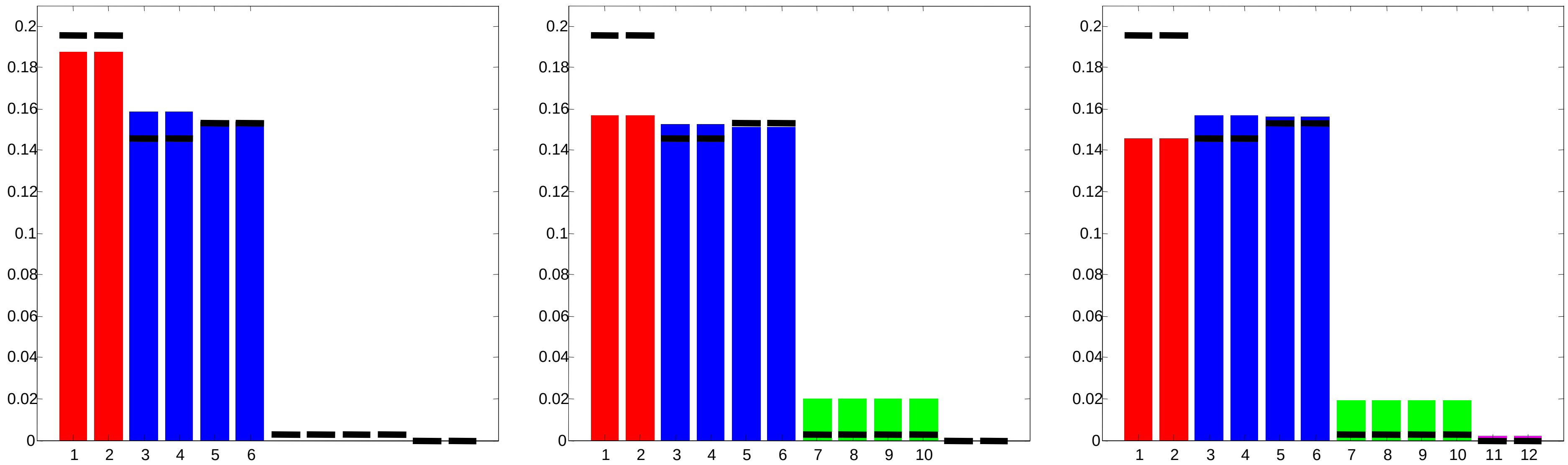
12 routes

$$\sigma_{ij}^{rs} = \frac{L_{ij}^{rs}}{\sqrt{L_i^{rs} L_j^{rs}}}$$

Paired combinatorial logit model

calibrated on 6 routes

$$\psi_i^{rs} = \sum_{j \in C^{rs}, j \neq i} \frac{\exp\left(\frac{\mu V_i^{rs}}{1-\sigma_{ij}^{rs}}\right)}{\exp\left(\frac{\mu V_i^{rs}}{1-\sigma_{ij}^{rs}}\right) + \exp\left(\frac{\mu V_j^{rs}}{1-\sigma_{ij}^{rs}}\right)} \cdot \frac{\left(1-\sigma_{ij}^{rs}\right)\left(\exp\left(\frac{\mu V_i}{1-\sigma_{ij}^{rs}}\right) + \exp\left(\frac{\mu V_j}{1-\sigma_{ij}^{rs}}\right)\right)^{1-\sigma_{ij}^{rs}}}{\sum_{k=1}^{|C^{rs}|-1} \sum_{m=k+1}^{|C^{rs}|} \left(1-\sigma_{km}^{rs}\right)\left(\exp\left(\frac{\mu V_k}{1-\sigma_{km}^{rs}}\right) + \exp\left(\frac{\mu V_m}{1-\sigma_{km}^{rs}}\right)\right)^{1-\sigma_{km}^{rs}}},$$

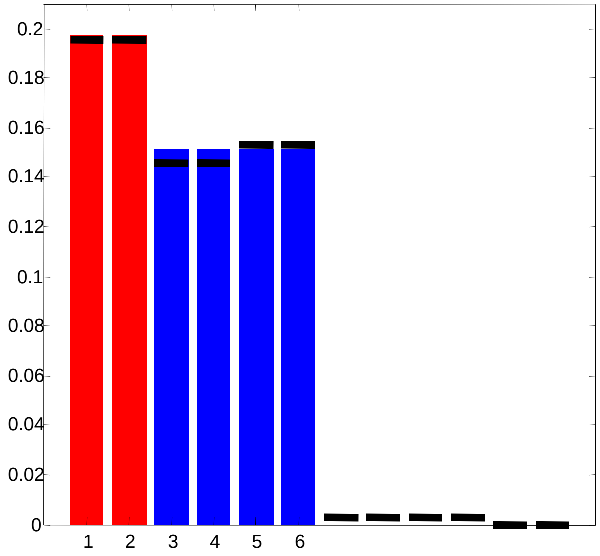


6 routes | 10 routes | 12 routes

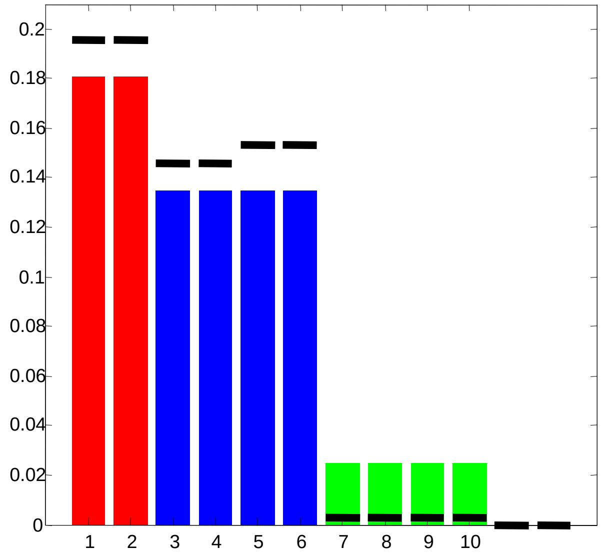
Cross-nested logit model

calibrated on 6 routes

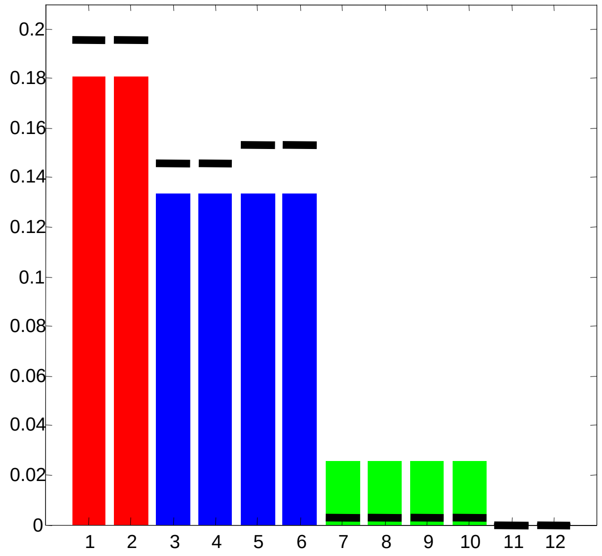
$$\psi_i^{rs} = \sum_a \frac{\left(\kappa_{ai}^{rs}\exp\left(\mu V_i\right)\right)^{1/\beta}}{\sum_{j\in C^{rs}}\left(\kappa_{aj}^{rs}\exp\left(\mu V_j\right)\right)^{1/\beta}} \frac{\left(\sum_{j\in C^{rs}}\kappa_{aj}^{rs}\exp\left(\mu V_j\right)\right)^{\beta}}{\sum_b\left(\sum_{j\in C^{rs}}\kappa_{bj}^{rs}\exp\left(\mu V_j\right)\right)^{\beta}}, \quad \kappa_{aj}^{rs}=\frac{l_a}{L_j^{rs}}\delta_{aj}^{rs}$$



6 routes



10 routes



12 routes

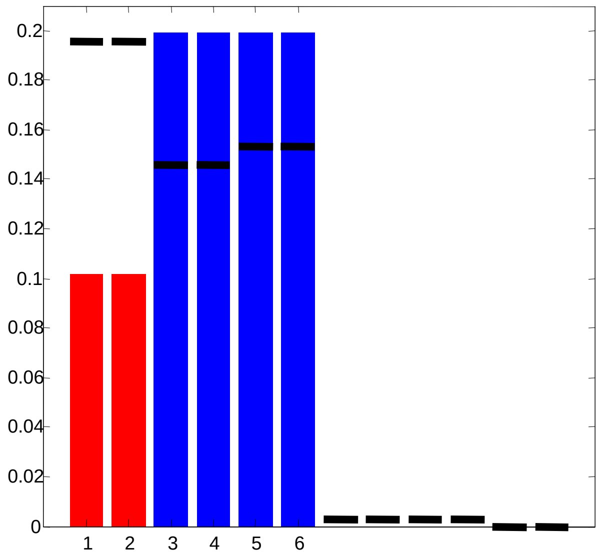
Case II

Calibration on all 12 routes

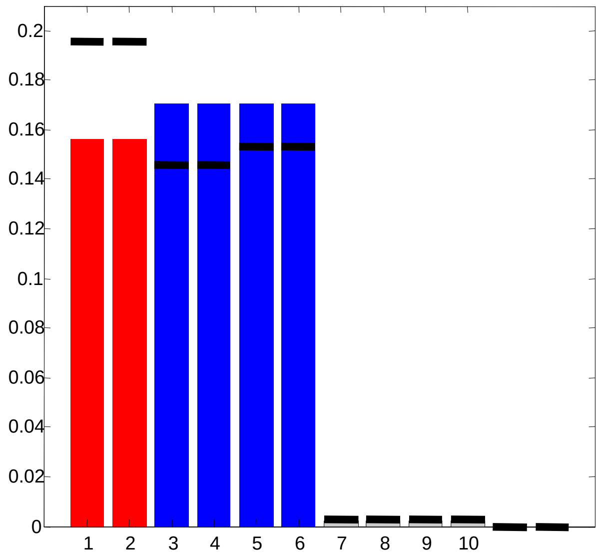
Path-size logit model

calibrated on 12 routes

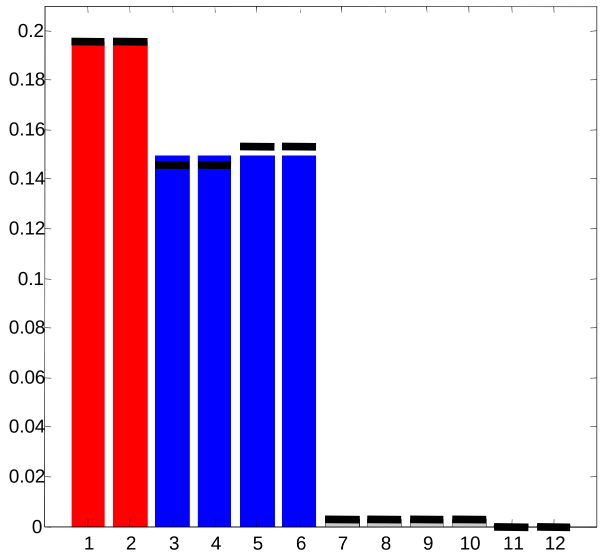
$$\psi_i^{rs} = \frac{\exp\left(\mu\left(V_i^{rs} + \beta \ln \mathrm{PS}_i^{rs}\right)\right)}{\sum_{j \in C^{rs}} \exp\left(\mu\left(V_j^{rs} + \beta \ln \mathrm{PS}_j^{rs}\right)\right)}, \qquad \mathrm{PS}_i^{rs} = \sum_{a \in \Gamma_i^{rs}} \left(\frac{l_a}{L_i^{rs}}\right) \frac{1}{\sum_{j \in C^{rs}} \delta_{aj}^{rs} \left(\frac{\bar{L}^{rs}}{L_j^{rs}}\right)^{\gamma}}.$$



6 routes



10 routes

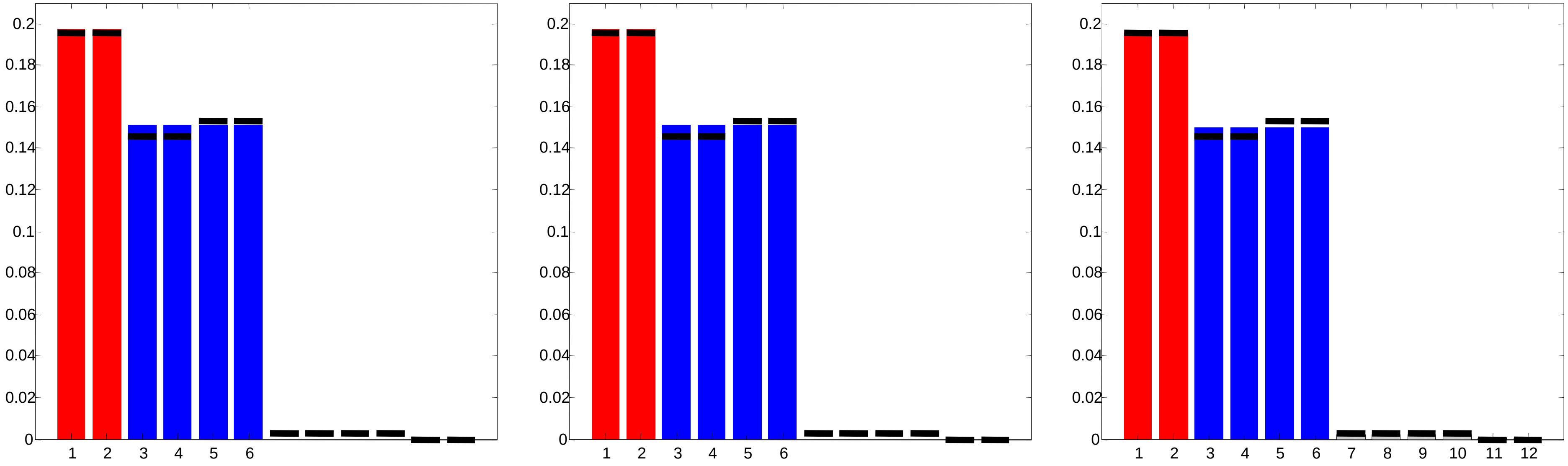


12 routes

Cross-nested logit model

calibrated on 12 routes

$$\psi_i^{rs} = \sum_a \frac{\left(\kappa_{ai}^{rs} \exp\left(\mu V_i\right)\right)^{1/\beta}}{\sum_{j \in C^{rs}} \left(\kappa_{aj}^{rs} \exp\left(\mu V_j\right)\right)^{1/\beta}} \frac{\left(\sum_{j \in C^{rs}} \kappa_{aj}^{rs} \exp\left(\mu V_j\right)\right)^{\beta}}{\sum_b \left(\sum_{j \in C^{rs}} \kappa_{bj}^{rs} \exp\left(\mu V_j\right)\right)^{\beta}}, \quad \kappa_{aj}^{rs} = \frac{l_a}{L_j^{rs}} \delta_{aj}^{rs}$$



6 routes | 10 routes | 12 routes

The University of Sydney

TU Delft

Route overlap: correlations

TABLE 1 Route correlations in the various route sets

(a) *Route correlation matrix*

	1	2	3	4	5	6	7	8	9	10	11	12
1	1	0.00	0.50	0.00	0.25	0.25	0.61	0.20	0.20	0.61	0.53	0.53
2	0.00	1	0.00	0.50	0.25	0.25	0.20	0.61	0.61	0.20	0.53	0.53
3	0.50	0.00	1	0.00	0.50	0.50	0.20	0.41	0.41	0.20	0.18	0.18
4	0.00	0.50	0.00	1	0.50	0.50	0.41	0.20	0.20	0.41	0.18	0.18
5	0.25	0.25	0.50	0.50	1	0.00	0.00	0.61	0.00	0.61	0.00	0.35
6	0.25	0.25	0.50	0.50	0.00	1	0.61	0.00	0.61	0.00	0.35	0.00
7	0.61	0.20	0.20	0.41	0.00	0.61	1	0.00	0.33	0.33	0.72	0.29
8	0.20	0.61	0.41	0.20	0.61	0.00	0.00	1	0.33	0.33	0.29	0.72
9	0.20	0.61	0.41	0.20	0.00	0.61	0.33	0.33	1	0.00	0.72	0.29
10	0.61	0.20	0.20	0.41	0.61	0.00	0.33	0.33	0.00	1	0.29	0.72
11	0.53	0.53	0.18	0.18	0.00	0.35	0.72	0.29	0.72	0.29	1	0.50
12	0.53	0.53	0.18	0.18	0.35	0.00	0.29	0.72	0.29	0.72	0.50	1

(b) *Average route correlations in each of the route sets.*

	1	2	3	4	5	6	7	8	9	10	11	12
C_1^{rs}	0.20	0.20	0.30	0.30	0.30	0.30	--	--	--	--	--	--
C_2^{rs}	0.29	0.29	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	--	--
C_3^{rs}	0.34	0.34	0.28	0.28	0.28	0.28	0.34	0.34	0.34	0.34	0.39	0.39

Further research

- Somehow the route overlap correction factor needs to be probability weighted in order to decrease the impact of irrelevant routes
- This may lead to a new type of logit choice model, which is subject of further research

Conclusions

- Stochastic route choice set generation is able to produce all relevenant routes and mostly omits irrelevant routes
- The MNL model does not take route overlap into account, but is robust
- None of the models that try to correct for route overlap are robust; route probabilities heavily depend on the route choice set (including at least all relevant alternatives)
- The cross-nested logit model is most robust and seems to be the best model (without requiring simulations)
- When applying the C-logit, PS(C)L, PCL, or CNL model, make sure that the same route choice set is used for both model estimation and prediction (e.g., set with only relevant routes)