Identification of local detouredness and correlation effects in a bounded route choice model: solution & estimation in large-scale networks

Lawrence Christopher DUNCAN [a] *, Thomas Kjær RASMUSSEN [a], David Paul WATLING [b], Laurent CAZOR [a], Otto Anker NIELSEN [a]

[a] Department of Technology, Management and Economics, Technical University of Denmark
Akademivej Bygning 358, 2800 Kgs. Lyngby, Denmark.
[b] Institute for Transport Studies, University of Leeds
36-40 University Road, Leeds, LS2 9JT, United Kingdom.

* corresponding author:
Department of Technology, Management and Economics, Technical University of Denmark
Akademivej Bygning 358, 2800 Kgs. Lyngby, Denmark.
E-mail: lawdun@dtu.dk

Highlights

- Formulation of a Bounded Path Size Local Detour Threshold (BPS-LDT) route choice model
- Bounds imposed on full-route & sub-route detours and used route overlap captured
- Solution method with pre-processed sub-route information from pre-generated routes
- Estimation in a real-life case study with BPS-LDT model outperforming relevant models
- Experiments show statistical significance, forecastability, and MLE solution uniqueness

Abstract

The Bounded Choice Model with Local Detour Threshold (BCM-LDT) route choice model (Rasmussen et al., 2024) proposes that local detouredness (the extent to which a route detours on its subparts) is an influential factor upon route choice probability and choice set formation. The current paper addresses three unresolved challenges regarding the BCM-LDT model: i) accounting for correlations between overlapping routes defined as realistic, ii) developing a solution approach for applying the model to large-scale networks, and, iii) developing, testing, and applying a procedure for estimating the model. Addressing i), appropriate path size correction factors are integrated within the BCM-LDT probability relation, to adjust probabilities to capture correlations between routes defined as realistic by the local and global bounds. Addressing ii), due to the high computational burden of current methods for generating all routes below the local and global bounds, the current paper develops a heuristic solution approach that works from representative universal choice sets and pre-processes the necessary segment information for computing local detour measures. Solution tricks are developed to considerably improve computation times. Addressing iii), a modified maximum likelihood estimation procedure utilising tracked route observation data is developed, tested in a simulation study, and applied in a real-life large-scale case study. The local detour models are successfully estimated and found to outperform associated non-local detour models. Experiments provide evidence that the parameter estimates can be identified, are statistically significant and unique, and that the model is suitable for forecasting and can be estimated in feasible computation times.

Key Words: bounded route choice model, path size, local detouredness, choice set formation, parameter estimation

1 Introduction

1.1 Background

Route choice models are widely used by transport researchers and policy makers for an array of purposes, such as appraising travellers' perceptions of route characteristics (e.g. Hood et al. 2011; Toledo et al., 2020; Zhong & Miao, 2024), assessing transport policies (e.g. Chen et al., 2018; Tsai & Li, 2019; Mardan et al., 2024), predicting the impact of future changes in demand (e.g. Martens & Hurvitz, 2009; Wei et al., 2020), traffic assignment (e.g. Prashker & Bekhor et al., 2004; Gentile, 2012; Lim & Kim, 2016; Brederode et al., 2018; Duncan et al. 2023,2024), and designing transport networks (e.g. Joksimovic et al., 2005; Jiang & Szeto, 2014; Cadarso & Marín, 2016).

Although other decision rules are considered in the literature (e.g. regret minimisation (Chorus, 2012; Li & Huang, 2016) and non-compensatory decision rules (Chorus & van Cranenburgh, 2024; Cazor et al., 2024), by far the most commonly used route choice decision rule is utility maximisation. There are two types of utility-maximisation-based route choice model: path-based and link-based. Path-based models such as Multinomial Logit (MNL) and its many

different variants use route utilities to determine route choice probabilities, while link-based models such as Recursive Logit (Fosgerau, 2013) and the Perturbed Utility Route Choice (PURC) model (Fosgerau et al., 2022) operate at a link level using link utilities and the network structure to determine link usage. The latter are attractive as they avoid the need to generate route choice sets, which is not straightforward, and the latest development of the PURC model allows for irrelevant parts of the network to be unused. As a route choice model, however, link-based models are hampered by the requirement that route attributes must be link-additive, meaning that route-related attributes such as some transit fare schemes, transfer penalties, travel time reliability measures, tolling schemes, and local detouredness (Rasmussen et al. 2024), cannot currently be accounted for. In this study, we focus on path-based models.

Generally, path-based route choice modelling has two components: i) generating a set of realistic routes that travellers choose between (the route choice set), and ii) applying a route choice probability model to these routes to determine their choice probabilities. Numerous route choice set generation methods have been developed (see Prato (2009), Bovy (2009), Rieser-Schüssler et al. (2013) for reviews) and numerous route choice probability models have been developed (see Prato (2009), Duncan et al. (2020,2022), Jing et al. (2018) for reviews). We highlight, however, two common issues.

The first issue is that choice set generation and route choice probability computation are often conducted in two independent, sequential steps, where the choice set formation and choice probability criteria are not consistent. Consequently, a route identified by the choice set formation criteria may be considered unrealistic by the choice probability criteria, and *vice versa*. This is supported by the work of Horowitz & Louviere (1995) who found in an empirical analysis that choice set formation and choice from the choice set are often driven by the same preferences, and thereby choice need not be modelled as a two-step process.

The second issue is that the attractiveness of a route, both in choice set generation and choice probability computation, is often judged solely by qualities of the complete route, which we term the *global* properties of the route. These global properties include total length, travel time, travel time (un)reliability, and direct monetary costs such as might be imposed through road pricing. However, it has recently been shown in Rasmussen et al. (2024) that it is insufficient to judge the attractiveness of a route solely by its global properties. It is contended that the *local* properties of a route should also be considered; namely, its *local detouredness*, which is the extent to which it detours on subsections of the route. Through an empirical and theoretical analysis, it was shown that it is important to consider local detouredness both when determining realistic and tractable route choice sets and when determining route choice probabilities. For example, analysis of observed route choice data showed that route usage tends to decay with local detouredness, and that there is an apparent limit on the amount of local detouredness seen as acceptable.

Several choice probability models have been developed in the following studies that aim to address the first issue: Swait (2001), Elrod et al. (2004), Gilbride & Allenby (2004), Martínez et al. (2009), Paleti (2015), Watling et al. (2018), Duncan et al. (2022), and Tan et al. (2024).[1] In the application of route choice, one can generalise the overall approach as penalising the probability of a route if it has an attribute value or utility beyond a 'cutoff' value (bound/threshold), where it is then deemed to be unrealistic/unused (Cascetta & Papola, 2001). Thus, only routes with all attributes / utilities within the cutoff(s) are considered realistic/used, and the probabilities of these routes relates in some way to the likelihood they are used, thereby ensuring consistency between choice set formation and choice probability criteria. However, the models in Swait (2001), Martínez et al. (2009), and Paleti (2015) impose 'soft' cutoffs rather than 'hard' cutoffs, meaning that routes with attributes/utilities beyond the cutoff only receive reduced probabilities and not zero probabilities. The model in Elrod et al. (2004) imposes absolute cutoffs, which is inflexible and may be difficult to specify suitably given different Origin-Destination (OD) movements have different attribute value ranges (e.g. lengths). And, the model in Gilbride & Allenby (2004) has a non-continuous choice probability function, resulting in 'abrupt changes' in the choice probabilities as alternatives enter and exit the used route choice set. The Bounded Choice Model (BCM) developed in Watling et al. (2018), and its adaptations/extensions (Duncan et al., 2022; Tan et al., 2024), on-the-other-hand: a) imposes a hard cutoff so that routes with utilities below the bound receive zero choice probability, b) can be stipulated so that the bound is relative to the route utilities (e.g. φ times worse than best utility), and c) has a continuous choice probability function, including as routes enter and exit the used route set (cross from below to above the bound, and *vice versa*).

To develop a model that both provides consistency between choice set formation and choice probability criteria and judges route attractiveness based on both its global and local properties (i.e. addresses both issues highlighted), Rasmussen et al. (2024) recently developed the BCM with Local Detour Threshold (BCM-LDT) model. The BCM-LDT model is derived by combining the BCM with a conjunctive choice model (Jedidi & Kohli, 2005; Gilbride & Allenby, 2004; Swait, 2001; Shin & Ferguson, 2017; Kohli & Jedidi, 2005), so that separate bounds are imposed on different route 'aspects' (Tversky, 1972), which in this case are total route travel cost and local detouredness. A route thus receives a zero choice probability (excluded implicitly from the choice set) if it has a cost above the global bound or a local detouredness above the local bound. Moreover, the probability of a route with both a cost and detouredness below the respective bounds, is determined according to its relative attractiveness in cost and detouredness compared to those of others.

[1] We note that an interesting relevant link-based route choice model is the perturbed utility model (Fosgerau et al. (2022)), which assigns links zero choice probabilities, thereby implicitly determining the implied set of used routes.

1.2 Paper contributions

The paper is based on an extension of the BCM-LDT model to account for correlations between overlapping used routes. Numerous models have been proposed for capturing correlations between overlapping routes. Detailed reviews of these models can be found in Duncan et al. (2020,2022), where we review their strengths and weaknesses. Among them, correction term models (Ben-Akiva & Ramming, 1998; Cascetta et al., 1996; Duncan et al., 2020) are particularly attractive due to their simple closed-form choice probability functions that are quick and easy to compute and estimate. In Duncan et al. (2022) we combined the Path Size Logit (PSL) correction-term model with the BCM to formulate a Bounded Path Size (BPS) route choice model. Using analogous methodology, we first in the current paper combine concepts from the BPS model with the BCM-LDT to formulate a Bounded Path Size Local Detour Threshold (BPS-LDT) model. To ensure that correlations are captured between only routes defined as realistic by both the local and global bounds, and continuity of the choice probability function is maintained, careful consideration is given to the formulation of the path size term. This is therefore a new model variant that we have created, though as it can be quite straightforwardly deduced from our previous work, we do not make a claim that this is a major contribution.

Having formulated a BPS-LDT route choice model that accounts for both local detouredness and route correlation, the three major contributions of the paper are as follows.

Contribution 1: Developing a solution approach for efficiently applying the model to large-scale networks

The computational burden of current methods for generating all possible routes below the local and global bounds – such as the initial approach proposed in Rasmussen et al. (2024) – means it is not yet feasible to fully apply the BCM-LDT/BPS-LDT models to large-scale networks. The first major contribution of the current paper is thus to develop a solution approach for application to large-scale networks. The heuristic approach involves working from representative universal choice sets. The approach is to pre-generate large enough working choice sets so that one can be fairly certain enough realistic alternatives are present, regardless of how many unrealistic routes are generated, and then apply the local detour model to exclude generated unrealistic routes. Although with this approach there may be some routes not generated that may be considered realistic by the choice probability criteria, at least unrealistic routes should be dealt with.

Given the pre-generated representative universal choice sets, all necessary information regarding route segments and segment alternative choice sets required to calculate local detouredness is pre-processed. However, if not done intelligently, this still has the potential to be computationally demanding, as there are for example many possible segments a route can detour. In the paper we propose some tricks to pre-process useful information that can be used to improve the efficiency of computing local detour model choice probabilities, which are shown to considerably improve computation times.

Contribution 2: Developing and testing a procedure for estimating the model

The second major contribution of the paper is to develop and test a Maximum Likelihood Estimation (MLE) procedure for estimating the BPS-LDT model with tracked route observations. The focus of Rasmussen et al. (2024) was on the derivation of the BCM-LDT model and its application to traffic user equilibrium; model estimation has yet to be explored. Estimating the BCM-LDT/BPS-LDT models with MLE is, however, not straightforward, as: a) the Likelihood function can be zero meaning that the Log-Likelihood function has constraints, and b) the Likelihood function is non-differentiable, meaning that gradient approaches cannot be directly adopted to solve the objective function, and standard errors of parameter estimates cannot be calculated analytically. Addressing these complications, we: a) develop a simple-to-implement method for estimating the BPS-LDT model working round the constraints on the Log-Likelihood objective function, and b) approximate standard errors of estimates using a computationally-efficient resampling method.

Contribution 3: An empirical analysis of the model

The third major contribution of the paper is conducting an empirical analysis of the model with real route choice data, to explore whether:

i) Local detouredness is an influential factor upon route choice probability / choice set formation,

ii) Local detouredness and correlation effects are identifiable as distinct route choice attributes,

iii) Parameter estimates are statistically significant,

iv) MLE solutions are unique,

v) The model can be estimated in feasible computation times on a large-scale network,

vi) The model is suitable for forecasting,

vii) Model estimates are sensitive to the assumed representative universal choice set.

To do this, we estimate the BPS-LDT model using the proposed MLE procedure in a real-life large-scale GPS case study, and analyse the results. For ii), we conduct a simulation study to assess whether model parameters can be successfully reproduced and identified. For iv), we visualise the Log-Likelihood surface to search for multiple maxima, and test for multiple MLE solutions numerically. And, for vi) we conduct out-of-sample validation.

The paper is structured as follows. Section 2 introduces general network notation, and Section 3 formulates the BPS-LDT model. The three major contributions of the study are then addressed in the three following sections. Section 4 introduces the proposed solution method for computing local detour model choice probabilities, Section 5 introduces the proposed estimation procedure, and Section 6 conducts the empirical analysis of the model estimation results. Finally, Section 7 summarises the paper conclusions and provides thoughts on future research.

2 General network notation

The model developed in this paper is applicable to general networks with multiple OD movements and flow-dependent link costs. However, without compromising the model derivation, we simplify notation by considering a single OD movement with fixed link costs. The network consists of link set A and node set B. For the OD movement, R is the choice set of routes, having size $N = |R|$. This could be the universal choice set of all routes, or a representative universal choice set of routes. $A_i \subseteq A$ is the set of links in route $i \in R$, and $\delta_{a,i} = \begin{cases} 1 & if\ a \in A_i \\ 0 & otherwise \end{cases}$. $B_i \subseteq B$ is the set of nodes belonging to route $i \in R$. Suppose that the generalised travel cost t_a of each link $a \in A$ is a weighted sum (by parameter vector $\boldsymbol{\alpha}$) of variables $\boldsymbol{w}_a$, i.e. $t_a = t_a(\boldsymbol{w}_a; \boldsymbol{\alpha})$, and that the generalised travel cost for route $i \in R$, c_i, can be attained through summing up the total cost of its links so that $c_i\big(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha})\big) = \sum_{a \in A_i} t_a(\boldsymbol{w}_a; \boldsymbol{\alpha})$, where $\boldsymbol{t}$ is the vector of all link travel costs and $\boldsymbol{w}$ is the vector of all link variables.[2] To simplify notation $c_i\big(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha})\big)$ is denoted just as c_i.

3 Bounded Path Size Local Detour Threshold model

In this section we combine the BCM-LDT model developed in Rasmussen et al. (2024) with the BPS model developed in Duncan et al. (2022), to formulate a Bounded Path Size Local Detour Threshold (BPS-LDT) model. For details on how the BCM-LDT model is derived and formulated we direct the reader to Sections 3&4 of Rasmussen et al. (2024), and for details on how the BPS model is derived and formulated we direct the reader to Section 3.2 of Duncan et al. (2022). In order to avoid repeating content, we shall henceforth assume readers are familiar with both models.

Measure of local detouredness

Rasmussen et al. (2024) proposed a new attribute influencing route choice: local detouredness, which in essence is the extent to which a route detours on its sub-routes. The measure of local detouredness is formulated as follows. Define a *set of segments* S_i for route $i \in R$ as the set of ordered node pairings:

$$S_i = \{(u, v): u \in B_i, v \in B_i \text{ and node } u \text{ precedes node } v \text{ when traversing route } i \in R\}.$$

Note that S_i includes not just ordered node pairings of adjacent nodes, but all ordered node pairs in the route. The *universal set of segments* is defined as $S = \bigcup_{i \in R} S_i$.

Define the set $K_{u,v}$ of *segment alternatives* for segment $(u, v) \in S$ as an index set of simple sub-routes from node u to node v. Furthermore, define the *used segment alternative* for segment $(u, v) \in S_i$ of route $i \in R$ as the element $k_{u,v,i} \in K_{u,v}$, denoting the index of the segment alternative actually used by route $i \in R$ from node u to node v.

The *measure of local detouredness* ϕ_i of route $i \in R$ is defined as:

$$\phi_i = \max\left\{\frac{\omega_{k_{u,v,i}} - \min\big(\omega_l: l \in K_{u,v}\big)}{\min\big(\omega_l: l \in K_{u,v}\big)}: (u, v) \in S_i\right\}, \tag{1}$$

where ω_l is the travel cost on segment alternative $l \in K_{u,v}$ for segment $(u, v) \in S_i$ of route $i \in R$. The measure of local detouredness identifies the maximum relative detour from each of its segments, by comparing, for each of the route's segments, the used segment alternative of the route against the minimum costing segment alternative for the segment. Note that the measure of local detouredness is a relative measure, i.e. a detour measure of $\phi_i = y$ corresponds to the worst detouring segment of the route being $(100 \times y)\%$ or $y + 1$ times greater than the cost of the minimum segment alternative cost for that segment. A demonstration of the local detour measure can be found in Section 4.2 of Rasmussen et al. (2024) as well as in Appendix C.1 of the current paper.

BPS-LDT model

[2] Note that it is possible to include route-based attributes that are not link-additive within the route choice models in this paper, by including these within the total route travel cost component.

The idea of the BCM-LDT model is that two separate bounds are imposed upon total route travel cost and local detouredness. For the global cost bound, a route receives zero choice probability if it has a total cost as great as or greater than φ times the minimum cost route, i.e. if $c_i \geq \varphi \min(\boldsymbol{c})$. For the local detour bound, a route receives zero choice probability if, at its most detouring segment, the used segment alternative has a travel cost as great as or greater than $\eta + 1$ times the minimum costing segment alternative for that segment, i.e. if it has a local detour measure $\phi_i \geq \eta$. Let $\bar{R}(\boldsymbol{c}, \boldsymbol{\phi}; \varphi, \eta) \subseteq R$ therefore be the restricted choice set of all routes $i \in R$ where both $c_i < \varphi \min(\boldsymbol{c})$ and $\phi_i < \eta$, i.e. the set of all used routes.

The BPS-LDT model is derived by integrating path size correction factors directly within the BCM-LDT probability relation (see equation (17) in Rasmussen et al. (2024)).[3] This is a common approach, adopted for example by Kitthamkesorn & Chen (2013), Xu et al. (2015), and Duncan et al. (2022) to formulate Path Size Weibit, Path Size Hybrid, and Bounded Path Size route choice models, respectively, where the correction factor is multiplied to the choice probability kernel for each route (i.e. the numerator of the choice probability function).

Given $\bar{R}$, the choice probability relation for route $i \in R$ is:

$$P_i = \begin{cases} \dfrac{(\bar{\gamma}_i)^{\beta}\left(\exp\left(-\theta_1(c_i - \varphi \min(\boldsymbol{c}))\right) - 1\right)\left(\exp\left(-\theta_2(\phi_i - \eta)\right) - 1\right)}{\sum_{j \in \bar{R}} (\bar{\gamma}_j)^{\beta}\left(\exp\left(-\theta_1(c_j - \varphi \min(\boldsymbol{c}))\right) - 1\right)\left(\exp\left(-\theta_2(\phi_j - \eta)\right) - 1\right)} & if\ i \in \bar{R} \\ 0 & if\ i \notin \bar{R} \end{cases}, \tag{2}$$

where $(\bar{\gamma}_i)^{\beta}$ is the path size correction factor for used route $i \in \bar{R}$. $\bar{\gamma}_i \in (0,1]$ is the path size term for used route $i \in \bar{R}$ measuring its distinctiveness (noting that unused routes do not have path size terms), and $\beta \geq 0$ is the path size scaling parameter measuring sensitivity to distinctiveness. A completely distinct route not overlapping at all with any other used route has a path size term equal to 1, resulting in no penalisation. Less distinct routes overlapping with other used routes have smaller path size terms and incur greater penalisation. The BPS-LDT path size term for the BPS-LDT model is as follows for used route $i \in \bar{R}$ is:

$$\bar{\gamma}_i = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{k \in \bar{R}} \left(\dfrac{\left(\exp\left(-\theta_1(c_k - \varphi \min(\boldsymbol{c}))\right) - 1\right)\left(\exp\left(-\theta_2(\phi_k - \eta)\right) - 1\right)}{\left(\exp\left(-\theta_1(c_i - \varphi \min(\boldsymbol{c}))\right) - 1\right)\left(\exp\left(-\theta_2(\phi_i - \eta)\right) - 1\right)} \right) \delta_{a,k}}. \tag{3}$$

To dissect the BPS-LDT path size term: each link a in route i is penalised (in terms of decreasing the path size term and hence the probability of the route) according to the number of <u>used</u> routes in the choice set that also use that link $\left(\sum_{k \in \bar{R}} \delta_{a,k}\right)$, where each contribution of a link-sharing route is weighted (i.e. $\sum_{k \in \bar{R}} \left(\frac{W_k}{W_i}\right) \delta_{a,k}$), and the significance of the penalisation for link a relates to how prominent link a is in route i, i.e. the cost of route a in relation to the total cost of route i $\left(\frac{t_a}{c_i}\right)$. The BPS-LDT path size contribution factor is:

$$\frac{W_k}{W_i} = \frac{\left(\exp\left(-\theta_1(c_k - \varphi \min(\boldsymbol{c}))\right) - 1\right)\left(\exp\left(-\theta_2(\phi_k - \eta)\right) - 1\right)}{\left(\exp\left(-\theta_1(c_i - \varphi \min(\boldsymbol{c}))\right) - 1\right)\left(\exp\left(-\theta_2(\phi_i - \eta)\right) - 1\right)}.$$

One can see in Appendix A how this path size contribution factor relates to those from other path size choice models. The BPS-LDT path size term function in (3) is formulated as such so that unused routes with costs/detourednesses above the bound/threshold do not contribute to reducing the path size terms of used routes with costs/detourednesses below the bound/threshold. And, so that the path size term function is continuous.

A key formatting difference between the BCM-LDT choice probability function in equation (17) in Rasmussen et al. (2024) and the BPS-LDT probability function in equation (2) here, is the use of the restricted choice set $\bar{R}$ to determine the zero probability routes (rather than $(.)_+$ functions). The reason for this, as proposed for the BPS model in Duncan et al. (2022) (see equations (6)-(7) in that paper), is so that the path size term can succinctly sum over the used routes $k \in \bar{R}$ to avoid cases of $\frac{0}{0}$.

There are five standard BPS-LDT model parameters, which are as follows: $\theta_1 > 0$ is the travel cost scaling parameter, $\theta_2 > 0$ is the local detouredness scaling parameter, $\beta \geq 0$ is the path size scaling parameter, $\varphi > 1$ is the relative surplus cost bound parameter, and $\eta > 0$ is the local detour threshold parameter. In Appendix C.3 we discuss and demonstrate the features / behavioural interpretations of these parameters.

Fig. 1 illustrates how the BPS-LDT model can collapse into numerous different route choice models in the literature. By setting $\beta = 0$, the BPS-LDT model collapses to the BCM-LDT model, which as shown in Rasmussen et al. (2024)

[3] Note that the BPS-LDT model can also be derived in a more theoretically rigorous manner by combining the conjunctive choice model (see Section 3.1 in Rasmussen et al. (2024)) with the conjunctive bounded choice model (see Section 3.3 in Rasmussen et al. (2024)), where distinctiveness is a conjunctive component and cost and detouredness are conjunctive bounded components.

can collapse into the BCM and MNL model. In the same way that the BCM-LDT model approaches the BCM as $\theta_2 \to 0$ and $\eta \to \infty$ under the condition that η tends to ∞ faster than θ_2 tends to zero, such that $\lim_{\substack{\theta_2 \to 0 \\ \eta \to \infty}} \theta_2 \eta = \infty$ (see Section 4.3 of Rasmussen et al. (2024)), the BPS-LDT model approaches the Bounded Path Size (BPS) model (see Appendix A.5) under the same conditions. Moreover, since the BPS model approaches the alternative Generalised Path Size Logit (GPSL′) model (see Appendix A.3) as $\varphi \to \infty$, the BPS-LDT model can also collapse into GPSL′. In Appendix C.2 we demonstrate the benefits of the BPS-LDT model compared the BCM, BPS model, and BCM-LDT model.

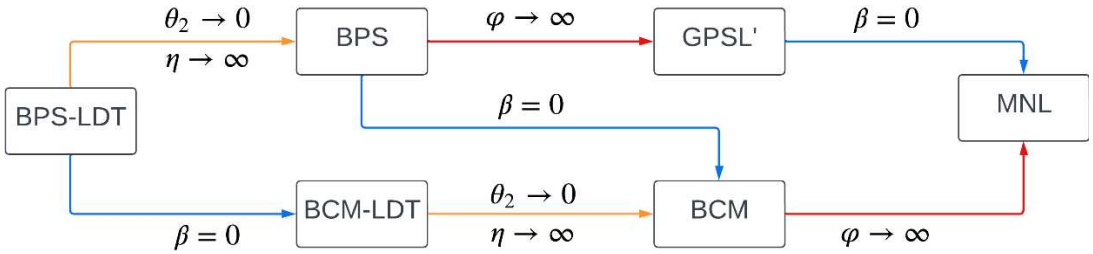

Fig. 1. Schematic diagram of how local detour threshold models, bounded models, and path size logit models collapse into one another.

4 Solution method

As demonstrated in Rasmussen et al. (2024), despite imposing both a bound upon total route travel cost and a threshold upon local detouredness, the number of all possible routes from the full network that will be assigned non-zero probabilities by these criteria may still be considerably large, especially for large-scale networks. Therefore, with existing methods for generating all routes below the cost bound and detour threshold to solve local detour threshold models – such the initial branch-and-bound-based algorithm proposed in Rasmussen et al. (2024) – applying local detour models to large-scale networks is not yet feasible computationally. In the current paper, we therefore adopt a heuristic approach, by working from a set of pre-generated routes that represent the universal set of possible routes. From these routes, a smaller set of routes will be assigned non-zero probabilities by the choice model that satisfy the cost-bound/detour-threshold conditions. This also allows for us in Section 6 to directly and consistently compare estimation results between related local detour and non-local detour models.

In this section, we propose a solution method for computing BPS-LDT model choice probabilities with pre-generated representative universal choice sets. This involves first pre-processing the route / network data to obtain the necessary segment information, and then using this information to compute BPS-LDT probabilities without requiring any further knowledge of the network.

Now, computing local detour measures is not straightforward/quick, even when only doing it for a pre-generated representative universal choice set of routes, it still has the potential to be computationally demanding. Computing the local detour measure of a route entails finding the segment of the route that has the greatest relative detour from the best segment alternative. Since there are many segments of a route, it can be computationally demanding to iterate through every segment of the route and calculate its relative detour at each segment. We thus propose here two main tricks for improving the efficiency of computing local detour measures and calculating choice probabilities.

The first trick is to identify during the pre-processing of the route-segment information of the generated routes, which segments are redundant. With some logical thinking regarding which segments will always lead to a greater relative detour than others in the route, one can identify which segments are redundant and exclude them, retaining only *essential* segments. This avoids storing and then operating with a lot of redundant information, thus improving efficiency. In Section 4.1 below we demonstrate essential/redundant segments.

The second trick is to harness the feature that if a segment violates the local detour threshold, all routes using that segment will receive zero choice probability. Thus, when it is found that a segment violates the threshold when checking the segments of one route, one can simultaneously assign all routes using that segment a zero choice probability, and one does not need to then iterate through those routes to test for threshold violation. This reduces the number of routes that need to be iterated through to compute local detour measures, thus improving efficiency.

In Section 4.1 and Section 4.2 below we introduce and discuss the proposed pre-processing method and solution method, respectively. Then, in Section 4.3 we assess the efficiency of our proposed approach, comparing it with more rudimentary methods that do not utilise the tricks of filtering out redundant segments or simultaneous route removal.

4.1 Pre-processing method

Algorithm 1 presents the proposed method for pre-processing segment information for solving the BPS-LDT model. The general idea of the method is to in Steps 1-4 identify the segments of each generated route, and then in Step 5 construct segment alternative choice sets for each segment from the routes. The segment alternatives are identified from the sub-

routes taken by the generated routes between the segment nodes. Information is stored regarding the links of each segment alternative so that (during the solution method) the costs of all segment alternatives can be computed, and thereby the best segment alternatives can be identified and local detour measures computed.

Now, operationalising the first trick discussed above, Step 3 in Algorithm 1 filters out 'redundant' segments and identifies just the 'essential' segments in each route. Moreover, operationalising the second trick, Step 5.4 stores information on which routes use each identified essential segment. The solution method in the following section uses this information to remove multiple routes at once when a segment violates the local detour threshold.

Step 1. Generate a new unique route from the network (or select a route from the pre-generated set of routes).

Step 2. Store the link-route information (e.g. add the route to a link-route incidence matrix). If first unique route generated, return to Step 1.

Step 3. Compare the new generated route with each previously generated route to filter out 'redundant' segments and identify just the 'essential' segments in each route. This can be done, for example, as follows for two generated routes:

> **Step 3.1.** Identify the common segments (ordered pairings of nodes) between the two routes.
>
> **Step 3.2.** For each common segment, check to see if the used segment alternative from each route (i.e. the links used by the route between the initial and end segment nodes) share any links. If no links are shared, add this segment (if not already present) to the lists for both routes of the essential segments.

Step 4. If the universal choice set of routes has been generated, or an approximation of the universal choice set has been obtained, continue to Step 5. Otherwise, return to 1.

Step 5. By iterating through each route and its associated essential segments in turn, construct segment alternative choice sets and obtain other relevant information for computing local detour measures. This can be done, for example, as follows. For each of the generated routes, iterate through each essential segment in turn and conduct the following:

> **Step 5.1.** Identify the used segment alternative of the current route, i.e. the links used by the route between the initial and end segment nodes.
>
> **Step 5.2.** If a new segment alternative for that segment has been identified, add the segment alternative to the segment alternative choice set for that segment, and store the link-segment alternative information (e.g. add the segment alternative to a link-segment alternative incidence matrix).
>
> **Step 5.3.** Note which segment alternative in the segment alternative choice set is used by this route during this segment.
>
> **Step 5.4.** Add the route to a list of routes that use that segment alternative.

Algorithm 1. Proposed method for pre-processing segment information for solving the BPS-LDT model.

To demonstrate what is meant by redundant/essential segments and how these can be identified, consider Fig. 2 which displays the step-by-step generation and adjoining of three routes from an underlying network where the OD movement is from node 1 to node 6. First, route 1 is generated which has node sequence: 1→2→3→4→5→6, then route 2 is generated with node sequence: 1→2→7→3→4→8→5→6, then route 3: 1→9→4→5→6.

Upon the generation of route 2, step 3 in the pre-processing method above compares the segments / used segment alternatives of routes 1&2: redundant segments of both routes are filtered out to identify just the essential segments that each route can eliminate the other with (in terms of a local detour at that segment above the threshold). The first obvious filtering of redundant segments is to discard all segments that are not shared by routes 1&2: (2,7), (7,3), (4,8), & (8,5). The common segments between the two routes are thus: (1,2), (1,3), (1,4), (1,5), (1,6), (2,3), (2,4), (2,5), (2,6), (3,4), (3,5), (3,6), (4,5), (4,6), & (5,6). The next obvious redundant segments are those where the used segment alternative is the same for both routes, which are segments: (1,2), (3,4), & (5,6).

This leaves just the segments where routes 1&2 have different used segment alternatives, and therefore have non-zero local detourednesses. However, if for a given segment all of the segment alternatives share a common link, then the segment(s) consisting of the non-overlapping part(s) will always have a greater detour measure than the full segment. For example, supposing that each link in Fig. 2 costs 1, then the route 2 detour measure at segment (1,3) is $\frac{3-2}{2} = 0.5$ (1→2→7→3 against 1→2→3), but 1→2→7→3 and 1→2→3 share link 1→2 and therefore the segment (2,3) will always have a greater detour measure, which in this case is $\frac{2-1}{1} = 1$ (2→7→3 against 2→3). This means that all segments where the segment alternatives overlap are redundant, which in the case of comparing routes 1&2 means that (1,3), (1,4), (1,5), (1,6), (2,4), (2,5), (2,6), (3,5), (3,6), & (4,6) are also redundant segments. This just leaves (2,3) and (4,5) as the essential segments which will dominate over all the others above in the local detour measure. (2,3) and (4,5) are thus added to the lists for routes 1&2 of essential segments to test.

Now, upon the generation of route 3, route 3 must be compared to both routes 1&2 to search for redundant/essential segments. Compared to route 1, the only segment in common where there are no shared links between segment alternatives is (1,4), and so (1,4) is added to the essential segments list of routes 1&3. Compared to route 2, the segments in common where there are no shared links are (1,4) and (4,5). These are thus both added to the essential segments lists for route 2&3, if not there already.

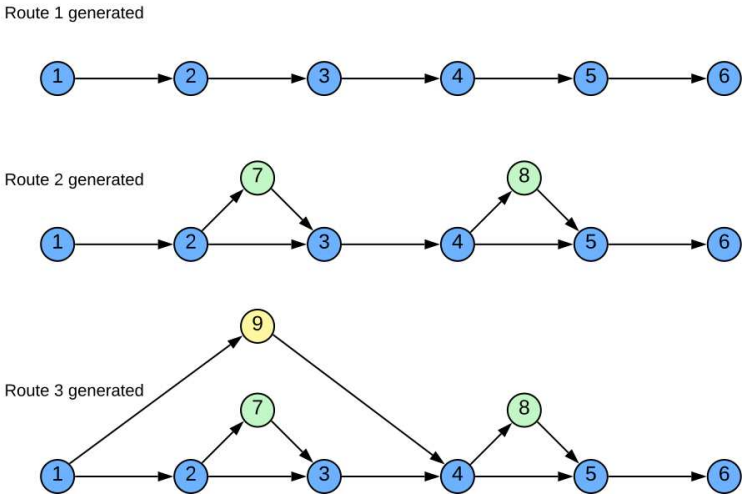

Fig. 2. Demonstrating redundant/essential segments.

4.2 Solution method

For a given setting of the link costs and given values of the relative cost bound parameter φ and local detour threshold parameter η, Algorithm 2 presents pseudo-code for solving the BPS-LDT model with representative universal choice sets and pre-processed segment information. In Step 1, the relative surplus total travel cost of each generated route is checked against the global cost bound; violating routes are then removed from the used route choice set. In Step 2, the remaining routes are iterated through, calculating the local detourednesses of the segment alternatives for each essential segment identified during the pre-processing. Step 2.2.3 operationalises the second trick discussed above: for segment alternatives that violate the local detouredness threshold, all routes that use that segment alternative are removed from the used route choice set (potentially removing multiple routes at once). If none of the segment alternatives of a route violate the detour threshold, then the local detour measure is calculated and it remains in the used route choice set. In Step 3, the BPS-LDT choice probabilities are computed for the used routes given the local detour measures calculated in Step 2.

Step 1: *Cost bound route elimination phase*. Compute the route costs then check the cost of each route against the cost bound. Remove all violating routes from the used route choice set.

Step 2: *Local detour threshold route elimination phase*.

Step 2.1: Compile an ordered list of the routes remaining in the used route choice set.

Step 2.2: For the next route in the ordered list, go through each (essential) segment of the route in turn:

Step 2.2.1: If the segment has been analysed by a previous route, note the relative surplus cost of the segment alternative used by the current route (already computed, see below), and continue to next segment.

Step 2.2.2: Otherwise, compute the segment alternative costs and then the relative surplus cost of each segment alternative compared to the lowest costing segment alternative for that segment.

Step 2.2.3: For each segment alternative that violates the detouredness bound, remove all the routes that use that segment alternative from the ordered list and the used route choice set.

Step 2.2.4: If during this segment, the current route violates the detouredness bound, move on to the next route in the ordered list and return to Step 2.2. If this segment is the last segment of the current route, continue to Step 2.3, otherwise move on to the next segment of the route and return to Step 2.2.1.

Step 2.3: Given no segments of the current route violate the detouredness bound, compute the measure of local detouredness for the route given the computed relative surplus costs of its segment alternatives during Step 2.2. If at the end of the ordered list of routes, continue to Step 3. Otherwise, move on to the next route in the ordered list and return to Step 2.2

Step 3: *Choice probability computation*. Given the remaining choice set of used routes, and the costs and measures of detouredness of the used routes computed in Step 1 and Step 2, respectively, compute the BPS-LDT choice probabilities.

Algorithm 2. Solution method for computing BPS-LDT model choice probabilities given representative universal route choice sets and pre-processed segment information.

4.3 Demonstration

Here we shall demonstrate the efficacy of the proposed solution method compared to other more rudimentary methods one could adopt. As discussed, there are two main tricks that we employ to speed up computation times. The first involves filtering out redundant segments during the pre-processing stage, which dramatically reduces the number of segments being considered when computing the local detour measures. The second involves removing multiple routes at once from the used route choice set when checking if a segment alternative violates the local detour threshold (and thus not needing to compute detour measures for every route). The combination of these two tricks significantly improves BPS-LDT choice probability computation times.

To demonstrate, we shall compute the BPS-LDT choice probabilities using four different methods:

- Method 1: During the pre-processing stage, exhaustively identify all segments of a route, and then during the solution stage consider each route in turn, calculating the local detouredness of every segment of the route and thus identifying the maximum for the local detour measure.
- Method 2: During the pre-processing stage, filter out all redundant segments of a route, and then during the solution stage consider each route in turn, calculating the local detouredness of each essential segment of the route and thus identifying the maximum for the local detour measure.
- Method 3: During the pre-processing stage, exhaustively identify all segments of a route, and then during the solution stage, if a segment alternative violates the local detour threshold, remove all routes that use that segment alternative from the used route set.
- Method 4 (our proposed method): During the pre-processing stage, filter out all redundant segments of a route, and then during the solution stage, if a segment alternative violates the local detour threshold, remove all routes that use that segment alternative from the used route set.

Table 1 displays the computation times for each method when computing the BPS-LDT choice probabilities for a single OD movement of the real-life case study in Section 6.2. The OD movement has 100 routes in the representative universal choice set, and the BPS-LDT model parameters are those calibrated in Table 4 but with the local detour threshold parameter set to $\eta = 0.5$, i.e. where 88% of the routes are cut out by the local detour threshold. As can be seen, Method 1 is very slow, having to consider each route in turn and every segment of the route. Either filtering out redundant segments (Method 2) or removing multiple routes at once (Method 3) reduces computation times considerably, and a combination of the two, our proposed approach, results in much faster computation times. It is also worth noting that filtering out redundant segments dramatically reduces the memory required to store the pre-processed segment information. In this example, the unfiltered segment information required 120 times more memory.

Fig. 3 displays how the BPS-LDT choice probability computation time using our proposed method varies as the local detour threshold parameter η varies. As can be seen, the lower the threshold, the greater the number of violating routes, and thus the greater the number of routes that are being cut out simultaneously, speeding up computation times.

Method	Method 1	Method 2	Method 3	Method 4
Computation Time [s]	22.19	0.384	3.74	0.022

Table 1. BPS-LDT model choice probability computation times using our proposed method (Method 4), and three more rudimentary methods (Methods 1-3).

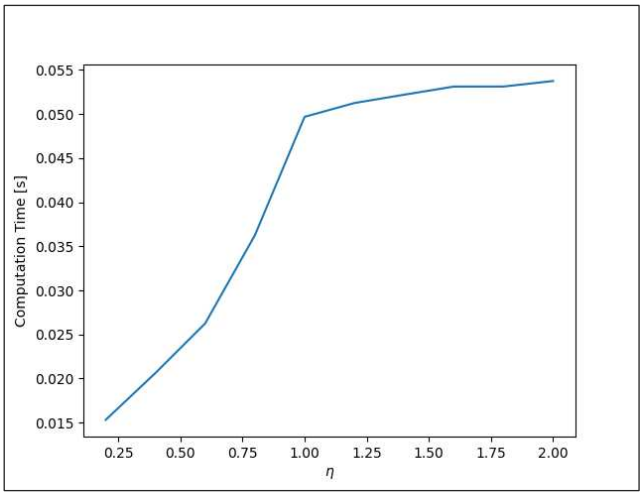

Fig. 3. BPS-LDT model choice probability computation times using our proposed method, for different settings of the local detour threshold parameter η.

5 Estimation method

In this section, we provide a modified Maximum Likelihood Estimation (MLE) procedure for estimating the BPS-LDT model with tracked route observations, as well as a computationally tractable method for evaluating estimate precision.

5.1 Likelihood formulation

Suppose that we have available a set of observed routes Z, e.g. collected through GPS units or smart phones, and consider a situation where it is not needed to distinguish individuals in their preferences (the approach is, of course, readily generalised to permit multiple user classes differing in their parameters). Let R_z be the choice set of route alternatives for observation $z \in Z$. Suppose that the observation data is contained in a vector $\boldsymbol{x}$ of size $|Z|$ where:

$$x_z = i \quad \text{if alternative } i \in R_z \text{ is chosen,} \qquad \forall z \in Z.$$

The BPS-LDT model Likelihood, L, for a sample of size $|Z|$ is:

$$L(\boldsymbol{\alpha}, \theta_1, \theta_2, \beta, \varphi, \eta | \boldsymbol{x}) = \prod_{z \in Z} P_{x_z}(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha}); \theta_1, \theta_2, \beta, \varphi, \eta), \tag{4}$$

where $P_{x_z}(\boldsymbol{t})$ is the BPS-LDT model choice probability function given by (4) for route $x_z \in R_z$.

If for a given setting of the travel cost, cost bound, and local detour threshold parameters $\tilde{\boldsymbol{\alpha}}$, $\tilde{\varphi}$, & $\tilde{\eta}$, there is an observation z such that either $c_{x_z}\big(\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}})\big) \geq \tilde{\varphi} \min\Big(\boldsymbol{c}\big(\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}})\big)\Big)$ or $\phi_{x_z}\big(\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}})\big) \geq \tilde{\eta}$, the BPS-LDT Likelihood value is zero. This means that the maximum likelihood estimates $\big(\hat{\boldsymbol{\alpha}}, \hat{\theta}_1, \hat{\theta}_2, \hat{\beta}, \hat{\varphi}, \hat{\eta}\big)$ will always be such that $c_{x_z}\big(\boldsymbol{t}(\boldsymbol{w}; \hat{\boldsymbol{\alpha}})\big) < \hat{\varphi} \min\Big(\boldsymbol{c}\big(\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}})\big)\Big)$ and $\phi_{x_z}\big(\boldsymbol{t}(\boldsymbol{w}; \tilde{\boldsymbol{\alpha}})\big) < \tilde{\eta}$ for all $z \in Z$ (see Duncan et al. (2022)).

The BPS-LDT model Log-Likelihood function, LL, to be maximised is thus:

$$LL(\boldsymbol{\alpha}, \theta_1, \theta_2, \beta, \varphi, \eta | \boldsymbol{x}) = \ln\left(\prod_{z \in Z} P_{x_z}(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha}); \theta_1, \theta_2, \beta, \varphi, \eta)\right) = \sum_{z \in Z} \ln\Big(P_{x_z}(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha}); \theta_1, \theta_2, \beta, \varphi, \eta)\Big), \tag{5}$$

$$subject\ to \quad \begin{cases} c_{x_z}\big(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha})\big) < \varphi \min\Big(\boldsymbol{c}\big(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha})\big)\Big) \\ \phi_{x_z}\big(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha})\big) < \eta \end{cases}, \qquad \forall z \in Z, \tag{6}$$

where $P_{x_z}(\boldsymbol{t})$ is the BPS-LDT model choice probability relation in (2) for route $x_z \in R_z$.

Solution Existence. In Duncan et al. (2022), it was shown that MLE solutions are guaranteed to exist for the BPS model. Moreover, that these MLE solutions are guaranteed to exist in the parameter space where the likelihood is not zero. In an analogous manner, MLE solutions are guaranteed to exist for the BPS-LDT model, in the parameter space where the likelihood is not zero. In order to avoid unnecessarily repeating what would be very similar material, we refer the reader to Section 7.2.2 of Duncan et al. (2022) for the arguments for these results.

Solution Uniqueness. The BPS-LDT Log-Likelihood function is not guaranteed to be globally convex. Standard proofs of solution uniqueness thus cannot be applied to prove that BPS-LDT model MLE solutions are unique. That is

not to say though that solutions are not / cannot be unique, and in the experiments in this paper we have not experienced any cases of multiple solutions. In Section 6.2.6 we conduct an empirical analysis of the uniqueness of MLE solutions.

5.2 Estimation procedure

The proposed procedure for estimating the BPS-LDT model is a modification of a standard MLE procedure. As can be seen in (5)-(6), maximising the BPS-LDT Log-Likelihood is complicated by the constraints requiring all chosen routes to have both a cost below the bound and detour measure below the threshold, otherwise the Log-Likelihood function is undefined. Like as discussed for the BPS model in Duncan et al. (2022), it is possible to predetermine the parameter space for MLE where the Log-Likelihood will always be defined, by identifying (for a given closed-bounded range for the cost parameters) the lower limits for the bound parameter φ and local detour threshold parameter η before it is possible for any chosen route to violate the cost bound / local detour threshold. Or, one could also incorporate corresponding constraints for the optimisation algorithm, like those in (6) but adjusted to include equivalence.

However, since identifying the parameter space / incorporating the corresponding constraints is far from straightforward, we adopt an easier to implement approach. For a closed-bounded parameter space Ω where for some settings of the parameters the Likelihood is zero, the MLE solution will always lie in the parameter subspace $\overline{\Omega}$ where all observed routes have travel costs within the bound / local detour measures within the threshold, i.e. where the Likelihood is non-zero. Thus, similar to as done for the BPS model in Duncan et al. (2022), the idea is simply to tell the Log-Likelihood maximisation algorithm to search for solutions within $\overline{\Omega}$ only, by setting nonoptimal values for the objective function when testing parameters not in $\overline{\Omega}$.

To do this, the estimation procedure includes a step that checks if any observed route violates the current cost-bound/detour-threshold, and if so, sets the Log-Likelihood value as an appropriately large and negative value. The proposed BPS-LDT estimation procedure is as follows:

Step 1: For each route observation $z \in Z$, generate a representative universal choice set and pre-process the link attributes, link-route information, and segment information (see Algorithm 1). Set an initial set of parameters to test.

Step 2: Given the current set of parameters to test, calculate the travel cost of every route and the local detour measure of each observed route.

Step 3: For each route observation, check if its cost violates the cost bound or its detour measure violates the local detour threshold. If so, set the Log-Likelihood value as an appropriate large and negative value (see below), choose a new set of parameters to test, and return to Step 2. Otherwise, continue to Step 4.

Step 4: Given the current set of parameters being tested, use the solution method detailed in Algorithm 2 to compute the BPS-LDT model choice probability of each route observation.

Step 5: Given these choice probabilities, compute the Log-Likelihood value.

Step 6: If Log-Likelihood converged, stop. Otherwise, choose a new set of parameters to test and return to Step 2.

Initial conditions. Although it is not necessarily a requirement, our recommendation is that initial conditions are set such that none of the chosen route cost bounds or local detour thresholds are violated. This can be done simply by setting intuitively large bound/threshold values, or by calculating the lower limits of φ and η before any chosen route violates the bound/threshold. In Section 6.2.6 we test solving MLE with different initial conditions.

Cost-bound/detour-threshold violation check. Note that for the cost-bound/detour-threshold violation check, one only needs to compute the local detour measures for the route observations. Moreover, by performing the cost-bound violation check first, one may not need to compute any detour measures, which is the more computationally demanding part. For the experiments in this paper, supposing that Z' is the set of observations that violate either the bound or the threshold, we set the appropriate large and negative value as

$$LL = \sum_{z \in Z} \begin{cases} -999 & if\ z \in Z' \\ \ln\left(P_{x_z}^{MNL}\right) & otherwise' \end{cases}$$

where $P_{x_z}^{MNL}$ is the MNL choice probability for route $x_z \in R_z$. Setting the appropriate large and negative value in this way, rather than as some constant arbitrary number, means that some information can be gathered on the relevance of the parameters even when bound-violating parameters are tested. MNL choice probabilities are chosen as these are quick and easy to compute. Note that for the violation

Direction searching. In general, one can apply procedures from standard numerical optimisation methods to identify the parameters to evaluate at the next iteration. Since the BPS-LDT choice probability function is non-differentiable for parameter values that lead to either: i) a route's travel cost being equal to the bound value, ii) a route's detour measure being equal to the detour threshold value, or iii) two or more routes having the same minimum travel cost (i.e. a point at which the minimum cost route changes), caution must be given when using gradient approaches such as Newton-Raphson or BHHH for the minimisation algorithm. While the non-differentiability at i) and ii) are not an issue as

we are only concerned with parameter settings within the subspace $\overline{\Omega}$ (the Log-Likelihood value will be large and negative anyway at those parameter values), the non-differentiability at iii) may be problematic if the minimum cost route is likely to change a lot for different parameter values. One possibility could be to replace the min functions in (1)-(3) with smooth approximations of the minimum function, such as the Boltzmann operator or Mellowmax operator. In this study though, we use the L-BFGS-B bound-constraint quasi-Newton minimisation algorithm (Byrd et al., 1995) (where we minimise $-LL$), which approximates Log-Likelihood differentials using finite difference. We found that this converged well to MLE solutions, as we demonstrate in Section 6.2.6. The L-BFGS-B algorithm was implemented using the scipy.optimize.minimize package in Python. The parameter bounds and initial conditions are given in each study.

5.3 Evaluating estimate precision

For the BPS-LDT model it is not possible to calculate standard errors for the estimates analytically. This is because the model is non-differentiable at certain points and therefore violates the regularity conditions that establish asymptotic standard errors of the maximum likelihood estimates as the inverse of the Fisher information. Instead, we detail here a method for evaluating estimate precision numerically through resampling.

Resampling approaches that could be utilised include, among others, the JackKnife, Bootstrap, and Subsampling methods (Efron, 1979). For a sample size of $|Z|$ observations, the JackKnife method involves estimating the models for $|Z|$ subsamples of $|Z|-1$ observations, where each of the $|Z|$ observations are sequentially removed one at a time from the sample set (with replacement). The Bootstrap method involves estimating the models for H samples of $|Z|$ observations drawn randomly from the full sample set Z, with replacement. The Subsampling method involves estimating the models for H subsamples of $G < |Z|$ observations drawn randomly from the full sample set, with or without replacement. With the $|Z|$ estimates of the parameters obtained from the JackKnife method, and H estimates from the Bootstrap and Subsampling methods, estimate precision statistics can be approximated such as standard errors and confidence intervals, given mean estimates. Indeed, Lim Jr et al. (2009), Tilahun et al. (2007), Jánošíková et al. (2014), and Fosgerau et al. (2023) all utilise Bootstrap to assess estimate precision in route choice parameter estimation studies.

However, while estimating the BPS-LDT model once for a dataset of $|Z|$ observations is computationally feasible (see estimation times for our real-life case study in Section 6.2.3), re-estimating the model many times for different samples of $|Z|$ / $|Z|-1$ observations for the Boostrap/Jacknife methods can be computationally onerous. The Subsampling method has a lower computational burden, but it violates sample size properties. We thus instead suggest using the Bag of Little Bootstraps (BLB) method (Kleiner et al, 2014), which was proposed as a more computationally efficient alternative to Bootstrap. Although it is to the best of our knowledge yet to be utilised in the context of transport route choice, the BLB method has been used to explore parameter estimate precision in other contexts, see e.g. Allahviranloo et al. (2017), Sharma & Kuma (2021), Covington et al. (2021). The BLB resampling method is as follows for a full sample set size of Z observations:

Step 1: Randomly draw without replacement $G < |Z|$ observations from the full sample set.

Step 2: Repeat Step 1 H times so that there are H subsamples of G unique observations.

Step 3: From each of the H subsamples of G unique observations, randomly draw with replacement $|Z|$ observations.

Step 4: Repeat Step 3 Y times so that for each of the H subsamples of G unique observations, there are Y samples of $|Z|$ (not-all-unique) observations.

The result is $H \times Y$ samples of $|Z|$ observations, G of them unique in each sample. The models are estimated for each of the $H \times Y$ samples and estimate precision statistics are evaluated from the set of $H \times Y$ estimates.

The computational advantage of the BLB method compared to the Bootstrap method is that each of the $H \times Y$ samples only have G unique values, and therefore the computational demand scales in G instead of $|Z|$. In our case study, each observation is a unique OD movement with its own choice set of routes. Thus, for JackKnife, one must compute the route choice probabilities for $|Z|-1$ OD movements. For Bootstrap, the number of OD movements varies depending on how the observations are randomly drawn; the average according to some simple and standard calculations done by Efron & Tibshirani (1993) is approximately $0.632 \cdot |Z|$, which is large if $|Z|$ is large, and there could be as many as $|Z|$ in the unlikely scenario every observation is drawn without repetition. The BLB method, however, allows for one to control the number of OD movements, and thus the computational burden, with G. Moreover, as Kleiner et al. (2014) discuss/demonstrate, the BLB method is well suited for modern parallel and distributed computing architects (more so than Bootstrap), and one can also use an iterative algorithm to seek the minimal values of the hyperparameters H and Y that are sufficiently large to yield good statistical performance. Once more, the BLB method does not violate the sample size properties (each sample has $|Z|$ observations (that are not all unique)).

6 Empirical analysis

In this section, we first conduct a simulation study to assess whether assumed true parameters can be successfully reproduced and identified. Then, the BPS-LDT model is estimated on a large-scale network using real route choice observation data tracked by GPS units. Results are compared with BCM-LDT model as well as other relevant non-local detour route choice models (see Appendix A). For this estimation work, we work with pre-generated representative universal choice sets, utilising the pre-processing and solution method proposed in Section 4.

6.1 Simulation study

Here we investigate the procedure proposed in Section 5.2 for estimating the BPS-LDT model in a simulation study, assessing the possibility of estimating reasonable parameters that reproduces observed behaviour. This is particularly important in the case of the BPS-LDT model as: i) there is some uncertainty over whether the heuristic estimation procedure will work effectively, ii) it is not guaranteed that the BPS-LDT solutions are unique, and iii) there is a complex interaction between the parameters of the BPS-LDT model, with possibilities for identification issues. For example, θ_1 and φ both to some extent control sensitivity to travel cost, and θ_2 and η both to some extent control sensitivity to local detouredness. We can also explore whether local detouredness and correlation are distinct, independent route attributes.

6.1.1 Experiment setup

A similar approach is adopted to that utilised for the Adaptive Path Size Logit model in Duncan et al. (2020) and BPS model in Duncan et al. (2022). In general, the approach is to sample observations according to an assumed 'true' model, and then use these in combination with the Log-Likelihood function to evaluate the ability to reproduce the assumed 'true' parameters. The simulation estimation experiment consists of four steps:

Step 1: Postulate a set of true parameters of the BPS-LDT model.

Step 2: Given these assumed true parameters, calculate the BPS-LDT choice probability of each route alternative in the representative universal choice sets.

Step 3: Given these route choice probabilities, sample $|Z|$ observed route choices.

Step 4: Given these observed route choices, apply the estimation procedure discussed in Section 5.2 to obtain BPS-LDT parameter estimates.

Steps 3 & 4 are replicated many times to obtain a set of parameter estimates. Then, by analysing the Bias and Standard Error of these sets of parameter estimates we can assess whether the assumed true BPS-LDT parameters can be successfully retrieved.

6.1.2 Sioux Falls application

The Sioux Falls network consists of 76 links and 528 OD movements with non-zero travel demands. Details of the network were obtained from https://github.com/bstabler/TransportationNetworks. The travel cost of link a is specified as the free-flow travel time $w_{a,1}$ only, such that:

$$t_a(\boldsymbol{w}_a; \boldsymbol{\alpha}) = w_{a,1} \cdot \alpha_1,$$

where $\alpha_1 > 0$ is the free-flow travel time parameter, and thus the travel cost for route $i \in R_m$ is:

$$c_{m,i}\big(\boldsymbol{t}(\boldsymbol{w}; \boldsymbol{\alpha})\big) = \sum_{a \in A_{m,i}} t_a(\boldsymbol{w}_a; \boldsymbol{\alpha}) = \alpha_1 \sum_{a \in A_{m,i}} w_{a,1}.$$

The BPS-LDT models require in this case the specification of six parameters: α_1, θ_1, θ_2, β, φ, and η but to ensure identification θ_1 is fixed at $\theta_1 = 1$ throughout.

To generate the representative universal choice sets, we utilised a simulation approach (Sheffi & Powell, 1982) where the link costs were drawn randomly from a truncated normal distribution with mean value being free-flow travel time and standard deviation being 0.6 times the mean. The link costs were simulated 100 times for each OD movement and for each simulation a shortest path search was conducted to generate a route, so that a maximum of 100 unique routes were generated for each choice set.

For the Log-Likelihood maximisation algorithm (see Section 5.2) we set the initial conditions to $\big(\tilde{\alpha}_1^{(0)}, \tilde{\theta}_1^{(0)}, \tilde{\beta}^{(0)}, \tilde{\varphi}^{(0)}, \tilde{\eta}^{(0)}\big) = (\alpha_1^{true} + 0.2, \theta_1^{true} + 0.2, \beta_{true} + 0.2, \varphi_{true} + 0.2, \eta_{true} + 0.2)$, and the bounds to $\tilde{\alpha}_1 \in [0.01,2]$, $\tilde{\theta}_1 \in [0.01,8]$, $\tilde{\beta} \in [0,2]$, $\tilde{\varphi} \in [1.01,3]$, $\tilde{\eta} \in [0.01,4]$.

Table 2 reports, for various settings of the true parameters, the mean Bias (true parameter minus mean estimate), Standard Error (standard deviation of estimates), and Route Mean Squared Error $\left(RMSE = \sqrt{(Bias)^2 + (S.E)^2}\right)$, of the estimates across 250 experiment replications, each with $|Z| = 5000$ simulated observations. As shown, the mean bias of

the estimates of α_1, θ_2, β, and φ are relatively small for nearly all settings of the true parameters tested. For example, the maximum absolute percentage biases for α_1, θ_2, β, and φ are 6.5% ($\frac{0.013}{0.2} \times 100\%$), 5.2%, 2.4%, and 0.6%, respectively. For η, the maximum absolute percentage bias is 10.9%, from the third setting of the true parameters where θ_2^{true} is set the highest. As one can see, the precision of estimating η (as measured by the RMSE) decreases as θ_2^{true} increases and θ_2 becomes more influential. And, the precision of estimating θ_2 decreases as η_{true} decreases and η becomes more influential. This makes sense as θ_2 and η both have a similar scaling effect upon local detouredness and the parameters have some correlation (see below). Thus, the 10.9% bias for η is due to a dominating θ_2 parameter, where the standard error for η is also its largest.

Table 3 displays the estimated covariances between the α_1, θ_2, β, φ, and η parameters, from a single simulation experiment with the true parameters set as $\alpha_1^{true} = 0.2, \theta_2^{true} = 2, \beta_{true} = 0.7, \varphi_{true} = 1.5, \eta_{true} = 1$. As shown, there is no evidence of strong correlation between the parameters. But, the greatest correlation is between the $\hat{\theta}_2$ and $\hat{\eta}$ estimates, where there is a logical positive correlation: an increase in θ_2 or a decrease in η gives more probability to lower detouring routes.

Overall, the simulation study results suggest that the parameters of the BPS-LDT model can in general be successfully estimated and identified, evident from the low bias and standard error of the estimates, as well as the low covariance between parameters. There is though perhaps some confounding between the local detouredness scaling parameter θ_2 and local detour threshold parameter η. Importantly, the results suggest that local detouredness and correlation are distinct, independent route attributes.

α_1^{true}	θ_2^{true}	β_{true}	φ_{true}	η_{true}	$\hat{\alpha}_1$			$\hat{\theta}_2$			$\hat{\beta}$			$\hat{\varphi}$			$\hat{\eta}$		
					Bias	S.E	RMSE	Bias	S.E	RMSE	Bias	S.E	RMSE	Bias	S.E	RMSE	Bias	S.E	RMSE
0.2	2	0.7	1.5	1	0.013	0.047	0.049	0.048	0.582	0.584	-0.017	0.132	0.133	0.007	0.010	0.013	-0.037	0.167	0.171
0.2	2.5	0.7	1.5	1	0.013	0.050	0.052	-0.004	0.556	0.556	-0.016	0.138	0.139	0.009	0.011	0.014	-0.072	0.206	0.218
0.2	3	0.7	1.5	1	0.005	0.051	0.051	0.045	0.509	0.511	-0.014	0.143	0.144	0.008	0.012	0.014	-0.109	0.325	0.344
0.2	2	0.7	1.35	0.75	0.005	0.076	0.076	0.104	0.892	0.892	0.001	0.168	0.168	-0.002	0.008	0.008	-0.006	0.100	0.100
0.1	3	0.8	1.7	1.2	0.001	0.040	0.040	0.062	0.363	0.368	0.012	0.115	0.116	0.004	0.027	0.028	-0.006	0.220	0.220

Table 2. Sioux Falls simulation study: Stability of estimated BPS-LDT parameters across 250 experiment replications ($|Z| = 5000$ observations).

	$\hat{\alpha}_1$	$\hat{\theta}_2$	$\hat{\beta}$	$\hat{\varphi}$	$\hat{\eta}$
$\hat{\alpha}_1$	-	-0.020	0.000	0.000	-0.001
$\hat{\theta}_2$	-0.020	-	**-0.018**	0.001	0.047
$\hat{\beta}$	0.000	**-0.018**	-	-0.000	0.006
$\hat{\varphi}$	0.000	0.001	-0.000	-	-0.000
$\hat{\eta}$	-0.001	0.047	0.006	-0.000	-

Table 3. Sioux Falls simulation study: Estimated covariances between BPS-LDT model parameters from 250 experiment replications ($\alpha_1^{true} = 0.2, \theta_2^{true} = 2, \hat{\beta} = 0.7, \hat{\varphi} = 1.5, \hat{\eta} = 1$).

6.2 Real-life large-scale case study

In this section we estimate the BPS-LDT model in a real-life case study using the procedure discussed in Section 5 with observed route choices tracked by GPS units.

6.2.1 Setup

The GPS data has been collected among drivers in the eastern part of Denmark in 2011, and includes a total of 17,115 observed routes. The dataset is the same as used in Prato et al. (2014), Rasmussen et al. (2017), and Duncan et al. (2020,2022).

The GPS traces were map matched to a network, for which corresponding time-of-day-dependent travel times were available on the entire network. See more details in Prato et al. (2014). The network is large-scale, representing all of Eastern Denmark, and thus includes 34,251 links. With current route generation techniques, it is not feasible to enumerate the universal choice set for such a large network, and even enumerating all alternatives with a cost below a rather large relative bound (e.g. $\varphi = 2$) is not feasible. Instead, we generated representative universal choice sets by generating a choice set for each observed route by applying the doubly stochastic approach also applied in Prato et al. (2014). This approach is based on repeated shortest path searches in which the network attributes and parameters of the cost function are perturbated between searches (Nielsen, 2000; Bovy & Fiorenzo-Catalano, 2007). To reduce the risk of bias in estimation, care was taken to ensure a large variety of alternatives with different characteristics were generated, by assuming large variance in the parameters of the cost function. Up to 100 unique paths were generated for each observation, and the observation was matched to the most similar generated route, with a requirement of the overlap in

length being at least 80% for the observation to be included in the estimation. This removed 841 routes, i.e. the coverage was 95%. We also removed trips where the sum of travel time (in minutes) and length (in km) was less than 10, as well as observations where only one route was generated, leaving a total of 8,105 observations.

For the estimation in this study, like in Duncan et al. (2020,2022), the travel cost of link a is specified as a weighted sum of congested travel time $w_{a,1}$ (in minutes), and length $w_{a,2}$ (in kilometres), such that:

$$t_a(\boldsymbol{w}_a; \boldsymbol{\alpha}) = w_{a,1} \cdot \alpha_1 + w_{a,2} \cdot \alpha_2$$

where $\alpha_1 \geq 0$ and $\alpha_2 \geq 0$ are the congested travel time, and length parameters, respectively. The generalised travel cost for route $i \in R_m$ is thus:

$$c_{m,i}\big(\boldsymbol{t}(\boldsymbol{w};\boldsymbol{\alpha})\big) = \sum_{a \in A_{m,i}} t_a(\boldsymbol{w}_a;\boldsymbol{\alpha}) = \sum_{a \in A_{m,i}} \big(w_{a,1} \cdot \alpha_1 + w_{a,2} \cdot \alpha_2\big) = \alpha_1 \sum_{a \in A_{m,i}} w_{a,1} + \alpha_2 \sum_{a \in A_{m,i}} w_{a,2}.$$

The BPS-LDT model requires in this case the specification of seven parameters: α_1, α_2, θ_1, θ_2, β, φ, and η, but to ensure identification, the θ_1 parameter is fixed at $\theta_1 = 1$.

There were a few outlying observations where the relative surplus cost to the minimum cost alternative was rather high. These may be a result of erroneous processing of the GPS data (e.g. stop not classified) or behaviour beyond the intended scope of the model such as the travellers getting lost. After visual inspection of the most deviating observations, we removed 10 observations from the original dataset, which had the largest relative deviation from the minimum cost alternative. Further interrogation of the data revealed 102 observations with unrealistic local detours, leaving a total of 7,993 observations.

Using the calibrated travel cost parameters from the BPS-LDT model (see estimation results below), Fig. 4A-B display the cumulative distributions of the relative surplus travel costs and local detour measures, respectively, of the observed routes compared to all routes generated. As can be seen, a large variety of routes were generated. While most routes have low relative surplus costs and/or low local detour measures, some routes are very relatively costly (maximally 2.95), and some have very large local detour measures (maximally 98.84). The observed routes did not take the most costly / most detouring routes, however. 46.9% of observations took the lowest costing route (relative surplus costs of 1), and 45.6% of observations took a route with no local detours (detour measure of 0). Moreover, as can be seen, route usage decays with cost/detouredness, where the maximum observed relative surplus cost was 1.58 and the maximum observed local detour measure was 3.78.

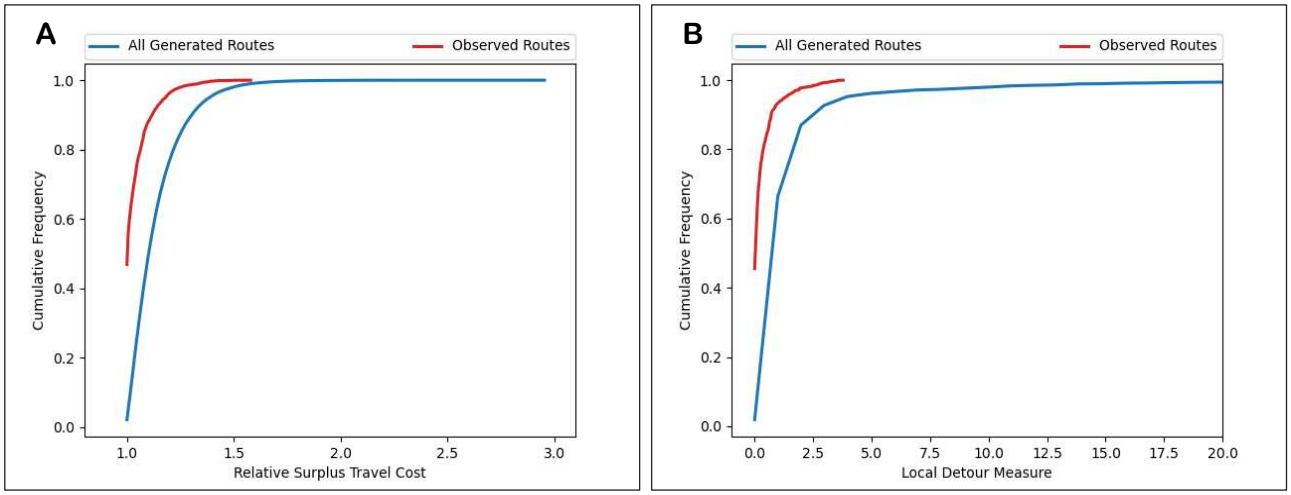

Fig. 4. Cumulative distributions of relative surplus travel cost (**A**) and local detour measures (**B**), of all generated routes and of the observed routes. Travel cost was calculated using calibrated travel cost parameters from BPS-LDT model in Table 4.

6.2.2 Estimation results

Here we present results from estimating the BPS-LDT model, and compare estimated parameters and goodness-of-fit with those for other relevant models. As shown again in Fig. 5, the BPS-LDT model acts as a unified model for several route choice models in the literature, which can be found in Appendix A. By comparing the BPS-LDT estimation results with other models we assess the main hypotheses of the paper that:

i) Local detouredness is an influential route choice attribute,
ii) Travellers impose a threshold on the local detouredness of routes to determine their route choice sets,
iii) Capturing correlations between used routes will lead to more realistic route choice probabilities.

The different models were all estimated utilising the same Log-Likelihood maximisation algorithm (L-BFGS-B, see Section 5.2), initial conditions, and parameter bounds. The initial conditions were: $\big(\tilde{\alpha}_1^{(1)}, \tilde{\alpha}_2^{(1)}, \tilde{\theta}_2^{(1)}, \tilde{\beta}^{(1)}, \tilde{\varphi}^{(1)}, \tilde{\eta}^{(1)}\big) =$

(1,0.7,1,0.8,1.7,5), with bounds: $\tilde{\alpha}_1, \tilde{\alpha}_2 \in [0.01,3]$, $\tilde{\theta}_2 \in [0.01,5]$, $\tilde{\beta} \in [0,3]$, $\tilde{\varphi} \in [1.01,5]$, $\tilde{\eta} \in [2,30]$. In Section 6.2.6 we estimate the BPS-LDT model with different initial conditions to test for multiple solutions.

Table 4 shows the estimated parameters, Log-Likelihood values, and penalised-likelihood criteria (in this case BIC statistic), and adjusted rho squared measures for the different models. Note that standard errors of the estimates are not provided in Table 4, as these cannot be calculated analytically for the bounded / local detour models. We instead in Section 6.2.4 explore estimate precision numerically through a resampling approach.

Local detouredness. As evident from the better Log-Likelihood, BIC, and adjusted rho squared statistics, the BCM-LDT provides better fit to the data than the BCM, and the BPS-LDT model provides better fit to the data than the BPS model. These two results imply that either or both hypotheses i) and ii) are true, that local detouredness is an influential route choice attribute and/or travellers impose a threshold upon local detouredness. By inspecting the estimated BPS-LDT model parameters, however, it is evident that only hypothesis i) is true. The estimated θ_2 parameter that scales sensitivity to local detouredness is estimated significantly different from zero (see Section 6.2.4), but the estimated local detour threshold parameter is estimated at the upper limit of 30 set for the L-BFGS-B algorithm, which approximates infinity. The latter is a surprising result, that when given the opportunity to define routes as unrealistic and assign them zero choice probabilities the local detour models opted not to. We explore this result in greater detail in Section 6.2.8. There is strong empirical evidence, however, that local detouredness is an influential route choice attribute.

Correlation. The BPS-LDT model provides considerably better fit to the data than the BCM-LDT model, and the path size scaling parameter β is estimated significantly different from zero (see Section 6.2.4), thus providing strong empirical evidence to support hypothesis iii) that capturing correlations between used routes leads to more realistic route choice probabilities.

Cost-bounds. Shadowing similar results in Duncan et al. (2022), as can be seen for all of the bounded models, the cost-bound parameters $\hat{\varphi}$ were estimated to be around 1.5-1.6, which corresponds suitably with the observation with the maximum relative surplus cost, which is around 1.53-1.6 (see Fig. 4 and Fig. 12), depending on the estimated cost parameters. This cuts off a small proportion of routes (see Fig. 4), resulting in the BCM and BPS model providing marginally better fit to the data (as measured by BIC) than their corresponding non-bounded models MNL and GPSL′, respectively. This implies that the cost-bounds within the BCM-LDT and BPS-LDT models result in better fit to the data.

Value of Time. Comparing the ratio between the travel time preference parameter α_1 and distance preference parameter α_2 gives a ratio between 2 and 3, inferring that travellers are willing to travel an extra 2-3 kilometres to save 1 minute of travel time. This ratio range seems plausible and aligns well with our findings in a separate study of the same area in Duncan et al. (2025).

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\theta}_2$	$\hat{\beta}$	$\hat{\varphi}$	$\hat{\eta}$	LL	BIC	Adj. ρ^2
MNL	1.221	0.698					-21443	42904	0.399
PSL	2.077	0.599		2.132			-20155	40336	0.435
GPSL′	1.243	0.378		2.269			-18547	37121	0.480
BCM	1.203	0.696			1.537		-21437	42901	0.399
BPS	1.235	0.377		2.270	1.607		-18540	37116	0.480
BCM-LDT	0.855	0.473	1.581		1.530	30	-19940	39925	0.441
BPS-LDT	**1.053**	**0.316**	**0.357**	**2.031**	**1.600**	**30**	**-17862**	**35778**	**0.499**

Table 4. Real-life case-study estimation results.

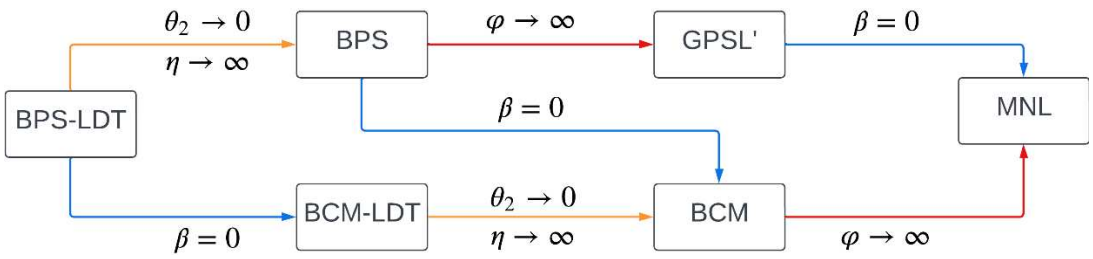

Fig. 5. Schematic diagram of how the BPS-LDT model collapses into other models.

6.2.3 Estimation times

Table 5 displays the estimation times for each model. As shown, generally, the more complex the model and number of parameters to estimate, the longer the estimation times. The local detour models take longer to estimate, partly because it takes longer to compute the choice probabilities, but mainly because more function evaluations / iterations are required to reach convergence. MLE convergence patterns for the BPS-LDT model can be seen in Fig. 9. The BPS-LDT model can though be estimated in feasible computation times. As we shall discuss in Section 6.2.8, one can speed up estimation

times for the BPS-LDT model considerably by setting the relative cost bound and local detour threshold parameters to marginally above the relative cost / detour measure of the worst respective observations.

Model	MNL	PSL	GPSL′	BCM	BPS	BCM-LDT	BPS-LDT
Estimation Time [min]	1.15	8.64	5.11	6.67	24.96	480.61	**406.55**

Table 5. Estimation times for each model.

6.2.4 Estimate precision

Here we evaluate the precision at which the parameters of the different models are estimated, thereby assessing whether the parameters are statistically significant and thus influential factors upon route choice.

To evaluate estimate precision, we adopt the BLB resampling method discussed in Section 5.3. We do not, however, adopt an iterative algorithm for seeking sufficient H and Y as we wish to compare goodness-of-fit and estimate precision across different models, where the same samples will be used for each model. In simulation experiments, Kleiner et al (2014) investigate standard sufficient values for H and Y, and conclude that $Y = 100$ and $H = 10$ were sufficient for $G \geq |Z|^{0.6}$. After some preliminary experiments of our own, we concluded that – wanting to select as large a value of G as computational resources would allow – $G = |Z|^{0.8} = 1415$ would be feasible in our case for $Y = 100$ and $H = 10$. We therefore generated 10 'bags' of $G = 1415$ observations (without replacement from the full $|Z| = 7993$ observation sample set), and then randomly sampled with replacement $|Z|$ observations from each of the 10 bags 100 times, to obtain a total of 1000 samples. Then, for each of the 1000 samples, we estimated the parameters for each of the models in Table 4. In order to optimise efficiency, in the Python code for performing the BLB method for each model, we used parallel processing to split the 1000 estimations across 15 logical processors.

Table 6 displays for each model the mean estimates μ, standard deviations of the estimates σ, mean Log-Likelihood values $\overline{LL}$, and mean BIC values. We do not include the local detour threshold parameter η estimates and standard deviations as the estimates are all 30 (the upper algorithm limit) with zero standard deviation. As can be seen, the mean estimates in Table 6 are all similar to the estimates on the full dataset in Table 4. The standard deviations are also all appear reasonable, and there is no evidence that the parameters of the BPS-LDT model are estimated less precisely than the other models. Testing whether the $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\theta}_2$, and $\hat{\beta}$ parameter estimates are significantly different from zero and the $\hat{\varphi}$ parameter estimates are significantly different from 3 (a reasonable proxy for infinity, see Fig. 4/Fig. 8), the p-values for every parameter estimate of every model are all smaller than can be computed in Python, and thus the parameter estimates are statistically significant. This is of course apart from the η parameter estimates, which are not statistically significant.

In terms of goodness-of-fit, although the mean Log-Likelihood / BIC values in Table 6 are slightly different from the Log-Likelihood / BIC values on the full dataset in Table 4, the comparisons in fit are roughly the same. The BPS-LDT model provides by far the best mean fit. Fig. 6 plots for each of the models the BIC from each of the 1000 BLB samples, ordered from lowest to highest BPS-LDT BIC. As can be seen, for all samples, the BPS-LDT model performs the best, highlighting the robustness of the result.

	$\hat{\alpha}_1$		$\hat{\alpha}_2$		$\hat{\theta}_2$		$\hat{\beta}$		$\hat{\varphi}$		$\overline{LL}$	$\overline{BIC}$
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ		
MNL	1.21	0.11	0.73	0.12							-21526	43070
PSL	2.06	0.13	0.63	0.09			2.13	0.09			-20250	40527
GPSL′	1.24	0.06	0.39	0.06			2.28	0.08			-18599	37225
BCM	1.19	0.11	0.72	0.12					1.52	0.07	-21515	43057
BPS	1.22	0.06	0.39	0.06			2.29	0.08	1.56	0.10	-18582	37200
BCM-LDT	0.81	0.09	0.48	0.08	1.70	0.28			1.50	0.06	-19922	39880
BPS-LDT	1.03	**0.06**	0.32	**0.05**	0.40	**0.09**	2.04	**0.08**	1.54	**0.08**	**-17851**	**35747**

Table 6. BLB results for each model (from 1000 estimations), mean μ and standard deviation σ of estimates.

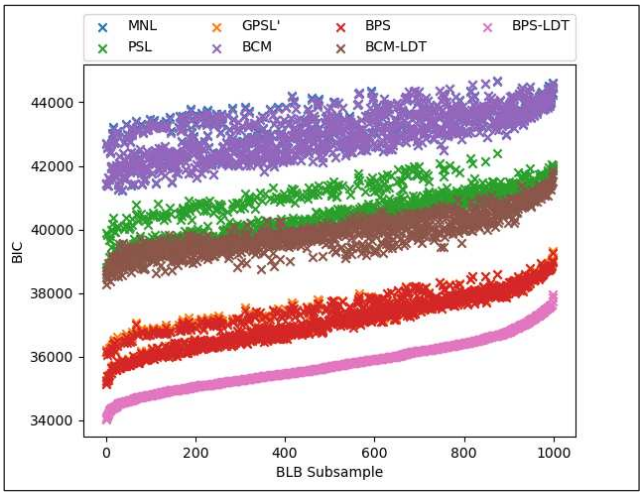

Fig. 6. Plotting for each model the BIC penalised likelihood criteria from each of the 1000 BLB samples, in order of lowest to highest BPS-LDT model BIC.

6.2.5 Out of sample validation

Here we assess the forecastability of the BPS-LDT model, assessing whether the model overfits the data, and whether the model is suitable for transferring estimates to predict choice probabilities in other datasets. To do this we perform Monte Carlo cross-validation, following the same process as Cazor et al. (2024). For each model, we repeated the following steps 10 times:

Step 1: Randomly split the full dataset Z into a training set Z_t and validation set Z_v, where $|Z_t| = |Z_v| = 0.5 \cdot |Z|$.

Step 2: Estimate the model on the training set Z_t, obtaining a set of training parameter estimates.

Step 3: Given these training parameters, calculate the Log-Likelihood on the validation set Z_v.

Fig. 7 displays the validation set Log-Likelihood of each model from each of the 10 experiments. As can be seen, the comparative fits to the data of the different models are the same as in Table 4 and Table 6. The BPS-LDT model outperforms the BPS model supporting the hypothesis that it is important to consider local detouredness as a route attribute, and considerably outperforms the BCM-LDT model supporting the hypothesis that it is important to consider route correlation. The result is consistent across all experiments. There is no evidence of overfitting; the BPS-LDT model is suitable for forecasting, and moreover provides the most realistic choice probabilities for doing so.

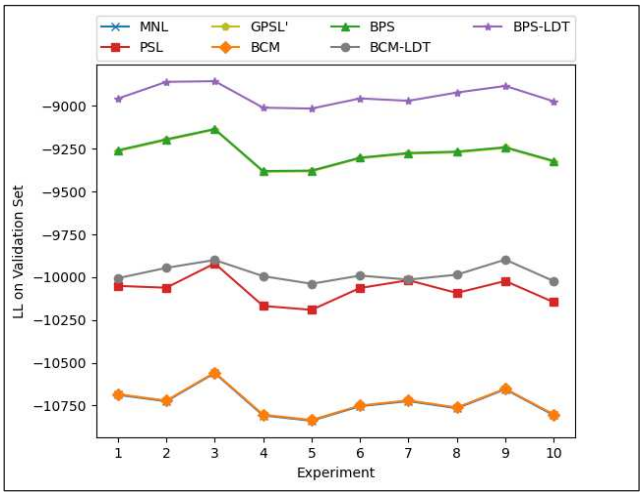

Fig. 7. Validation set Log-Likelihood of each model from each of the 10 experiments.

6.2.6 Uniqueness of MLE solutions

Here we explore numerically whether BPS-LDT MLE solutions are unique. This is important as non-uniqueness can undermine the interpretability, stability, and reliability of results, as well as cause issues with convergence of the maximisation algorithm.

Fig. 8 visualises the BPS-LDT Log-Likelihood surface around the 6 parameter estimates in Table 4, i.e. varying each of the parameters around the estimate with the other parameters fixed. Note that we only visualise the Log-Likelihood surface in parameter ranges where the Likelihood is non-zero. As can be seen, the Log-Likelihood surface is not globally concave, evident from the surface around the φ parameter. There is though no evidence of multiple maxima.

To test for multiple solutions, Fig. 9 and Fig. 10 display the model parameters and Log-Likelihood value, respectively, at each iteration of the MLE solution algorithm with three different initial conditions. As can be seen, each initial condition converges to the same MLE solution.[4] Additional evidence for uniqueness is provided in the simulation study, where for each set of drawn observations we obtained parameter estimates in the neighbourhood of the true parameters.

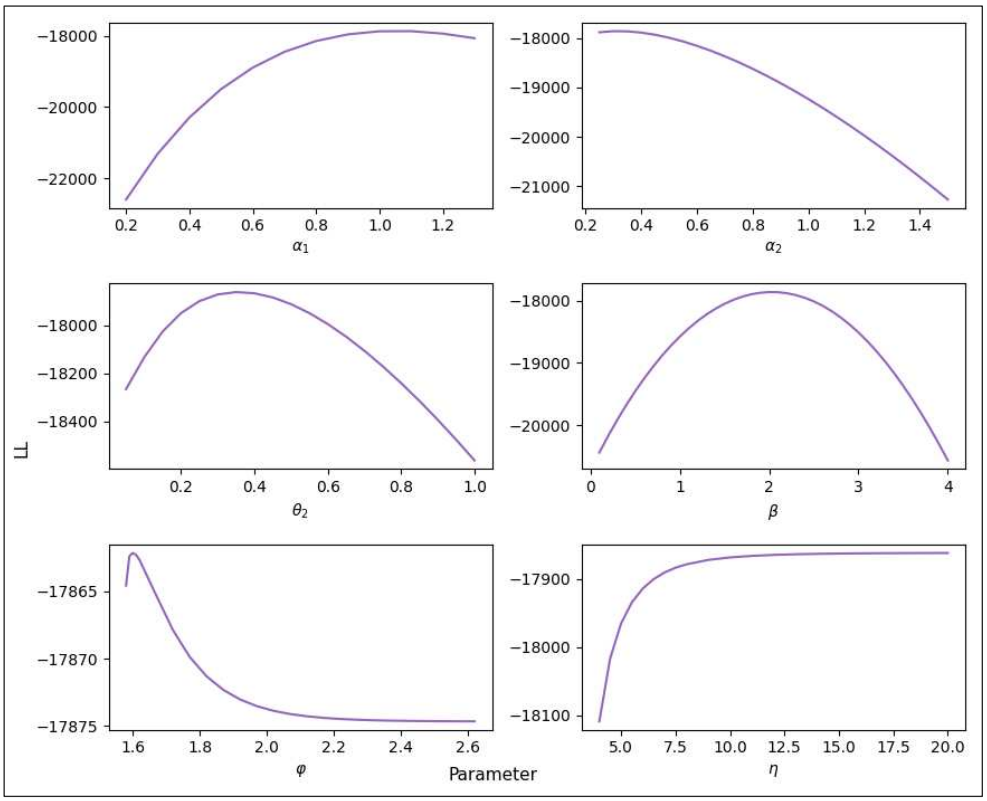

Fig. 8. Visualising the BPS-LDT model Log-Likelihood surface around the parameter estimates in Table 4, varying one parameter and fixing the others to the estimates.

[4] Initial condition 2 converges to an η estimate of 26.35, but this is only because the Log-Likelihood is very flat in this parameter range where η is approximating infinity. A slightly greater Log-Likelihood value is achieved with $\eta = 30$.

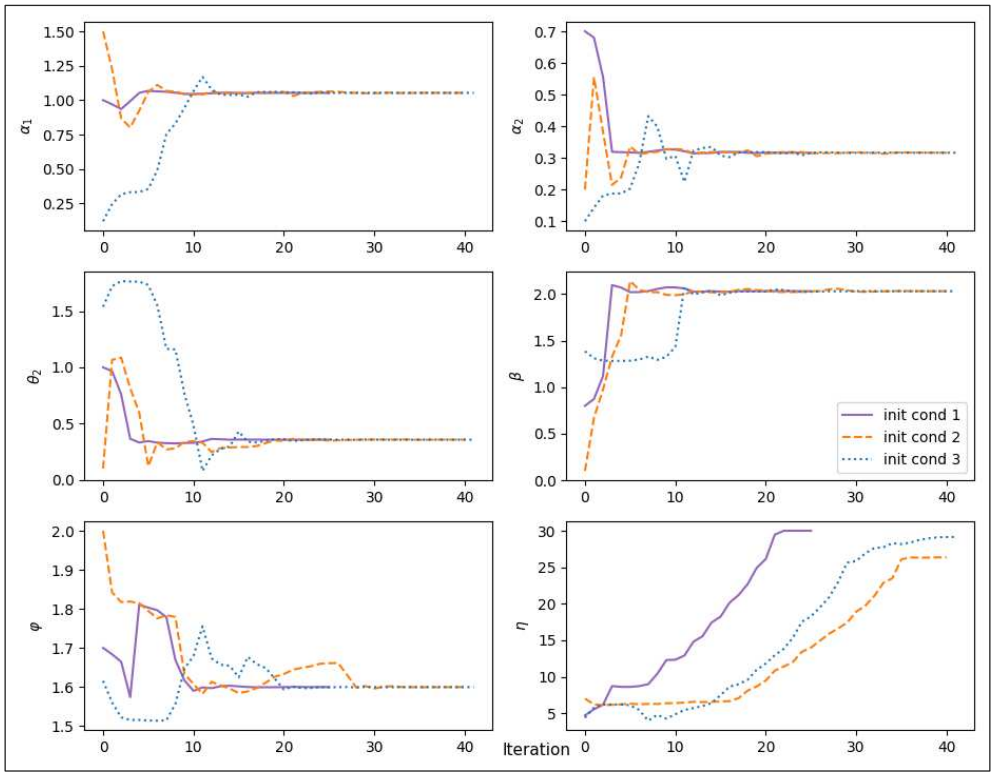

Fig. 9. BPS-LDT model parameters at each iteration of MLE solution algorithm, with different initial conditions.

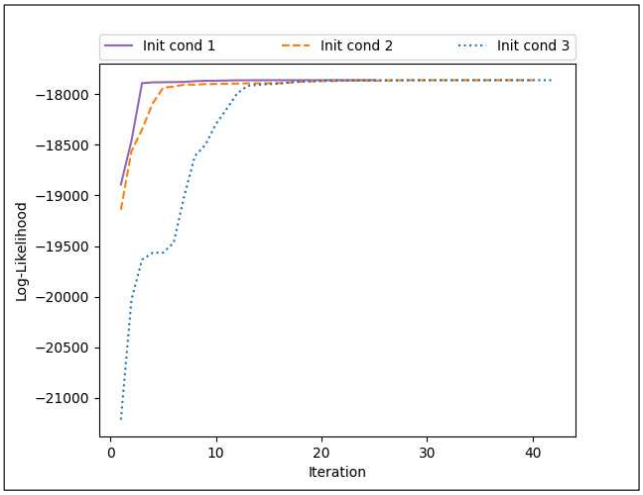

Fig. 10. BPS-LDT Log-Likelihood value at each iteration of MLE solution algorithm, with different initial conditions.

6.2.7 Sensitivity to the representative universal choice sets

The BPS-LDT choice probably solution method proposed in this paper operates from pre-generated representative universal choice sets rather than the full universal route set. While the BPS-LDT model deals with unintentionally generated circuitous routes with costs / local detours above the bounds (by assigning them zero choice probabilities), routes that were not pre-generated with costs / local detours below the bound are not accounted for. Here we explore how sensitive results are to the pre-generated representative universal choice sets adopted, to assess how robust results are to missing out relevant routes.

Using the same doubly-stochastic route generation method discussed in Section 6.2.1, we generated additional routes to obtain choice sets with a maximum of 100, 200, and 300 routes, and then estimated the different models for each choice set composition. Table 7 displays the estimated parameters for the BPS-LDT model for each maximum choice set size. As shown, the estimated parameters remain relatively stable, though there is a decrease in the β path size scaling parameter and an increase in the θ_2 detouredness scaling parameter. This is likely because the newly generated routes overlap with the observed routes, as well as detour from them. Fig. 11 displays for each model the difference in Log-Likelihood between estimation with the original choice sets and estimation with the expanded choice sets. As can be

seen, the Likelihood worsens less for the BPS-LDT model than for the other models, and remains by far the model with the best fit. These results indicate that the BPS-LDT model is relatively robust to missing out relevant routes from the pre-generated representative universal choice sets. The BPS-LDT model is also of course robust to unintentionally generated irrelevant routes.

Maximum Choice Set Size	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\theta}_2$	$\hat{\beta}$	$\hat{\varphi}$	$\hat{\eta}$
100	1.053	0.316	0.357	2.031	1.600	30
200	1.070	0.321	0.415	1.844	1.620	30
300	1.094	0.319	0.460	1.745	1.625	30

Table 7. BPS-LDT estimated parameters for each choice set composition.

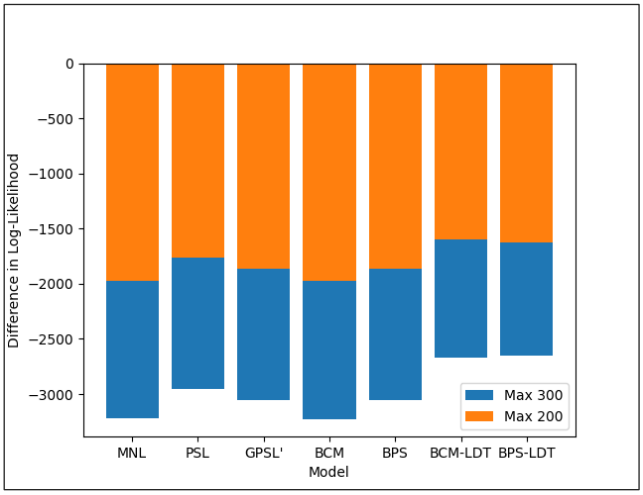

Fig. 11. Difference in Log-Likelihood of each model between estimation with original choice sets and estimation with expanded choice sets.

6.2.8 Exploring the local detour threshold

Since the greatest local detour measure from the observed routes was 3.78, it was anticipated that the BPS-LDT local detour threshold parameter would be estimated close to (but greater than) 3.78. This would define 18,961 routes / 5.06% of routes as unrealistic. However, by estimating an infinite bound, no routes were defined as unrealistic by the local detour threshold. It is surprising that when given the opportunity to define routes as unrealistic and assign them zero choice probabilities the local detour models opted not to. A possible reason for this is explained as follows.

The cost-bound and detour-threshold in local detour models do not only assign zero probabilities to routes with travel costs / detour measures above the bound/threshold, but also impact the probabilities of routes below the bound/threshold. Route probabilities decrease towards zero if their cost approaches the cost bound from below, or if their detour measure approaches the detour threshold from below. This ensures the choice probability function is continuous. However, it has the potential to cause conflictions. If one considers a standard MNL model where the route utility function is a linear combination of travel cost and local detouredness, then a route receives a small probability if it has an unattractive combination of both cost and detouredness. If it has a relatively attractive cost but a relatively unattractive detouredness, or vice versa, then it may receive a middling choice probability. For local detour models, however, a route receives a low choice probability if it has either a relatively unattractive cost that is just below the bound or an unattractive detouredness that is just below the threshold, regardless of how attractive the route may be in terms detouredness or cost, respectively.

Therefore, given that the MLE procedure is trying to find the parameter values that overall produce the highest probabilities for all route observations, it could be the case that setting a low detour threshold (but still above the worst observed detour measure) reduces the probabilities for some observed routes, that otherwise may receive considerably better probabilities due to their relatively attractive travel cost. And, that this gain in probability for these observations is not cancelled out by the assigning of non-zero probabilities to generated routes with large detour measures, which under lower settings of the detour threshold would receive zero probabilities and thus take no probability away from observed route probabilities. Fig. 12 plots relative surplus cost against local detour measure for the observed routes. As shown, it is not the case that there is a positive correlation between cost and detouredness. In fact, as evident from the population of datapoints towards the left of the figure, and particular top left, there are many observations that have a low relative

surplus cost but a large local detour measure. In contrast, there are few routes that have large costs and small detour measures, especially since at worst the largest local detour of a route is the global detour.

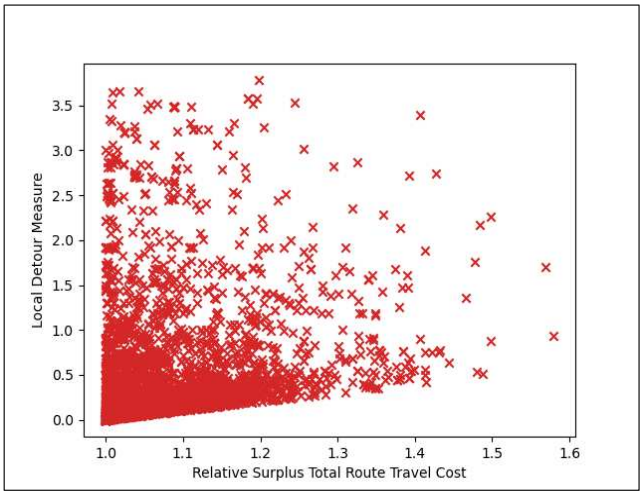

Fig. 12. Plotting relative surplus cost against local detour measure for the observed routes.

As discussed, the cost bound and detour threshold do impact the choice probabilities of routes with costs / detour measures below the bound/threshold, but, principally, this is a mathematical construct to ensure the choice probability function is continuous. Although the parameters do scale in a non-direct way sensitivity to travel cost / local detouredness (as demonstrated in Appendix B.3), there are other parameters present in the model to do exactly such, i.e. the θ_1 and θ_2 parameters. The key role of the bound/threshold parameters is to identify the bound/threshold travellers have on travel cost / local detouredness, and thereby use this to consistently identify and define routes as unrealistic, i.e. that have costs / detour measures greater than the bound/threshold. One could therefore suppose that one need not actually estimate the bound/threshold parameters, and instead simply set them to marginally above the relative cost / detour measure of the worst respective observations (which may vary depending on the travel cost parameters).

Doing exactly such, Table 8 displays estimation results from estimating the local detour models when setting the cost bound parameter φ marginally greater (+0.01) than the most relatively costly observation, and the detour threshold parameter η marginally greater (+0.01) than the most detouring observation. As shown, the local detour threshold parameter values are all around the worst detour observation (around 3.7-3.8). As anticipated though, the Log-Likelihood and BIC values are worse than when estimating free bounds/thresholds (see Table 4). Both models do still all outperform their associated non-local detour model (BCM and BPS model), however, by a sizeable margin. But, while the parameter estimates for $\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\beta}$ are quite similar for the BPS-LDT model with free and fixed bound/threshold parameters, the $\hat{\theta}_2$ parameter estimate when fixing the bound/threshold parameters is very different, approximating zero. This indicates that the observations with low travel costs but high local detour measures are highly influential in the Log-Likelihood function, where a small θ_2 parameter is required to increase the probability of these routes.

One benefit of setting the bound/threshold parameters marginally above the worst observations is that it reduces the number of parameters to be estimated, improving estimation times. This is especially pertinent for the estimation of bounded / local detour models, which have an unusual MLE pattern due to the heuristic setting of the Log-Likelihood when the bound or threshold are violated, which can consequently mean more function evaluations / iterations are required to find the MLE solution. The estimation times for the results in Table 8 were 65 minutes for the BCM-LDT model and 133 minutes for the BPS-LDT model, which are considerably faster than those given in Table 5.

	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\theta}_2$	$\hat{\beta}$	φ	η	LL	BIC
BCM-LDT	0.814	0.492	1.652		1.518	3.735	-19998	40023
BPS-LDT	1.070	0.315	0.000	1.996	1.591	3.796	-18103	36241

Table 8. Estimation results from estimating local detour models with bound/threshold parameter set marginally above values from worst respective observations.

7 Conclusions and future research

Local detour route choice models suppose that local detouredness is an influential factor upon both route choice probability and route choice set formation. This paper has addressed three priorly unresolved challenges regarding local detour models: i) accounting for correlations between overlapping routes defined as realistic, ii) developing a solution approach for applying the model on large-scale networks, and, iii) developing, testing, and applying a procedure for estimating the model.

- Addressing i), path size terms were integrated within the local detour model probability relation to adjust probabilities to capture correlations between overlapping routes defined as realistic by local and global bounds. Careful consideration has been given to the formulation of the path size term to ensure that the choice probability function maintains continuity, consistency, and existence.
- Addressing ii), due to the high computational burden of current methods for generating all routes below the local and global bounds, the current paper has developed a heuristic solution approach that operates from pre-generated representative universal choice sets and pre-processes the necessary segment information for computing local detour measures. Tricks have also been developed that considerably improve computation times.
- Addressing iii), a maximum likelihood estimation procedure utilising tracked route observation data was developed, tested in a simulation study where it was shown that assumed true parameters can generally be reproduced, and applied in a real-life large-scale case study with GPS-tracked route observations. The local detour models were successfully estimated and found to outperform associated non-local detour models.

In response to the research questions listed in the introduction, the main findings of the empirical analysis were that:

- Local detouredness is an influential factor upon route choice probability,
- Local detouredness and correlation effects are identifiable as distinct route choice attributes,
- Parameter estimates are statistically significant (other than the local detour threshold parameter),
- No evidence was uncovered that MLE solutions are not unique,
- The model can be estimated in feasible computation times on a large-scale network,
- The model is suitable for forecasting,
- Model estimates were relatively insensitive to the assumed representative universal choice set.

A surprising finding from the real-life estimation work was that the local detour models estimated local detour threshold parameter approximating infinity. One might expect that the local detour threshold parameter would be estimated marginally above the largest local detour measure from the route observations, thereby maximising the number of generated routes the local detour models assign zero probabilities to (i.e. define as unrealistic). However, the local detour threshold parameters estimated as approximating infinity meant that no generated routes were assigned zero probabilities. Our hypothesis is that this result is, in part, a symptom of working from pre-generated representative universal choice sets, or at least the ones in this case study. As shown in Fig. 4B, a very small proportion of the generated routes have large local detour measures, and thereby have the potential to be assigned zero probabilities, to give probability to observed routes. Clearly, the working choice sets do not even come close to the universal set of all possible routes, which would include an enormous number of routes with large local detours. As shown in Rasmussen et al. (2024), when generating all routes with a local detour measure less than the detour threshold from the full network, the choice set size grows exponentially and becomes extremely large very quickly. If we were to estimate the local detour threshold models on the full network with the actual universal choice set of routes available, it thus seems significantly more likely that the cost bound / local detour threshold would try to exclude as many routes as possible.

Future research could therefore explore developing a more computationally efficient method (than that developed in Rasmussen et al. (2024)) for generating all possible routes below the local and global bounds. For example, developing an efficient branch-and-bound route generation method utilising violating segments to efficiently branch, and re-uses information across OD movements. This will mean that one can apply local detour models with full consistency between the choice set generation criteria and choice probability criteria, where no route defined as realistic by the route choice probability criteria will be unused (i.e. not generated).

However, as discussed, the set of all possible routes below the local and global bounds has the potential to be enormous. Thus, perhaps further measures need to be introduced that define routes as unrealistic. For example, by imposing more bounds upon individual route 'aspects', e.g. length, travel time, number of left/right turns. After some investigation, it appears the main issue is that there are many routes with costs/detours below the global/local bounds that only deviate from each other in a minor way on very small subsections. It appears then that limiting this should be the next focus of the model. There are several approaches one could take. One approach could be to continue with the choice set pre-generation approach and incorporate certain constraints on the similarity of routes generated, so that the approximated universal choice sets contain plenty of variability in the routes, but do not include many routes that are essentially the same. Alternatively, perhaps one could develop a measure that can be used to bound route over-similarity in some way, perhaps by scaling the local detour measure by segment cost compared to total route cost, or by developing a C-Logit-type measure penalising over-similarity with a lesser costing route.

In the present paper we have adopted a path-size approach for capturing correlations. In future research it would be interesting to explore to what extent many of the other existing correlation-based route choice models could be adapted to deal with bounds and local detours.

8 Acknowledgements

We gratefully acknowledge the financial support of the Independent Research Fund Denmark to the projects “Next-generation route choice models for behavioural realism and large-scale applications” [Grant ID: 0136-00242B] and “Optimal tolls for reaching emissions goals: A novel behaviourally realistic and national-scale applicable road transport modelling system” [Grant ID: 0217-00173B].

9 References

Allahviranloo M, Regue R, & Recker W, (2017). Modeling the activity profiles of a population. Transportmetrica B: Transport Dynamics, 5(4), p.426-449.

Ben-Akiva M & Ramming S, (1998). Lecture notes: discrete choice models of traveler behavior in networks. Prepared for Advanced Methods for Planning and Management of Transportation Networks. Capri, Italy.

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

Bovy P & Fiorenzo-Catalano S, (2007). Stochastic Route Choice Set Generation: Behavioral and Probabilistic Foundations. Transportmetrica, 3, p.173–189.

Brederode L, Pel A, Wismans L, de Romph E & Hoogendoorn S, (2019). Static Traffic Assignment with Queuing: model properties and applications. Transportmetrica A: Transport Science, 15(2), p.179-214.

Byrd R, Lu P, Nocedal J, & Zhu C (1994). A Limited Memory Algorithm for Bound Constrained Optimization. Technical Report NAM-08, Northwestern University, Department of Electrical Engineering and Computer Science.

Cadarso L & Marín Á, (2016). Combining robustness and recovery in rapid transit network design. Transportmetrica A: Transport Science, 12(3), p.203-229.

Cascetta E, Nuzzolo A, Russo F, & Vitetta A, (1996). A modified logit route choice model overcoming path overlapping problems: specification and some calibration results for interurban networks. In: Proceedings of the 13th International Symposium on Transportation and Traffic Theory, Leon, France, p.697–711.

Cascetta E & Papola A, (2001). Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand. Transportation Research Part C: Emerging Technologies, 9(4), p.249-263.

Cazor L, Watling D, Duncan L, Nielsen O, & Rasmussen T, (2024). A novel choice model combining utility maximization and the disjunctive decision rules, application to two case studies. Journal of choice modelling, 52, p.100510.

Chen Y, Li Z, & Lam W, (2018). Cordon toll pricing in a multi-modal linear monocentric city with stochastic auto travel time. Transportmetrica A: Transport Science, 14(1-2), p.22-49.

Chorus C, (2012). Regret theory-based route choices and traffic equilibria. Transportmetrica, 8(4), p.291-305.

Chorus C & van Cranenburgh S, (2024). Alternative decision rules in (travel) choice models: a review and critical discussion. Handbook of Choice Modelling, p.339-371.

Covington C, He X, Honaker J, & Kamath G, (2021). Unbiased statistical estimation and valid confidence intervals under differential privacy. arXiv preprint arXiv:2110.14465.

Duncan L, Watling D, Connors R, Rasmussen T, & Nielsen O, (2020). Path Size Logit Route Choice Models: Issues with Current Models, a New Internally Consistent Approach, and Parameter Estimation on a Large-Scale Network with GPS Data. Transportation Research Part B: Methodological, 135, p.1-40.

Duncan L, Watling D, Connors R, Rasmussen T, & Nielsen O, (2022). A bounded path size route choice model excluding unrealistic routes: formulation and estimation from a large-scale GPS study. Transportmetrica A: Transport Science, 18(3), p.435-493.

Duncan L, Watling D, Connors R, Rasmussen T, & Nielsen O, (2023). Choice set robustness and internal consistency in correlation-based logit stochastic user equilibrium models. Transportmetrica A: transport science, 19(3), p.2063969.

Duncan L, Watling D, Connors R, Rasmussen T, & Nielsen O, (2024). Formulation and solution method of bounded path size stochastic user equilibrium models–consistently addressing route overlap and unrealistic routes. Transportmetrica A: transport science, 20(2), p.2178240.

Duncan L, Rasmussen T, Watling D, & Nielsen O, (2025). Dynamic multi-region MFD stochastic user equilibrium: Formulation and parameter estimation in a large-scale case study. Transportation Research Part C: Emerging Technologies, 173, 105008.

Efron B, (1979). Bootstrap methods: Another look at the jackknife. Annals of Statistics, 7(1), p.1-26.

Efron B & Tibshirani R, (1993). An Introduction to the Bootstrap. Chapman and Hall.

Elrod T, Johnson R, & White J, (2004). A new integrated model of noncompensatory and compensatory decision strategies. Organizational Behavior and Human Decision Processes, 95(1), p.1-19.

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

Fosgerau M, Paulsen M, & Rasmussen T, (2022). A perturbed utility route choice model. Transportation Research Part C: Emerging Technologies, 136, p.103514.

Fosgerau M, Łukawska M, Paulsen M, & Rasmussen T, (2023). Bikeability and the induced demand for cycling. Proceedings of the National Academy of Sciences, 120(16), p.e2220515120.

Gentile G, (2014). Local User Cost Equilibrium: a bush-based algorithm for traffic assignment. Transportmetrica A: Transport Science, 10(1), p.15-54.

Gilbride T & Allenby G, (2004). A choice model with conjunctive, disjunctive, and compensatory screening rules. Marketing Science, 23(3), p.391-406.

Hood J, Sall E, & Charlton B, (2011). A GPS-based bicycle route choice model for San Francisco, California. Transportation letters, 3(1), p.63-75.

Horowitz J & Louviere J, (1995). What is the role of consideration sets in choice modeling? International Journal of Research in Marketing, 12(1), p.39-54.

Jánošíková Ľ, Slavík J, & Kohání M, (2014). Estimation of a route choice model for urban public transport using smart card data. Transportation planning and technology, 37(7), p.638-648.

Jedidi K & Kohli R, (2005). Probabilistic subset-conjunctive models for heterogeneous consumers. Journal of Marketing Research, 42(4), p.483-494.

Jiang Y & Szeto W, (2015). Time-dependent transportation network design that considers health cost. Transportmetrica A: transport science, 11(1), p.74-101.

Jing P, Zhao M, He M, & Chen L, (2018). Travel mode and travel route choice behavior based on random regret minimization: A systematic review. Sustainability, 10(4), p.1185.

Joksimovic D, Bliemer M, & Bovy P, (2005). Optimal toll design problem in dynamic traffic networks with joint route and departure time choice. Transportation Research Record, 1923(1), p.61-72.

Kitthamkesorn S, & Chen A, (2013). Path-size weibit stochastic user equilibrium model. Transportation Research Part B, 57, p.378-397.

Kleiner A, Talwalkar A, Sarkar P, & Jordan M, (2014). A scalable bootstrap for massive data. Journal of the Royal Statistical Society: Series B: Statistical Methodology, p.795-816.

Kohli R & Jedidi K, (2005). Probabilistic subset conjunction. Psychometrika, 70, p.737-757.

Li M & Huang H, (2017). A regret theory-based route choice model. Transportmetrica A: Transport Science, 13(3), p.250-272.

Lim Y & Kim H, (2016). A combined model of trip distribution and route choice problem. Transportmetrica A: Transport Science, 12(8), p.721-735.

Lim Jr H, Lim M, & Piantanakulchai M, (2019). Modeling route choice behavior of evacuees in highly urbanized area: a case study of Bagong Silangan, Quezon City, Philippines. Asia Pacific Management Review, 24(2), p.98-105.

Mardan A, Qi Z, Zhu Z, & Zhu S, (2024). A theoretic analysis of expressed toll lane pricing policy targeting for a guaranteed level of service. Transportmetrica A: Transport Science, p.1-30.

Martens K & Hurvitz E, (2011). Distributive impacts of demand-based modelling. Transportmetrica, 7(3), p.181-200.

Martínez F, Aguila F, & Hurtubia R, (2009). The constrained multinomial logit: A semi-compensatory choice model. Transportation Research Part B: Methodological, 43(3), p.365-377.

Nielsen O, (2000). A Stochastic Transit Assignment Model Considering Differences in Passengers Utility Functions. Transportation Research Part B Methodological, 34 (5), p.377–402. Elsevier Science Ltd.

Paleti R, (2015). Implicit choice set generation in discrete choice models: Application to household auto ownership decisions. Transportation Research Part B: Methodological, 80, p.132-149.

Prashker J & Bekhor S, (2004). Route choice models used in the stochastic user equilibrium problem: a review. Transport reviews, 24(4), p.437-463.

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

Prato C, Rasmussen T, & Otto N, (2014). Estimating Value of Congestion and of Reliability from Observation of Route Choice Behavior of Car Drivers. Transportation Research Record: Journal of the Transportation Research Board, 2412, p.20–27.

Rasmussen T, Nielsen O, Watling D, & Prato C, (2017). The Restricted Stochastic User Equilibrium with Threshold model: Large-scale application and parameter testing. European Journal of Transport and Infrastructure Research, 17(1).

Rasmussen T, Duncan L, Watling D, & Nielsen O, (2024). Local Detouredness: A New Phenomenon for Modelling Route Choice and Traffic Assignment. Transportation Research Part B: Methodological, 190, p.103052.

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

Sharma S & Kumar S, (2021). Fast and accurate bootstrap confidence limits on genome-scale phylogenies using little bootstraps. Nature computational science, 1(9), p.573-577.

Sheffi Y & Powell W, (1982). An algorithm for the equilibrium assignment problem with random link times. Network: An International Journal, 12(2), p.191-207.

Shin J & Ferguson S, (2017). Exploring product solution differences due to choice model selection in the presence of noncompensatory decisions with conjunctive screening rules. Journal of Mechanical Design, 139(2), 021402.

Swait J, (2001). A non-compensatory choice model incorporating attribute cutoffs. Transportation Research Part B: Methodological, 35(10), p. 903-928.

Tan H, Xu X, & Chen A, (2024). On endogenously distinguishing inactive paths in stochastic user equilibrium: A convex programming approach with a truncated path choice model. Transportation Research Part B: Methodological, 183, 102940.

Tilahun N, Levinson D, & Krizek K, (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. Transportation Research Part A: Policy and Practice, 41(4), p.287-301.

Toledo T, Atasoy B, Jing P, Ding-Mastera J, Santos J, & Ben-Akiva M, (2020). Intercity truck route choices incorporating toll road alternatives using enhanced GPS data. Transportmetrica A: Transport Science, 16(3), p.654-675.

Tsai J & Li S, (2019). Cordon tolling for mixed traffic flow. Transportmetrica A: Transport Science, 15(2), p.1662-1687.

Tversky A, (1972). Elimination by aspects: A theory of choice. Psychological review, 79(4), p.281.

Watling D, Rasmussen T, Prato C, & Nielsen O, (2018). Stochastic user equilibrium with a bounded choice model. Transportation Research Part B: Methodological, 114, p.254-280.

Wei X, Wang W, Yu W, Hua X, & Xiang Y, (2020). A stochastic user equilibrium model under traffic rationing based on mode shifting rate. Sustainability, 12(13), p.5433.

Xu X, Chen A, Kitthamkesorn S, Yang H, & Lo H.K, (2015). Modeling absolute and relative cost differences in stochastic user equilibrium problem. Transportation Research Part B, 81, p.686-703.

Zhong Q & Miao L, (2024). Capturing impacts of travel preference on connected autonomous vehicle adoption of risk-averse travellers in multi-modal transportation networks. Transportmetrica A: Transport Science, p.2396921.

10 Appendix

10.1 Appendix A – Relevant non-local detour route choice models

10.1.1 A.1: Multinomial Logit

The Multinomial Logit (MNL) choice model is derived from random utility theory. The deterministic utility of alternative $i \in R$ is V_i, and the random utility of alternative $i \in R$ is U_i such that $U_i = V_i + \varepsilon_i$, where the ε_i terms are the individually and identically distributed random variable error terms. Assuming individuals seek the alternative with highest utility, the probability that an individual selects alternative $i \in R$ is:

$$P_i = \Pr\big(U_i \geq U_j, \forall j \in R, j \neq i\big) = \Pr\big(V_i + \varepsilon_i \geq V_j + \varepsilon_j, \forall j \in R, j \neq i\big).$$

The defining characteristic of Logit models is that the random variable error terms assume a Gumbel distribution. Consequently, the choice probability relation for alternative $i \in R$ is:

$$P_i = \frac{e^{\theta V_i}}{\sum_{j \in R} e^{\theta V_j}},$$

where $\theta > 0$ is the Logit scaling parameter. The MNL model in the context of route choice supposes that the deterministic utility of route $i \in R$ is given by the negative of travel cost: $V_i = -c_i$, and thus:

$$P_i = \frac{e^{-\theta c_i}}{\sum_{j \in R} e^{-\theta c_j}}.$$

The MNL model assumes the route utilities are independent from one another, however routes with overlapping links share unobserved attributes, and the assumption that the random error terms are all independently and identically distributed is no longer valid.

10.1.2 A.2: Path Size Logit

The Path Size Logit (PSL) model (Ben-Akiva & Bierlaire, 1998) was developed to address the deficiency of the MNL model in its inability to capture the correlation between routes. To do this, PSL incorporates path size correction terms within the MNL choice probability function to penalise routes for sharing links with other routes. The PSL choice probability for route $i \in R$ is hence:

$$P_i = \frac{(\gamma_i)^\beta e^{-\theta c_i}}{\sum_{j \in R} (\gamma_j)^\beta e^{-\theta c_j}},$$

where $(\gamma_i)^\beta$ is the path size correction term for route $i \in R$. $\gamma_i \in (0,1]$ is the path size term for route $i \in R$ measuring route distinctiveness, and $\beta \geq 0$ is the path size scaling parameter scaling sensitivity to route distinctiveness. A distinct route with no shared links has a path size term equal to 1, resulting in no penalisation. Less distinct routes have smaller path size terms and incur greater penalisation. The PSL path size is as follows for route $i \in R$:

$$\gamma_i = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{k \in R} \delta_{a,k}}.$$

To dissect the path size term: each link a in route i is penalised (in terms of decreasing the path size term and hence the choice probability of the route) according to the number of routes in the choice set that also use that link $\left(\sum_{k \in R} \delta_{a,k}\right)$, and the significance of the penalisation for link a is weighted according to how prominent link a is in route i, i.e. the cost of route a in relation to the total cost of route i $\left(\frac{t_a}{c_i}\right)$.

10.1.3 A.3: Alternative Generalised Path Size Logit

An issue with the PSL model that was raised by Ramming (2002) and later explored further by Duncan et al. (2020) is that all routes in the choice set contribute equally to path size terms, regardless of how unrealistic they may be. As such, the correction terms of realistic routes and thus their choice probabilities are negatively affected by link sharing with unrealistic routes. To combat this, Ramming (2002) proposed the Generalised Path Size Logit (GPSL) model where the contribution of route k to the path size term of route i is weighted according to the ratio of travel cost between the two routes. As such, routes with excessively large travel costs have a diminished impact upon the correction terms of routes with small travel costs, and consequently the choice probabilities of those routes. Duncan et al. (2020) reformulated the GPSL model (proposing the alternative GPSL model (GPSL′)) so that the contribution weighting resembles the probability relation, providing internal consistency. The GPSL′ path size term is as follows for route $i \in R$:

$$\gamma_i = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{k \in R} \left(\frac{e^{-\theta c_k}}{e^{-\theta c_i}}\right) \delta_{a,k}}.$$

Therefore, if the cost of route k is greater than the cost of route i then $\left(\frac{e^{-\theta c_k}}{e^{-\theta c_i}}\right)$ is less than 1, and thus the contribution of route k to the path size term of route i is diminished. Note that the path size contribution factor can be generalised to be $\left(\frac{e^{-\lambda c_k}}{e^{-\lambda c_i}}\right)$, where $\lambda \geq 0$ is a path size contribution scaling parameter, providing flexibility and also collapsibility to the PSL model. In this study though we set $\lambda = \theta$ for improved internal consistency.

10.1.4 A.4: Bounded Choice Model

The Bounded Choice Model (BCM) (Watling et al., 2018) route choice principle is that travellers choose a route based on the probability of it having the best utility relative to a reference utility. By setting this reference utility equal to the maximum deterministic utility of all alternatives, the attractiveness of a route depends on the utilities of all routes, meaning that the BCM falls within the class of relative random utility theory.

The BCM is derived as follows. Define V_{r^*} as the reference utility of the reference alternative r^*, which in this case is the maximum utility alternative, i.e. $V_{r^*} = \max(V_r : r \in R)$. Thus, if U_i and V_i are the random and deterministic utilities for route $i \in R$, respectively, the difference in random utility for route $i \in R$ relative to the reference utility is:

$$U_{r^*} - U_i = V_{r^*} + \varepsilon_{r^*} - V_i - \varepsilon_i = V_{r^*} - V_i + \epsilon_i = \max(V_r : r \in R) - V_i + \epsilon_i,$$

where ε_i is the individually and identically distributed random variable error term for route $i \in R$, and ϵ_i is the difference random variable for route $i \in R$ with the reference alternative. The MNL model can be derived by assuming the ε_i error terms are Gumbel distributed and thus the ϵ_i difference random error terms assume the logistic distribution. The BCM, however, proposes that the difference random variable error terms ε_i assume a truncated logistic distribution, obtained by left-truncating a logistic distribution with mean 0 and scale θ^{-1} at a lower bound of $-\psi$ for some $\psi \geq 0$. As such, a bound is applied to the difference in utility to the reference utility, so that if a route has a utility below the bound, it

receives zero choice probability. This means that routes with utilities below the bound have zero probability of being the best alternative (relative to the reference utility).

Given the above, it follows from the derivation of the BCM in Duncan et al. (2022, Supplementary Material, Appendix A) that the choice probability relation for route $i \in R$ is:

$$P_i = \frac{\left(\exp\left(\theta\left(V_i - \max(V_r : r \in R) + \psi\right)\right) - 1\right)_+}{\sum_{j \in R}\left(\exp\left(\theta\left(V_j - \max(V_r : r \in R) + \psi\right)\right) - 1\right)_+}.$$

Supposing that the deterministic utility of route $i \in R$ is given by the negative of travel cost: $V_i = -c_i$, then:

$$P_i = \frac{\left(\exp\left(-\theta\left(c_i - \min(\boldsymbol{c}) - \psi\right)\right) - 1\right)_+}{\sum_{j \in R}\left(\exp\left(-\theta\left(c_j - \min(\boldsymbol{c}) - \psi\right)\right) - 1\right)_+}.$$

In this study, we suppose the bound is a relative bound on surplus total route travel cost, achieved by setting $\psi = (\varphi - 1) \cdot \min(c_r : r \in R)$, so that:

$$P_i = \frac{\left(\exp\left(-\theta\left(c_i - \varphi \min(\boldsymbol{c})\right)\right) - 1\right)_+}{\sum_{j \in R}\left(\exp\left(-\theta\left(c_j - \varphi \min(\boldsymbol{c})\right)\right) - 1\right)_+}.$$

where $\varphi > 1$ is the relative surplus cost bound parameter. In this case, a route receives zero probability if it has a cost as great as or greater than φ *times* the minimum route cost.

10.1.5 A.5: Bounded Path Size model

Like MNL, the BCM does not capture route correlations. To capture such, the BPS model (Duncan et al., 2022) incorporates path size correction factors within the BCM choice probability function. The BPS model is formulated as follows.[5] Let $\bar{R}(\boldsymbol{c}; \varphi) \subseteq R$ be the restricted choice set of all routes $i \in R$ where $c_i < \varphi \min(\boldsymbol{c})$. Given $\bar{R}$, the choice probability function for route $i \in R$ is:

$$P_i = \begin{cases} \dfrac{(\bar{\gamma}_i)^\beta \left(\exp\left(-\theta\left(c_i - \varphi \min(\boldsymbol{c})\right)\right) - 1\right)}{\sum_{j \in \bar{R}} (\bar{\gamma}_j)^\beta \left(\exp\left(-\theta\left(c_j - \varphi \min(\boldsymbol{c})\right)\right) - 1\right)} & if\ i \in \bar{R} \\ 0 & if\ i \notin \bar{R} \end{cases},$$

where the BPS path size term for route $i \in \bar{R}$ is:

$$\bar{\gamma}_i = \sum\nolimits_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{k \in \bar{R}} \left(\dfrac{\left(\exp\left(-\theta\left(c_k - \varphi \min(\boldsymbol{c})\right)\right) - 1\right)}{\left(\exp\left(-\theta\left(c_i - \varphi \min(\boldsymbol{c})\right)\right) - 1\right)} \right) \delta_{a,k}}.$$

The BPS path size term is formulated as such to ensure that the path size term function is continuous, including as a route enters and exits $\bar{R}$ as its cost crosses from below-to-above and above-to-below the bound. And, to avoid occurrences of $\frac{0}{0}$. See Duncan et al. (2022) for more details.

10.2 Appendix B – Demonstrations of model features

In this section, we conduct some numerical experiments on a small example network to demonstrate the features of the BPS-LDT model, including comparing results with relevant non-local detour models (which can be found in Appendix A).

For these demonstrations, consider example network 1 in Fig. 13 where there are 5 routes: route 1: $1 \to 2 \to 9$, route 2: $1 \to 3 \to 9$, route 3: $1 \to 4 \to 5 \to 8 \to 9$, route 4: $1 \to 4 \to 6 \to 8 \to 9$, and route 5: $1 \to 4 \to 7 \to 8 \to 9$. Routes 1&2 are completely distinct while routes 3-5 overlap to a high degree since they all share a large proportion of their travel cost. Routes 3-5 differ only according to a single local detour: if route 3 is for example the main road, then routes 4&5 offer a detour from the main road, where route 5 has a greater detour than route 4. As shown in Fig. 13, route 1 has a total travel cost of 3, route 2 has a cost of 1, and routes 3-5 have costs 1.01, $1+\rho$ ($0.01 \leq \rho \leq 0.05$), and 1.05, respectively.

[5] Note that in Duncan et al. (2022) two versions of a BPS model are proposed: the Bounded Bounded Path Size model and the Bounded Adaptive Path Size model, but in this paper we focus on the Bounded Bounded Path Size model as it is conveniently closed-form, and refer to it just as the BPS model.

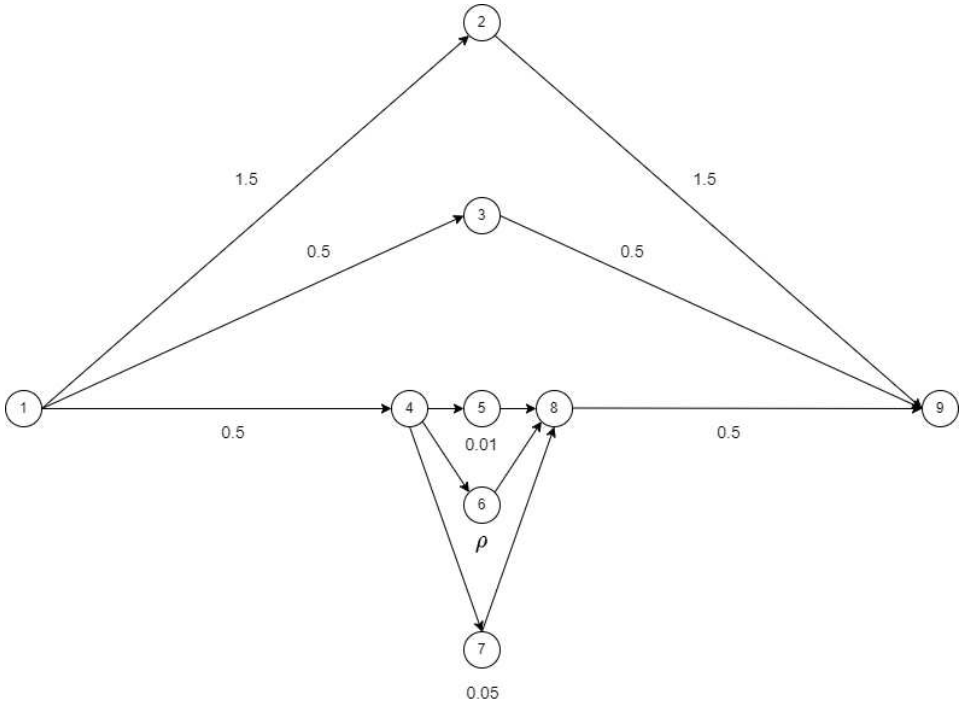

Fig. 13. Example network 1.

10.2.1 B.1: Demonstration 1 – the local detour measure

To make the results presented later easier to understand, we shall begin by walking the reader through the computation of the local detour measures for example network 1, (see Rasmussen et al. (2024) for other demonstrations and examples on this). In example network 1, there are 4 segments (u, v) that have multiple segment alternatives: (1,9), (1,8), (4,8), & (4,9). Segments (1,8) & (4,9) are redundant, however, as (4,8) will always be a dominant segment, i.e. it will always have a greater detour measure than (1,8) & (4,9). This is because all segment alternatives for segments (1,8) & (4,9) share a common link, and thus the segment (4,8) consisting of the non-overlapping part will always have a greater detour measure than the full segment, see the discussion around Fig. 2 for more details. We thus disregard segments (1,8) & (4,9). (1,9) has all 5 routes as the segment alternatives and (4,8) have routes 3-5 as the segment alternatives.

There are five segment alternatives for segment (1,9) that consist of routes 1-5, respectively. The index set of segment alternatives for segment (1,9) is thus $K_{1,9} = \{1,2,3,4,5\}$ and the used segment alternatives of each route 1-5 at (1,9) are $k_{1,9,1} = 1$, $k_{1,9,2} = 2$, $k_{1,9,3} = 3$, $k_{1,9,4} = 4$, & $k_{1,9,5} = 5$. The costs of the segment alternatives $l \in K_{1,9}$, ω_l, are: $\omega_1 = 3$, $\omega_2 = 1$, $\omega_3 = 1.01$, $\omega_4 = 1 + \rho$, & $\omega_5 = 1.05$. Denote $\phi_{u,v,i}$ as the local detouredness of route i at segment $(u, v) \in S_i$. The local detourednesses of routes 1-5 at segment (1,9) are thus:

$$\phi_{1,9,1} = \frac{\omega_{k_{1,9,1}} - \min(\omega_l : l \in K_{1,9})}{\min(\omega_l : l \in K_{1,9})} = \frac{\omega_1 - \min(\omega_l : l \in \{1,2,3,4,5\})}{\min(\omega_l : l \in \{1,2,3,4,5\})} = \frac{3 - \min(\{3,1,1.01,1.05\})}{\min(\{3,1,1.01,1.05\})} = \frac{3-1}{1} = 2,$$

$$\phi_{1,9,2} = \frac{1-1}{1} = 0, \qquad \phi_{1,9,3} = \frac{1.01-1}{1} = 0.01, \qquad \phi_{1,9,4} = \frac{1+\rho-1}{1} = \rho, \qquad \phi_{1,9,5} = \frac{1.05-1}{1} = 0.05.$$

Notice that in the context of global bounds of total route travel cost, the relative costliness of routes 1-5 compared to the cheapest alternative are 3,1, 1.01, 1+ρ, & 1.05, respectively, i.e. route 1 is 3 times more costly than the cheapest route. Local detouredness is always equal to relative costliness minus 1.

For segment (4,8), there are three segment alternatives: segment alternative 1 is 4→5→8 (used only by route 3), segment alternative 2 is 4→6→8 (used only by route 4), and segment alternative 3 is 4→7→8 (used only by route 5). The index set of segment alternatives between segment (4,8) is thus $K_{4,8} = \{1,2,3\}$ and the used segment alternatives of each route 3-5 at segment (4,8) are $k_{4,8,3} = 1$, $k_{4,8,4} = 2$, & $k_{4,8,5} = 3$. The costs of the segment alternatives $l \in K_{4,8}$ are: $\omega_1 = 0.01$, $\omega_2 = \rho$, $\omega_3 = 0.05$. The local detourednesses of routes 3-5 at segment (4,8) are thus:

$$\phi_{4,8,3} = \frac{\omega_{k_{4,8,3}} - \min(\omega_l : l \in K_{4,8})}{\min(\omega_l : l \in K_{4,8})} = \frac{\omega_1 - \min(\omega_l : l \in \{1,2,3\})}{\min(\omega_l : l \in \{1,2,3\})} = \frac{0.01 - \min(\{0.01,0.05\})}{\min(\{0.01,0.05\})} = \frac{0.01-0.01}{0.01} = 0,$$

$$\phi_{4,8,4} = \frac{\rho - 0.01}{0.01} = \frac{\rho}{0.01} - 1, \qquad \phi_{1,9,3} = \frac{0.05-0.01}{0.01} = 4.$$

The measure of local detouredness for each route 1-5 are thus:

$$\phi_1 = \max\left\{\phi_{u,v,1}: (u,v) \in \{(1,2),(2,9),(1,9)\}\right\} = \max\{\phi_{1,2,1}, \phi_{2,9,1}, \phi_{1,9,1}\} = \max\{0,0,2\} = 2,$$

$$\phi_2 = \max\{0\} = 0$$

$$\phi_3 = \max\{0, \phi_{1,9,3}, \phi_{4,8,3}\} = \max\{0,0.01,0\} = 0.01,$$

$$\phi_4 = \max\{0, \phi_{1,9,4}, \phi_{4,8,4}\} = \max\left\{0, \rho, \frac{\rho}{0.01} - 1\right\} = \begin{cases} \rho & if\ \rho \leq \frac{1}{99} \\ \frac{\rho}{0.01} - 1 & otherwise \end{cases},$$

$$\phi_5 = \max\{0, \phi_{1,9,5}, \phi_{4,8,5}\} = \max\{0,0.05,4\} = 4.$$

10.2.2 B.2: Demonstration 2 – comparison of features with other models

We shall now demonstrate the features of the BPS-LDT model compared to other relevant models. Upon inspection of the route travel costs, it would appear that route 1 is potentially an unrealistic alternative as it is 3 times more costly than the cheapest alternative, but there is little difference between routes 2-5. Fig. 14A displays BCM choice probabilities for $\rho \in [0.01,0.05]$, with $\theta = 1$ & $\varphi = 2$. As shown, since the relative cost of route 1 compared to the cheapest route (route 2) is greater than the bound (i.e. a relative costliness of 3 compared to the bound 2), route 1 receives a zero probability. And, since the total costs of routes 2-5 are negligibly different, they all receive approximately $\frac{1}{4}$ probability each.

However, due to the fact that: a) routes 3-5 are highly correlated, and b) route 5 (and potentially route 4 depending on ρ) has a large local detour (a measure of 4, see above) and thus may be considered unrealistic, these BCM probabilities are potentially inaccurate.

Addressing deficiency a), the BPS model was developed to capture correlations between overlapping routes with costs below the BCM cost bound. Fig. 14B displays the BPS model choice probabilities for $\rho \in [0.01,0.05]$, with $\theta = 1$, $\beta = 0.8$, & $\varphi = 2$. As shown, route 1 still receives a zero probability as it has a travel cost above the bound, but now routes 3-5 have reduced probabilities (yet still relatively equal) compared to route 2 as they have been penalised to capture their correlation.

Addressing deficiency b), the BCM-LDT was developed to exclude routes with large local detours, in a way that is consistent and mathematically well-defined (i.e. with a continuous probability function). Fig. 14C displays BCM-LDT choice probabilities for $\rho \in [0.01,0.05]$, with $\theta_1 = 1$, $\theta_2 = 0.1$, $\varphi = 2$, & $\eta = 3.5$. As shown, route 1 has zero probability, and, since the local detouredness of route 5 is above the local detour threshold (i.e. a local detouredness measure of 4 compared to the threshold 3.5), route 5 now also receives a zero probability. As also shown, at $\rho = 0.01$, routes 2-4 have approximately $\frac{1}{3}$ probability each since they all have similar travel costs and route 4 has an equal detour to route 3 at segment $(4,8)$, i.e. $\phi_{4,8,3} = \phi_{4,8,4} = 0$. However, as ρ increases, the local detouredness of route 4 increases; therefore, although the total travel cost of route 4 remains similar to routes 2&3, the probability of route 4 decreases, up until $\rho = 0.045$ where the local detouredness measure of route 4 is equal to the threshold: $\phi_4 = \frac{0.045}{0.01} - 1 = 3.5$ (see above), where it then receives zero probability.

So, the BPS model addresses deficiency a) and the BCM-LDT addresses deficiency b), but neither model addresses both deficiencies. The BPS-LDT model is thus developed in this paper to address both deficiencies. Fig. 14D displays the BPS-LDT choice probabilities for $\rho \in [0.01,0.05]$, with $\theta_1 = 1$, $\theta_2 = 0.1$, $\beta = 0.8$, $\varphi = 2$, & $\eta = 3.5$. As shown, routes 3&4 have equal probability at $\rho = 0.01$ but route 2 has a greater probability. This is because routes 3&4 have the same cost and local detour measure, but are highly correlated compared to the similar costing but distinct route 2. As ρ increases and the probability of route 4 decreases and tends towards 0 (due to its local detouredness increasing towards the threshold), the contribution of route 4 to the realistic route path size terms of route 3 decreases. At $\rho = 0.045$, route 4 becomes defined as an unrealistic route due to its local detouredness being exactly at the threshold; it is thus assigned a zero probability and no longer contributes to the path size terms of route 3. Thus, for $\rho \geq 0.045$, since routes 1,4,5 are defined as unrealistic by the cost bound / local detour threshold criteria, the network is in-effect reduced to just two routes (2&3) which are distinct and similar costing.

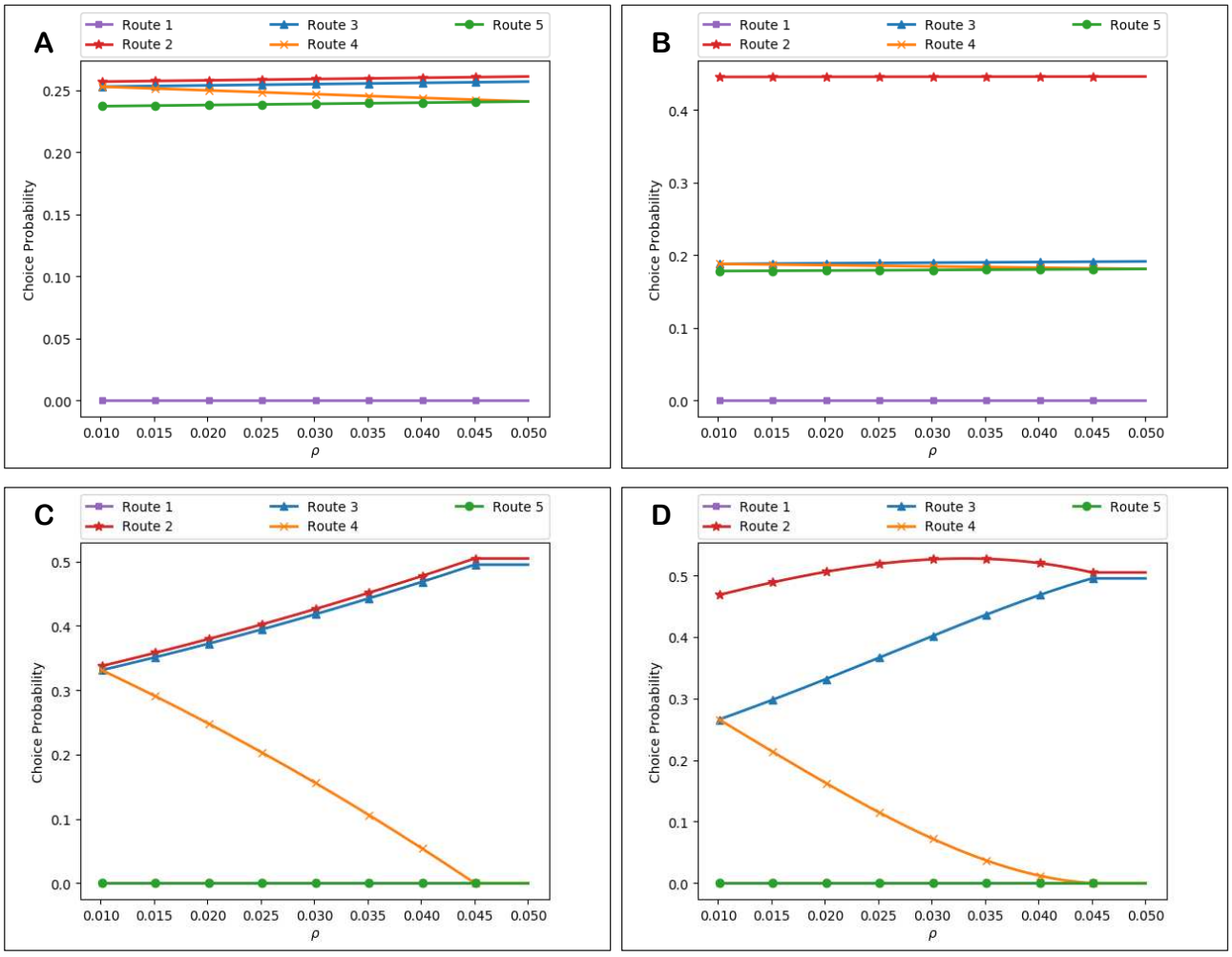

Fig. 14. Example network 1: Choice probabilities from the different models as ρ is varied ($\theta_1 = 1$, $\theta_2 = 0.1$, $\varphi = 2$, & $\eta = 3.5$). **A:** BCM. **B:** BPS. **C:** BCM-LDT. **D:** BPS-LDT.

10.2.3 B.3: Demonstration 3 – model parameters

The BPS-LDT model has five standard parameters (there will likely be more for multiple travel cost attributes, see e.g. Section 6.2.1), which have complex interactions. In this section, we shall thus demonstrate how the model behaves according to the different parameters. To help the reader, Table 9 provides a summary description of the five standard BPS-LDT model parameters.

Parameter	Name	Description
θ_1	Travel cost scaling parameter	Controls sensitivity to travel cost in the cost-BCM component. A small θ_1 value corresponds to drivers being less aware of / less sensitive to differences in route travel cost, and a large θ_1 corresponds to the opposite.
θ_2	Local detouredness scaling parameter	Controls sensitivity to local detouredness in the detour-BCM component. A small θ_2 value corresponds to drivers being less aware of / less sensitive to differences in local detouredness, and a large θ_2 corresponds to the opposite.
β	Path size scaling parameter	Scales the path size correction factor. A small β value corresponds to drivers being less aware of route correlation / less sensitive to route distinctiveness, and a large β corresponds to the opposite.

φ	Relative surplus total route travel cost bound parameter	The bound that drivers have in terms of the maximum excess travel cost they are willing to consider compared to the cheapest route. A route is only considered realistic if it has a cost less than φ times the cost on the cheapest route.
η	Local detour threshold parameter	The threshold that drivers have in terms of the maximum local detour they are willing to consider. A route is considered unrealistic if it has a used segment alternative that costs greater than $100 \cdot \eta\%$ more than cheapest segment alternative.

Table 9. Summary description of the BPS-LDT model parameters.

Fig. 15 displays BPS-LDT choice probabilities for varying η, with $\rho = 0.03$, $\theta_1 = 1$, $\theta_2 = 0.1$, $\beta = 0.8$, & $\varphi = 2$. As shown, route 1 has a zero probability for all η since the cost bound criterion defines it as an unrealistic route. As η is decreased (from 10) the local detouredness measures of routes 4&5 (=2 & =4 respectively) approach the local detour threshold from below. Their probabilities as well as path size contributions thus decrease until their detouredness meets the threshold where they then – in as continuous manner – are assigned zero probabilities as well as zero path size contributions to overlapping realistic routes (i.e. the contribution of route 5 to the path size terms of routes 3&4 become eliminated, and the contribution of route 4 to the path size terms of route 3 become eliminated.

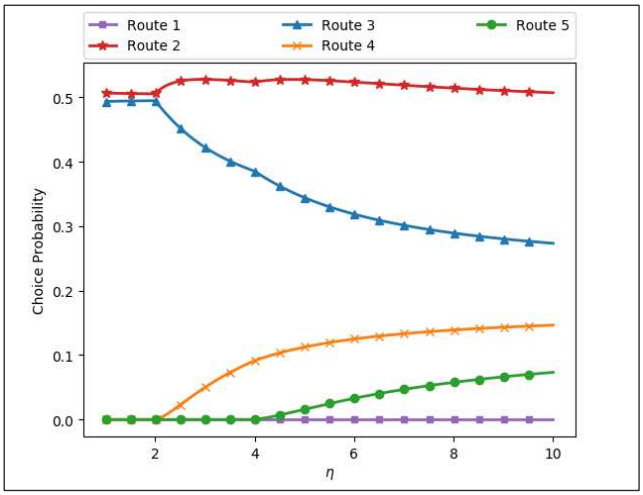

Fig. 15. Example network 1: Choice probabilities from the BPS-LDT model as the local detour threshold η parameter is varied ($\rho = 0.03$, $\theta_1 = 1$, $\theta_2 = 0.1$, $\beta = 0.8$, & $\varphi = 2$).

Fig. 16 displays BPS-LDT choice probabilities for varying φ, with $\rho = 0.02$, $\theta_1 = 0.1$, $\theta_2 = 0.1$, $\beta = 0.8$, & $\eta = 2.5$. As shown, route 5 has a zero probability for all φ as it has a detouredness measure greater than local detour threshold (i.e. a measure of 4 compared to the 2.5 threshold). Route 4, on-the-other-hand, has a non-zero probability as it has a detouredness measure below the threshold (i.e. a measure of 1 compared to the 2.5 threshold). Routes 3-5 have total travel costs well below the cost bound in this range (i.e relative costs of 1.01, 1.02, & 1.05, respectively, compared to the bound $\varphi \in [2.5,6]$), and therefore are negligibly affected by φ. Route 1, however, is affected. In this case, route 1 has a detouredness measure below the threshold, i.e. the measure taken from the global detour (from origin to destination) is 2 compared to the 2.5 threshold – it is therefore not defined as an unrealistic route (and assigned a zero probability) by the local detour criteria. However, the cost bound criteria can still define the route as unrealistic, depending on φ. As shown in Fig. 16, as φ is decreased, the relative cost of route 1 (=3) tends toward the bound and therefore decreases in probability up until $\varphi = 3$ where route 1's relative cost meets the bound and it is assigned zero probability. Note here that route 1 is a distinct route and therefore does not contribute to the path size terms of other routes, but if it did its contribution to those overlapping routes would decrease as its cost approaches the bound (to then be eliminated at the bound). Suppose alternatively that the local detour threshold was instead $\eta = 1.5$, then route 1 would have a global detour (=2) above the threshold. Route 1 would thus be assigned a zero probability regardless of the relative cost bound φ.

In terms of defining routes as unrealistic, it does not make sense for φ to be greater than $\eta + 1$. This is because a local detour threshold of $\eta + 1$ is equivalent to a cost bound of φ in terms of the routes that will be defined as unrealistic (the local detour measure will always consider the global detour). Therefore, φ values greater than $\eta + 1$ will not define

any more routes as unrealistic than the $\eta + 1$ threshold will. However, φ values greater than $\eta + 1$ will still affect the choice probabilities of routes defined as realistic (by the local detour threshold criteria).

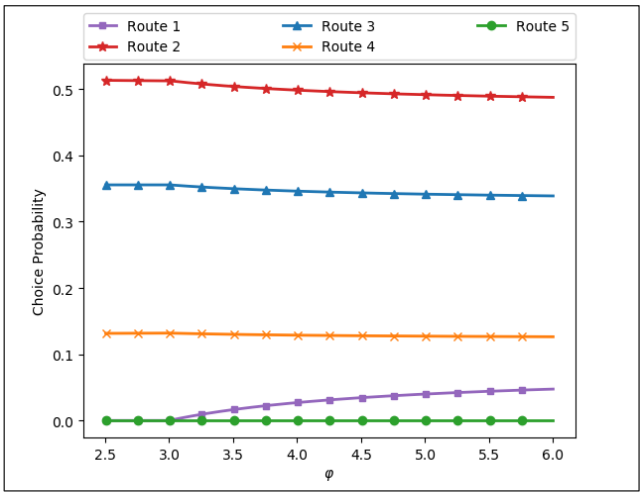

Fig. 16. Example network 1: Choice probabilities from the BPS-LDT model as the relative surplus travel cost bound parameter φ is varied ($\rho = 0.03$, $\theta_1 = 0.1$, $\theta_2 = 0.1$, $\beta = 0.8$, & $\eta = 2.5$).

Fig. 17 display BPS-LDT choice probabilities for varying θ_1, with $\rho = 0.02$, $\theta_2 = 0.1$, $\beta = 0.8$, & $\varphi = \eta = 3.5$. As shown, route 5 has zero probability as it has a detouredness measure (=4) above the threshold (=3.5), but route 1 has a non-zero probability as it has a relative cost (=3) below the bound (=3.5). As per the parameter description in Table 9, for smaller θ_1 drivers are less sensitive to / aware of the differences in travel cost and route 1 (the route with the only/most significant cost difference) increases in probability. For larger θ_1, drivers are more sensitive to / aware of the differences and route 1 decreases in probability (since it has the greatest travel cost among the realistic routes), and approaches zero probability (but does not reach it).

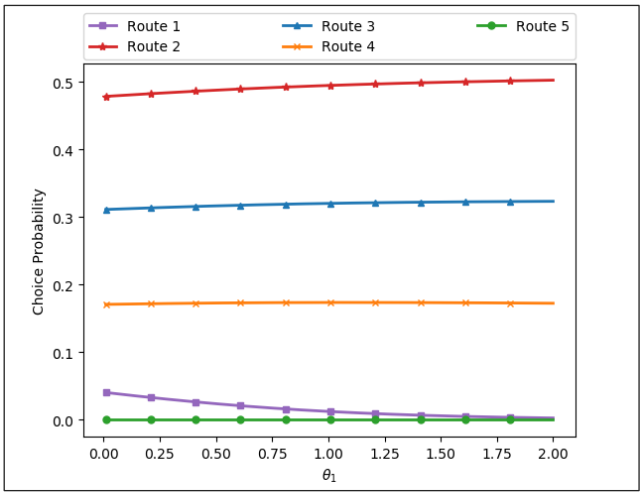

Fig. 17. Example network 1: Choice probabilities from the BPS-LDT model as the travel cost scaling parameter θ_1 is varied ($\rho = 0.02$, $\theta_2 = 0.1$, $\beta = 0.8$, & $\varphi = \eta = 3.5$).

Fig. 18 displays BPS-LDT choice probabilities for varying θ_2, with $\rho = 0.02$, $\theta_1 = 1$, $\beta = 0.8$, $\varphi = 2$, & $\eta = 8$. As shown, route 1 has zero probability as it has a relative cost (=3) above the bound (=2), but route 5 has detouredness measure (=4) below the threshold (=8). As per the parameter description in Table 9, for smaller θ_2 drivers are less sensitive to / aware of the differences in local detouredness and routes 3-5 (the main detour routes) have closer probabilities. For larger θ_2, drivers are more sensitive to / aware of the differences and routes 4&5 decrease in probability since they have greater detour measures than route 3.

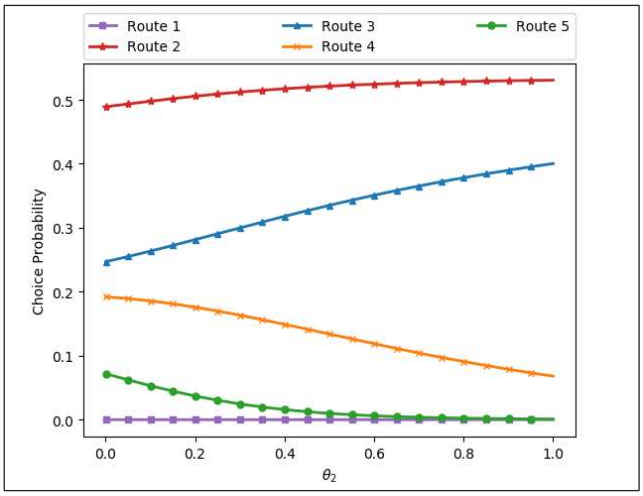

Fig. 18. Example network 1: Choice probabilities from the BPS-LDT model as the local detouredness scaling parameter θ_2 is varied ($\rho = 0.02, \theta_1 = 1, \beta = 0.8, \varphi = 2, \& \eta = 8$).

Lastly, Fig. 19 displays BPS-LDT choice probabilities for varying β, with $\rho = 0.02, \theta_1 = 1, \theta_2 = 0.1, \varphi = 2, \& \eta = 3.5$. As shown, routes 1&5 have zero probability as route 1 has a relative cost (=3) above the bound (=2) and route 5 has a local detour measure (=4) above the threshold (=3.5). At $\beta = 0$, the BCM-LDT probabilities are plotted, where route 3 is not penalised for overlapping with route 4 and thus has a similar probability to route 2 (due to similar travel costs and local detour measures). As β increases, however, route 3 becomes increasingly penalised for overlapping with route 4.

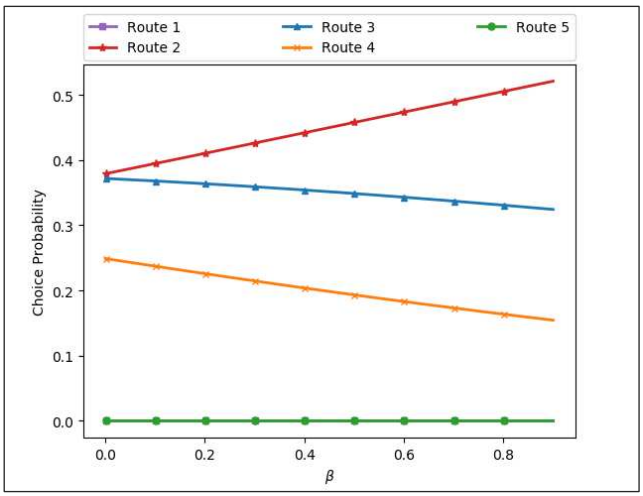

Fig. 19. Example network 1: Choice probabilities from the BPS-LDT model as the path size scaling parameter β is varied ($\rho = 0.02$, $\theta_1 = 1, \theta_2 = 0.1, \varphi = 2, \& \eta = 3.5$).