EJTIR

Issue 16(3), 2016
pp. 467-489
ISSN: 1567-7141
tlo.tbm.tudelft.nl/ejtir

# Timetable-based simulation method for choice set generation in large-scale public transport networks

**Thomas Kjær Rasmussen**[1]
DTU Management Engineering, Technical University of Denmark, Denmark.

**Marie Karen Anderson**[2]
DTU Management Engineering, Technical University of Denmark, Denmark.

**Otto Anker Nielsen**[3]
DTU Management Engineering, Technical University of Denmark, Denmark.

**Carlo Giacomo Prato**[4]
School of Civil Engineering, The University of Queensland, Australia.

The composition and size of the choice sets are a key for the correct estimation of and prediction by route choice models. While existing literature has posed a great deal of attention towards the generation of path choice sets for private transport problems, the same does not apply to public transport problems. This study proposes a timetable-based simulation method for generating path choice sets in a multimodal public transport network. Moreover, this study illustrates the feasibility of its implementation by applying the method to reproduce 5131 real-life trips in the Greater Copenhagen Area and to assess the choice set quality in a complex multimodal transport network. Results illustrate the applicability of the algorithm and the relevance of the utility specification chosen for the reproduction of real-life path choices. Moreover, results show that the level of stochasticity used in choice set generation should be high in order to provide stable parameter estimates when the choice sets are used for estimation regardless of the initial parameters for choice set generation. Last, results illustrate that adding heterogeneity across travellers should be required because coverage increases significantly, a relevant result considering that models are becoming more disaggregate in nature in real-life applications.

***Keywords***: *choice set generation, public transport networks, path choice, simulation-based approach, timetable-based simulation.*

## 1. Introduction

In order to understand the determinants of choice of public transport modes and to optimise the yield of investments in public transport systems, it is essential to have available a transport model. This should be able to capture the travellers' behaviour sensitivity to public transport system's attributes and to predict demand and path choices on public transport networks in a realistic manner.

[1] A: Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark T: +45 4525 6562 F: +45 4593 6533 E: tkra@transport.dtu.dk
[2] A: Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark T: +45 4525 6546 F: +45 4593 6533 E: tkra@transport.dtu.dk
[3] A: Bygningstorvet 116B, 2800 Kgs. Lyngby, Denmark T: +45 4525 1514 F: +45 4593 6533 E: tkra@transport.dtu.dk
[4] A: Brisbane 4072, Queensland, Australia T: +07 3365 1569 F: +07 3365 4599 E: c.prato@uq.edu.au

Modelling path choice essentially consists of two parts, namely the generation of a choice set and the representation of the choice between the generated paths (see, e.g., Bovy 2009; Prato 2009). The available paths can be generated either explicitly prior to the choice process or implicitly in the choice process, but explicit choice set generation allows full control over desired properties of the generated paths, size and composition of the choice sets, and flexibility of the model specification. Travellers are assumed to maximise their utility (i.e., minimise their cost) and hence to choose their preferred path in the set of available paths.

Recent studies have given increasing attention towards the importance of the size and the composition of choice sets for path choice (see, e.g., Prato and Bekhor 2007; Bliemer and Bovy 2008), whether they are to be used for model estimation or for prediction purposes. When used for model estimation, choice sets should facilitate statistical consistency and efficiency, while when used for prediction, they should contain all scenario-relevant alternatives (Van Nes et al. 2008). As a result, it is crucial to generate a choice set including alternatives that are considered relevant by travellers (Prato and Bekhor 2007; Bliemer and Bovy 2008). However, there exists no objective definition of what constitutes a relevant path, and hence the assessment of the generated path choice sets relies upon the experience of the analyst rather than objective measures of choice set quality.

The literature in path choice shows that choice set generation has been extensively investigated for car users and small synthetic networks, and has drawn much less attention for public transport users and large-scale networks. Deterministic and stochastic techniques have been implemented to the generation of alternative paths for car users: variations of shortest path algorithms (e.g., Akgün et al. 2000; Hunt and Kornhauser 1997; Lombard and Church 1993; Van der Zijpp and Fiorenzo-Catalano 2005); application of heuristic rules (e.g., Ben-Akiva et al. 1984; Azevedo et al. 1993; De la Barra et al. 1993); branch and bound algorithms (Hoogendoorn-Lanser et al. 2006; Prato and Bekhor 2006); single and doubly stochastic simulation approaches (e.g., Nielsen 2000; Bekhor et al. 2006; Bovy and Fiorenzo-Catalano 2007); biased random walk algorithm (Frejinger et al. 2009); breadth first search with network reduction (Rieser-Schüssler et al. 2013); Metropolis-Hastings sampling (Flötteröd and Bierlaire 2013). Choice set generation techniques have also been applied for public transport users: in metro networks, a heuristic approach pooling observations for the same origin-destination pair was applied in Santiago (Raveau et al. 2011); in multimodal networks, constrained enumeration was applied to a multimodal interregional hub-and-spoke transport corridor in the Netherlands (Hoogendoorn-Lanser et al. 2007) and a simulation-based doubly stochastic choice set generation method was tested on the same corridor (Bovy and Fiorenzo-Catalano 2007); aggregation of the network into "route segments" with consequent approximation of travel time and waiting time calculations was applied to evaluate existing choice set generation methods with smart card data in Singapore (Tan et al. 2014); a Google Map procedure was used to generate alternative routes in the public transport network of Montreal (Eluru et al. 2012).

The current study contributes to the literature on public transport path choice by proposing, implementing and testing a timetable-based simulation approach for the choice set generation of paths in large-scale multimodal networks. The proposed approach has its foundation in the doubly stochastic method which has previously been applied with success as a choice set generation method for car users, bicycle users and in multi-modal public transport networks where timetables are not considered explicitly (e.g., Bekhor et al. 2006; Bovy and Fiorenzo-Catalano 2007; Halldórsdóttir et al. 2014). Explicit treatment of the timetables is however important, since state-of-the-art models facilitating this allow a much more accurate and realistic representation of the performance of the network and the behaviour of travellers. The importance of the current study lies not only in the solution of the challenges of generating paths in a complex multi-layered, timetable-based public transport system, but also in the implementation in a large-scale network with multiple public transport modes. Moreover, the current study extends existing literature in public transport path choice by assessing the quality of path choice

sets via their comparison with real-life path choices in the public transport system of the Greater Copenhagen Area as well as the evaluation of the ability to produce stable parameter estimates in model estimation with respect to the parameter values used for choice set generation.

Recent literature has tackled the issue of consistent estimates after instances of importance sampling (e.g., Frejinger et al. 2009; Flötteröd and Bierlaire 2013; Guevara and Ben-Akiva 2013a, 2013b), but limitations in the implementation to large-scale networks emerge when considering the following: (i) the random walk (Frejinger et al. 2009) has convergence problems when tested on large networks with two-way links and the original application was on a small network with one-way links obviating loop formation; (ii) the Metropolis-Hastings algorithm (Flötteröd and Bierlaire 2013) has computational requirements as shown by its application that also lacks comparison with observed routes; (iii) the sampling and the importance sampling correction proposed by Guevara and Ben-Akiva (2013a, 2013b) performed excellently with Monte Carlo simulation, but was tested for a real data set with an extremely low number of observations and consequently rather large standard errors that facilitated a positive comparison for very large samples. Although a couple of recent studies have succeeded in estimating models while correcting for importance sampling in medium-size networks (Mai et al. 2015a; Vacca et al. 2015), recent literature has obviated the problem by introducing recursive models that can be consistently estimated for logit-type choice probabilities (Fosgerau et al. 2013; Mai et al. 2015b). These models rely on a dynamic specification of link choices and hence avoid a priori choice set generation and are consistently estimated or used for prediction in a computationally efficient way. However, their specification in contexts other than private transport has not been considered because of the specificity and complexity of multimodal networks.

This study tests the proposed method on 5,131 observations of actual path choices collected as part of the Danish Travel Survey, which is a one-day travel diary with high level of detail for the collection of public transport paths. For each observation, corresponding choice sets are generated and their coverage is assessed for various configurations of the generation (utility) function, thereby enabling recommendations of good configurations. Notably, the traditional coverage measures cannot be applied due to the temporal dimension of the public transport network and the possibility of several different line variants using the same segments of the network. Accordingly, different levels of coverage measures between generated and observed paths (i.e., line level, stop level) are defined.

The next section introduces the proposed timetable-based simulation method to generate choice sets for public transport path choice and describes how the generated choice sets can be evaluated. Then, the case study is presented including the configurations tested and how the generated choice sets are evaluated in the study. Next, the results of the assessment are presented, followed by a discussion and conclusions summarising the main findings of the study.

## 2. Proposed choice set generation in multimodal public transport networks

This section presents the timetable-based simulation method used to generate the path choice sets. Subsequently follows the introduction of methods for evaluating the choice set generation method and the choice sets generated.

### *2.1 Timetable-based simulation method to generate path choice sets*

The proposed method generates choice sets by repeated shortest path searches in a timetable-based public transport time-space network graph (i.e., a diachronic graph with spatial as well as temporal component of edges and nodes, cfr. Cascetta 2001) with the aim of having a more accurate representation of the network and a more realistic representation of the travellers' behaviour. The simulation concerns the attributes on the edges of the graph, the individual preferences of the travellers, and the departure time from the origin (within a certain interval of a pre-specified desired departure time). This approach implies that different unique paths may be

generated by repeated application of the simulation method. For example, one instance of the simulation with a high nuisance towards bus travel may generate a route with a long detour by train that arrives at the destination later than a direct bus route, and then another instance may generate the direct bus route itself. Also, one instance of the simulation may have a departure time that allows the traveller to board an early fast train, but another instance may have a later departure time that excludes the fast option leaving the traveller with a slower train connection. The union of the unique paths constitutes the choice set. The method uses generation (cost) functions, and the utility (cost) on path $i$ for individual $m$ is expressed as:

$$C_{im} = V_{im} + \varepsilon_{im} = \sum_j \beta_{jm} \cdot x_{ijm} + \varepsilon_{im} \tag{1}$$

where $V_{im}$ is the systematic utility of path $i$ for individual $m$, $\varepsilon_{im}$ accounts for perception errors as well as elements not accounted for in the systematic part of the generation function, $x_{ijm}$ is an attribute $j$ of path $i$ for individual $m$, and $\beta_{jm}$ is a parameter that expresses the preference of individual $m$ for the attribute $j$ of path $i$. It should be noted that it is assumed that for each attribute $j$ there exists a randomly distributed parameter $\beta_j$ accounting for taste heterogeneity across individuals $m$. The parameters are drawn once from the distribution of $\beta_j$ for each individual $m$ in each iteration of the path search.

The generation of the paths is based on the assumption that the cost on path $i$ is the sum of the costs on the edges belonging to path $i$ in the time-space network graph:

$$C_{im} = \sum_{l \in \Gamma_{im}} \left( \sum_j \beta_{jm} \cdot x_{ljm} + \varepsilon_{lm} \right) \tag{2}$$

where $\Gamma_{im}$ represent the set of edges belonging to path $i$ for individual $m$, $x_{ljm}$ is the attribute $j$ on edge $l$ for individual $m$, and $\varepsilon_{lm}$ is a random variable on edge $l$ for individual $m$ accounting for perception errors as well as elements not accounted for in the systematic part of the cost function. The assumption of additivity allows the paths to be consistently generated via shortest path searches in the graph: the individual specific parameters $\beta_{jm}$ are initially drawn before the search (in each iteration), and the impedances of the edges are then drawn as the shortest path tree is built (i.e., not all elements of the graph have to be simulated). In order to ensure consistency in the aggregation of the costs from edge- to path-level that allows for the costs of the path to follow a known distribution, it is necessary that (i) the error term follows a distribution which is additive in mean and variance, (ii) the error term has a variance proportional to the mean of the cost on the edge, and (iii) the cost function is specified as linear-in-parameters (Nielsen and Frederiksen 2006). We use a gamma distribution for condition (i), we define a proportionality factor $\gamma$ for condition (ii), and hence we have a distribution with mean $\sum_j \beta_{mj} \cdot x_{lmj}$ and variance $\gamma \cdot \sum_j \beta_{mj} \cdot x_{lmj}$. We note that the gamma distribution guarantees not only condition (i), but also that negative travel times are avoided. Also, we note that this specification implies that longer paths have higher variation (since the errors are summed over the link) and induces the link error term to depend on individual $m$ (since the mean of the cost is dependent on the random variables $\beta_{jm}$).

It should be noted that filtering after stochastic choice set generation has been previously suggested to increase the realism of the generated paths (see Bovy and Fiorenzo-Catalano 2007). However, filtering would require the definition of constraints with the drawbacks of (i) relying entirely on the discretion of the analyst and (ii) adhering to the spirit of the method by considering heterogeneity across travellers and hence introducing additional computation in drawing numbers from an arbitrary number of distributions of the thresholds. It should be also noted that the simulation of the edge costs does not influence the network graph, but only the path search, namely the graph is the same from realisation to realisation. The network graph is

based on the full (deterministic) timetable, and no vehicles, runs or passengers are thus simulated for the creation of the graph, only edge costs and preferences are simulated prior to the path search.

### *2.2 Methods for evaluation of choice sets*

In lack of a direct objective measure of what constitutes a relevant path, choice set generation methods can be evaluated based on a combination of (i) the size of the choice sets generated, (ii) the ability to generate choice sets containing at least one path having high similarity to a corresponding observed path, and (iii) the ability to generate choice sets which facilitate stable parameter estimates when used in the estimation of route choice models.

The first evaluation criterion relies on the analysis of the evolvement and size of the choice sets defined as the sets of unique paths generated by the repeated application of the simulation-based path-generation method. As timetable-based multimodal public transport networks are very detailed, the distinction between unique paths can be done at various levels of detail: (i) departure level, where a path is only considered unique if no other paths use the same departures of the same lines to and from the same stops; (ii) line level, where a path is unique only if no other paths use the same lines to/from the same stops; (iii) stop level, where a path is unique only if no other paths use lines with the same stopping pattern between the same to/from stops; (iv) trip leg mode sequence, where a path is unique only if no other path uses the same sequence of modes for the different trip legs used.

The second evaluation criterion relies on the analysis of whether the applied choice set generation method is able to generate paths similar to the observed path of an individual. This is performed through a measure of coverage, equal to the share of observations for which at least one path within the generated choice set has an overlap with the observed path equal to or above a certain threshold. As for the identification of unique paths, the overlap can be calculated according at various levels of detail. Note that using stop level rather than line level can avoid the possible bias introduced by what is known as the common line problem: in a segment served by many lines, people might not remember which line they used for trips between stops in a segment, and may consequently report the wrong line.

Another dimension to consider when specifying the overlap is the unit of measure, which could be overlap-in-time, overlap-in-utility or overlap-in-length. Using stop level as the level of detail and overlap-in-length as the unit of measure, the overlap $O_{i,stop,m}$ of the generated path $i$ with the observed path of observation $m$ can be computed as (Ramming 2002):

$$O_{i,stop,m} = \frac{L_{mi}}{L_m} \tag{3}$$

where $L_{mi}$ is the sum of length of overlapping elements between path $i$ and the observed path for observation $m$, and $L_m$ is the length of the observed path used by observation $m$.

This overlap measure can be computed for each generated path $i$ for observation $m$, and let $O^{\max}_{stop,m}$ denote the best overlap (measured on stop level using overlap-in-length as unit of measure) among the paths generated for observation $m$. Then the coverage for an overlap-threshold equal to $\delta$ can be computed as (Ramming 2002):

$$Cov_{stop}(\delta) = \frac{\sum_{n=1}^{M} I\left(O^{\max}_{stop,m} \geq \delta\right)}{M} \tag{4}$$

where $M$ is the number of observations and $I(\cdot)$ is an indicator equal to 1 when the criteria is fulfilled and 0 otherwise.

A path choice set generation method should produce an array of relevant paths within a reasonable amount of iterations, and the observed path should be among these. Visual inspection combined with network knowledge is one possible approach to use when evaluating whether counterintuitive paths are generated[5]. Such a procedure however could become tedious and infeasible when having many observations and large-scale networks. Alternatively, whether possibly counterintuitive paths and/or redundant paths being only minor deviations to existing paths are generated, could be evaluated at the aggregate level by comparing the increase in coverage to the increase in average choice set size. Large increases in average choice set size combined with low improvements in coverage would indicate that redundant paths similar to already existing paths or counterintuitive paths were generated. An efficient algorithm is characterised by fast increases in coverage as well as average choice set sizes.

As the composition of the choice set influences the parameter estimates when used for model estimation purposes (e.g., Train 2002; Van Nes et al. 2008), obtaining good coverage does not necessarily imply that parameters can be consistently estimated. For example, alternatives might or might not be similar to each other and choice sets might be different although reproducing at least once the same chosen path. Consequently, the third evaluation criterion relies on the analysis of whether the proposed choice set generation method is able to generate path sets which include relevant alternatives and allow obtaining statistically significant estimates of model parameters. This analysis relies on the estimation of the same choice model on the choice sets generated at different stochasticity levels, and enables also the analysis of whether and to what extent the estimated parameters vary across stochasticity levels.

## 3. Case study: Greater Copenhagen Area

This section presents the case study. Section **Error! Reference source not found.** introduces the data sources used in the study and the data preparation, while section **Error! Reference source not found.** presents the different tested configurations of the path choice set generation method. Section **Error! Reference source not found.** describes how the generated choice sets were evaluated.

### *3.1 Data*

*Observed paths*

The current study uses revealed preference data collected as part of the Danish Travel Survey, and the dataset consists of 5,131 observed paths in the multimodal public transport network of the Greater Copenhagen Area. The survey is an ongoing questionnaire-based collection of one-day travel diaries and associated respondents' and households' socio-economic characteristics. The respondents are a representative sample of the Danish population between 10 and 84 years of age who provide detailed information on all their trips during the day, and since February 2009 answer specific questions investigating the path choice of trips using public transport. Respondents fill in information at a level of detail enabling the path to be reproduced by the analyst while still being fairly easy to fill in by the respondent. Addresses and purposes at start points, change points and end points of the trips, as well as detailed information about the modes used *en route*, are collected in the survey:

- Walk, bike, car, etc.
    - Length and travel time
- Bus

[5] In this study, a counterintuitive path is a path that is clearly less attractive than an alternative path connecting the same origin and destination because it has a considerably larger travel cost.

- Waiting time, bus line, length and travel time

- Suburban train (S-train)
  - Waiting time, boarding station, S-train line, alighting station, length and travel time

- Train, Metro
  - Waiting time, boarding station, alighting station, length and travel time

In order to perform comparisons, the observed paths are map matched to the same digital network representation as the one used for the choice set generation. The paths are mapped at the line level identifying the line and the stops where boarding and alighting the different public transport modes occurred, but omitting identification of which actual departure was used. This map-matching level is chosen due to the uncertainty the stated travel times and departure time from the origins (as the departure time from the origin is reported only in 5 minute intervals) as well as possible delays on the day the paths were observed. The mapping has been documented in Rasmussen (2010) (for documentation in English, see Anderson and Rasmussen, 2010).

Figure 1 presents the characteristics of the observed paths in the data set, namely trip purpose, trip length, and number and mode of trip legs for multimodal paths. Most observations are either commuting or leisure trips, whereas only a few business trips have been observed in this representative sample of trips of the Danish population. Notably, trip characteristics appear similar between commuting and business trips, while education trips are similar with only a higher share of shorter trips below 10 km. In the Greater Copenhagen Area, the average commuting distance for public transport users is 21.0 km per direction, which is slightly higher than the average 17.1 km for the trips in the dataset. Leisure trips are generally shorter and

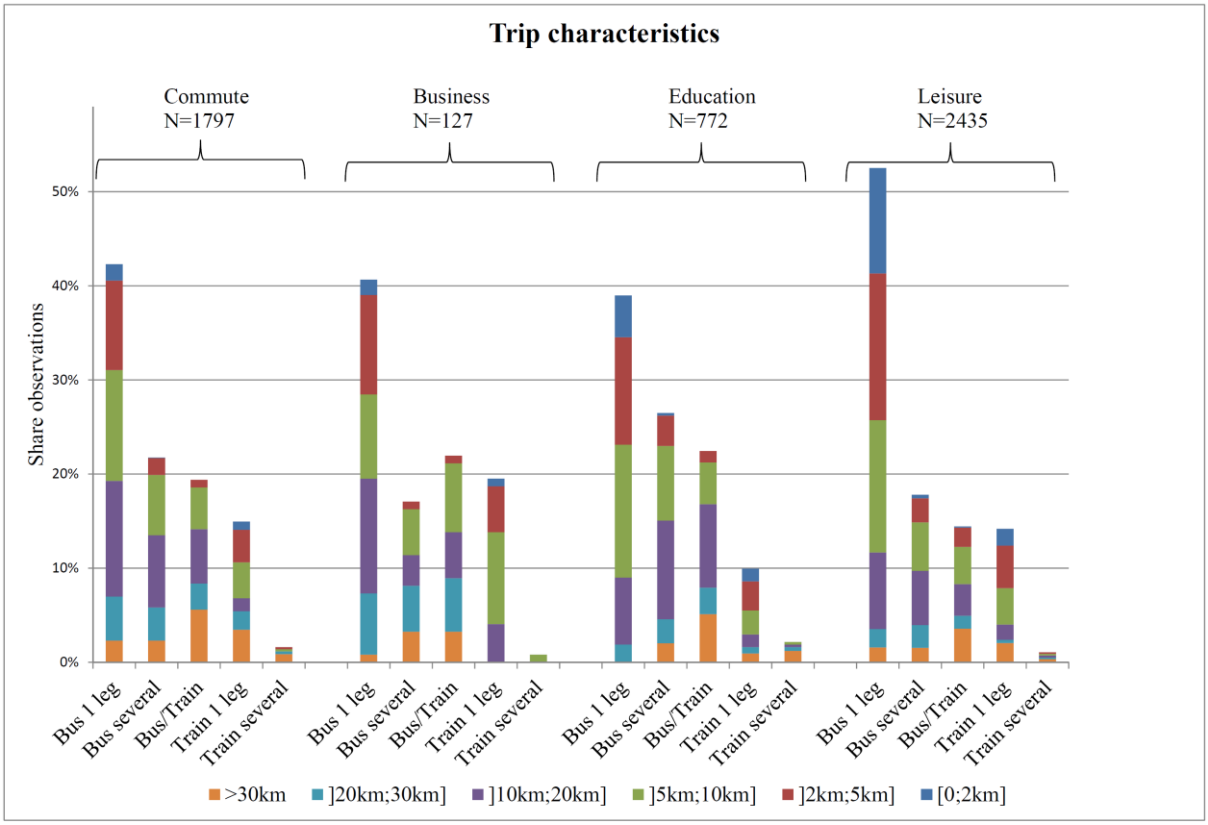


*Figure 1. Characteristics of the observed public transport trips. Share of observations by trip purpose, trip length and trip mode composition. Note: Bus 1 leg & Train 1 leg - trip consists of one trip leg using bus or train, respectively. Bus several/Train several - trip consists of several trip legs using bus or train, respectively. Bus/Train - trip consists of several trip legs using a combination of bus and train.*

consist of only one bus trip leg (53% of the observations), while other purposes observe less bus-only one-trip-leg trips (38%-42% of the observations).

*Network data*

The digital network represents the Greater Copenhagen Area, in which approximately 2 million people live and which covers an area of approximately 2300 km$^2$. The area is served by an array of different public transport modes, including numerous bus lines with different levels of service (i.e., regular, frequent, express, rapid), two metro lines, several Intercity and Regional train lines, seven S-train lines and various local train lines (for details, see Kaplan et al. 2014). The digital network representation is timetable-based and includes the departures of all the public transport in the area, namely 479 lines, 1,677 line variants, 5,652 stops and 635,027 daily stop departures. The data originates from The Danish National Transport Model (currently under development at DTU Transport), and the schedule is a digital representation of how the real-life public transport network was scheduled on November 10, 2010. Note that this represents a typical day and that the network structure has not changed significantly between this date and the period of the data collection. Transfers are available between lines at every stop, but the most important transfers (e.g., between bus and train at the Copenhagen Central Station) are also represented in the graph through 560 transfer edges that are defined via a length-dependent impedance expressing walking time between the connected stops.

The analysis evaluates the proposed choice set generation method through the generation of choice sets corresponding to the observed paths. Therefore the start and end locations (addresses) of the observations are introduced into the network, and linked to relevant public transport stops by connectors. The simplest approach to generate connectors would be to generate connectors to the nearest stop only. However, in order to facilitate the possibility of a wide array of alternative paths, a new approach to generating connectors between trip start and end locations and public transport stops is developed as a part of this study. The approach aims at generating connectors to all stops considered relevant by travellers, and thus connectors are generated according to the following criteria: (i) the 5 nearest bus stops served by bus lines with low service level within a distance of 2,500m; (ii) the 5 nearest bus stops served by bus lines with high service level within a distance of 5,000m; (iii) the 5 nearest train stations within a distance of 20km; (iv) the nearest bus stop on each of the A-, S or E- bus lines (high service level bus lines) within 20km if not already generated by step (ii). The travel time on the connectors is calculated by using actual network distances. Summarising, the multimodal public transport network used for choice set generation consists of 5,652 stops, 560 transfer edges, 202,035 connector edges, 635,027 public transport run edges between stops.

### *3.2 Configuration of the timetable-based simulation method to generate path choice sets*

In the current study, the detailed generation (cost) function used for generating the paths specifies the cost of alternative path *i* for observation *m* at the path level as:

$$\begin{aligned} C_{im} &= V_{im} + \varepsilon_{im} = \beta_{walktime,m} \cdot TT_{walktime,im} + \beta_{waittime,m} \cdot TT_{waittime,im} + \beta_{changepen,m} \cdot N_{change,im} \\ &+ \beta_{conntime,m} \cdot TT_{conntime,im} + \beta_{waitzone,m} \cdot TT_{waitzone,im} + \beta_{IVT,train,m} \cdot TT_{IVT,train,im} \\ &+ \beta_{IVT,IC-train,m} \cdot TT_{IVT,IC-train,im} + \beta_{IVT,S-train,m} \cdot TT_{IVT,S-train,im} + \beta_{IVT,bus,m} \cdot TT_{IVT,bus,im} \\ &+ \beta_{IVT,metro,m} \cdot TT_{IVT,metro,im} + \varepsilon_{im} \end{aligned} \quad (5)$$

where for path *i* and observation *m* $TT_{walktime,im}$ and $TT_{waitime,im}$ are walking and waiting time when transferring, $N_{change,im}$ is the number of transfers, $TT_{conntime,im}$ is the time spent travelling between the origin/destination and the first/last public transport stop, and $TT_{waitzone,im}$ is the schedule delay representing the difference between the desired departure time from the origin and the departure time from the origin in order to arrive at the first public transport stop at the departure of the first public transport leg. $TT_{IVT,train,im}$, $TT_{IVT,ICtrain,im}$, $TT_{IVT,S\text{-}train,im}$ and $TT_{IVT,bus,im}$ are in-vehicle times spent respectively in regional trains, IC-trains, S-trains and buses. The corresponding

parameters are distributed with mean $\beta_j$ and variance $a \cdot \beta_j$ where $a$ is a scale parameter, while $\varepsilon_{im}$ is the error term constituted as the sum of error terms drawn at the edge level.

In order to be able to recommend good formulations, three formulations of the generation (cost) function *(5)* were tested. To also be able to recommend levels of stochasticity introduced, each of these formulations is tested with nine levels of the variances of the distributions of error components and/or error term. This induces in total 27 configurations to test for each of the 5,131 observations. The choice of formulations to investigate is based on findings by Rasmussen (2010), who tested six different formulations on a limited number of observations. Consequently, in this present study, path choice sets are generated for three different formulations of the generation function *(5)*:

- *ErrTermOnly*: all $\beta$'s not randomly distributed across the population, and $\varepsilon_{im}$ Gamma distributed.
- *ErrCompAll*: all $\beta$'s Log-Normal distributed, and no consideration of $\varepsilon_{im}$ in the generation function.
- *ErrCompErrTerm*: all $\beta$'s Log-Normal distributed, and $\varepsilon_{im}$ Gamma distributed.

The distributions are chosen in order to avoid counterintuitive draws while still maintaining the theoretical assumptions: (i) negative values cannot be drawn from the Log-Normal distribution, securing to avoid counterintuitive cases where longer travel time generates lower cost; (ii) the Gamma distribution is additive in mean and variance and the variance is proportional to the mean; (iii) the consistency between the edge and the path level is maintained. Furthermore, the Gamma distribution has a finite support, whereby the risk of some alternative to have negative cost due to the error term can be avoided.

The nine levels of the scale parameters are defined based on starting values found in Rasmussen (2010) (see also Larsen et al. 2010) and are presented in Table 1. Consequently, *ErrCompErrTerm_7* refers to a configuration where all parameters and the error term are distributed with a scale of $a = \gamma = 1.5$. Rasmussen (2010) tested the levels denoted by _1, _2 and _3, and found that coverage increases with increasing size of the scale parameters. The present study additionally test cases where the scale parameters are considerably higher in order to find the level from which the coverage does not continue to improve and possibly becomes worse by increasing the level of stochasticity. The parameters are drawn for each observation in each iteration, whereas the error terms are drawn at the edge level for each observation in each iteration. The respondents could report the departure time only in 5 minute intervals. Accordingly, to account for this to allow different connections, a random departure time within a 10 minute interval around the recorded departure time is drawn before each path search.

**Table 1. Levels of the parameters that scale the variance of the distribution**

| Levels of the parameters | _1 | _2 | _3 | _4 | _5 | _6 | _7 | _8 | _9 |
|---|---|---|---|---|---|---|---|---|---|
| $\gamma$ | 0.05 | 0.10 | 0.15 | 0.20 | 0.50 | 1.00 | 1.50 | 2.00 | 5.00 |
| $a$ | 0.025 | 0.05 | 0.10 | 0.20 | 0.50 | 1.00 | 1.50 | 2.00 | 5.00 |

The means of the parameter values are based upon results estimated in Nielsen (2000), and are shown in Table 2.

**Table 2. Parameter values in the generation function (source: Nielsen, 2000)**

| $\beta_{walktime}$ | $\beta_{waittime}$ | $\beta_{changepen,i}$ | $\beta_{conntime}$ | $\beta_{waitzone}$ |
|---|---|---|---|---|
| 38.0 DKK/h | 38.0 DKK/h | 7 DKK/change | 45.0 DKK/h | 16.0 DKK/h |
| $\beta_{IVT,train}$ | $\beta_{IVT,IC\text{-}train}$ | $\beta_{IVT,S\text{-}train}$ | $\beta_{IVT,bus}$ | $\beta_{IVT,metro}$ |
| 27.0 DKK/h | 27.0 DKK/h | 27.0 DKK/h | 35.4 DKK/h | 21.6 DKK/h |

For each observation and configuration, 200 paths (i.e. 200 iterations of the path search) are generated between the corresponding origin and destination points of the observation. In total, 27,707,400 shortest path searches are conducted in the large-scale network (200 iterations, 27 configurations, 5,131 observations).

### *3.3 Evaluation of choice sets*

The generated choice sets are evaluated by their ability to generate unique paths, reproduce the observed paths, and produce stable parameter estimates when used in model estimation. As described, there are several levels of detail on which the paths can be distinguished and the overlap computed, and this section describes the choices made for the current case study.

The multimodal public transport network in the Greater Copenhagen Area is complex by often providing numerous different alternatives using the same sequence of modes for the trip legs used. These alternatives might however differ considerably from each other in terms of attributes such as e.g. travel times, and will need to be distinguished as different unique possibilities. Consequently, distinguishing between unique paths on the level of trip leg mode sequence is not considered attractive in this present study. By being timetable-based, the available digital representation of the network allows distinguishing paths at the departure level. The generated paths are however to be compared to observed paths mapped at the line level, and so there is no need to distinguish between generated paths at the departure level. In this study, the distinction between paths has thus been done at the line level.

This study calculates the overlap between paths on the aggregate stop level for the public transport trip legs (excluding access/egress), as this is less sensitive to the correctness of the input data. As an example of the common line problem, several S-train lines share the same alignment and stopping pattern through the city of Copenhagen, and people might not remember which line they used for trips between stops in the segment, and may consequently report the wrong line. Additionally, using the stop level would lower the sensitivity towards delays experienced on the day of the reported trip, as such delay is not represented in the digital network representation used for the choice set generation. This study adopts length as the unit of measure for the overlap, and the coverage can thus be computed as in equation (4).

Using the observed paths and the corresponding choice sets generated, a choice model is estimated for each of the levels of stochasticity used. This is done in order to evaluate the ability to produce statistically efficient parameter estimates and to test whether these are stable across stochasticity levels. The study does not explore several different specifications of the model to be estimated, but rather uses a Path Size Logit model formulation which in Anderson et al. (2014) was found to perform well. The utility function includes in-vehicle travel time in different modes of transport, access/egress times, walking and waiting times when transferring, number and type of transfers, headway between departures dependent on time-of-day as well as correction for path overlapping using the PSC correction term presented in Bovy et al. (2009). Biogeme (Bierlaire 2003) is used to conduct the maximum likelihood estimations.

## 4. Results

### *4.1 Choice set size*

Ideally, the number of unique paths would stabilise after generating a variety of paths, indicating that no counterintuitive and redundant paths were added to the choice sets. The path generation would then be terminated when this 'stable' situation was reached. With a high level of stochasticity, the simulation however seems to continue to generate new unique paths even after 200 iterations, whereas the choice set composition seems to stabilise for the smallest levels of stochasticity. This is expected, as introducing more randomness in terms of larger variance

around the mean might produce paths with minor deviations to the actually most attractive path as well as cause some obviously unattractive paths to become attractive.

The size of the generated choice sets is highly dependent on the formulation and the size of the stochasticity. This is indicated in Table 3, which lists various key figures describing the number of unique paths in the choice sets. As can be seen, the combined formulation generates the largest choice sets. Comparing formulations ErrTermOnly and ErrCompAll, the latter seems to generate the largest choice sets for the lowest stochasticity levels. The opposite is observed in the cases with high level of stochasticity.

**Table 3. Choice set size characteristics at iteration 40 and 200**

| | | ErrTermOnly | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Iteration | _1 | _2 | _3 | _4 | _5 | _6 | _7 | _8 | _9 |
| Min | 40 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 |
| | 200 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 5 |
| Mean | 40 | 2.9 | 3.5 | 4.5 | 5.9 | 9.2 | 13.6 | 17.2 | 20.0 | 28.9 |
| | 200 | 3.8 | 5.0 | 6.9 | 10.2 | 19.7 | 36.0 | 51.6 | 65.7 | 117.8 |
| Median | 40 | 2 | 3 | 4 | 5 | 8 | 12 | 16 | 20 | 31 |
| | 200 | 3 | 4 | 5 | 8 | 15 | 29 | 45 | 60 | 126 |
| Max | 40 | 18 | 20 | 28 | 32 | 37 | 40 | 40 | 40 | 40 |
| | 200 | 31 | 39 | 61 | 84 | 132 | 181 | 192 | 198 | 200 |

| | | ErrCompAll | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Iteration | _1 | _2 | _3 | _4 | _5 | _6 | _7 | _8 | _9 |
| Min | 40 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | 200 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 2 | 4 |
| Mean | 40 | 3.7 | 4.7 | 5.5 | 6.1 | 8.4 | 10.2 | 11.2 | 11.8 | 13.6 |
| | 200 | 5.3 | 7.5 | 9.3 | 10.9 | 16.9 | 22.5 | 26.2 | 28.3 | 34.4 |
| Median | 40 | 3 | 4 | 5 | 5 | 8 | 10 | 11 | 11 | 13 |
| | 200 | 4 | 6 | 8 | 9 | 15 | 20 | 24 | 26 | 32 |
| Max | 40 | 20 | 22 | 25 | 27 | 30 | 32 | 34 | 33 | 36 |
| | 200 | 35 | 51 | 54 | 69 | 86 | 105 | 121 | 128 | 143 |

| | | ErrCompErrTerm | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Iteration | _1 | _2 | _3 | _4 | _5 | _6 | _7 | _8 | _9 |
| Min | 40 | 1 | 1 | 1 | 1 | 1 | 1 | 4 | 3 | 3 |
| | 200 | 1 | 1 | 1 | 1 | 2 | 4 | 7 | 8 | 9 |
| Mean | 40 | 4.6 | 6.3 | 8.0 | 9.9 | 15.4 | 20.4 | 23.1 | 25.0 | 29.2 |
| | 200 | 7.4 | 11.3 | 16.0 | 22.5 | 43.5 | 67.6 | 82.6 | 93.2 | 118.7 |
| Median | 40 | 4 | 5 | 7 | 9 | 15 | 20 | 23 | 25 | 29 |
| | 200 | 6 | 9 | 13 | 19 | 39 | 65 | 80 | 92 | 118 |
| Max | 40 | 23 | 31 | 32 | 34 | 39 | 40 | 40 | 40 | 40 |
| | 200 | 47 | 76 | 100 | 135 | 172 | 187 | 190 | 200 | 200 |

Note: the four characteristics (minimum, mean, median, maximum size) are presented for each formulation and each configuration of the parameters that scale the variance of the distribution

In general, higher levels of the stochasticity imply larger choice sets, especially for the two formulations including a distributed error term. This indicates that, when adding high stochasticity, the method becomes efficient in terms of generating unique paths to the choice sets. Larger choice sets for higher stochasticity are also observed in Figure 2, which is an example of a commute trip between a suburb and the city centre of Copenhagen. As can be seen, the observed path is represented in the choice set for the different levels of stochasticity shown (formulation ErrCompErrTerm).

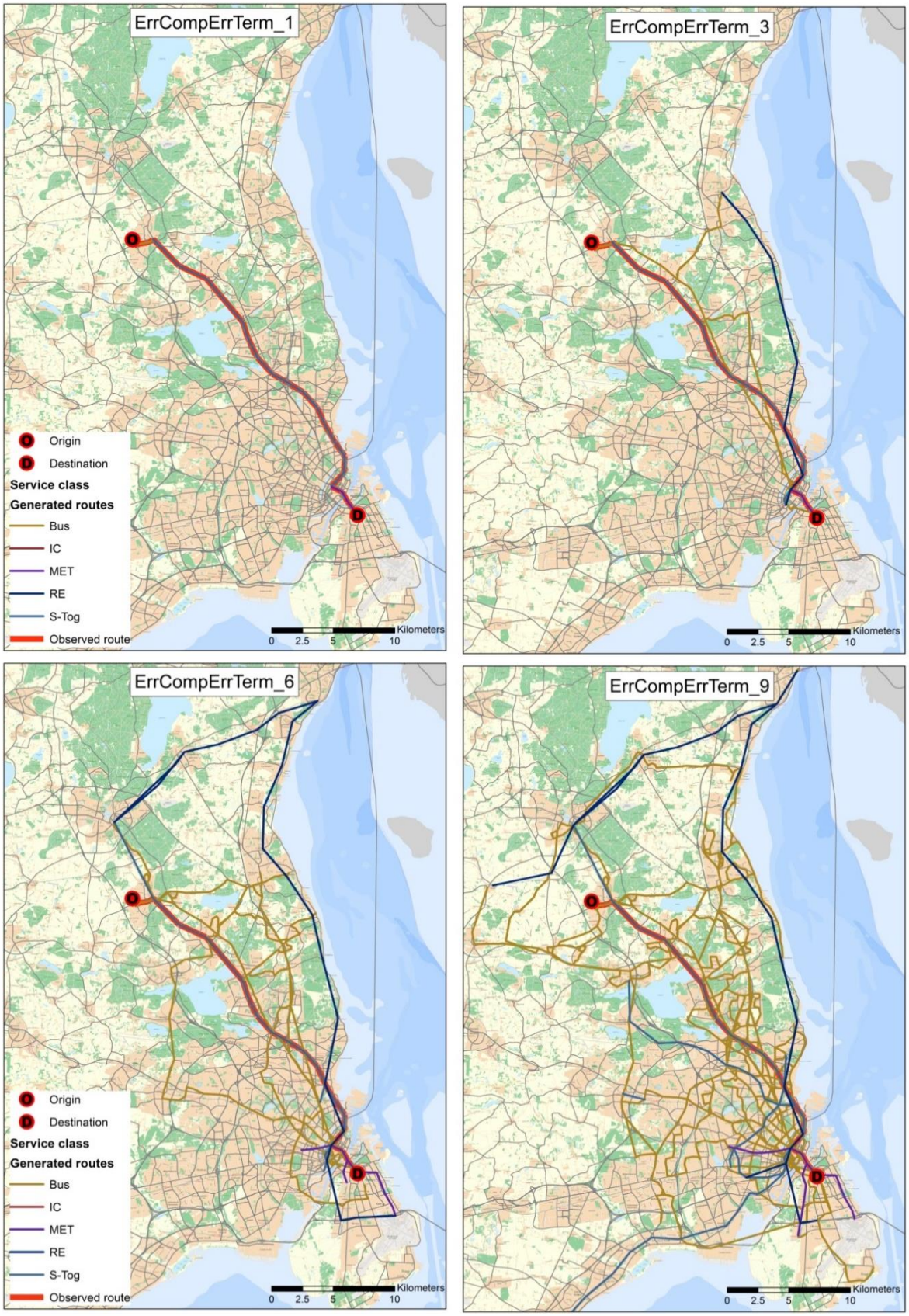


*Figure 2. Example of generated choice set for various levels of stochasticity for the ErrCompErrTerm formulation*

Figure 2 indicates a general tendency that has been verified by visual inspection in a Geographic Information System of the choice sets generated for numerous observations: when using a high

level of the stochasticity and after a number of iterations, new unique paths generated are redundant or counterintuitive. This suggests that it is undesirable to iterate until the number of unique paths stabilises.

### *4.2 Coverage*

The observed path should be represented among the set of generated paths. Therefore, the improvement in coverage could supplement the choice set size as an additional indicator of performance. Applying an overlap threshold of 80%, a value often used in the literature focusing on private transport (see, e.g., Ramming 2002; Prato and Bekhor, 2007), induces the results illustrated in Figure 3.

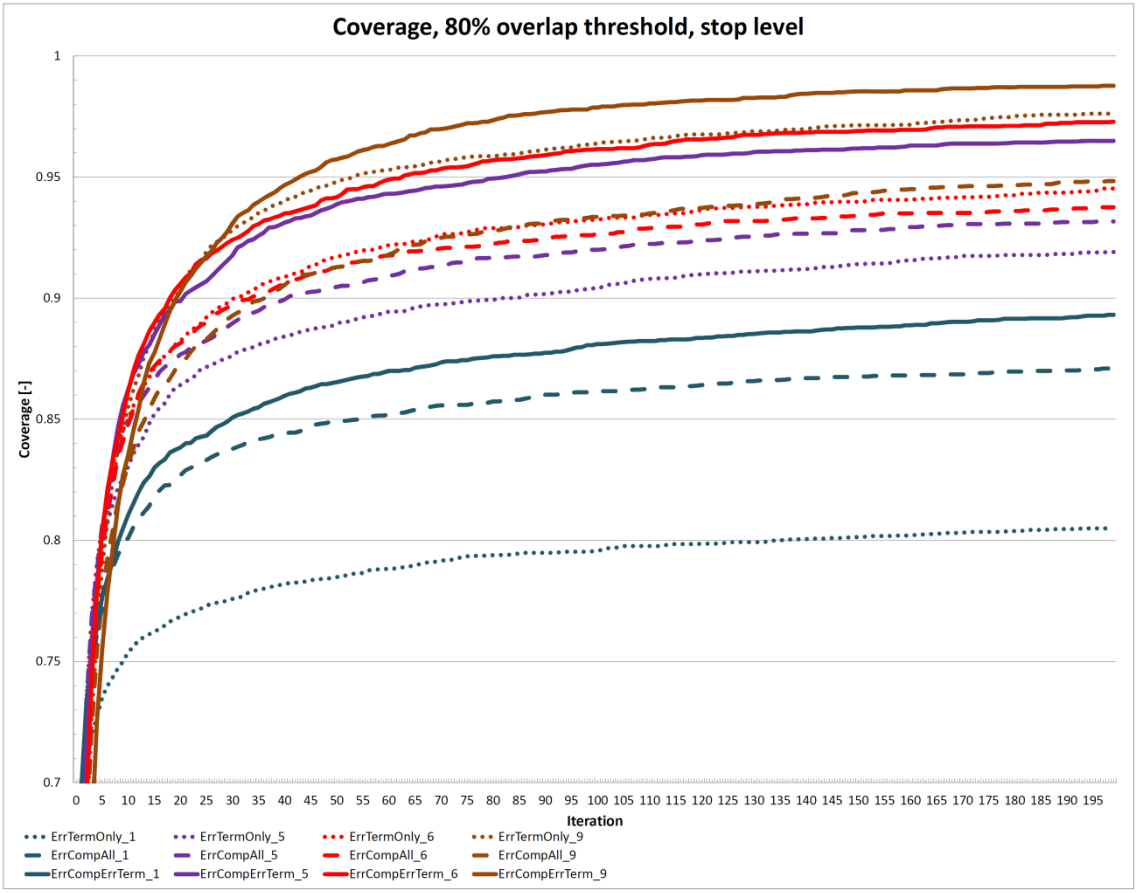


*Figure 3. Coverage (at the stop level) as a function of the iteration number with overlap threshold of 80% for all three formulations with configurations _1, _5, _6 and _9*

The timetable-based simulation method produces in general high coverage, especially when performing 40 iterations or more. Complete convergence is not seen within the 200 iterations, however the increment is rather small after 40 iterations. Figure 4 shows that setting a higher threshold of the overlap reduces the coverage, as expected, but the levels are still high, even for a threshold of 100%.

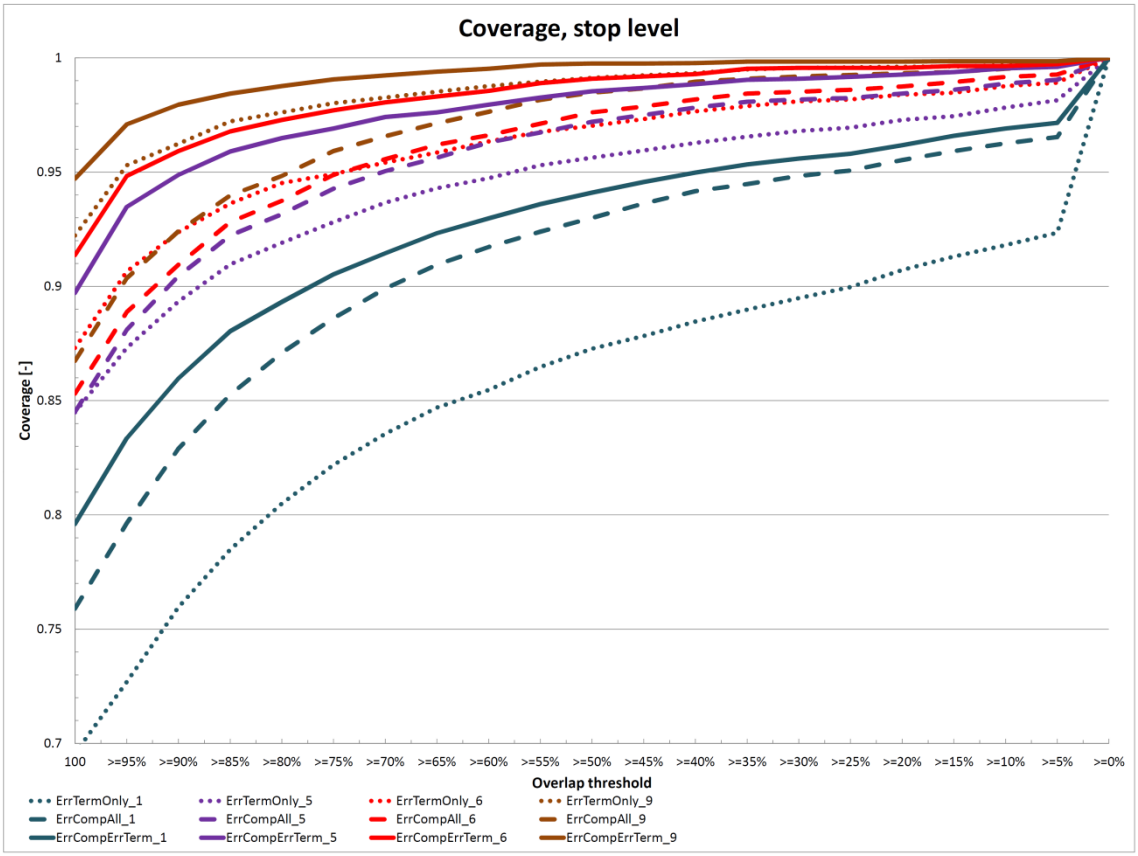


*Figure 4. Coverage (at the stop level) as a function of the overlap threshold (200 iterations) for all three formulations with configurations _1, _5, _6 and _9*

Rasmussen (2010) found that increasing the size of the scale parameter of the variance of the distributed terms does also increase the coverage. This is verified in the present analysis, and found valid even when applying large scale parameters. However, the increase in coverage by increasing the scale parameters from 1 (configurations _6) to 5 (configurations _9) only induce approximately a 2 percentage-point increase in the coverage. This increase is at the cost of generating counterintuitive paths when using higher levels of stochasticity.

When comparing the coverage across the formulations, it can be seen that the formulation ErrCompErrTerm outperforms the two other formulations at low as well as high levels of stochasticity (see Table 4). This indicates that accounting for taste heterogeneity by adding distributed parameters to the single stochastic formulation (widely used in car choice set generation) improves the coverage, and confirm previous findings by Bovy and Fiorenzo-Catalano (2007). By comparing Tables 3 and 4, it can be seen that formulation ErrCompAll not only generates more unique paths, but also produces better coverage than formulation ErrTermOnly when applying low stochasticity (after 40 as well as 200 iterations). Accordingly, if only low variance is to be applied to either the taste parameters or a distributed error term, the best results in terms of coverage are generated by applying distributions to the parameters. The opposite is seen when applying high stochasticity.

**Table 4. Coverage levels obtained at the 80% overlap threshold**

| | | Iteration 40 | | Iteration 200 | |
|---|---|---|---|---|---|
| | | Low stochasticity | High stochasticity | Low stochasticity | High stochasticity |
| Formulation | ErrorTermOnly | 78-83% | 92-94% | 81-86% | 95-97% |
| | ErrCompAll | 84-88% | 90-91% | 87-91% | 93-95% |
| | ErrCompErrTerm | 86-90% | 94-95% | 89-93% | 97-99% |

Note: the comparison is across the three formulations, low stochasticity indicates levels _1 through _5, high stochasticity indicates levels _6 through _9

The coverage grows, dependent on formulation and size of variance, between 2.7 percentage points and 4.5 percentage points when doing 200 iterations rather than 40 iterations. This gain is however at the cost of a 5 fold increase in computation time and, in cases with very high variance, larger choice sets including counterintuitive paths.

### *4.3 Model estimation results*

The Path Size Logit choice model with the utility function verbally described in section **Error! Reference source not found.** is estimated for the choice sets generated by the different variants of formulation ErrCompErrTerm. The observed path is added to the choice set if not generated by the choice set generation method. The focus is on the ErrCompErrTerm formulation, as it was found to perform best in terms of coverage. In section 4.1 it was established that the configurations with high stochasticity produced counterintuitive paths. In order to further investigate the influence of the presence of counterintuitive paths on the model estimation results, an additional level of stochasticity, denoted as ErrCompErrTerm_10, with $a = \gamma = 10$ is tested. Table 5 presents the parameter estimates for the models estimated with the different choice sets.

**Table 5. Comparison of the estimates of Path Size Logit models from choice sets generated with the ErrCompErrTerm formulation and different stochasticity levels**

| Parameters | _1 ($a$=0.025;$\gamma$=0.05) | | _3 ($a$=0.10;$\gamma$=0.15) | | _5 ($a$=0.50;$\gamma$=0.50) | | _7 ($a$=1.50;$\gamma$=1.50) | | _9 ($a$=5.00;$\gamma$=5.00) | | _10 ($a$=10.00;$\gamma$=10.00) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | est. | std.err. | est. | std.err. | est. | std.err. | est. | std.err. | est. | std.err. | est. | std.err. |
| Headway | | | | | | | | | | | | |
| Up to 6 min | 0.668* | 0.028 | 0.596* | 0.026 | 0.564* | 0.027 | 0.665* | 0.029 | 0.671* | 0.029 | 0.587* | 0.028 |
| Above 6 min | -0.053* | 0.008 | -0.053* | 0.007 | -0.058* | 0.006 | -0.047* | 0.006 | -0.043* | 0.006 | -0.067* | 0.006 |
| In-vehicle time | | | | | | | | | | | | |
| IVT Bus | -0.044* | 0.005 | -0.120* | 0.007 | -0.189* | 0.007 | -0.178* | 0.007 | -0.192* | 0.007 | -0.194* | 0.007 |
| IVT Local train | -0.050* | 0.012 | -0.132* | 0.012 | -0.178* | 0.014 | -0.165* | 0.017 | -0.180* | 0.020 | -0.186* | 0.019 |
| IVT Metro | -0.017* | 0.011 | -0.069* | 0.011 | -0.099* | 0.011 | -0.089* | 0.011 | -0.102* | 0.011 | -0.096* | 0.011 |
| IVT Regional/IC-train ≤ 20 km | -0.190* | 0.015 | -0.283* | 0.016 | -0.358* | 0.018 | -0.296* | 0.019 | -0.279* | 0.019 | -0.288* | 0.019 |
| IVT Regional/IC-train > 20 km | -0.011 | 0.010 | -0.072* | 0.009 | -0.131* | 0.011 | -0.132* | 0.012 | -0.144* | 0.013 | -0.144* | 0.013 |
| IVT S-train | -0.037* | 0.006 | -0.101* | 0.007 | -0.157* | 0.007 | -0.141* | 0.007 | -0.150* | 0.008 | -0.153* | 0.007 |
| TT Access/Egress | 0.027* | 0.006 | -0.133* | 0.017 | -0.308* | 0.016 | -0.276* | 0.014 | -0.316* | 0.013 | -0.367* | 0.014 |
| Transfer | | | | | | | | | | | | |
| Walking time | 0.170* | 0.018 | 0.053* | 0.016 | -0.057* | 0.015 | -0.078* | 0.014 | -0.128* | 0.015 | -0.146* | 0.016 |
| Waiting time | 0.216* | 0.009 | 0.074* | 0.005 | -0.028* | 0.006 | -0.079* | 0.006 | -0.102* | 0.006 | -0.092* | 0.005 |
| Bus to bus penalty | -1.170* | 0.100 | -1.960* | 0.118 | -2.820* | 0.107 | -1.930* | 0.088 | -1.380* | 0.081 | -1.880* | 0.089 |
| Bus to train penalty | -1.590* | 0.108 | -2.250* | 0.122 | -2.930* | 0.117 | -1.750* | 0.107 | -0.960* | 0.102 | -1.330* | 0.105 |
| Train to bus penalty | -1.840* | 0.112 | -2.500* | 0.125 | -3.190* | 0.118 | -2.060* | 0.108 | -1.250* | 0.102 | -1.540* | 0.108 |
| Train to train penalty | -0.604* | 0.087 | -1.290* | 0.096 | -2.110* | 0.085 | -1.330* | 0.072 | -0.890* | 0.065 | -1.350* | 0.075 |
| Path Size factor | | | | | | | | | | | | |
| PSC | 0.145* | 0.046 | 0.525* | 0.040 | 0.612* | 0.033 | -0.065* | 0.025 | -0.643* | 0.027 | -0.353* | 0.031 |
| Number of estimated parameters: | 16 | | 16 | | 16 | | 16 | | 16 | | 16 | |
| Number of observations: | 5131 | | 5131 | | 5131 | | 5131 | | 5131 | | 5131 | |
| Null log-likelihood: | -9785.6 | | -13389.4 | | -18557.5 | | -22179.8 | | -24278.9 | | -24807.5 | |
| Final log-likelihood: | -8235.3 | | -11174.9 | | -13253.3 | | -15791.2 | | -13748.9 | | -10783.0 | |
| Likelihood ratio test: | 3100.5 | | 4428.8 | | 10608.3 | | 12777.1 | | 21060.1 | | 28049.1 | |
| Adjusted rho-square: | 0.157 | | 0.164 | | 0.285 | | 0.287 | | 0.433 | | 0.565 | |

Note: * statistically significant at the 0.05 level

For configurations _1 and _3 (i.e. low variance) some parameters are non-significant and/or non-reasonable, e.g. with a decrease in the cost for increasing walking time. From configuration _5 onwards the parameters are all highly significant and with logical signs. Comparing the rates between parameters indicates as reasonable to associate less nuisance to 1 min. of travel in train (local train, S-train, Metro) than by bus, as these have a higher level of service and e.g. provide better possibility to work while travelling. However, the estimates indicate that, for short trips, one minute of travel on a regional or IC train is associated with higher nuisance than when travelling by bus. One possible explanation of this could be the somewhat more difficult boarding/alighting of the trains and the lower accessibility of trains' platforms. Another finding is that travellers prefer, reasonably, to travel with a high frequency line, which is seen through a positive parameter value for short headways.

Though statistically significant, it seems that the parameter estimates associated to travel time and headway are not stable across the cases with lower stochasticity, but only become stable after a higher level of stochasticity is introduced according to statistical testing across the estimates presented in Table 5 (see, for a similar comparison, Guevara and Ben-Akiva, 2013a, 2013b).

From Table 5 it can be seen that the parameter estimate associated to the number of transfers do not appear stable across stochasticity levels, also for configurations _7, _9 and _10. However, in general these configurations find that changing to a bus (from train or bus) is associated with a higher nuisance (corresponds to 6.5-11.6 min. of in-vehicle bus travel time) than when changing to a train (from train or bus, corresponding to 4.6-9.8 min. of in-vehicle bus travel time). This seems reasonable, as train stations and train platforms typically provide better level-of-service than bus stops (e.g., better shelter for weather).

The PSC parameter is also highly significant in all cases and the negative sign of the parameter estimate (for all cases but _1, _3 and _5) is also expected. However, when observing Table 5 it can be seen that the parameter estimate are non-stable across stochasticity levels, which seems reasonable since the PSC by definition depends on the choice set composition.

It should be noted that the choice sets used for estimating the parameters reported in Table 5 are all generated using mean parameter values in the utility function found by Nielsen (2000), while the estimated parameters do not align with these initial parameter values. With the aim of investigating the effect of the difference in the parameter values, an additional round of estimation is proposed to verify the ability to reproduce the input parameter estimates as well as the sensitivity to the specification of the mean values of the parameters of the choice set generation method. Specifically, the values estimated while using the choice set with configuration _10 (refer to as 'round 1') are used into the generation function of the choice set generation method, new choice sets are generated and then new models are estimated while maintaining the level of stochasticity for that configuration _10 (refer to as 'round 2'). Table 6 reports the estimated parameters as well as indicates whether the confidence intervals of the parameters overlap across round 1 and round 2.

**Table 6. Estimates of the Path Size Logit models from choice sets generated with the ErrCompErrTerm formulation and the stochasticity level _10, with different mean parameter values in the choice set generation**

| | Estimates | | | | |
|---|---|---|---|---|---|
| | first round | | second round | | overlap conf. |
| | est. | std. err. | est. | std. err. | |
| Headway | | | | | |
| min btw departures | -0.059 | 0.003 | -0.057 | 0.003 | * |
| In-vehicle time | | | | | |
| IVT Bus | -0.180 | 0.004 | -0.174 | 0.004 | * |
| IVT Local train | -0.187 | 0.012 | -0.169 | 0.013 | * |
| IVT Metro | -0.099 | 0.009 | -0.109 | 0.010 | * |
| IVT Regional/IC-train ≤ 20 km | -0.275 | 0.014 | -0.241 | 0.015 | * |
| IVT Regional/IC-train > 20 km | -0.145 | 0.011 | -0.153 | 0.012 | * |
| IVT S-train | -0.162 | 0.005 | -0.155 | 0.006 | * |
| TT Access/Egress | -0.370 | 0.005 | -0.334 | 0.005 | |
| Transfer | | | | | |
| Walking time | -0.142 | 0.017 | -0.159 | 0.018 | * |
| Waiting time | -0.093 | 0.003 | -0.123 | 0.004 | |
| Bus to bus penalty | -1.880 | 0.070 | -1.680 | 0.069 | ** |
| Bus to train penalty | -1.080 | 0.081 | -0.841 | 0.087 | ** |
| Train to bus penalty | -1.290 | 0.082 | -1.030 | 0.086 | ** |
| Train to train penalty | -1.290 | 0.061 | -1.100 | 0.062 | ** |
| Path Size factor | | | | | |
| PSC | -0.171 | 0.026 | -0.154 | 0.027 | * |
| Number of estimated parameters: | 15 | | 15 | | |
| Number of observations: | 5131 | | 5131 | | |
| Null log-likelihood: | -24808 | | -24917 | | |
| Final log-likelihood: | -11145 | | -8463 | | |
| Likelihood ratio test: | 27326 | | 32907 | | |
| Adjusted rho-square: | 0.550 | | 0.660 | | |

* 80% confidence intervals overlapping; ** 90% confidence intervals overlapping

Table 6 shows that the parameter estimates are stable as the confidence intervals of the parameters are overlapping. This indicates that a "fixed" point has been reached in which the input parameters used for the choice set generation are aligned with the estimated parameters from the model estimation. It should be noted that this "fixed" point is reached within one iteration only, and the results suggest that the estimated parameter values are not highly dependent on the input parameters used for choice set generation. This has an important implication on the transferability of the model, since the results are not dependent on the initial specification of the value of the input parameters as long as the level of stochasticity is high.

## 5. Discussion

The current study investigates the generation of path choice sets in a complex real-life multimodal public transportation network. The analysis focuses on 5,131 actual choices of public

transport users in the Greater Copenhagen Area and the choice set quality is evaluated against these revealed preference data.

The study implements a timetable-based simulation method for choice set generation of public transport paths. The model is flexible regarding the configuration of the generation (cost) function, as it can capture similarities across alternatives and perception errors through a distributed error term as well as taste heterogeneity through distributed parameters. Various configurations are tested in order to be able to give recommendations. The cost function includes in-vehicle time for all the public transport modes available in the Greater Copenhagen Area (i.e., bus, metro, train), waiting and walking (connecting) time at the stations, the number of transfers as well as network-distance-based access/egress time from/to the origin/destination.

Results show that the proposed timetable-based simulation method for choice set generation in general produces high coverage, especially when using a high level of stochasticity. Adding parameters drawn from a log-normal distribution to account for taste heterogeneity improves the results considerably compared to the traditional single stochastic model with a gamma distributed error term. Distributing the parameters without having a distributed error term also generates good results, as it actually performs better than the traditional single stochastic formulations at low levels of stochasticity. The formulation is however outperformed by the doubly stochastic formulation, which generates the best results among the three formulations tested. When evaluating the coverage, it is important to bear in mind that while high coverage should be sought, it is usually not possible to obtain 100% coverage. This is due to possible (non-traceable) errors in the observed data as well as deviations between the real-life situation when collecting the observed data and the available network data. When comparing to coverage levels obtained elsewhere in the literature (e.g., Ramming 2002; Prato and Bekhor 2007), it is confirmed that high coverage levels are found using the doubly stochastic formulation of the timetable-based simulation method. It should be noted that results are not directly comparable across different studies as different data sources (networks, observations) and methods are used. Additionally, results are highly dependent on the chosen overlap threshold as well as aggregation level (e.g., departure level, trip leg mode sequence). The choice of this should depend on, among others, the level of detail as well as accuracy of the available data. However, the current analysis indeed shows high coverage levels even at high overlap thresholds and for both line level and stop level. Future research could seek to apply some of the numerous alternative methods for choice set generation proposed in the literature (section **Error! Reference source not found.**) on the network and observations, thereby facilitating consistent comparison across methods.

For all formulations, the coverage seems to increase when increasing the level of stochasticity. The improvements are small at high levels of stochasticity though, and our tests show that adding too much stochasticity generates large choice sets with counterintuitive paths. Consequently, adding stochasticity to improve coverage should be done with parsimony and controlled for by observing its increase with the number of iteration. For the lowest levels of stochasticity, the size of the choice sets does not grow fast, indicating that the same paths are generated over and over. For the highest levels of stochasticity, large choice sets are generated within a reasonable amount of iterations, corresponding to a high efficiency in terms of generating alternatives. The observed path is also often among the initial alternatives generated to the choice sets, which is seen through a high coverage level at 40 iterations. Actually, using 200 iterations rather than 40 iterations, at the cost of a 5-fold increase in calculation time, only improves the coverage marginally.

The study finds that large choice sets containing counterintuitive paths are generated when increasing the scale parameters above 1. Such paths are not only behaviourally unrealistic, but may also influence the subsequent step where the choice sets are typically used for either estimation or prediction purposes. When used for prediction, the large choice sets could potentially pose a computational challenge if a path-based solution algorithm is used. When

used for estimation, the study finds statistically significant and reasonable parameter values when using choice sets generated at high stochasticity levels. Furthermore, apparently adding counterintuitive paths do not change the estimates considerably for the rates of substitution related to time. However, the rates of substitution associated to the number and type of transfers, though highly significant and at a logical level, does not stabilise above a certain stochasticity level. Accordingly, the parameter estimates for transfers are apparently highly dependent on the composition of the choice set.

The value of the rate of substitution of transfers reported in other revealed preference studies also varies greatly between studies, ranging from 3.8 (relative to in-vehicle metro travel time) in Raveau et al. (2011) to 22.4 (relative to in-vehicle train travel time) in Vrtic and Axhausen (2002). None of the studies reporting evidence on the value of changes has investigated in detail the implication of varying choice set composition on the estimated values, often also because the choice set composition effect was simply not considered. An interesting future research direction would be to investigate whether the fluctuation found arises due to the discrete nature of the variable(s) with values typically in the lower end of the scale (0, 1 or 2). It would also be interesting in a future study to account for trip purpose in the specification of the generation function as well as in the choice model estimation, to see whether this would improve the coverage even further and possibly generate better model fit and more stable estimates for the parameters associated to transfers. Last, it would also be valuable in a future study to analyse the effect of the mentioned specifications and the choice set composition not only on the estimates, but also on the prediction performances.

The present study does not address the issue of consistency across the choice set generation component and the choice model component (estimation or prediction). Moreover, the specification of the generation function and the utility function of the estimation process are different: the functions contained different components (e.g. the utility function contained path-based attributes such as PSC correction), and the stochasticity and the size of this are specified differently. Some of these components and the defined stochasticity does not easily break down from path- to link-level, which are typically required by choice set generation methods as these adopt search-tree algorithms in the generation of paths. Arguably, theoretical consistency should be ensured across the choice set generation component and the choice model component by using the same specification of the utility and generation function. The 'hypothesis' about traveller preferences (used in the path generation) does thereby become consistent with the preferences actually estimated based on these. The development of methods which ensure this consistency across model components is an important future research direction.

## 6. Conclusions

This study investigates actual path choices of public transport users and assesses the choice set quality against these. The study illustrates that the timetable-based simulation method for choice set generation of public transport paths is applicable to large-scale networks, produces good results in terms of coverage and facilitates consistent estimation of model parameters with different choice sets. Adding variability across people improves the results considerably, and the best results are seen with the doubly stochastic formulation when the level of the stochasticity introduced was high, although not too high. It is found that adding stochasticity translates into the generation of redundant and counterintuitive paths after a certain level, but interestingly the estimation results are not affected considerably by the presence of these (with the exception of the parameters associated to transfers and access/egress time). Moreover, results show that the model is transferable across studies, since the estimation results are not highly sensitive to changes in the mean parameter values used in the choice set generation as long as the stochasticity level is high.

## Acknowledgements

The financial support of the Danish Council for Strategic Research to the project 'Analyses of activity-based travel chains and sustainable mobility' (ACTUM) is gratefully acknowledged. We also thank two anonymous reviewers for their insightful comments that helped us improve an earlier draft of this paper.

## References

Akgün, V., Erkut, E., and Batta, R. (2000). On finding dissimilar paths. European Journal of Operational Research, 121, 232-246.

Anderson, M.K., Nielsen, O.A., and Prato, C.G. (2014). Multimodal route choice models of public transport passengers in the Greater Copenhagen Area. DTU Transport working paper submitted for publication in the EURO Journal on Transportation and Logistics.

Anderson, M.K., and Rasmussen, T.K. (2010). Matching observed public route choice data to a GIS network. Selected proceedings of the Annual Transport Conference in Aalborg, Aalborg University, 2010.

Azevedo, J.A., Santos Costa, M.E.O., Silvestre Madera, J.J.E.R., and Vieira Martins, E.Q. (1993). An algorithm for the ranking of shortest paths. European Journal of Operational Research, 69 (1), 97-106.

Bekhor, S., Ben-Akiva, M.E., and Ramming, M.S. (2006). Evaluation of choice set generation algorithms for route choice models. Annals of Operations Research, 144 (1), 235-247.

Bekhor, S., Toledo, T., and Prashker, J.S. (2008). Effects of Choice Set Size and Route Choice Models on Path-Based Traffic. Transportmetrica, 4 (2), 117-133.

Ben-Akiva, M.E., Bergman, M.J., Daly, A.J., and Ramaswamy, R. (1984). Modeling interurban route choice behaviour. In J. Volmuller & R. Hamerslag (Eds.), Proceedings of the 9th international symposium on transportation and traffic theory. Utrecht, The Netherlands: VNU Science Press, pp. 299-330.

Bierlaire, M. (2003). BIOGEME: A free package for the estimation of discrete choice models. Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Schwitzerland.

Bliemer, M.C.J. and Bovy, P. H. L. (2008). Impact of route choice set on route choice probabilities. Transportation Research Record, 2076, 10-19.

Bovy, P.H.L. (2009). On modelling route choice sets in transportation networks: a synthesis. Transport Reviews, 29 (1), 43-68.

Bovy, P.H.L., Bekhor, S., and Prato, C.G. (2009). The Factor of Revisited Path Size: Alternative Deriviation. Transportation Research Record, 2076, 132-140.

Bovy, P.H.L., and Fiorenzo-Catalano, S. (2007). Stochastic route choice set generation: behavioural and probabilistic foundations. Transportmetrica, 3 (3), 173-189.

Cascetta, E. (2001). Transportation Systems Engineering: Theory and Methods. Applied Optimization Series, 49, Norwell, MA, USA: Kluwer Academic Publishers.

De la Barra, T., Perez, B., and Anez, J. (1993). Multidimensional path search and assignment. In Proceedings of the 21st PTRC summer annual meeting, Manchester, UK, 307–319.

Eluru, N., Chakour, V., and El-Geneidy, A. M. (2012). Travel mode choice and transit route choice behaviour in Montreal - Insights from McGill University members commute patterns. Public Transport, 4, 129-149.

Fosgerau, M., Frejinger, E., and Karlstrom, A. (2013). A link based network route choice model with unrestriced choice set. Transportation Research Part B: Methodological, 56, 70-80.

Frejinger, E. (2008). Route Choice Analysis: Data, Models, Algorithms and Applications. Ph.D. thesis, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland.

Flötteröd, G., and Bierlaire, M. (2013). Metropolis–Hastings sampling of paths. Transportation Research Part B: Methodological, 48, 53-66.

Frejinger, E., Bierlaire, M., and Ben-Akiva, M.E. (2009). Sampling of alternatives for route choice modeling. Transportation Research Part B: Methodological, 43 (10), 984–994.

Friedrich, M., Hofsass, I., and Wekeck, S. (2001). Timetable-based transit assignment using branch and bound. Transportation Research Record, 1752, 100-107.

Guevara, C.A., and Ben-Akiva, M. (2013a). Sampling of alternatives in Multivariate Extreme Value (MEV) models. Transportation Research Part B: Methodological, 48, 31-52.

Guevara, C.A., and Ben-Akiva, M. (2013b). Sampling of alternatives in Logit Mixture models. Transportation Research Part B: Methodological, 58, 185-198.

Halldórsdóttir, K., Rieser-Schüssler, N., Axhausen, K.W., Nielsen, O.A., and Prato, C.G. (2014). Efficiency of choice set generation methods for bicycle routes. European Journal of Transport and Infrastructure Research, 14(4), 332-348.

Hoogendoorn-Lanser, S. (2005). Modeling Travel Behavior for multi-Modal Transport Networks. Ph.D. Thesis. Delft University of Technology, Delft, Holland.

Hoogendoorn-Lanser, S., Bovy, P.H.L., and van Nes, R. (2007). Application of constrained enumeration approach to multimodal choice set generation. Transportation Research Record, 2014, 50-57.

Hoogendoorn-Lanser, S., van Nes, R., and Bovy, P.H.L. (2006). A Constrained Enumeration Approach to Multi-modal Choice Set Generation. In Proceedings, 11th International Conference on Travel Behaviour Research (CD-ROM), Kyoto, Japan.

Hunt, D.T., and Kornhauser, A.L. (1997). Assigning traffic over essentially-least-cost paths. Transportation Research Record, 1556, 1-7.

Kaplan, S., Popoks, D., Prato, C.G., and Ceder, A. (2014). Using connectivity for measuring equity in transit provision. Journal of Transport Geography, 37, 82-92.

Larsen, M.K., Nielsen, O.A., Prato, C.G., and Rasmussen, T.K. (2010). Generation and quality assessment of route choice sets in public transport networks by means of RP data analysis. In Proceedings of the European Transport Conference, 11–13 October 2010, Glasgow, Scotland.

Lombard, K., and Church, R. L. (1993). The gateway shortest path problem: Generating alternative routes for a corridor location problem. Geographical Systems, 1, 25-45.

Mai, T., Frejinger, E., and Bastin, F. (2015a). A misspecification test for logit based route choice models. Economics of Transportation, 4, 215-226.

Mai, T., Fosgerau, M., and Frejinger, E. (2015b). A nested recursive logit model for route choice analysis. Transportation Research Part B: Methodological, 75, 100-112.

Nielsen, O.A. (2000). A stochastic transit assignment model considering differences in passengers utility functions. Transportation Research Part B: Methodological, 34 (5), 377-402.

Nielsen, O.A., and Frederiksen, R.D. (2006). Optimization of timetable-based, stochastic transit assignment models based on MSA. Annals of Operations Research, 144(1), 263-285.

Prato, C.G. (2009). Route choice modeling: past, present and future research directions. Journal of Choice Modelling, 2 (1), 65-100.

Prato, C.G., and Bekhor, S. (2006). Applying branch-and-bound technique to route choice set generation. Transportation Research Record, 1985, 19-28.

Prato, C.G., and Bekhor, S. (2007). Modeling route choice behavior: how relevant is the composition of choice set? Transportation Research Record, 2003, 64-73.

Ramming, M.S. (2002). Network Knowledge and Route Choice. Ph.D. Thesis. Massachusetts Institute of Technology, Cambridge, USA.

Rasmussen, T.K. (2010). Public transport route choice in the Copenhagen Region – generation and assessment of route choice sets (in Danish). Master thesis, DTU Transport, Technical University of Denmark.

Raveau, S., Muñoz, J.C., and de Grange, L. (2011). A topological route choice model for metro. Transportation Research Part A: Policy and Practice, 45 (2), 138-147.

Rieser-Schüssler, N., Balmer, M., and Axhausen, K.W. (2013). Route choice sets for very high-resolution data. Transportmetrica A: Transport Science, 9 (9), 825-845.

Sheffi, Y. (1985). Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall, Inc., Englewood Cliffs, NJ.

Tan, R., Robinson, S., Lee, D.-H., and Ben-Akiva, M.E. (2014). Evaluation of choice set generation methods with smart card data in large-scale public transport network. Presented at 13th ITS Asia Pacific Forum, Auckland, New Zealand.

Train, K.E. (2002). Discrete Choice Models with Simulation. Cambridge University Press, Cambridge.

Tsui, S.Y.A., and Shalaby, A.S. (2006). Enhanced system for link and mode identification for personal travel surveys based on global positioning systems. Transportation Research Record, 1972, 38-45.

Vacca, A., Prato, C.G., and Meloni, I., 2015. Estimating route choice models from stochastically generated choice sets on large-scale networks: Correcting for unequal sampling probability. Transportation Research Record, 2493, 11-18.

Van der Zijpp, N. J., and Fiorenzo-Catalano, S. (2005). Path enumeration by finding the constrained K-shortest paths. Transportation Research Part B: Methodological, 39 (6), 545-563.

Van Nes, R., Hoogendoorn-Lanser, S., and Koppelman, F.S. (2008). Using choice sets for estimation and prediction in route choice. Transportmetrica, 4(2), 83-96.

Vrtic, M., and Axhausen, K.W. (2002). Impact of tilting trains in Switzerland: Route choice model of regional and long-distance public transport trips. Presented at the 82nd Annual Meeting of the Transportation Research Board, Washington, D.C.