A closed-form bounded route choice model accounting for heteroscedasticity, overlap, and choice set formation

Laurent Cazor[*,a], Lawrence Christopher Duncan[a], David Paul Watling[b], Otto Anker Nielsen[a], and Thomas Kjær Rasmussen[a]

[*] *Corresponding author (lauca@dtu.dk)*
[a] *Technical University of Denmark, Department of Technology, Management and Economics, Akademivej, Building 358, 2800 Kgs. Lyngby, Denmark*
[b] *University of Leeds, Institute for Transport Studies, 36-40 University Road, Leeds LS2 9JT, UK*

Abstract

The Multinomial Logit (MNL) model is widely used in route choice modelling due to its simple closed-form choice probability function. However, MNL assumes that the error terms are independently and identically distributed with infinite support. As a result, it imposes homoscedasticity, meaning that long and short trips share the same error variance, disregards correlations between overlapping routes, and assigns non-zero choice probabilities to all available routes, regardless of their cost. This paper addresses these limitations by developing a closed-form route choice model. We introduce the Bounded q-Product Logit (BqPL) model, which incorporates heteroscedastic error terms with bounded support. The parameter q controls the rate at which error term variance increases with trip cost, and routes that violate cost bounds receive zero choice probabilities, implicitly defining the route choice set. Furthermore, we extend the BqPL model to account for correlations between overlapping routes by integrating path size correction terms within the choice probability function, resulting in the Bounded Path Size q-Product Logit (BPSqPL) model. We illustrate the properties of the BPSqPL model on small-scale networks, contrasting it with a range of existing choice models into which it can collapse. We then present a method to estimate the model parameters and standard errors, using bootstrapping. Finally, we estimate the model using a large-scale bicycle route choice case study, comparing its goodness-of-fit, interpretability, and forecasting ability with relevant collapsing models. We also test the impact of the choice set size on the estimated parameters. The results underscore the importance of addressing the three key limitations of the MNL model and demonstrate the effectiveness of the BPSqPL model in doing so.

Keywords: route choice, heteroscedasticity, route overlap, choice set formation, large-scale application

Highlights

- Derivation of a heteroscedastic route choice model implicitly determining considered routes
- Extension to consistently account for route overlap through integrating path size terms
- Demonstration of model properties, contrasting with existing models
- Parameter estimation and interpretation in a large-scale bicycle route choice case study
- Comparison of fit, predictive ability, and behavioural implications with existing models

1 Introduction

The Multinomial Logit (MNL) is a Random Utility Model (RUM) (McFadden (1974)) which is widely used in route choice applications due to its simple closed-form choice probability function that is quick and easy to compute and estimate. MNL makes, however, several questionable assumptions of the random error terms, leading to some behavioural limitations:

- *Homoscedasticity of the error terms:* MNL assumes that the error terms for all routes of all OD movements are identically distributed with the same variance, i.e. the error terms are homoscedastic. However, as Bhat (1995) argues, the amount and variability of unobserved characteristics of routes is likely to differ from one route to another, both of routes *within a choice set* and *between choice sets* (Munizaga et al. 2000). Longer routes, for instance, are likely to have greater variability in their unobserved characteristics than shorter routes, meaning there should be greater variance in their error terms.

- *Independence of the error terms:* MNL assumes that the error terms for all routes are independently distributed. Error terms are likely to be correlated, however, as routes tend to be highly overlapping, sharing links and their unobserved characteristics with many other routes (Ben-Akiva and Bierlaire 1999; Cascetta et al. 1996). It is thus unrealistic to assume that there is no covariance between the error terms, as there is likely to be a correlation between the unobserved attributes of routes.

- *Non-bounded support of the error terms:* The MNL model assumes that error terms have infinite support, meaning that, theoretically, every possible alternative, no matter how improbable, has a non-zero probability of being chosen. This assumption is often unrealistic, especially in practical scenarios like route choice, where travellers typically avoid highly circuitous or inefficient routes. Real-life networks contain many such routes that are rarely, if ever, considered (Bovy 2009; Duncan et al. 2022). Consequently, the MNL struggles to account for the natural process of choice set formation, where only a subset of alternatives is genuinely considered by decision-makers.

Numerous route choice modelling approaches have been developed to address each of these limitations.

- *Accounting for heteroscedasticity:* Bhat (1995) developed a heteroscedastic logit model where the error term for each route has a different scale. However, the model is not closed-form and is computationally expensive to compute in real-life networks. Other models building on Mixed Logit or Multinomial Probit have also been developed to account for heterogeneity in error term variance (e.g., Ben-Akiva and Bierlaire (1999); Greene et al. (2006)). These models suffer from large estimation times and potential confounding effects between cost coefficient correlation and heteroscedasticity among alternatives (Hess and Train 2017). Among closed-form models, Chen et al. (2012) addresses the heteroscedasticity issue of MNL by modifying the Logit scaling parameter – which controls the variance of the error terms – to relate to the distance of the OD movement. While this model captures the heteroscedasticity of error terms across between choice sets, it cannot do so within alternatives of a choice set. Castillo et al. (2008) developed the closed-form Multinomial Weibit (MNW) model where the error term standard deviation grows linearly with travel cost. Thus, MNW captures the heteroscedasticity of route error terms across and within OD movements. Fosgerau and Bierlaire (2009) derive MNW in an alternative way through multiplicative random error terms. Chikaraishi and Nakayama (2016) developed the q-Product Logit (qPL) model that generalises both MNL and MNW, where one can control the rate at which error term variance increases with travel cost with a parameter q. Setting $q = 0$ results in the MNL model and $q = 1$ results in the MNW model. Yao and

Chen 2014 and Xu et al. 2015 also developed Hybrid Logit-Weibit models that balance MNL and MNW, but these lack a strong behavioural foundation.

- *Accounting for route overlap:* Many models have been developed to address the overlap issue, which is often categorised into Generalised Extreme Value (GEV) structure models (e.g. Nested or Cross Nested Logit (Vovsha 1997), Paired Combinatorial Logit (Koppelman and Wen 2000)), simulation models (e.g., Mixed Logit (Bekhor et al. 2002), Probit (Yai et al. 1997)), and correction-term models (C-Logit (Cascetta et al. 1996), Path Size Logit (Ben-Akiva and Bierlaire 1999)). A detailed review of correlation-based route choice models can be seen in Duncan et al. (2020). Correction-term models, like MNL, feature simple closed-form choice probability functions that are easy to compute and estimate, making them the most commonly used in practice.

- *Accounting for choice set formation:* The assumption of unbounded support for the MNL error terms presents two key issues. The first is computational: enumerating the universal set of routes between an origin and destination on a large-scale network is infeasible. As a result, route choice models typically rely on sampling protocols to generate a realistic subset of considered routes (see Prato (2009) for a review). However, these sampling algorithms often use criteria that are inconsistent with those used for calculating route choice probabilities, leading to biased estimates Frejinger et al. 2009; Frejinger and Bierlaire 2010. The second issue is behavioural: travellers cannot process the universal set of routes and are likely to use heuristics to build their consideration set Simon 1955. Failing to account for the effects of choice set formation is also known to bias estimates Williams and Ortuzar 1982. To address this, the concept of one-stage choice set formation models has been applied to route choice modelling (e.g., Cascetta and Papola 2001; Martínez et al. 2009 with the Implicit Availability/Perception and Constrained MNL models, respectively). These models introduce the concept of fuzzy choice sets and penalise the deterministic cost of alternatives that do not meet non-compensatory cutoffs. However, this approach is inconsistent with Horowitz and Louviere (1995)'s finding that the same preferences should drive both choice set formation and choice probability determination. Additionally, these models are 'soft', meaning violating routes only receive reduced probabilities rather than strictly zero. Their application to large-scale data has also struggled to produce significant network-related estimates Ramming 2002. A recent stream of models addresses choice set formation implicitly by modifying the distributional assumptions of error terms. Specifically, by bounding their support, Watling et al. (2018) with the Bounded Logit (BL) model[1] and Tan et al. (2024) with the truncated path choice model, these models exclude alternatives from the choice sets (assign them zero probabilities) if their systematic cost is not within a bound set on cost. Moreover, these approaches do not impose additional assumptions on the decision rule used for choice set formation, aligning with Horowitz and Louviere (1995)'s observation that the same preferences should drive both processes.

Some models have been developed to jointly address two of the three main shortcomings of the MNL model. For instance, the Path Size Weibit (Kitthamkesorn and Chen 2013), Nested Weibit (Gu et al. 2022), and Cross-Nested q-Product Logit models (Chikaraishi and Nakayama 2016) are designed to handle both heteroscedasticity and route correlation. Similarly, the Bounded Path Size Logit (BPSL) model (Duncan et al. 2022) accounts for route overlap and choice set formation, though it does not address heteroscedasticity. However, to the best of our knowledge, no existing route choice model addresses all three of these deficiencies simultaneously.

In this paper, we tackle this gap by developing a new closed-form route choice model that integrates several concepts from the literature. We begin by building on Chikaraishi and Nakayama

[1]This model was originally termed the Bounded Choice Model but is referred to here as the Bounded Logit model to distinguish it from other bounded models, such as the Bounded Weibit and Bounded q-Product Logit models.

(2016)'s q-Product Logit (qPL) model and Watling et al. (2018)'s Bounded Logit (BL) model to derive the Bounded q-Product Logit (BqPL) model. In this model, the error terms are assumed to follow a novel truncated q-log-logistic distribution, which is heteroscedastic and has bounded support. The parameter q controls the rate at which error term variance increases with route cost, and routes that exceed a cost bound receive zero choice probabilities, thus implicitly determining the route choice set.

Subsequently, drawing on Duncan et al. (2022)'s Bounded Path Size Logit (BPSL) model, we extend the BqPL model to account for correlations between overlapping considered routes by incorporating appropriately-defined path size correction terms into the choice probability function. The resulting Bounded Path Size q-Product Logit (BPSqPL) model is a unified framework that can collapse into various existing and new models, including path size and/or bounded logit, weibit, and qPL variants.

The structure of the paper is as follows: Section 2 presents the theoretical background and derivation of the BPSqPL model, along with a discussion of its collapsing properties. Section 3 demonstrates the model's theoretical properties by contrasting it with its collapsing models. In Section 4, we describe a method to estimate the BPSqPL parameters and their standard errors with observed discrete choice data. In Section 5, we estimate the models in a large-scale bicycle route choice case study, comparing goodness-of-fit, interpretability, and forecasting performance. We also test the robustness of the estimated parameters with respect to the choice set size. Finally, Section 6 concludes with a summary, discussion, and suggestions for future research.

Model acronyms used throughout the paper are presented in Table 1.

Acronym	**Meaning**
MNL/MNW	Multinomial Logit/Weibit
qPL	q-Product Logit
BL/BW	Bounded Logit/Weibit
BqPL	Bounded q-Product Logit
BPSL/BPSW	Bounded Path Size Logit/Weibit
BPSqPL	Bounded Path Size q-Product Logit
GPSL/GPSW	Generalized Path Size Logit/Weibit
GPSqPL	Generalized Path Size q-Product Logit

Table 1: Table of acronyms

2 Model derivation and properties

In this section, we will derive the Bounded q-Product Logit (BqPL) model and its extension to account for route overlap consistently: the Bounded Path Size q-Product Logit (BPSqPL) model. These models utilise a generalisation of the relation between the systematic and stochastic components of the random cost of alternatives, which may vary between the sum and the product depending on the value of a model parameter q. This assumption allows the variance of the random costs to increase with their expected value, and q influences this growth rate. We will first introduce in Section 2.1 some preliminary modelling definitions we will use to derive the BqPL model. In Section 2.2, we shall then derive the BqPL model, following a method adapted from Watling et al. (2018) and Duncan et al. (2022) for deriving the BL model. Special cases of the BqPL model are discussed in Section 2.3. Taking inspiration from how Duncan et al. (2022) formulate the BPSL model, in Section 2.5 we integrate analogous path size correction terms within the BqPL probability function to formulate the BPSqPL model. In Section 2.6, we summarise the collapsing properties of the different presented models and their properties.

2.1 Preliminaries

Here, we introduce some preliminary modelling definitions used to derive the BqPL model. In Section 2.1.1 we define a set of "q-operators", a set of functions used to formulate the qPL model in Chikaraishi and Nakayama (2016) that we shall also utilise here. These q-operators allow for one to establish a more complex relationship between the systematic and random components of route costs. In Section 2.1.2, we introduce the probability distribution we will assume of error terms to derive the BqPL model. distributions.

2.1.1 q-Operators

Here, we present some q-operators. First, the "q-product" (Borges 2004) can be seen as an in-between of the sum and the product. The $q \in [0, 1]$ parameter controls the closeness to the sum or product. It is defined for $a > 0$ and $b > 0$ such that $a^{1-q} + b^{1-q} - 1 > 0$ as:

$$a \otimes_q b = \left(a^{1-q} + b^{1-q} - 1\right)_+^{\frac{1}{1-q}} \tag{1}$$

where $(.)_+ = \max(0, .)$. Its limiting cases are $\lim_{q \to 1} a \otimes_q b = ab$ and $a \otimes_0 b = a + b - 1$. We can similarly define the "q-ratio" as the inverse operator of the q-product:

$$a \oslash_q b = \left(a^{1-q} - b^{1-q} + 1\right)_+^{\frac{1}{1-q}} \tag{2}$$

Its limiting cases are $\lim_{q \to 1} a \oslash_q b = a/b$ and $a \oslash_0 b = a - b + 1$. It is the inverse of the q-product because $a \otimes_q (1 \oslash_q a) = 1$. Additionally, for any positive real number x, we define the "q-logarithm" (Tsallis 1994) as:

$$\ln_q(x) = \begin{cases} \dfrac{x^{1-q} - 1}{1-q} & \text{if } q \neq 1 \\ \ln(x) & \text{if } q = 1 \end{cases} \tag{3}$$

Notably, $\ln_0(x) = x - 1$. We have that $\ln_q(a \otimes_q b) = \ln_q(a) + \ln_q(b)$. The inverse function of the q-logarithm is given by the "q-exponential":

$$\exp_q(x) = \begin{cases} (1 + (1-q)x)_+^{\frac{1}{1-q}} & \text{if } q \neq 1 \\ \exp(x) & \text{if } q = 1 \end{cases} \tag{4}$$

We have the property that, for any $x > 0$, $\exp_q(\ln_q(x)) = \ln_q(\exp_q(x)) = x$.

These functions retain the morphism properties of the classic exponential and logarithms, e.g., $\ln_q(a \oslash_q b) = \ln_q(a) - \ln_q(b)$, $\ln_q(a \otimes_q b) = \ln_q(a) - \ln_q(b)$, $\exp_q(a + b) = \exp_q(a) \otimes_q \exp_q(b)$ and $\exp_q(a - b) = \exp_q(a) \oslash_q \exp_q(b)$.

2.1.2 Probability distributions

The BqPL model is derived by assuming random error terms follow a new probability distribution named the Truncated q-Log-Logistic distribution. This distribution is derived from a series of other distributions. The Truncated q-log-logistic distribution is derived by truncating a q-log-logistic distribution. The q-log-logistic distribution is derived by taking the q-ratio of two q-Gumbel distributions. The q-Gumbel distribution is derived by introducing a q-logarithm operator to transform the dependent variable in the Gumbel distribution. It thus generalises both the Gumbel and Weibull distributions.

We shall thus begin by defining the q-Gumbel distribution as introduced in Chikaraishi and Nakayama (2016), which has the following Probability Distribution Function (PDF):

$$f_{qG}(x|\theta,\mu,q) = \theta x^{-q} \exp\left(-\theta(\ln_q(x) - \ln_q(\mu))\right) \exp\left[-\exp\left(-\theta(\ln_q(x) - \ln_q(\mu))\right)\right] \tag{5}$$

θ is the scale parameter, μ is the location parameter, and q is the heteroscedasticity parameter. Indeed, q parametrises the rate at which the variance of the q-Gumbel increases with μ (an illustration can be found in Chikaraishi and Nakayama (2016)). Its Cumulative Distribution Function (CDF) is given by:

$$F_{qG}(x|\theta,\mu,q) = 1 - \exp\left[-\exp\left(-\theta(\ln_q(x) - \ln_q(\mu))\right)\right] \tag{6}$$

As one can see, the q-Gumbel distribution is derived by replacing logarithm operators with q-logarithm operators in the Gumbel distribution. Note that, as shown in Chikaraishi and Nakayama (2016), the q-Gumbel distribution is equivalent to the Gumbel distribution for $q = 0$ and equivalent to the Weibull distribution for $q = 1$. As proved in Chikaraishi and Nakayama (2016), the q-Gumbel distribution variance increases with its mean μ, which allows the qPL model to account for heteroscedasticity.

Next, we shall derive a q-Log-Logistic distribution by taking the q-ratio of two independent q-Gumbel distributions. Let X and Y follow two independent q-Gumbel distributions with the same scale and q parameter, i.e., $X \sim q\text{Gumbel}(\theta, \mu_X, q)$ and $Y \sim q\text{Gumbel}(\theta, \mu_Y, q)$. Then, their q-ratio $Z = X \oslash_q Y$ follows a q-Log-Logistic distribution, whose CDF, denoted F_{qL}, is given by:

$$F_Z(x) = F_{qL}(x|\theta,\mu,q) := \frac{1}{1 + \exp\left[-\theta(\ln_q(x) - \ln_q(\mu_Z))\right]} \tag{7}$$

where $\mu_Z = \mu_X \oslash_q \mu_Y$. This is proven as follows:

Proof. Let us assume two variables $X \sim q\text{Gumbel}(\theta, \mu_X, q)$, $Y \sim q\text{Gumbel}(\theta, \mu_Y, q)$ as defined above. If $q = 1$, the result comes directly from the Weibit proof from Gu et al. (2022). If $q \neq 1$, we have that $\ln_q(X \oslash_q Y) = \ln_q(X) - \ln_q(Y)$. As proved in Chikaraishi and Nakayama 2016, $\ln_q(X) \sim \text{Gumbel}(\theta, \ln_q(\mu_X))$ and $\ln_q(Y) \sim \text{Gumbel}(\theta, \ln_q(\mu_Y))$. Their difference is thus Logistically distributed, i.e., $\ln_q(X \oslash_q Y) \sim \text{Logistic}(\theta, \mu)$, where $\mu = \ln_q(\mu_X) - \ln_q(\mu_Y) = \ln_q(\mu_X \oslash_q \mu_Y)$. It has the following CDF:

$$F_{\ln_q(X\oslash_q Y)}(x|\theta,\mu) = \Pr\left(\ln_q(X \oslash_q Y) \leq x\right) = \frac{1}{1 + \exp(-\theta(x - \mu))}$$

The function $x \to \exp_q(x)$ is monotonically increasing, meaning that:

$$\begin{aligned} F_{\ln_q(X\oslash_q Y)}(x) = \Pr\left(\ln_q(X \oslash_q Y) \leq x\right) &= \Pr\left(X \oslash_q Y \leq \exp_q(x)\right) \\ &= F_{X\oslash_q Y}(\exp_q(x)) \end{aligned}$$

Thus,

$$\begin{aligned} F_{X\oslash_q Y}(x) &= F_{\ln_q(X\oslash_q Y)}(\ln_q(x)) \\ &= \frac{1}{1 + \exp\left[-\theta(\ln_q(x) - \ln_q(\mu))\right]} \end{aligned}$$

where $\mu = \mu_X \oslash_q \mu_Y$ $\square$

The proposed Truncated q-Log-Logistic distribution is derived by left-truncating the q-Log-Logistic distribution at a lower bound $1 \oslash_q \phi$ for some $\phi \geq 1$. If $X \sim$ q-Log-logistic(θ, μ, q), we define the CDF of the Truncated q-Log-Logistic distribution as the distribution of the X given that its value is larger than the truncation threshold $1 \oslash_q \phi$:

$$
\begin{aligned}
F_{qT}(x|\theta,\mu,q,\phi) &= \mathbb{P}(X \leq x | X \geq 1 \oslash_q \phi) \\
&= \frac{\mathbb{P}(X \leq x; X \geq 1 \oslash_q \phi)}{\mathbb{P}(X \geq 1 \oslash_q \phi)} \\
&= \frac{\mathbb{P}(1 \oslash_q \phi \leq X \leq x)}{1 - \mathbb{P}(X \leq 1 \oslash_q \phi)} \\
&= \begin{cases} \dfrac{F_{qL}(x|\theta,\mu,q) - F_{qL}(1 \oslash_q \phi|\theta,\mu,q)}{1 - F_{qL}(1 \oslash_q \phi|\theta,\mu,q)} & \text{if} \quad x \geq 1 \oslash_q \phi \\ 0 & \text{if} \quad 0 \leq x < 1 \oslash_q \phi \end{cases}
\end{aligned} \tag{8}
$$

where F_{qL} is the CDF of the q-Log-Logistic distribution, given in Equation 7. The PDF of the Truncated q-Log-Logistic distribution can be found by differentiating the CDF with respect to x:

$$
f_{qT}(x|\theta,\mu,q,\phi) = \begin{cases} \dfrac{f_{qL}(x|\theta,\mu,q)}{1 - F_{qL}(1 \oslash_q \phi|\theta,\mu,q)} & \text{if} \quad x \geq 1 \oslash_q \phi \\ 0 & \text{if} \quad 0 \leq x < 1 \oslash_q \phi \end{cases} \tag{9}
$$

where $f_{qL} = \partial F_{qL}/\partial x$, i.e.,

$$
f_{qL}(x|\theta,\mu,q) = \theta x^{-q} \exp(\theta(\ln_q(x) - \mu)) \left[1 + \exp(\theta(\ln_q(x) - \mu))\right]^{-2}
$$

The Truncated q-Log-Logistic CDF can be written compactly as:

$$
F_{qT}(x|\theta,\mu,q,\phi) = \frac{(F_{qL}(x|\theta,\mu,q) - F_{qL}(1 \oslash_q \phi|\theta,\mu,q))_+}{1 - F_{qL}(1 \oslash_q \phi|\theta,\mu,q)}
$$

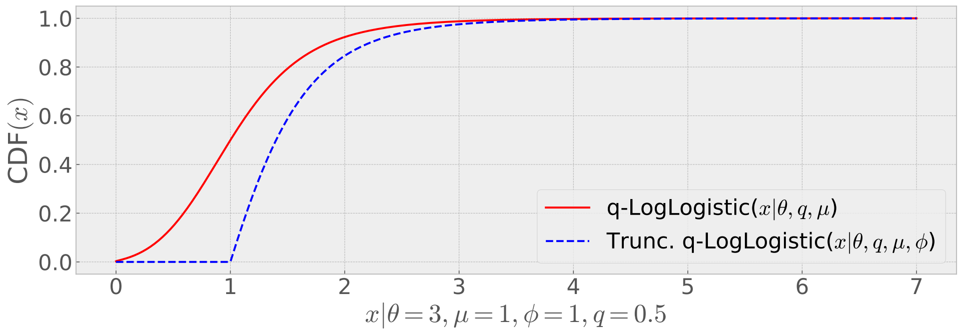
where $(.)_+$ is the operator $\max(0,.)$. The CDF of the Truncated q-Log-Logistic distribution for some fixed values of parameters is plotted in Figure 1. We observe that, for $\phi = 1$, the CDF equals zero for all $x < 1$, then tends to the q-Log-Logistic distribution CDF when x tends to $+\infty$. We also see that, for $\mu = 1$, the CDF gradient is steeper than for $\mu = 3$, which means that its values have less variance. This highlights the heteroscedasticity property of the q-Log-Logistic and its truncated version.

2.2 Derivation of the Bounded q-Product Logit Model

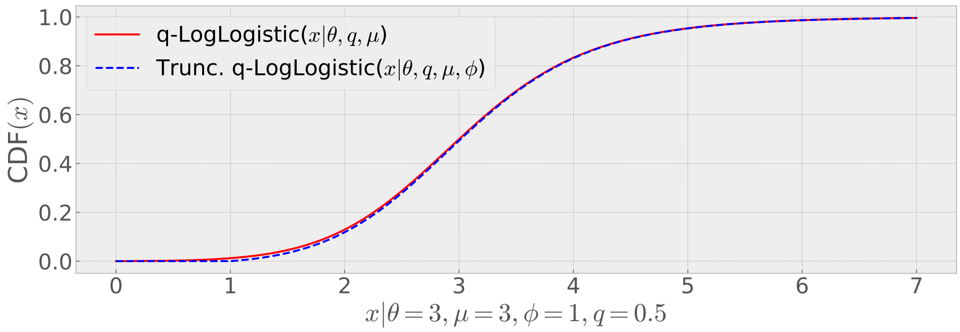
Here we derive the BqPL model, following a similar approach to how the BL model is derived in Duncan et al. (2022) (see Supplementary Material Appendix A of that paper). For pedagogical purposes, we shall briefly describe how the BL model is derived so that it is clear what is different in the derivation of the BqPL model. Suppose a decision-maker faces a route choice situation with a choice set $\mathcal{R} = \{1, ..., n\}$. Suppose that each alternative $i \in \mathcal{R}$ can be described by some observed attributes aggregated within a *positive* systematic cost function $c_i > 0$. To account for the analyst's lack of knowledge of the actual route cost that the decision-maker perceives, the costs have random errors. For the BL model, it is assumed that this random error is additive to the systematic cost: $C_i = c_i + \epsilon_i$, where ϵ_i is the error term for route i. Now, the BL model supposes that each route is compared with an imaginary reference alternative r^* in terms of the difference in random cost:

$$
C_i - C_{r^*} = (c_i + \epsilon_i) - (c_{r^*} + \epsilon_{r^*}) = c_i - c_{r^*} + \epsilon_i - \epsilon_{r^*} = c_i - c_{r^*} + \varepsilon_i
$$

where ε_i is the difference random error term for route i 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 BL model, however,



(a) $\mu = 1$



(b) $\mu = 3$

Figure 1: q-Log-Logistic and Truncated q-Log-Logistic CDFs for $\mu = 1$ and $\mu = 3$, all the other parameters being fixed.

proposes that the difference random error terms ε_i assume a Truncated Logistic distribution, obtained by left-truncating a Logistic distribution with mean 0 and scale θ at a lower bound of $-\phi$ for some $\phi > 0$. The consequent BL choice probability function imposes a bound ϕ on the difference in systematic cost with the reference alternative, with violating routes receiving zero choice probabilities.

For the BqPL model, rather than supposing that the random error is additive to the systematic cost, we suppose the random error relates via a q-product operator to the systematic cost: $C_i = c_i \otimes_q \epsilon_i$. The q-ratio in random cost between route i and the reference alternative is thus:

$$C_i \oslash_q C_{r^*} = (c_i \otimes_q \epsilon_i) \oslash_q (c_{r^*} \otimes_q \epsilon_{r^*}) = c_i \oslash_q c_{r^*} \otimes_q \epsilon_i \oslash_q \epsilon_{r^*} = c_i \oslash_q c_{r^*} \otimes_q \varepsilon_i$$

It is important to note that while we define ϵ_i as the random cost error term for alternative i, its distribution is not explicitly specified. The only distributional assumption is made on the relative error term, defined as $\varepsilon_i = \epsilon_{r^*} \oslash_q \epsilon_i$. The BqPL model can be derived without assuming a random utility for each alternative, but instead by considering only the relative random utility with respect to the reference alternative. This approach, along with further discussion on the error term distribution and heteroscedasticity, is presented in Appendix D. The qPL model can be derived by assuming $\varepsilon_i = \epsilon_i \oslash_q \epsilon_{r^*}$ follows a q-Log-Logistic distribution. For the BqPL model, the q-ratio of random error terms ε_i is assumed to follow a Truncated q-Log-Logistic distribution with mean 0, scale θ, q-parameter q, and bound ϕ. This implies that:

$$\Pr(\epsilon_i \oslash_q \epsilon_{r^*} \leq 1 \oslash_q \phi) = 0$$

or equivalently,

$$\Pr(C_i \oslash_q C_{r^*} \leq c_i \oslash_q (\phi \otimes_q c_{r^*})) = 0$$

meaning that the random cost q-ratio must be within a bound $\phi \geq 1$ of the systematic cost q-ratio. The probability of choosing an alternative $i \in \mathcal{R}$ over a reference alternative r^* is given by:

$$
\begin{aligned}
\Pr(i|\{i,r^*\}) &= \Pr(C_i \le C_{r^*}) \\
&= \Pr(c_i \otimes_q \epsilon_i \le c_{r^*} \otimes_q \epsilon_{r^*}) \\
&= \Pr(\epsilon_i \oslash_q \epsilon_{r^*} \le c_{r^*} \oslash_q c_i) \\
&= \Pr(\varepsilon_i \le c_{r^*} \oslash_q c_i) \qquad (10)
\end{aligned}
$$

Assuming that $\varepsilon_i = \epsilon_i \oslash_q \epsilon_{r^*}$ follows a truncated q-Log-Logistic distribution with scale parameter θ, a location parameter of 1, a shape parameter q and a bound parameter ϕ, the choice probabilities can be rewritten:

$$
\begin{aligned}
\Pr(i|\{i,r^*\}) &= F_{qT}(c_{r^*} \oslash_q c_i|\theta,1,q,\phi) \\
&= \begin{cases} \dfrac{F_{qL}(c_{r^*} \oslash_q c_i|\theta,1,q) - F_{qL}(1 \oslash_q \phi|\theta,1,q)}{1 - F_{qL}(1 \oslash_q \phi|\theta,1,q)} & \text{if } c_{r^*} \oslash_q c_i \ge 1 \oslash_q \phi \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{\dfrac{1}{1+e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)}} - \dfrac{1}{1+e^{-\theta \ln_q(1 \oslash_q \phi)}}}{1 - \dfrac{1}{1+e^{-\theta \ln_q(1 \oslash_q \phi)}}} & \text{if } c_{r^*} \oslash_q c_i \ge 1 \oslash_q \phi \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{\dfrac{e^{\theta \ln_q(\phi)} - e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)}}{(1+e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)})(1+e^{\theta \ln_q(\phi)})}}{\dfrac{e^{\theta \ln_q(\phi))}}{1+e^{\theta \ln_q(\phi)}}} & \text{if } c_{r^*} \oslash_q c_i \ge 1 \oslash_q \phi \\ 0 & \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{e^{\theta \ln_q(\phi)} - e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)}}{(1+e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)})e^{\theta \ln_q(\phi)}} & \text{if } c_{r^*} \oslash_q c_i \ge 1 \oslash_q \phi \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

We define the odds ratio as $\eta_i = \frac{\Pr(i|\{i,r^*\})}{1-\Pr(i|\{i,r^*\})}$. First, $1 - \Pr(i|\{i,r^*\})$ can be expressed as follows:

$$
\begin{aligned}
1 - \Pr(i|\{i,r^*\}) &= \begin{cases} 1 - \dfrac{e^{\theta \ln_q(\phi)} - e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)}}{(1+e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)})e^{\theta \ln_q(\phi)}} & \text{if } c_{r^*} \oslash_q c_i \ge 1 \oslash_q \phi \\ 1 & \text{otherwise} \end{cases} \\
&= \begin{cases} \dfrac{e^{-\theta \ln_q(c_{r^*} \oslash_q c_i \oslash_q \phi)} + e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)}}{(1+e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)})e^{\theta \ln_q(\phi)}} & \text{if } c_{r^*} \oslash_q c_i \ge 1 \oslash_q \phi \\ 1 & \text{otherwise} \end{cases}
\end{aligned}
$$

Consequently, we can simplify η_i as:

$$
\begin{aligned}
\eta_i &= \frac{\left(e^{\theta \ln_q(\phi)} - e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)}\right)_+}{e^{-\theta \ln_q(c_{r^*} \oslash_q c_i \oslash_q \phi)} + e^{-\theta \ln_q(c_{r^*} \oslash_q c_i)}} \\
&= \frac{\left(e^{-\theta \ln_q(c_i \oslash_q (\phi \otimes_q c_{r^*}))} - 1\right)_+}{(e^{\theta \ln_q(\phi)} + 1)}
\end{aligned}
$$

We retrieve the choice probabilities of i among the choice set $\mathcal{R}$ using the log-odds ratio formula:

$$
\Pr(i|\mathcal{R}) = \frac{e^{\ln(\eta_i)}}{\sum_{j \in \mathcal{R}} e^{\ln(\eta_j)}}
$$

i.e.,

$$P_i^{\text{BqPL}} := \Pr(i|\mathcal{R}) = \frac{\left(e^{-\theta \ln_q(c_i \oslash_q (\phi \otimes_q c_{r^*}))} - 1\right)_+}{\sum_{j\in\mathcal{R}} \left(e^{-\theta \ln_q(c_j \oslash_q (\phi \otimes_q c_{r^*}))} - 1\right)_+} \tag{11}$$

For the BL model, Watling et al. (2018) proposed defining the reference alternative as the alternative with the cheapest systematic cost, i.e., $c_{r^*} = \min_{j\in\mathcal{R}} c_j$. And, they specify the bound so that it is relative to the cheapest cost, i.e. a route receives a zero choice probability if it has a cost $c_i \geq \varphi \min_{j\in\mathcal{R}} c_j$. This is because the desired bounding condition is relative to the scale of costs for each OD movement.

To specify the BqPL model so that it has the same bounding condition as the BL model, we set $\phi = \varphi c_{r^*} \oslash_q c_{r^*}$, where $\varphi > 1$ is the relative cost bound parameter. The consequent BqPL choice probability function is as follows for route i:

$$\begin{aligned} P_i^{\text{rel. BqPL}} := \Pr(i|\mathcal{R}) &= \frac{\left(\exp(-\theta \ln_q(c_i \oslash_q (\varphi \min_{l\in\mathcal{R}} c_l))) - 1\right)_+}{\sum_{j\in\mathcal{R}} \left(\exp(-\theta \ln_q(c_j \oslash_q (\varphi \min_{l\in\mathcal{R}} c_l))) - 1\right)_+} \\ &= \frac{\left(\exp(-\theta(\ln_q(c_i) - \ln_q(\varphi \min_{l\in\mathcal{R}} c_l))) - 1\right)_+}{\sum_{j\in\mathcal{R}} \left(\exp(-\theta(\ln_q(c_j) - \ln_q(\varphi \min_{l\in\mathcal{R}} c_l))) - 1\right)_+} \end{aligned} \tag{12}$$

With this BqPL choice probability function, an alternative will receive a non-zero choice probability if and only if $c_i < \varphi \min_{l\in\mathcal{R}} c_l$. Although in this study we specify the BqPL bound as a relative cost bound to be consistent with the BL model, one can also specify the bound in other ways with different implications, see Appendix A for a discussion.

2.3 Special cases of the BqPL model

Under certain specifications of the model parameters, the BqPL model can collapse into existing and new models. In this section, we summarise them.

2.3.1 When $\varphi \to +\infty$: the qPL model

When φ tends to infinity, the BqPL model collapses to Chikaraishi and Nakayama (2016)'s qPL model. Rearranging the BqPL probability relation in Equation 12:

$$P_i^{\text{rel. BqPL}} = \frac{\exp(-\theta \ln_q(c_i)) \left(1 - \exp(\theta \ln_q(c_i \oslash_q (\varphi \min_{l\in\mathcal{R}} c_l)))\right)_+}{\sum_{j\in\mathcal{R}} \exp(-\theta \ln_q(c_j)) \left(1 - \exp(\theta \ln_q(c_i \oslash_q (\varphi \min_{l\in\mathcal{R}} c_l)))\right)_+} \tag{13}$$

As $\lim_{\varphi\to+\infty} \exp(\theta \ln_q(c_i \oslash_q (\varphi \min_{l\in\mathcal{R}} c_l))) = 0$, we have that, from Equation 13:

$$\lim_{\phi\to+\infty} \frac{\left(\exp(-\theta \ln_q(c_i \oslash_q (\varphi \min_{l\in\mathcal{R}} c_l))) - 1\right)_+}{\sum_{j\in\mathcal{R}} \left(\exp(-\theta \ln_q(c_j \oslash_q (\varphi \min_{l\in\mathcal{R}} c_l)) - 1\right)_+} = \frac{\exp(-\theta \ln_q(c_i))}{\sum_{j\in\mathcal{R}} \exp(-\theta \ln_q(c_j))} = P_i^{q\text{PL}} \tag{14}$$

2.3.2 When $q = 0$: the BL model

When $q = 0$, as $\ln_q(x) = x - 1$ for any $x > 0$, the BqPL collapses to Watling et al. (2018) Bounded Logit (BL), with the following choice probabilities:

$$P^{\text{rel. BqPL}}_{i,q=0} = \frac{\left(\exp(-\theta(c_i - \varphi \min_{l\in\mathcal{R}} c_l)) - 1\right)_+}{\sum_{j\in\mathcal{R}} \left(\exp(-\theta(c_j - \varphi \min_{l\in\mathcal{R}} c_l)) - 1\right)_+} \tag{15}$$

2.3.3 When $q = 1$: the Bounded Weibit (BW) model

When $q = 1$, as $\ln_q(x) = \ln(x)$ for any $x > 0$, the BqPL collapses to what we will refer to as the Bounded Weibit (BW) model, i.e., a bounded model which collapses to the Multinomial Weibit model (Castillo et al. 2008) when the bound tends to $+\infty$. Its choice probabilities are given by:

$$P^{\text{BW}}_i := P^{\text{BqPL}}_{i,q=1} = \frac{\left((c_i/\varphi \min_{l\in\mathcal{R}} c_l)^{-\theta} - 1\right)_+}{\sum_{j\in\mathcal{R}} \left((c_j/\varphi \min_{l\in\mathcal{R}} c_l)^{-\theta} - 1\right)_+} \tag{16}$$

2.4 Model variance

A key feature of the Bounded q-Product Logit model is its inherent heteroscedasticity, meaning the variance of the decision noise is not constant but varies with the deterministic cost structure. This property emerges naturally from the model's definition of the random cost q-ratio:

$$\begin{aligned} C_i \oslash_q C_{r^*} &= (c_i \oslash_q c_{r^*}) \otimes_q \varepsilon_i \\ &= \left((c_i \oslash_q c_{r^*})^{1-q} + \varepsilon_i^{1-q} - 1\right)^{\frac{1}{1-q}} \end{aligned}$$

where ε_i is an error term drawn from a truncated q-LogLogistic distribution with a mean of 1, which is equivalent to say that $C_i \oslash_q C_{r^*}$ is drawn from a truncated q-LogLogistic distribution with a mean of $c_i \oslash_q c_{r^*}$

The theoretical variance of the random cost q-ratio is therefore given by:

$$\begin{aligned} \text{Var}(C_i \oslash_q C_{r^*}) = &\int_{1\oslash_q\phi}^{\infty} \left(\left((c_i \oslash_q c_{r^*})^{1-q} + \varepsilon_i^{1-q} - 1\right)^{\frac{2}{1-q}}\right) f_{qT}(\varepsilon_i \mid \theta, 1, q, \phi)\, d\varepsilon_i - \\ &\left(\int_{1\oslash_q\phi}^{\infty} \left((c_i \oslash_q c_{r^*})^{1-q} + \varepsilon_i^{1-q} - 1\right)^{\frac{1}{1-q}} f_{qT}(\varepsilon_i \mid \theta, 1, q, \phi)\, d\varepsilon_i\right)^2 \end{aligned}$$

Where f_{qT} is the truncated q-LogLogistic distribution PDF, given in Equation 9. This expression for the variance has no closed-form solution and must hence be evaluated numerically. Figure 2 illustrates how the variance evolves with both the shape parameter q and the deterministic component of the cost ratio. Notably, when $q = 0$, corresponding to the Bounded Logit model, the variance remains relatively stable across different cost difference levels. As q increases, however, the variance becomes increasingly sensitive to the deterministic q-cost ratio $c_i \oslash_q c_{r^*}$, revealing the heteroscedastic nature of the model: variance grows with the cost disparity between alternatives.

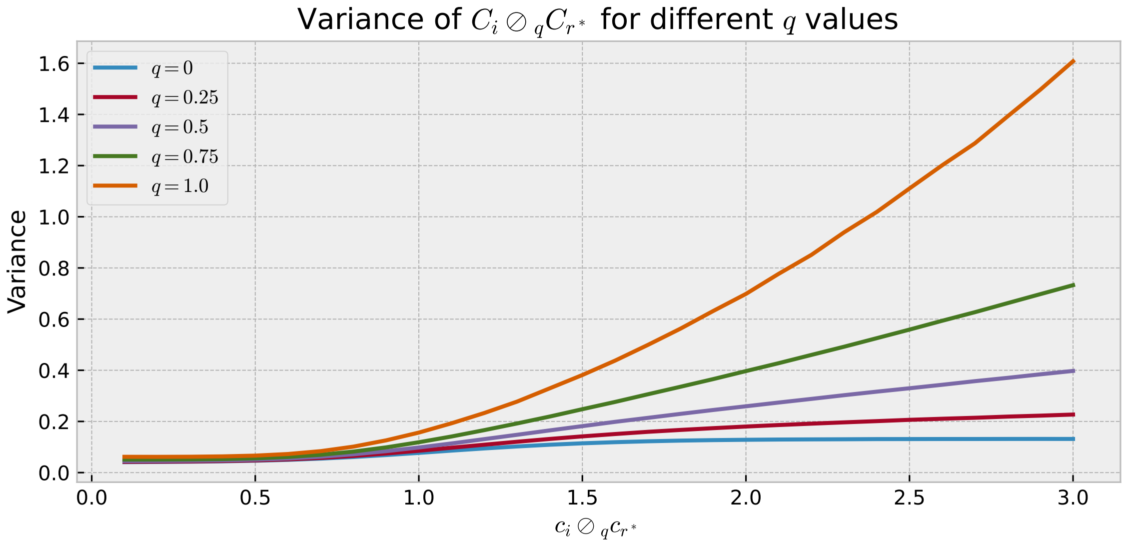


Figure 2: Variance of the random cost q-Ratio, for a fixed bound $\phi = 1.3$ and $\theta = 5$

2.5 Accounting for route correlation: The Bounded Path Size q-Product Logit model

In this section, we extend the BqPL model to account for route correlation. Path size correction models are widely used in route choice due to their applicability to large choice sets and the extensive work to accurately approximate the correlation structure of routes. Path size correction models capture correlations between overlapping routes by including heuristic correction factors within the probability relation to penalising routes for sharing links with other routes. These correction factors depend upon path size terms that measure the distinctiveness of a route: the less distinct a route is (i.e., the more it overlaps / shares links with other routes), the greater the penalisation. Path size correction terms have been incorporated, for example, within Logit (Ben-Akiva and Bierlaire (1999)), Weibit (Kitthamkesorn and Chen (2013)), and Hybrid Logit-Weibit (Xu et al. 2015) probability relations.

Recently, Duncan et al. (2022) explored how to incorporate path size correction terms within the BL model, developing a Bounded Path Size Logit (BPSL) model, where careful consideration was given over the definition of the path size term so that correlation is captured between only the routes defined as considered by the model through the cost bound. The contention is that penalising considered routes for overlapping with non-considered routes would be unrealistic. Careful consideration was also given to ensure the path size terms were continuous as routes cross from being considered to non-considered as costs cross from below to above the bound (and vice versa), thus ensuring a continuous choice probability function.

A key property of the BPSqPL model is that Path-Size correction terms are not included in route utilities (generalized route costs). As a result, route overlap affects route choice probabilities but does not influence the formation of the choice set. This approach is justified both behaviorally and practically. Behaviorally, we argue that a route is not inherently unrealistic simply because it overlaps significantly with others. Additionally, unlike traditional route attributes, Path-Size correction terms are neither directly observable nor comparable by travellers, making it unlikely that they would influence route consideration. To our knowledge, no empirical evidence supports excluding highly overlapping routes from choice sets based on behavioural factors.

From a mathematical perspective, incorporating Path-Size corrections into route costs leads to undesirable model properties. As discussed in Appendix C of Duncan et al. (2022), doing so violates key desirable properties, such as ensuring continuity of the probability function and preventing unrealistic routes from influencing the Path-Size terms of realistic routes. For these reasons, our model accounts for route overlap in choice probabilities but not in the bounding condition for choice set formation.

For the BqPL model, we take inspiration from the BPSL model to formulate a Bounded Path Size q-Product Logit (BPSqPL) model, where the path size terms are defined appropriately to capture correlations between considered routes only and a continuous choice probability function is maintained. Let us assume a route choice situation with the universal choice set $\mathcal{R}$ of routes. Each route $i \in \mathcal{R}$ consists of a set of links $A_i \subseteq A$, where A is the universal set of links in the network. These links are defined by attributes aggregated in positive cost functions $t_a, a \in A$, parameterised by a vector of parameters $\boldsymbol{\alpha}$. The total cost c_i of route i is link-additive, i.e. $c_i = \sum_{a \in A_i} t_a$. For fixed parameters $\boldsymbol{\alpha}, \varphi$, we define $\bar{\mathcal{R}}(\boldsymbol{\alpha}, \varphi) \subseteq \mathcal{R}$ as the subset of considered routes with a cost less than φ times the cheapest route cost:

$$\bar{\mathcal{R}}(\boldsymbol{\alpha}, \varphi) = \left\{ i \in \mathcal{R}, c_i(\boldsymbol{\alpha}) < \varphi \min_{l \in \mathcal{R}} c_l(\boldsymbol{\alpha}) \right\}$$

The BPSqPL choice probability function for route i is:

$$P_i^{\text{BPSqPL}} = \begin{cases} \dfrac{\left(\gamma_i^{\text{BPSqPL}}\right)^\eta \left(\exp(-\theta(\ln_q(c_i) - \ln_q(\varphi \min_{l \in \mathcal{R}} c_l))) - 1\right)}{\sum_{j \in \bar{\mathcal{R}}} \left(\gamma_j^{\text{BPSqPL}}\right)^\eta \left(\exp(-\theta(\ln_q(c_j) - \ln_q(\varphi \min_{l \in \mathcal{R}} c_l))) - 1\right)} & \text{if } i \in \bar{\mathcal{R}} \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

where $(\gamma_i^{\text{BPSqPL}})^\eta$ is the path size correction factor for considered route $i \in \bar{\mathcal{R}}$ (non-considered routes do not have path size terms). $\eta \geq 0$ is the path size scaling parameter scaling sensitivity to route distinctiveness, and $\gamma_i^{\text{BPSqPL}} \in (0, 1]$ is the path size term for considered route $i \in \bar{\mathcal{R}}$, calculated as follows:

$$\gamma_i^{\text{BPSqPL}} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{\left(\exp(-\theta(\ln_q(c_i) - \ln_q(\varphi \min_{l \in \mathcal{R}} c_l))) - 1\right)}{\sum_{j \in \bar{\mathcal{R}}} \left(\exp(-\theta(\ln_q(c_i) - \ln_q(\varphi \min_{l \in \mathcal{R}} c_l))) - 1\right) \delta_{aj}} \tag{18}$$

δ_{aj} is equal to 1 if route j uses link $a \in A_i$, and zero otherwise. γ_i^{BPSqPL} is specified so that a) non-considered routes with costs above the bound, i.e. routes $j \notin \bar{\mathcal{R}}$, do not contribute to reducing the path size terms of considered routes with costs below the bound, and b) the path size term function is continuous as routes enter and exit the considered route set $\bar{\mathcal{R}}$ (as costs cross from below to above the bound and vice versa). It is also formulated in terms of summing over $j \in \bar{\mathcal{R}}$ rather than with $(.)_+$ functions to avoid occurrences of 0/0. As special cases of the BPSqPL model, the following existing models can be obtained:

- The BPSL model (Duncan et al. 2022) when $q = 0$
- The alternative GPSL (GPSL') model (Duncan et al. 2020) when $q = 0$ and $\varphi \to +\infty$

Moreover, the following new models can be obtained:

- A Bounded Path Size Weibit (BPSW) model when $q = 1$
- An alternative Generalised Path Size qPL' (GPSqPL') model when $\varphi \to +\infty$
- An alternative Generalised Path Size Weibit (GPSW) model when $q = 1$ and $\varphi \to +\infty$

The choice probability functions for each of these models are provided in Appendix B.

2.6 Summary of model properties and collapsing

The BPSqPL model developed in the previous section accounts for the three deficiencies of the MNL model noted in the introduction:

- *Heteroscedasticity* is accounted for by assuming a distribution for the error terms where error variance increases with systematic cost. The rate at which variance increases with cost can be controlled by the heteroscedasticity parameter q, where $q = 0$ results in no increase, and $q = 1$ results in quadratic growth.

- *Route overlap* is accounted for by including path size correction terms within the probability function that penalise routes for sharing links with other routes. η is the path size scaling parameter that controls the penalisation.

- *Choice set formation* is accounted for by bounding the error term distribution so that the model allocates zero choice probabilities to routes with systematic costs greater than φ times the cheapest cost, where φ is the relative cost bound parameter.

In accounting for all three deficiencies, the BPSqPL model is an advantage over several existing and new models that only account for one or two deficiencies. The BPSqPL also provides a unified framework for these models, where, under different settings of the model parameters, it can collapse into different models. Figure 3 presents a schematic diagram of under which settings of the model parameters BPSqPL can collapse into different models, and Table 2 indicates which properties its collapsing models can account for.

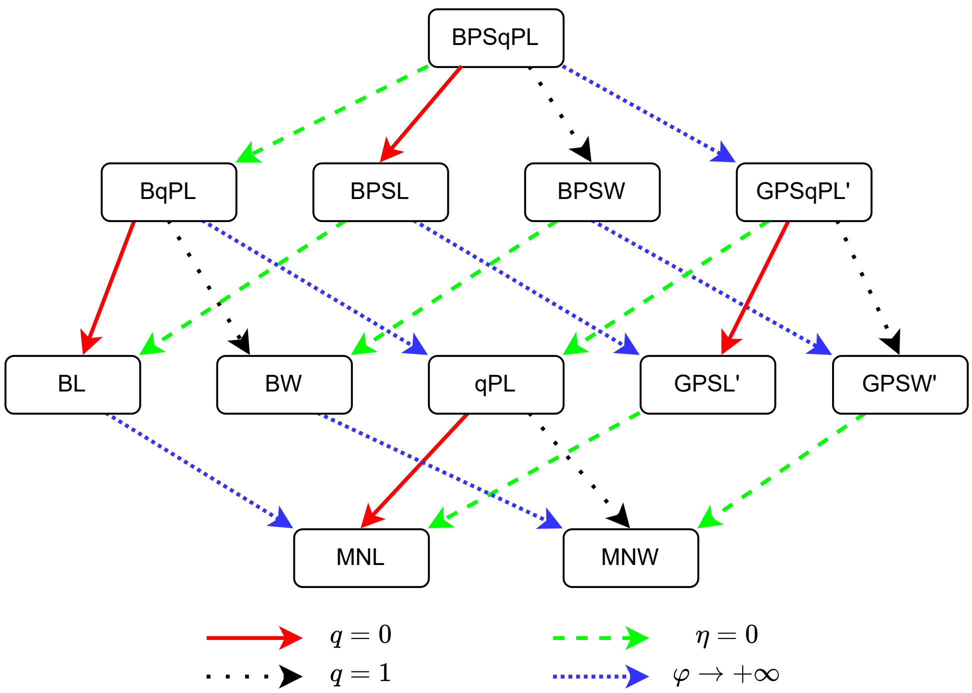


Figure 3: Model collapsing

Model	Route Correlation	Heteroscedasticity	Choice set formation
MNL			
GPSL'	✓		
MNW		✓	
qPL		✓	
BL			✓
GPSW'	✓	✓	
GPSqPL'	✓	✓	
BPSL	✓		✓
BqPL		✓	✓
BPSW	✓	✓	✓
BPSqPL	✓	✓	✓

Table 2: Properties of different models BPSqPL can collapse into

3 Demonstrations of model properties

In this section, we will demonstrate the properties of the developed BqPL and BPSqPL models in small-scale route choice examples, and contrast with associated collapsing models.

3.1 Accounting for heteroscedasticity

For the first example, we show how the different models' choice probabilities evolve in the binary route choice case illustrated in Figure 4. The two routes, denoted 1 and 2, have respective costs $c_1 = x$ and $c_2 = x + 1$, with $x > 0$. The choice probabilities of Route 1 with cost x according to the MNL, qPL and MNW are given in Figure 5. We observe that the MNL probability does not depend on x, as it is only sensitive to the difference in cost. Conversely, the MNW probability decreases with x, as it is sensitive to the cost ratio. This property is related to the random cost variance of both routes, which increases with x. Consequently, the likelihood that the perceived cost of Route 2 is smaller than that of Route 1 increases with x. The MNW probability tends to 0.5 as x goes to infinity because the $x/(x + 1)$ ratio tends to 1. The qPL is sensitive to the q-ratio of costs and is between the MNL and MNW. The parameter q influences how close the model would behave closer to an MNL or an MNW.

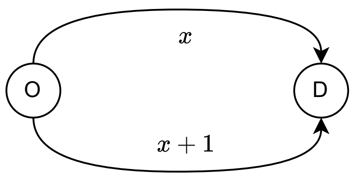


Figure 4: Two-alternative example, with costs x and $x + 1$

3.2 Accounting for heteroscedasticity and choice set formation

In this subsection, we compare the BL, BqPL and BW models with their non-bounded counterparts, using the same example as in Section 3.1 (Figure 4). The choice probabilities are plotted in Figure 6. Each bounded model gives a probability of one to Route 1 for all $x \leq 1$. Indeed, for $x < 1$, the cost ratio of the two routes exceeds the relative cost bound, i.e. $(x + 1)/x > \varphi = 2$.

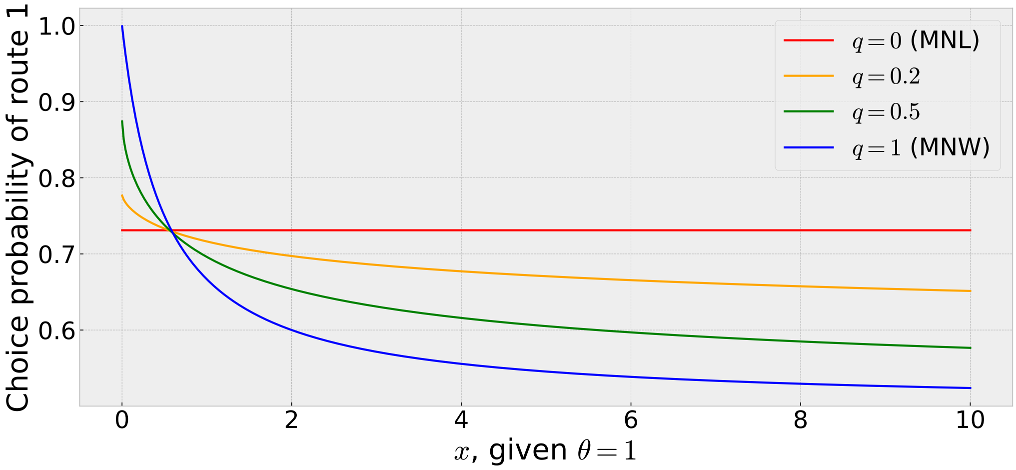


Figure 5: Binary choice probabilities of the route with cost x against route with cost $x+1$, according to the MNL ($q = 0$), qPL ($q = 0.2$ and $q = 0.5$) and MNW ($q = 1$) models

Thus, Route 2 has zero probability. This illustrates that all the relative models will have the same bounding condition. For $x > 1$, the choice probabilities become non-zero for both alternatives and collapse to their unbounded counterparts when $x \to +\infty$.

We finally notice that the higher the value of q, the closer BqPL is to the BW, and the lower the value of q, the closer the BqPL is to the BL.

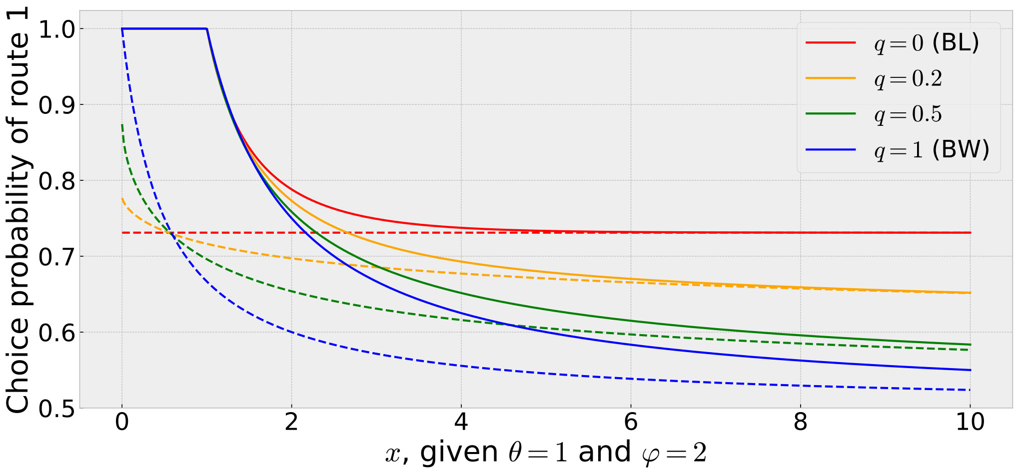


Figure 6: Binary choice probabilities of the route with cost x against route with cost $x+1$, according to the BL ($q = 0$), BqPL ($q = 0.2$ and $q = 0.5$) and BW ($q = 1$) models. The respective non-bounded counterparts are given in dashes.

3.3 Accounting for heteroscedasticity, choice set formation and route correlation

Next, we showcase the properties of the (BPSqPL) and its special cases (the BPSL when $q = 0$, the BPSW when $q = 1$, see Appendix B for their choice probabilities) on a test network. The network displayed in Figure 7 consists of seven links with costs parametrised by $\lambda > 0$ and $0 < \rho < \lambda$. These links form five routes connecting O and D: Route 1, using links 1 and 3; Route 2, using links 1 and 4; Route 3, using links 2 and 5; Route 4, using links 2 and 6; and Route 5, using link 7. Each route but Route 4 has a cost of λ, while Route 4 has a cost of $\lambda + \rho$. A distinct characteristic of this network is that Route 3 is correlated to Route 4, which becomes less desirable than the other routes as ρ increases. This property will highlight the difference between the BPSx (BPSW/BPSL/BPSqPL) model correction and the typically used Path size correction from Ben-Akiva and Bierlaire (1999). The route properties are summarised in Table 3.

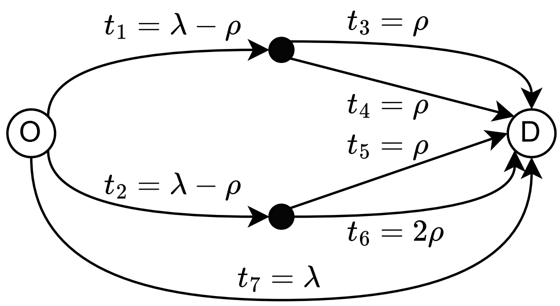


Figure 7: Toy network with 7 links and 5 routes

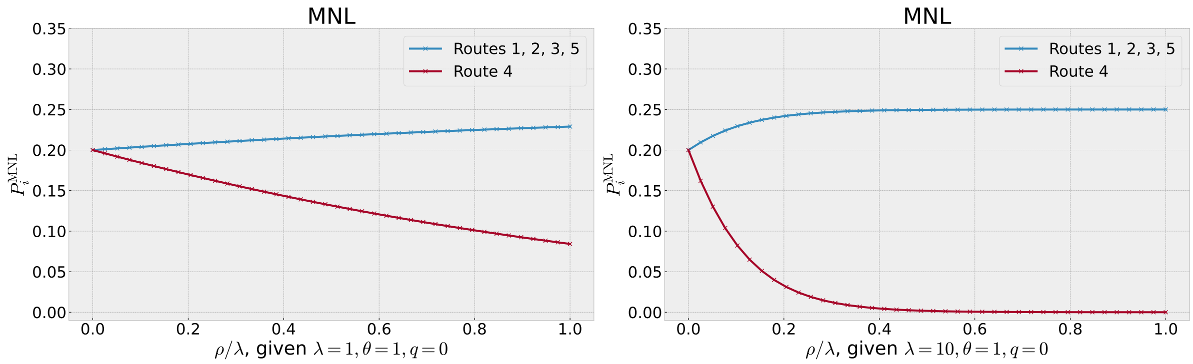
Route	**Links**	**Cost**
1	1-3	λ
2	1-4	λ
3	2-5	λ
4	2-6	$\lambda + \rho$
5	7	λ

Table 3: Set of routes using the network from Figure 7

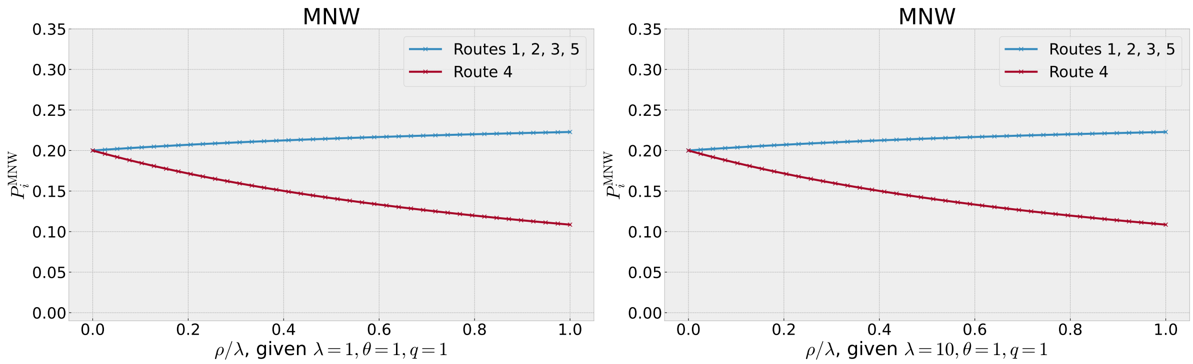
We plot the choice probabilities of each alternative in Figures 9 and 8 as a function of the ratio ρ/λ. The ratio ρ/λ varies between 0 and 1, and is an indicator of how distinct the routes 1 to 4 are (for $\rho/\lambda = 0$, they are confounded, for $\rho/\lambda = 1$, they are completely distinct). To mimic a short and long trip setting, we plot these for two different values of $\lambda \in \{1, 10\}$.

We first plot the choice probabilities of models that do not use any Path size correction nor bounding (Figure 8 (a), (b), and (c)) and models that use bounding but do not use any Path size correction (here, the BqPL, Figure 8 (d)). First, we can see that the MNL, MNW and qPL choice probabilities behave similarly for $\lambda = 1$ (left figures). For the three models, the probability of Route 4 decreases as it becomes less competitive when ρ increases. For $\lambda = 10$, we see that the three models have different behaviours: for the MNL (Figure 8 (a)), the decrease of the Route 4 choice probabilities is steeper (as a function of ρ/λ) than for $\lambda = 1$. This is because the MNL is sensitive to cost differences. For the MNW (Figure 8 (b)), we observe that the left and right plots are the same. That is because the MNW choice probabilities are only sensitive to the cost ratio between Route 4 and the other routes, which is equal to $(\lambda + \rho)/\lambda$ (i.e., the choice probabilities do not depend on the scale of the OD movement). The qPL (Figure 8 (c)) is an in-between to the MNL and MNW, as the choice probabilities are sensitive to the cost q-Ratio. Finally, the BqPL (Figure 8 (d)), while behaving similarly as the qPL, has Route 4 choice probabilities that equal zero when its relative cost to the other routes reach the relative cost bound (for $\rho = 0.8\lambda$). This impacts the choice probabilities' steepness, as seen on the left when compared to the qPL (Figure 8 (c)). Finally, we observe that regardless of their overlap with other routes, alternatives 1, 2, 3 and 5 have the same choice probabilities.

This shortcoming is tackled by the BPSqPL model, which is presented in Figure 9. This figure also displays the BqPL model with Path size correction (referred to as the BqPL-PS) according to the path size correction term of Ben-Akiva and Bierlaire (1999) (Figure 9 (a)). We chose to include this model as it uses the most commonly used correction term in the literature. The BqPL-PS choice probability expression is given in Appendix C. For the BqPL-PS (Figure 9 (a)), we observe that the correlation between routes is accounted for. Route 5 gets the largest choice probability for the BqPL-PS, as it does not overlap with other routes. Still, Route 3 has the same penalisation as Routes 1 and 2, even though it is correlated to a less realistic (nay non-considered) alternative. This is argued to be a main deficiency of using a standard path size correction term, according to Duncan et al. (2020), and is solved in the BPSqPL model (Figure 9 (b)). According to this Figure, Route 5 also gets the largest choice probability, as it is distinct to all the other routes for any value of ρ. For $\rho = 0$, Routes 1 and 2 and Routes 3 and 4 are confounded, so they get equal split probability that are half as large as the one of Route 5. On all the plots, we observe that the probability of choosing Route 3 collapses to the one of Route 5 when ρ increases, as Route 4 gets more unrealistic. For $\rho > 0.8\lambda$, the cost of Route 4 increases over the bound, so it is assigned a zero probability and Route 3 is thus assigned the same probability as Route 5 (which does not overlap with any other route).

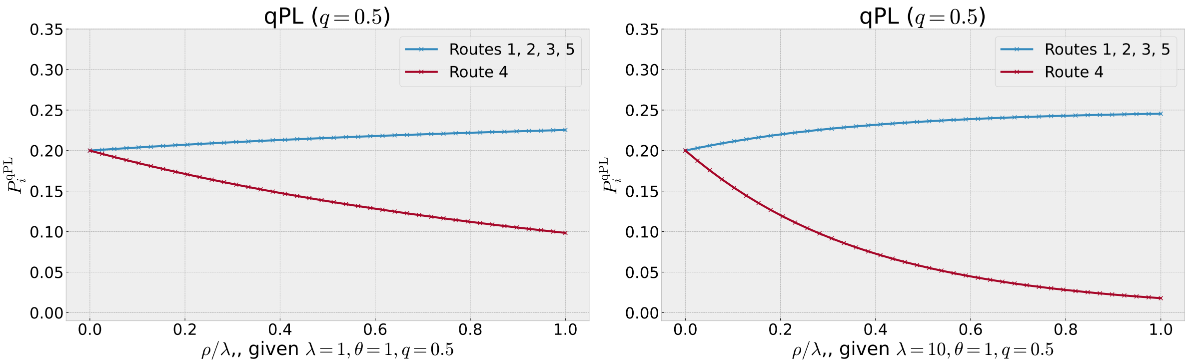


(a) MNL, Left: $\lambda = 1$, Right: $\lambda = 10$

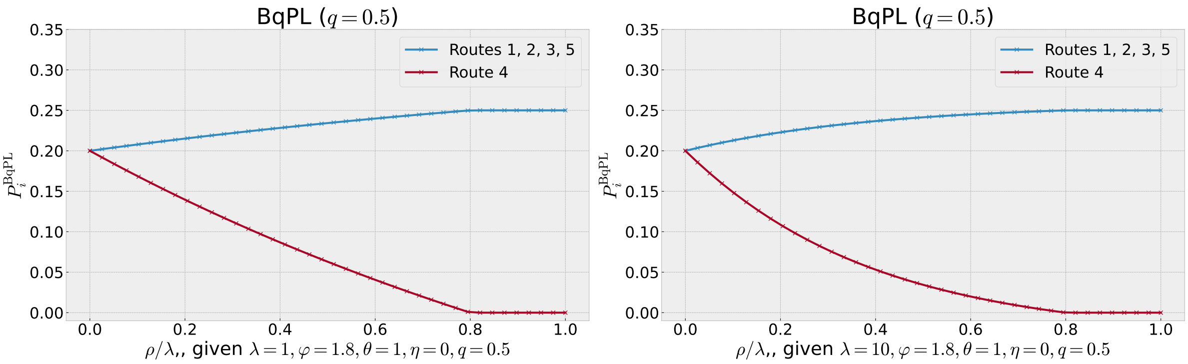


(b) MNW, Left: $\lambda = 1$, Right: $\lambda = 10$

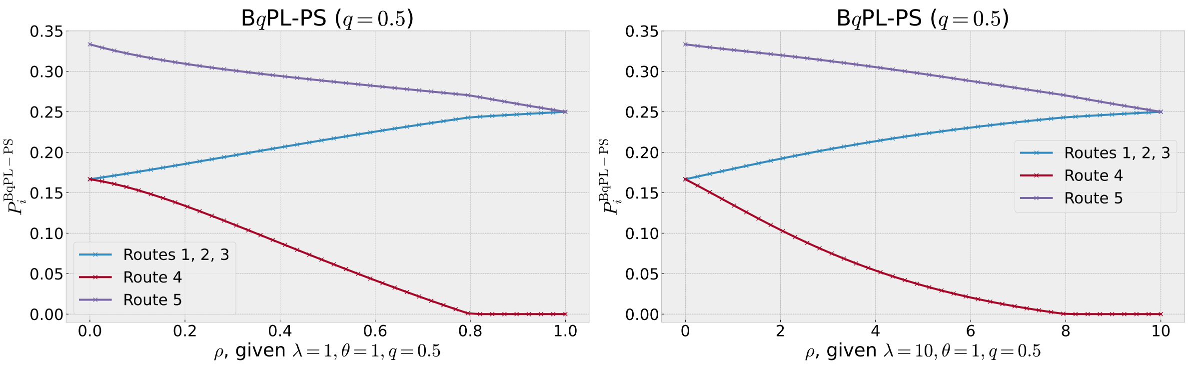
Figure 8: Choice probabilities as a function of ρ on the example network from Figure 7.



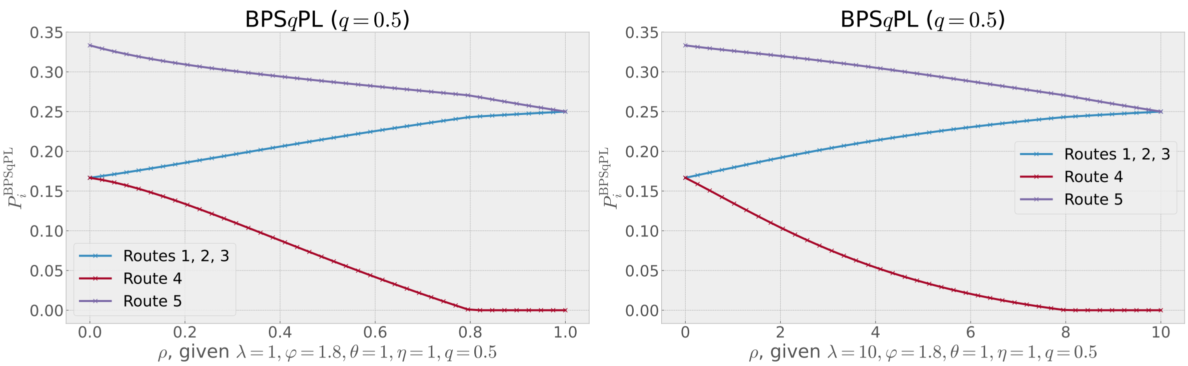
(a) qPL, Left: $\lambda = 1$, Right: $\lambda = 10$



(b) BqPL, Left: $\lambda = 1$, Right: $\lambda = 10$



(c) BqPL-PS, Left: $\lambda = 1$, Right: $\lambda = 10$



(d) BPSqPL, Left: $\lambda = 1$, Right: $\lambda = 10$

Figure 9: Choice probabilities as a function of ρ on the example network from Figure 7.

4 Parameter estimation approach

In this section, we discuss the estimation of the BPSqPL parameters from observed choice data. Let us assume that we observe N choices. An observation $n \in \{1, .., N\}$ has a choice set $\mathcal{C}_n$, and the index of the chosen alternative is given by $i_n \in \mathcal{C}_n$. To estimate the BPSqPL model and its special cases, we adopt a Constrained Maximum Likelihood Estimation (MLE) technique. The constraints on the parameter space exist to avoid the occurrence of zero probabilities for chosen alternatives, whose existence would raise zero likelihood. This is operationalised using a technique described in Duncan et al. (2022), originally developed for the BPSL model. If $\boldsymbol{\alpha}$ parametrises the cost function of any alternative and any choice set $c_i(\boldsymbol{\alpha})$. The parameter subspace is given by:

$$\Theta = \left\{ (\boldsymbol{\alpha}, \varphi), \forall n \in \{1, .., N\}, c_{i_n} \leq \varphi \min_{j \in \mathcal{C}_n} c_j \right\}$$

This also ensures that the likelihood function is continuous over the constrained parameter space. To ensure that the estimated parameters remain in Θ, we defined the following log-likelihood function. If β is the vector containing all the model parameters

$$LL(\beta) = \sum_{n=1}^{N} \begin{cases} \log(P_{i_n}^{\text{BPSqPL}}(\beta)) & \text{if } (\boldsymbol{\alpha}, \varphi) \in \Theta \\ -999 & \text{otherwise} \end{cases} \tag{19}$$

This formulation ensures that the couple $(\boldsymbol{\alpha}, \varphi)$ remains in the domain Θ when using maximisation algorithms. In this paper, we optimise the log-likelihood using the L-BFGS-B algorithm. A potential issue is that the BPSqPL likelihood function is not guaranteed to be concave. MLE solutions are not guaranteed to be unique. However, in our case study from Section 5, we plotted the likelihood surface around the estimated parameters and found no evidence of nonconcavity. We estimated the model using several randomly generated initial conditions and always found the same estimates. When choosing the initial conditions, it is important to ensure that the initial parameters $(\boldsymbol{\alpha}_0, \varphi_0)$ belong to the restricted space Θ. This can be done by first choosing the initial cost function parameters $\boldsymbol{\alpha}_0$, then calculating the highest relative cost over the observed choices, and choosing the initial bound as higher than this value, i.e.,

$$\varphi_0 \geq \max_{n \in \{1, \ldots, N\}} \frac{c_{i_n}(\boldsymbol{\alpha}_0)}{\min_{j \in \mathcal{C}_n} c_j(\boldsymbol{\alpha}_0)} := \varphi_{\min} \tag{20}$$

For instance, φ_0 can be drawn from a shifted exponential distribution at $\varphi_{\min}$.

The likelihood function of bounded models is non-differentiable (Cazor et al. 2024a); it is thus, in theory, impossible to derive its Hessian matrix analytically, which prevents deriving MLE standard errors. However, these standard errors can be derived numerically using methods such as bootstrapping (Efron and Tibshirani 1986), which we do in this paper.

5 Case study: Bicycle route choice in the Greater Copenhagen Area

This section benchmarks the various models presented in Table 2 against each other on a large-scale bicycle route choice dataset.

5.1 The Data

The case utilised a large-scale crowd-sourced data set of bicycle GPS trajectories received from Hövding[2]. The original dataset covers the entire Greater Copenhagen Area (see Figure 10) in the

[2] https://hovding.com

period from the 16th September 2019 until 31st May 2021. For a detailed description of the data, the bicycle network, and the algorithms applied for data processing, we refer to Łukawska et al. (2023). The final dataset for model estimation consists of a subset of this dataset containing 4,134 trips made by 4,134 cyclists.

Figure 10: Heatmaps of anonymised GPS trajectories from Hövding

The network can be represented as a directed graph $G = (V, E)$ where E is the set of links and V is the set of nodes. The network size is large, with $|E| = 420,973$ and $|V| = 324,492$. The network data was collected from Open Street Map (OSM[3]). The attributes of link $a \in E$ are as follows:

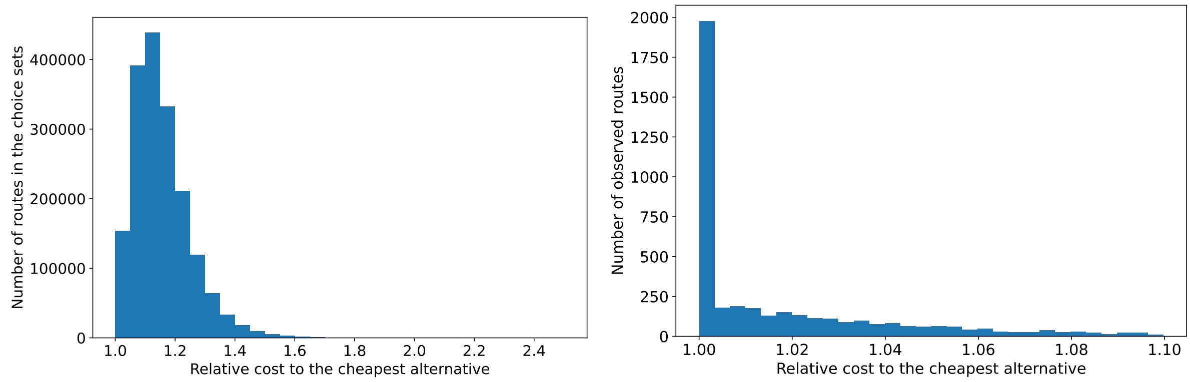
- L_a (km): Link length
- E_a (m): Link elevation gain when steepness $> 3.5\%$
- No_a (km): Link length without bicycle infrastructure
- S_a (km): Link length on a non-asphalt surface (i.e. gravel, cobblestones)
- W_a (km): Link length on wrong ways (cycling against traffic).

5.2 Choice set generation

Due to the extensive network size, it was not feasible to enumerate the universal set of routes between each set of observed origin and destination. We thus generated a representative universal choice set. This choice set generation algorithm uses a stochastic simulation approach (Nielsen 2004, Bovy and Fiorenzo-Catalano 2007), drawing a large number (10,000) of routes on the network between each OD, with random lengths. These lengths were normally distributed around the actual lengths, i.e., for each link $a \in A$ with length L_a, we drew a new value $\hat{L}_a \sim \mathcal{N}(L_a, \sigma L_a)$, where σ is a dispersion parameter. In this study, we used a value of $\sigma = 0.5$. These routes were then filtered with a local optimality criterion (Abraham et al. 2013, Fischer 2020), defined as the minimum length of a subpath that is not the shortest path. This criterion constrains the presence of small detours on routes and their mutual overlap. Figure 11 shows the relative cost distribution of the generated and observed routes. We observe that these

[3] www.openstreetmap.org

distributions differ, indicating that travellers may exclude certain generated routes from their choice sets, particularly those with high relative costs that are never chosen.



(a) Generated routes relative cost distribution (b) Observed routes relative cost distribution

Figure 11: Relative costs distribution of the generated choice set and observed routes, using the MNL estimates from Table 4.

5.3 Estimation results

In this case study, we compare the estimation results of 18 models:

- MNL, MNW, and qPL, which assume additive, multiplicative, and q-multiplicative relations, respectively, between the systematic and stochastic parts of random cost.
- MNL-PS, MNW-PS, and qPL-PS, which incorporate a simple path size correction term from Ben-Akiva and Bierlaire (1999) to account for route overlap. These "regular" path size models are estimated to benchmark with the more advanced path size corrected models (see next bullet). The models are presented in Appendix C.
- GPSL', GPSW', and GPSqPL', which incorporate more advanced path size correction terms that associated bounded models collapse into, see Appendix B.
- Bounded versions of the nine models above, i.e., BL, BW, BqPL, BL-PS, BW-PS, BqPL-PS, BPSL, BPSW, and BPSqPL. The difference between the BL-PS, BW-PS, BqPL-PS models and the BPSL, BPSW, BPSqPL models is that the former use the Ben-Akiva and Bierlaire (1999) path size term where overlap is captured between all route regardless of whether they are used or unused, whereas the latter use an appropriately-defined path size term to address this issue. The BL-PS, BW-PS, BqPL-PS models are given in Appendix C and the BPSL, BPSW, BPSqPL models are given in Apprndix B.

Every model is estimated in a "willingness-to-detour" space (analogous to the "willingness-to-pay" space), as the cost coefficient associated with Length is fixed to 1. Every other cost coefficient can be interpreted as the expected amount of detour travellers will take to avoid the attribute (e.g., wrong ways). For a network link a, we thus define its costt_a as:

$$t_a = L_a + \alpha_E \times E_a + \alpha_{\mathrm{No}} \times \mathrm{No}_a + \alpha_S \times S_a + \alpha_W \times W_a \tag{21}$$

The cost of a route i using the set of links $A_i \subseteq E$ is given by:

$$c_i = \sum_{a \in A_i} t_a$$

The objective is to study to what extent each generalisation of the logit impacts the model fit and interpretation. We will also investigate the number of alternatives cut off by each model's relative cost bound.

The results are presented in Tables 4 and 5. All the estimated parameters were significant at the 0.01 level. Each extension of the MNL and its one additional parameter significantly affects model fit, which can be seen in a decrease in the BIC for the more complex models. For the models for which the likelihood concavity was not guaranteed, we estimated the model with several sets of initial conditions, which were drawn from a uniform distribution within a reasonable range of values, using the method described in Equation 20 from Section 4. The final estimates were always the same, regardless of the initial conditions. The likelihood profiles were plotted to verify the local behaviour of functions around the model estimates, as shown in Figure 12 for the BPSqPL model. We observe that the optimization did not suffer from the potential non-concavity and converged to an optimum.

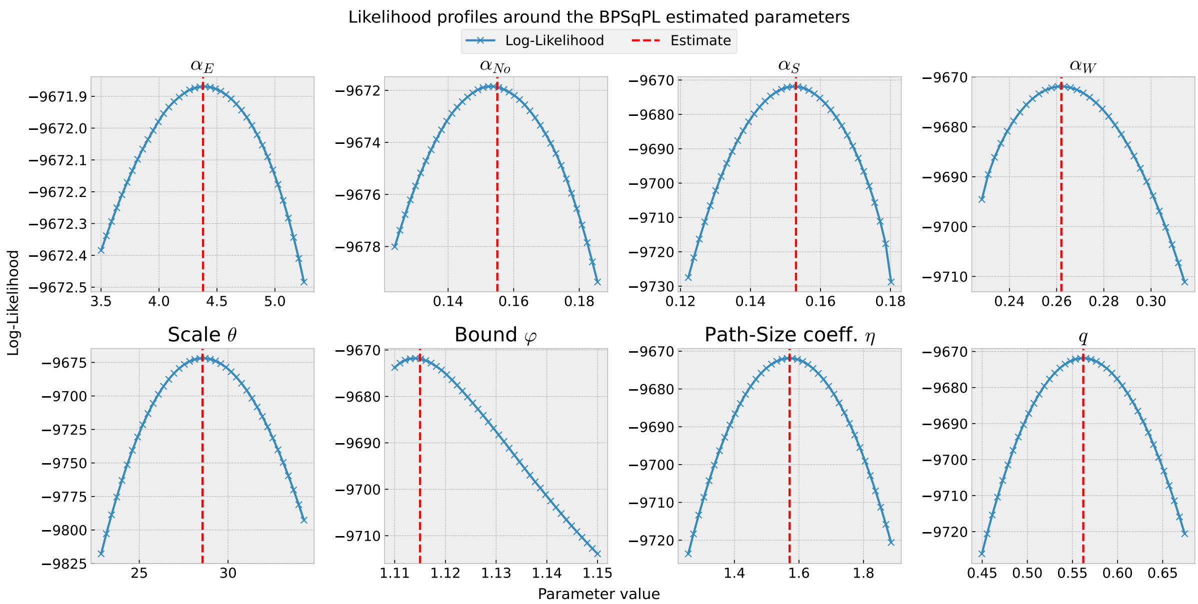
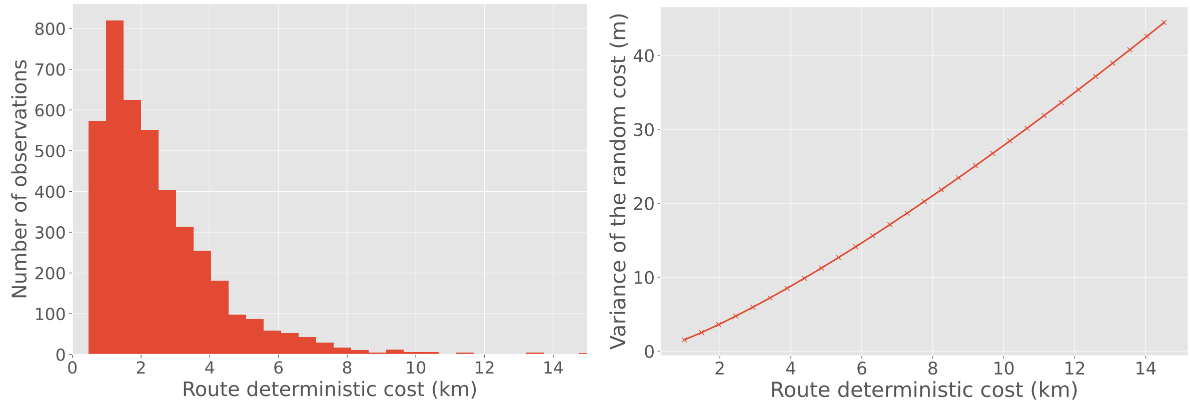


Figure 12: Likelihood profile around the BPSqPL model estimates

We observe that Weibit models fit our dataset better than their corresponding Logit models. The q-Product Logit, which generalises both models, also fits the data significantly better than the Weibit. The parameter q is estimated to be between 0.562 and 0.706, indicating that the perceived cost variance increases for longer trips. This increase is, however, lower than quadratic with systematic cost. In Figure 13a, we plot a histogram of the trip costs of the chosen routes in the estimation dataset, using the length-normalised cost function from the qPL model estimates from Table 4. In Figure 13b, we plot how the variance of the route costs increases with their systematic value, again using the qPL estimated parameters and the variance computation method from Chikaraishi and Nakayama (2016). This plot shows a non-linear increase of the variance with deterministic cost.

For the bounded models, which account for choice set formation (Table 5), we see that the relative cost bound is rather stable across models, estimated between 1.102 and 1.162. This suggests that cyclists never considered using routes that were more than 10.2% to 16.2% more costly than the cheapest in the choice set. We also observe that non-considered routes (allocated zero probability) represented between 41.9 and 68.5% of the pre-generated routes, suggesting that our choice-set generation method included a significant amount of behaviorally unrealistic routes.

We also observe that the scale is the only parameter that varies between models, while the



(a) Histogram of the normalised trip costs in the estimation dataset

(b) Simulated variance of the qPL random costs in function of the deterministic cost

Figure 13: Demonstration of the increasing variance for longer trips on the observed route, using the qPL estimated parameters

taste parameters are rather stable. The different scales are linked to the different assumed distributions for the error terms (Gumbel, q-Gumbel or Weibull), whose variance behaves differently when route length increases. For instance, the models output a willingness to detour between 15.1% and 18.4% to ride on bicycle infrastructure rather than in mixed traffic on the car lanes. We also see that, for every model, the willingness to detour to avoid one meter of steep elevation is rather low and fluctuates relatively more (between 0.26% and 0.78% per meter of elevation). This is most likely due to the flat topography of the Greater Copenhagen Area. Overall, This implies that accounting for heteroscedasticity, choice set formation, and alternative correlation does not change the taste interpretation (at least in this dataset) but is of good value for getting closer to realistic substitution patterns between routes and more reliable forecasting.

Notably, when comparing Path size specifications, the more advanced and consistent one greatly improves the fit (GPSx' vs. x-PS models and BPSx models vs. Bx-PS models). This suggests that this weighting is more behaviorally realistic. It is also worth noticing that, while there is only a small improvement in fit between the qPL and BqPL (respectively, MNW and BW), it is much larger between the GPSqPL' and its bounded version, the BPSqPL (respectively, GPSW' and BPSW). This suggests that combining these extensions, in this case, is more powerful than using only some of them.

Model	MNL	qPL	MNW	MNL-PS	MNW-PS	qPL-PS	GPSL'	GPSW'	GPSqPL'
Length	-	-	-	-	-	-	-	-	-
Elevation gain	0.00352	0.00658	0.0075	0.0027	0.0065	0.0047	0.00354	0.00785	0.00537
No Bike infrastructure	0.182	0.183	0.184	0.153	0.156	0.154	0.15	0.182	0.153
Non-smooth surface	0.194	0.2	0.202	0.156	0.168	0.162	0.143	0.198	0.152
Wrong way	0.332	0.358	0.368	0.278	0.313	0.3	0.242	0.387	0.262
Scale (θ)	-28.5	-51.76	-64.13	-33.61	-78.6	-58.69	-19	-53.9	-32.06
Path size term (η)	-	-		1.105	2.344	1.09	1.608	1.09	1.522
q	-	0.706	-	-	-	0.671	-	-	0.624
Final LL	**-11,087**	**-10,363**	**-10,495**	**-10,881**	**-10,300**	**-10,172**	**-10,315**	**-10,206**	**-9,747**
BIC	**22,192**	**20,748**	**21,008**	**21,784**	**20,622**	**20,369**	**20,652**	**20,434**	**19,519**
Adj. ρ^2	**0.513**	**0.545**	**0.539**	**0.522**	**0.548**	**0.553**	**0.547**	**0.552**	**0.572**
N. of parameters	**5**	**6**	**5**	**6**	**6**	**7**	**6**	**6**	**7**

Table 4: Model estimates, unbounded models

Model	BL	B-qPL	BW	BL-PS	BW-PS	BqPL-PS	BPSL	BPSW	BPSqPL
Length	-	-	-	-	-	-	-	-	-
Elevation gain	0.0031	0.00616	0.0071	0.0026	0.0065	0.0045	0.0036	0.0045	0.00438
No Bike infrastructure	0.17	0.181	0.183	0.152	0.155	0.151	0.159	0.157	0.155
Non-smooth surface	0.171	0.198	0.201	0.158	0.167	0.16	0.153	0.159	0.153
Wrong way	0.31	0.356	0.367	0.291	0.312	0.296	0.267	0.274	0.262
Scale (θ)	-25.42	-50.5	-63.77	-27.89	-78.59	-58.68	-14.71	-42.28	-28.56
Path size term (η)	-	-	-	1.088	2.34	1.122	1.643	1.4	1.571
q	-	0.688	-	-	-	0.67	-	-	0.562
Relative cost bound φ	1.111	1.149	1.159	1.102	1.162	1.126	1.105	1.115	1.115
Final LL	**-10,791**	**-10,357**	**-10,493**	**-10,557**	**-10,299**	**-10,163**	**-9,910**	**-9,876**	**-9,683**
BIC	**21,604**	**20,739**	**21,008**	**21,139**	**20,623**	**20,355**	**19,845**	**19,777**	**19,395**
Adj. ρ^2	**0.526**	**0.545**	**0.539**	**0.536**	**0.548**	**0.554**	**0.565**	**0.566**	**0.575**
N params	**6**	**7**	**6**	**7**	**7**	**8**	**7**	**7**	**8**
% of routes cut by φ	**66.3%**	**53.4%**	**49.9%**	**68.5%**	**41.9%**	**56.8%**	**65.7%**	**61.5%**	**60.3%**

Table 5: Model estimates, bounded models

5.4 Estimate precision

For the bounded models, the likelihood function is non-differentiable at certain points. This may prevent calculating the Hessian if the estimates are located near these points. We thus do not assess estimate precision by analytically computing standard errors of estimates using the inverse Hessian, as is commonly done. We instead adopt a bootstrapping approach, which is also a common procedure (see e.g., Duncan et al. (2022); Petrin and Train (2003)). An example of the bootstrapping results has been plotted in Figure 14 for the BqPL model. As can be seen, the histograms of the bootstrapping estimates resemble normal distributions, which are fitted and plotted in red. This suggests asymptotic normality of the maximum likelihood estimates.

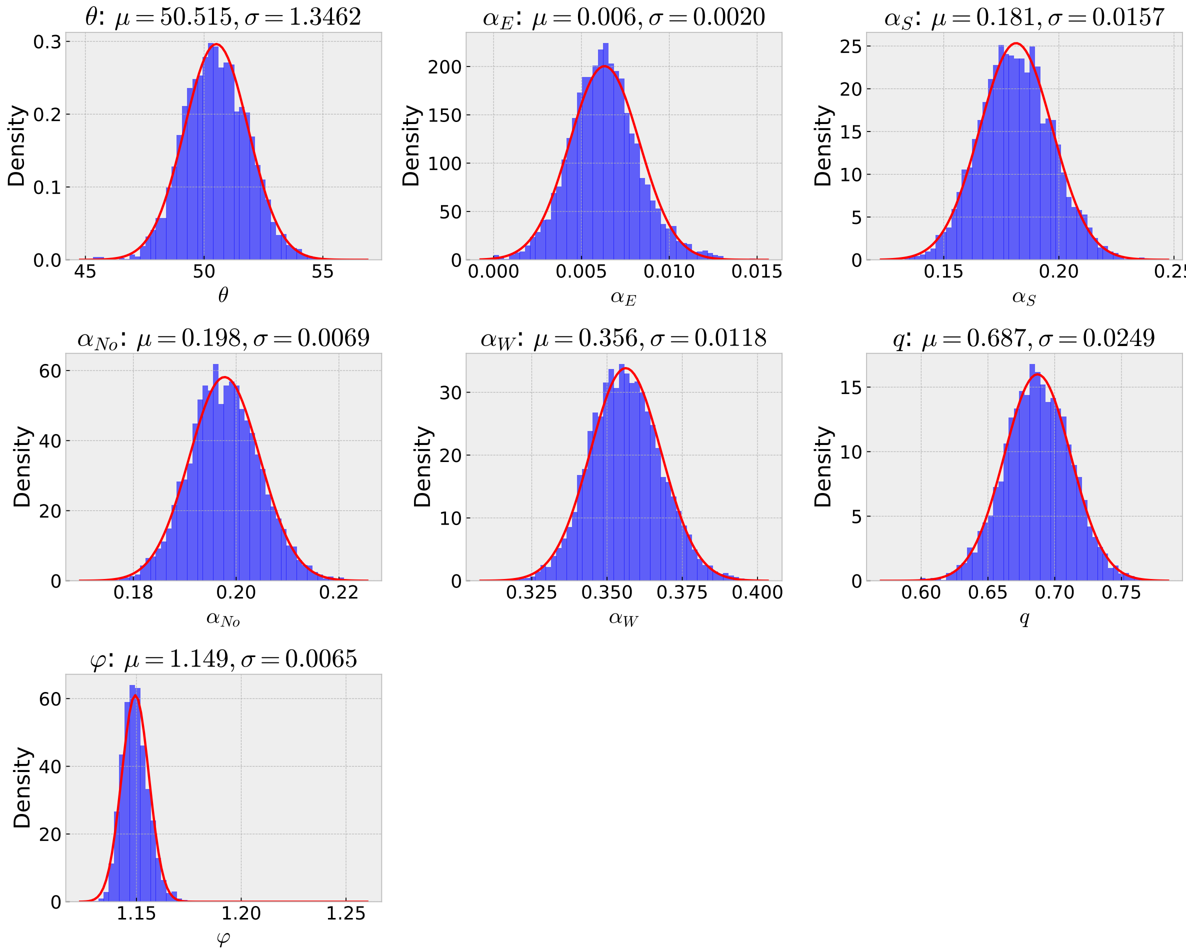


Figure 14: Distributions of the bootstrapping estimates for the BqPL model over $N = 5000$ random draws of observations

The bootstrapping estimates and standard errors are presented in Tables 6 and 7. These estimates facilitate conducting t-tests on the model parameters. The results indicate that, for all models, every parameter of the cost function—except the elevation gain parameter—is significantly different from zero at the 0.1% confidence level. The elevation gain parameter is only significant at the 5% level across all models. Both the scale and path-size terms are also significantly different from zero at the 0.1% confidence level. When estimated, the parameter q is significantly different from both zero and one at the 0.1% confidence level. Additionally, the relative cost bound is significantly different from infinity at the 0.1% confidence level. Note that to test the relative cost bound φ against infinity is equivalent to testing its inverse φ^{-1} against zero. From Daly et al. (2012), the standard error of φ^{-1} is given by $\mathrm{se}(\varphi^{-1}) = \mathrm{se}(\varphi)/\varphi^2$, which allows for running the t-test against zero.

Model	MNL	qPL	MNW	GPSL'	GPSW'	GPSqPL'
Length	-	-	-	-	-	-
Elevation gain	0.0038 (0.0019)	0.0067 (0.0021)	0.0075 (0.0021)	0.00369 (0.0016)	0.00667 (0.0023)	0.00547 (0.0018)
No Bike infrastructure	0.182 (0.0179)	0.183 (0.0165)	0.182 (0.0171)	0.153 (0.0134)	0.152 (0.0124)	0.150 (0.0124)
Non-smooth surface	0.194 (0.0078)	0.2 (0.0069)	0.201 (0.0070)	0.143 (0.0062)	0.159 (0.0056)	0.152 (0.0046)
Wrong way	0.332 (0.0123)	0.358 (0.0121)	0.366 (0.0129)	0.243 (0.0095)	0.387 (0.0105)	0.262 (0.0101)
Scale (θ)	-28.6 (0.6975)	-51.77 (1.2761)	-64.42 (1.3027)	-19.02 (0.3346)	-45.20 (0.7080)	-32.01 (0.6483)
Path size term (η)	-	-		1.614 (0.0042)	1.308 (0.0374)	1.519 (0.0357)
q	-	0.706 (0.0222)	-	-	-	0.625 (0.0153)

Table 6: Bootstrapping estimates, unbounded models

Model	BL	BqPL	BW	BPSL	BPSW	BPSqPL
Length	-	-	-	-	-	-
Elevation gain	0.0045 (0.0023)	0.0063 (0.0020)	0.0075 (0.0022)	0.0034 (0.0007)	0.0050 (0.0016)	0.0044 (0.0023)
No Bike infrastructure	0.164 (0.0103)	0.181 (0.0157)	0.181 (0.0166)	0.149 (0.0080)	0.156 (0.0112)	0.151 (0.0109)
Non-smooth surface	0.176 (0.0062)	0.198 (0.0069)	0.201 (0.0071)	0.150 (0.0032)	0.158 (0.0045)	0.153 (0.0048)
Wrong way	0.321 (0.0121)	0.356 (0.0118)	0.365 (0.0128)	0.263 (0.0070)	0.272 (0.0100)	0.258 (0.0081)
Scale (θ)	-24.96 (0.6844)	-50.5 (1.346)	-63.99 (1.371)	-14.69 (0.3415)	-42.45 (1.025)	-28.55 (0.7305)
Path size term (η)	-	-	-	1.655 (0.0396)	1.405 (0.0367)	1.497 (0.0419)
q	-	0.688 (0.0249)	-	-	-	0.656 (0.0160)
Relative cost bound φ	1.110 (0.0024)	1.149 (0.0065)	1.159	1.102 (0.0011)	1.114 (0.0024)	1.117 (0.0024)

Table 7: Bootstrapping estimates, bounded models

5.5 Experiments on choice sets

As it is computationally infeasible to generate the universal set of routes, even when restricting their relative cost to the cheapest alternative to 1.1, our model estimation relied on pre-generated choice sets, known as the "representative universal choice set". In this section, we investigate the influence of the composition of this representative choice set on the parameter estimates. To do this, we follow a similar experiment to that conducted by Prato and Bekhor (2007). Rather than using choice sets of up to 10,000 alternatives (as described in 5.2), which becomes too large to efficiently estimate the Path-Size corrected models, we chose to operate with a maximum size of 1,000 alternatives for this experiment. The procedure was as follows:

- A maximum of 1,000 unique routes were randomly drawn for each OD from the 10,000 routes (or fewer routes if less than 1,000 unique routes among the 10,000 generated routes). We denote these the *initial choice sets.*
- We estimated the BPSqPL model using these initial choice sets.
- For each OD, we randomly drew respectively 25, 50, and 75% of the alternatives from their initial choice sets, making sure we included the chosen alternative, to form new choice sets.
- We re-estimated the BPSqPL with these new choice sets.
- We repeated the last two steps 5 times to assess the variability of the estimates

Table 8 presents the average estimates for the model using subsets of the initial choice sets, with standard deviations between experiments shown in italics. Figure 15 illustrates the average relative parameter estimates compared to their initial values (estimated using the initial choice sets), along with the standard deviations across the 5 repetitions (also listed in Table 8). The low standard deviations in Table 8 indicate that the estimates and final likelihoods remained stable across experiments (α_E is the least stable parameter as it is the least significant). The taste parameters of the cost function ($\boldsymbol{\alpha}$, as given in Equation 21) are generally stable across the choice set sizes. The relative cost bound is the most stable parameter, suggesting that a large choice set is not essential for its identification. Among the parameters, the Path-Size term (η)

exhibits the most variation across the choice set sizes, consistent with Prato and Bekhor (2007) for the PSL model. Meanwhile, the scale parameter (θ) increases in absolute value as the choice set grows. This suggests that when fewer routes are included, choice probabilities become more dispersed across the choice set.

Relative size to full choice sets		α_E	α_{No}	α_S	α_W	η	θ	φ	q	Final LL
100%		3.835	0.128	0.130	0.234	1.936	-39.33	1.107	0.641	-10,874.2
75%	Mean	4.133	0.129	0.130	0.235	1.998	-38.37	1.107	0.647	-9,966.9
	SD	*0.279*	*0.0013*	*0.0006*	*0.0008*	*0.0045*	*0.0880*	*0.0007*	*0.0025*	*7.04*
50%	Mean	4.100	0.130	0.131	0.238	2.084	-37.02	1.105	0.657	-8,762.4
	SD	*0.316*	*0.0016*	*0.0014*	*0.0023*	*0.0093*	*0.391*	*0.0008*	*0.0093*	*8.23*
25%	Mean	3.966	0.134	0.134	0.244	2.263	-34.05	1.104	0.666	-6,889.0
	SD	*0.383*	*0.0040*	*0.0020*	*0.0007*	*0.0071*	*0.533*	*0.0009*	*0.0126*	*16.79*

Table 8: Comparison of the initial BPSqPL estimates on the whole choice set and the mean estimates of the BPSqPL using the reduced choice sets. Standard deviation values across experiments are shown in italics.

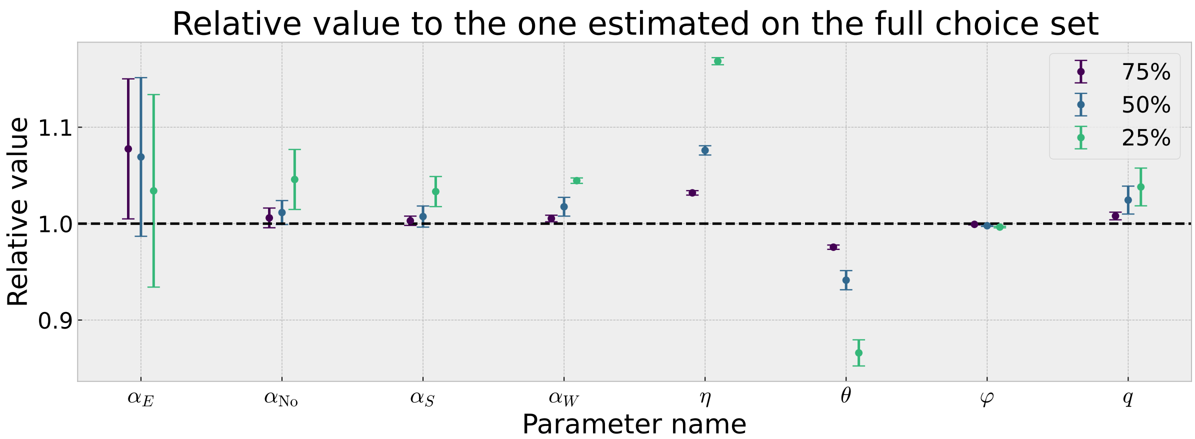


Figure 15: Relative parameter estimates for the different choice set subsets. The error bars output the standard deviation of the relative parameter estimates across experiments.

5.6 Model validation

We performed Monte-Carlo cross-validation on our dataset to assess the model's predictive performance and test for overfitting. To do so, we repeated $N = 10$ (as in Cazor et al. (2024b)) times the following steps:

1. Randomly split the original dataset S into a training S_t and validation set S_v ($|S_t| = 0.7|S|; |S_v| = 0.3|S|$).

2. Estimate all the models on the training set S_t, obtain, for each model m, the training parameters β_m^t.

3. Calculate, for each model m, the log-likelihood on the validation set, $\text{LL} = \sum_{x \in S_v} \log P_{i_x}^m(\beta_m^t)$, where i_x is the index of the chosen alternative for observation $x \in S_v$

The cross-validation results are shown in Figures 16 and Table 9. We observe that these results corroborate the fit values found in Tables 4 and 5, i.e., that the BPSqPL is the best-performing model in every cross-validation experiment, followed by the BPSL, the BqPL, the qPL, the BL and the MNL. This indicates that these models did not overfit the data. Moreover, we observed that the model estimates across the experiments were stable.

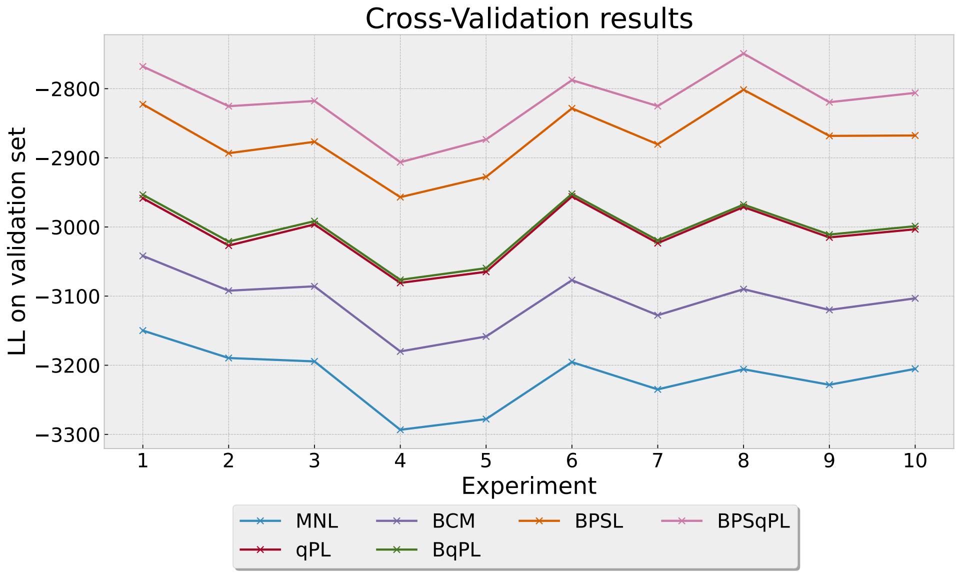


Figure 16: Cross-validation results

Model	MNL	qPL	BL	BqPL	BPSL	BPSqPL
Average LL	-3,217.48	-3,009.63	-3,107.61	-3,005.12	-2,872.35	-2,817.81

Table 9: Average log-likelihood on the cross-validation sets

6 Conclusion

In this paper, we have developed new closed-form choice models that generalise the MNL to account for heteroscedasticity of the error terms, correlations between overlapping routes, and choice set formation effects through setting a bound on the relative cost distribution. We first developed a model that combines Chikaraishi and Nakayama (2016)'s q-Product Logit (qPL) model with Watling et al. (2018)'s Bounded Logit (BL) model, to derive a Bounded q-Product Logit (BqPL) model. It is derived by assuming a Truncated q-Log-Logistic distribution for random error term differences and can be seen as a one-stage choice set formation model. We then extended the BqPL model to account for route overlap in a fashion similar to Duncan et al. (2022)'s BPSL model.

The BPSqPL model remains parsimonious and easy to estimate by introducing only one additional parameter per property. We also proposed simplified versions of this model that omit one or more properties (e.g., a Bounded Path Size Weibit model, which has a less flexible form of heteroscedasticity). Model properties were first demonstrated using small-scale examples and then benchmarked against existing models in a large-scale bicycle route choice case study. We presented a method to estimate the BPSqPL using a constrained version of the Maximum Likelihood Estimation technique. As the likelihood function of bounded models is non-differentiable, we used a bootstrapping method to compute standard errors of the estimates.

Our findings from the large-scale application show that accounting for heteroscedasticity, route overlap, and choice set formation considerably improves model fit and predictive accuracy. The estimate for the q parameter also indicates that error term variance does increase with expected route cost but is slightly slower than the quadratic increase assumed by a Weibit model. These enhancements were achieved without a substantial increase in model complexity. Additionally, these models offer more realistic substitution patterns between alternatives due to consistent path size corrections and the exclusion of unrealistic alternatives, which account for about 50% of the generated representative universal choice set of routes. The values of the taste coefficients for the observed attributes also remained stable across models. In the route choice context, while it is computationally intractable to generate the universal choice sets for large-scale case studies, we studied the choice set robustness of the estimated parameters. The results indicate that the choice set size has little influence on the estimated taste parameters and relative cost bound while having some influence on the path-size correction and scale parameters.

In this study, the focus has been on route choice modelling. There is scope, however, for applying the BPSqPL model to other choice situations. For example, activity-based models (Danalet and Bierlaire 2015; Pougala et al. 2023) adopt a similar network representation, where the correlation between alternatives can be captured through path size correction terms. One could develop a cross-nested version of the BqPL model to capture correlations between alternatives for choice situations that are not network-based. Without the correlation, the BqPL model is still a useful model for capturing heteroscedasticity and cost bounds on alternatives, such as mode or destination choice. Another potential stream of research could be to include route correlation in the choice set formation process, possibly by devising a choice model which sets a bound on the overlap between routes. This could perhaps be done by adopting Rasmussen et al. (2024)'s conjunctive bounded logit model.

Acknowledgements

We thank the Independent Research Fund Denmark (DFF) for funding this research through 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).

Appendix A Bounding conditions

As can be seen in the BqPL model choice probabilities, the parameter q not only influences how the variance of error terms increases with systematic cost but also on the condition on which an alternative receives a non-zero probability, which we refer to as the *bounding condition*.

According to the BqPL, the choice probability of an alternative $i \in \mathcal{R}$ will be greater than 0 if and only if, for some $\phi > 1$, $c_i \oslash_q \phi \otimes_q V_{r^*}$ is larger than 0. This happens when c_i is lower than $V_{r^*} \otimes_q \phi$, i.e., the systematic cost is lower than the q-product between the reference alternative systematic cost and the bound. When $q = 0$, this happens when c_i is lower than $c_{r^*} + (\phi - 1)$, meaning that the difference of costs is bounded: it is an *absolute bounding condition*. When $q = 1$, it happens when c_i is lower than ϕc_{r^*}, meaning that the cost ratio is bounded: it is a *relative bounding condition*. While relative absolute bounding conditions are intuitive, it is harder to understand what the q-product bounding condition means. To illustrate it, we plot in Figure 17 how the q-Ratio translates into an absolute bounding condition for different values of q. To do so, we plot, for a fixed value of ϕ, the value φ_{abs} so that

$$c_i \oslash_q c_{r^*} = \phi \iff c_i - c_{r^*} = \varphi_{abs}$$

We plot the maximum difference $c_i - c_{r^*}$ so that i has a non-zero probability as a function of c_i. We see that, for $q = 0$, it is constant to $\phi - 1 = 1$. This means that, whatever the value of c_{r^*}, c_i must be at most one unit higher than c_{r^*} to be given a non-zero probability. For $q = 1$, as expected, the absolute bound increases linearly with c_{r*} (for $c_{r^*} = 1$, we need $c_i < 2$, i.e. a difference of 1, for $c_{r^*} = 10$, we need $c_i < 20$, i.e. a difference of 10). The absolute bound grows with c_i, but slower than linearly for any value in between.

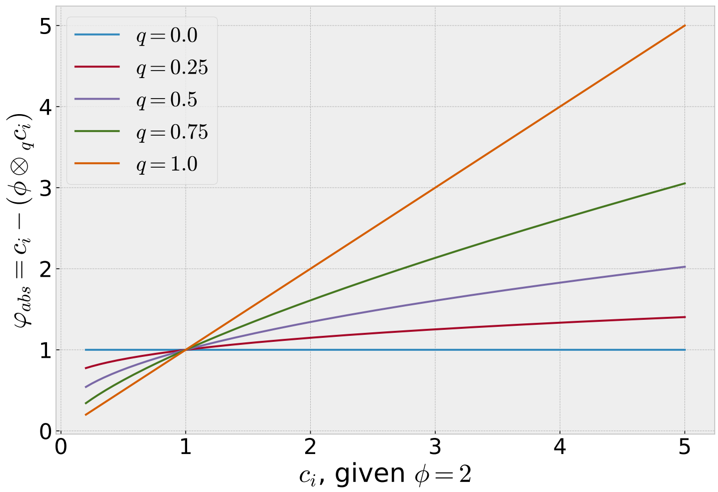


Figure 17: q-Ratio bounding condition expressed as an absolute bounding condition.

Similarly, in Figure 18, we plot how a q-Relative could translate into a variable relative bound by plotting the value φ_{rel} so that

$$c_i \oslash_q c_{r^*} = \phi \iff c_i / c_{r^*} = \varphi_{rel}$$

This plot shows a constant relative bound for $q = 1$, as the MNW already has a relative bounding condition. For $q < 1$, the relative bound decreases with c_i, which means the bound on the ratio c_i / c_{r^*} gets smaller with larger c_i. For the extreme case $q = 0$, ϕ_{rel} is inversely proportional to c_i (for $c_i = 11$, i will have zero probability when $c_{r^*} \leq 10$, i.e. $c_i / c_{r^*} \geq 1.1$, while when $c_i = 2$, i will have zero probability when $c_{r^*} \leq 1$, i.e. $c_i / c_{r^*} \geq 2$).

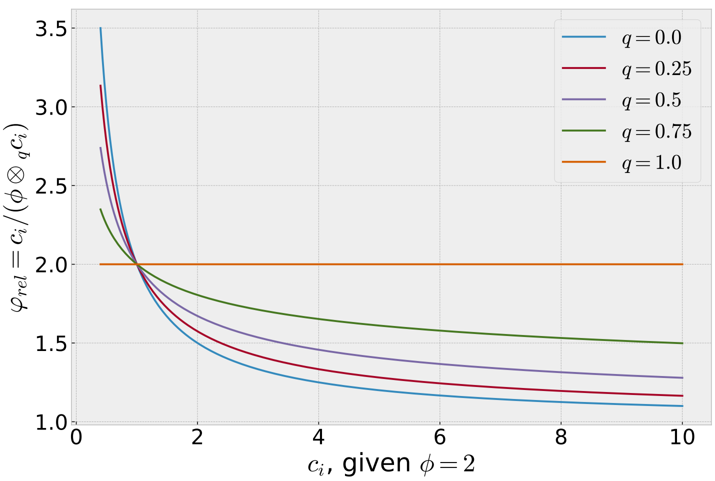


Figure 18: q-Ratio bounding condition expressed as a relative bounding condition.

Appendix B Special cases of the BPSqPL model

In this section, we detail the choice probabilities of the special cases (i.e., the models that can be obtained by fixing one or more parameters) of the BPSqPL model developed in this paper. Some of these models have been already developed in the literature, some others are new. Again, let us assume a route choice situation with a choice set $\mathcal{R}$. Each route $i \in \mathcal{R}$ consists of a set of links $A_i \subseteq A$, where A is the set of links in the studied network. These links are defined by attributes aggregated in positive cost functions $t_a, a \in A$, parameterised by a vector of parameters $\boldsymbol{\alpha}$. The cost of route i is defined by some observed attributes aggregated in a cost function c_i. This cost of route i is link-additive, i.e. $c_i = \sum\limits_{a \in A_i} t_a$. For fixed parameters $\boldsymbol{\alpha}, \varphi$, we define $\bar{\mathcal{R}}(\boldsymbol{\alpha}, \varphi) \subseteq \mathcal{R}$ the subset of routes with a cost less than $\varphi > 1$ times the cheapest route cost:

$$\bar{\mathcal{R}}(\boldsymbol{\alpha}, \varphi) = \left\{ i \in \mathcal{R}, c_i \leq \varphi \min_{l \in \mathcal{R}} c_l \right\}$$

B.1 When $q = 0$: the BPSL model (Duncan et al. 2022)

The BPSL model is obtained when the heteroscedasticity parameter q is fixed to 0.

$$P_i^{\text{BPSL}} = \begin{cases} \dfrac{(\gamma_i^{\text{BPSL}})^\eta \left(\exp(-\theta(c_i - \varphi \min\limits_{l \in \mathcal{R}} c_l)) - 1 \right)}{\sum\limits_{j \in \bar{\mathcal{R}}} \left(\gamma_j^{\text{BPSL}}\right)^\eta \left(\exp(-\theta(c_j - \varphi \min\limits_{l \in \mathcal{R}} c_l)) - 1 \right)} & \text{if } i \in \bar{\mathcal{R}} \\ 0 & \text{otherwise} \end{cases} \tag{22}$$

where $(\gamma_i^{\text{BPSqPL}})^\eta$ is the path size correction factor for considered route $i \in \bar{\mathcal{R}}$ (non-considered routes do not have path size terms). $\eta \geq 0$ is the path size scaling parameter scaling sensitivity to route distinctiveness, and $\gamma_i^{\text{BPSqPL}} \in (0, 1]$ is the path size term for considered route $i \in \bar{\mathcal{R}}$, calculated as follows:

$$\gamma_i^{\text{BPSL}} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{\left(\exp(-\theta(c_i - \varphi \min\limits_{l \in \mathcal{R}} c_l)) - 1 \right)}{\sum_{j \in \bar{\mathcal{R}}} \left(\exp(-\theta(c_i - \varphi \min\limits_{l \in \mathcal{R}} c_l)) - 1 \right) \delta_{aj}} \tag{23}$$

B.2 When $q = 1$: the BPSW model

The BPSW model is obtained when the heteroscedasticity parameter q is fixed to 1.

$$P_i^{\text{BPSW}} = \begin{cases} \dfrac{(\gamma_i^{\text{BPSW}})^\eta \left((c_i / \varphi \min\limits_{l \in \mathcal{R}} c_l)^{-\theta} - 1 \right)}{\sum\limits_{j \in \bar{\mathcal{R}}} \left(\gamma_j^{\text{BPSW}}\right)^\eta \left((c_j / \varphi \min\limits_{l \in \mathcal{R}} c_l)^{-\theta} - 1 \right)} & \text{if } i \in \bar{\mathcal{R}} \\ 0 & \text{otherwise} \end{cases} \tag{24}$$

where $(\gamma_i^{\text{BPSW}})^\eta$ is the path size correction factor for considered route $i \in \bar{\mathcal{R}}$ (non-considered routes do not have path size terms). $\eta \geq 0$ is the path size scaling parameter scaling sensitivity to route distinctiveness, and $\gamma_i^{\text{BPSW}} \in (0, 1]$ is the path size term for considered route $i \in \bar{\mathcal{R}}$, calculated as follows:

$$\gamma_i^{\text{BPSW}} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{\left((c_i / \varphi \min\limits_{l \in \mathcal{R}} c_l)^{-\theta} - 1 \right)}{\sum_{j \in \bar{\mathcal{R}}} \left((c_j / \varphi \min\limits_{l \in \mathcal{R}} c_l)^{-\theta} - 1 \right) \delta_{aj}} \tag{25}$$

B.3 When $\varphi \to +\infty$: the GPSqPL' model

The GPSqPL' model is obtained when the relative cost bound tends to infinity.

$$P_i^{\text{GPSqPL'}} = \frac{\left(\gamma_i^{\text{GPSqPL'}}\right)^{\eta} \exp(-\theta \ln_q(c_i))}{\sum\limits_{j \in \mathcal{R}} \left(\gamma_j^{\text{GPSqPL'}}\right)^{\eta} \exp(-\theta \ln_q(c_j))} \tag{26}$$

where $(\gamma_i^{\text{GPSqPL'}})^{\eta}$ is the path size correction factor for route $i \in \mathcal{R}$. $\eta \geq 0$ is the path size scaling parameter scaling sensitivity to route distinctiveness, and $\gamma_i^{\text{GPSqPL'}} \in (0, 1]$ is the path size term for route $i \in \mathcal{R}$, calculated as follows:

$$\gamma_i^{\text{GPSqPL'}} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{\exp(-\theta \ln_q(c_i))}{\sum_{j \in \mathcal{R}} \exp(-\theta \ln_q(c_j)) \delta_{aj}} \tag{27}$$

B.4 When $\varphi \to +\infty$ and $q = 0$: the GPSL' model (Duncan et al. 2020)

The GPSL' model is obtained when the relative cost bound tends to infinity.

$$P_i^{\text{GPSL'}} = \frac{\left(\gamma_i^{\text{GPSL'}}\right)^{\eta} \exp(-\theta c_i)}{\sum\limits_{j \in \mathcal{R}} \left(\gamma_j^{\text{GPSL'}}\right)^{\eta} \exp(-\theta c_j)} \tag{28}$$

where $(\gamma_i^{\text{GPSL'}})^{\eta}$ is the path size correction factor for route $i \in \mathcal{R}$. $\eta \geq 0$ is the path size scaling parameter scaling sensitivity to route distinctiveness, and $\gamma_i^{\text{GPSL'}} \in (0, 1]$ is the path size term for route $i \in \mathcal{R}$, calculated as follows:

$$\gamma_i^{\text{GPSqPL'}} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{\exp(-\theta c_i)}{\sum_{j \in \mathcal{R}} \exp(-\theta c_j) \delta_{aj}} \tag{29}$$

B.5 When $\varphi \to +\infty$ and $q = 1$: the GPSW' model

The GPSW' model is obtained when the relative cost bound tends to infinity and the heteroscedasticity parameter q is fixed to 1.

$$P_i^{\text{GPSW'}} = \frac{\left(\gamma_i^{\text{GPSW'}}\right)^{\eta} c_i^{-\theta}}{\sum\limits_{j \in \mathcal{R}} \left(\gamma_j^{\text{GPSW'}}\right)^{\eta} c_j^{-\theta}} \tag{30}$$

where $(\gamma_i^{\text{GPSW'}})^{\eta}$ is the path size correction factor for route $i \in \mathcal{R}$. $\eta \geq 0$ is the path size scaling parameter scaling sensitivity to route distinctiveness, and $\gamma_i^{\text{GPSW'}} \in (0, 1]$ is the path size term for route $i \in \mathcal{R}$, calculated as follows:

$$\gamma_i^{\text{GPSW'}} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{c_i^{-\theta}}{\sum_{j \in \mathcal{R}} c_j^{-\theta} \delta_{aj}} \tag{31}$$

Appendix C Path size and Bounded Path Size models

We benchmarked the "consistent" path size models in the case study to the typically considered Path size models. The Path size models use the Path size correction below (Ben-Akiva and Bierlaire 1999):

$$\gamma_i^{\mathrm{PS}} = \sum_{a \in A_i} \frac{t_a}{c_i} \frac{1}{\sum_{j \in \mathcal{R}} \delta_{aj}} \tag{32}$$

It is then possible to obtain the following models:

- The MNL-PS model, combining the choice probabilities from Equation 28 with the correction from Equation 32
- The MNW-PS model, combining the choice probabilities from Equation 30 with the correction from Equation 32
- The qPL-PS model, combining the choice probabilities from Equation 26 with the correction from Equation 32
- The BL-PS model, combining the choice probabilities from Equation 22 with the correction from Equation 32
- The BW-PS model, combining the choice probabilities from Equation 24 with the correction from Equation 32
- The BL-PS model, combining the choice probabilities from Equation 17 with the correction from Equation 32

Appendix D The BqPL: a relative utility model

In the derivation of the BqPL model, we introduced a random utility U_i for each alternative $i \in \mathcal{C}$, along with an imaginary reference alternative U_{r^*}. However, these utilities were not explicitly defined in terms of random distributions. Instead, we only assumed that their q-ratio, $U_i \oslash_q U_{r^*}$, followed a truncated log-logistic distribution. This approach primarily helped to establish parallels between the BL and MNL models, as well as between the BqPL and qPL models.

In this section, we present an alternative derivation that does not require explicit utility definitions, leading to a more direct interpretation of the model.

D.1 The relative utility framework

The concept of *relative utility* has been widely discussed in the choice modeling literature (see, e.g., Duesenberry 1949; Stadt et al. 1985; Zhang et al. 2004) and is based on the hypothesis that utilities are only meaningful in comparison to other alternatives. In most discrete choice models, choice probabilities remain unchanged under transformations such as shifting (adding a constant to all utilities in additive RUMs), scaling (multiplying all utilities by a constant in multiplicative RUMs), or *q-multiplication* (in more recent qRUMs, e.g., Nakayama (2013)). This invariance implies that these models do not require defining an absolute utility function for each alternative; instead, utilities are defined *relative* to a reference (or pivot) alternative.

Behaviorally, this aligns with the idea that decision-makers compare alternatives rather than evaluating them against some absolute metric. In this appendix, we present the concept of relative utility through the lens of qRUMs, which generalize both additive and multiplicative RUMs. Rather than assigning absolute utilities to alternatives, we introduce the notion of *random relative utility*. For an alternative i in a choice set $\mathcal{C}$, we define its random relative utility RU_i as:

$$RU_i = RV_i \otimes_q \varepsilon_i$$

where RV_i represents the *deterministic relative utility*, based on the observed attributes. In this study, we specify:

$$RV_i = V_i \oslash_q V_{r^*},$$

where the deterministic utility of alternative i is expressed relative to the utility of a reference alternative r^* using a q-ratio transformation. We further specify the reference alternative as the alternative with the highest deterministic utility, i.e.,

$$V_{r^*} = \max_{j \in \mathcal{C}} V_j.$$

To determine the choice outcome, the random relative utility is compared to 1: if $RU_i > 1$, the alternative i is preferred over the reference alternative; otherwise, it is not. The probability of selecting alternative i over the reference alternative is then given by:

$$\begin{aligned}
\mathbb{P}(i|\{i, r^*\}) &= \mathbb{P}(RU_i \geq 1) \\
&= \mathbb{P}(\varepsilon_i \geq 1 \oslash_q RV_i) \\
&= 1 - F_{\varepsilon_i}(1 \oslash_q RV_i)
\end{aligned}$$

Where F_{ε_i} denotes the CDF of ε_i. Depending on the assumed distribution of $\boldsymbol{\varepsilon} = (\varepsilon_i)_{i \in \mathcal{C}}$, several choice models can be derived. In the following section, we derive the BqPL model introduced in Section 2 of the paper.

D.2 Alternative derivation of the BqPL model

To derive the BqPL model, we assume that the error terms ε_i follow independent Truncated q-LogLogistic distributions defined in Equation 8 with mean 1, scale θ, heteroscedasticity parameter q and truncation parameter ϕ:

$$\varepsilon_i \sim \text{Trunc. q-LogLogistic}(1, \theta, q, \phi)$$

Then, we similarly define the odds ratio as:

$$\begin{aligned}
\eta_i &= \frac{\mathbb{P}(i|\{i, r^*\})}{1 - \mathbb{P}(i|\{i, r^*\})} \\
&= \frac{1 - F_{\varepsilon_i}(1 \oslash_q RV_i)}{F_{\varepsilon_i}(1 \oslash_q RV_i)}
\end{aligned}$$

Thus, the choice probability follows:

$$\mathbb{P}(i|\mathcal{C}) = \frac{\eta_i}{\sum_{j \in \mathcal{C}} \eta_j}$$

Some algebraic manipulations show that this formulation output the BqPL choice probabilities from Equation 11, if we additionally define $RV_i = V_i \oslash_q V_{r^*}$.

D.3 Heteroscedasticity illustration

The distributional assumption on ε_i implies the following distribution for RU_i:

$$RU_i \sim \text{Trunc. q-LogLogistic}(RV_i, \theta, q, \phi \otimes_q RV_i)$$

Figures 19, 20 and 21 illustrate how the probability density function (PDF) of RU_i changes with RV_i (denoted as μ in the figure). The figure considers two different values of RV_i and three different values of q, corresponding to the BL ($q = 0$), BqPL (with $q = 0.5$), and BW ($q = 1$) cases. The PDF of the non-truncated distribution, which corresponds to the MNL, qPL, and MNW models, is also shown.

We observe that, for $q \neq 0$, the variance of the relative utility distribution (RU_i) increases with its deterministic component (RV_i). This means that when the utility q-ratio with the reference alternative increases, the uncertainty around the value of this q-ratio also increases.

In practice, this heteroscedasticity implies:

- **Within a choice set:** Alternatives with larger deterministic utility differences from the reference alternative exhibit greater variance in relative utility.
- **Across choice situations:** RV_i is defined by the q-ratio between V_i and V_{r*}. For $q \in [0, 1]$, $a/b \leq a \oslash_q b \leq a - b + 1$, which implies that, for a fixed value of $V_i - V_{r*}$, $V_i \oslash_q V_{r*}$ increases with V_i (for instance $1 \oslash_q 2 < 2 \oslash_q 3$). This implies that, when the scale of choices increases (e.g., OD pairs with larger travel times), the q-ratio between utilities and the reference alternative increases, leading to higher variance in relative random utility.

This property, which can also be found for the qPL model, distinguishes the BqPL model from the homoscedastic models such as the MNL and BL.

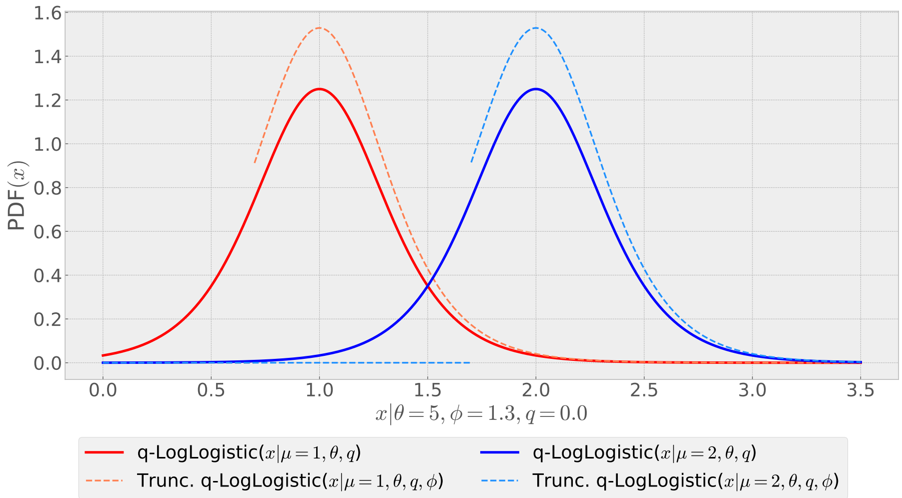


Figure 19: Relative utility distributions for two values of RV_i (1 and 2), $q = 0$ (corresponding to the MNL and BL models)

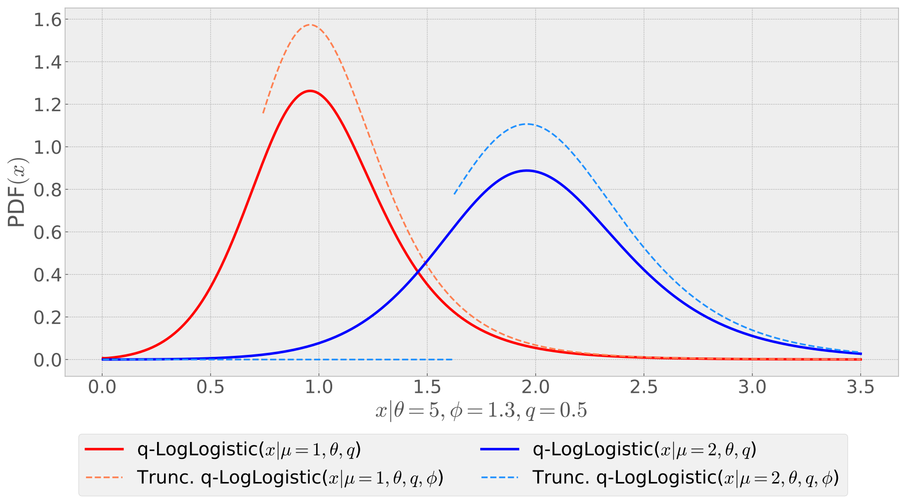


Figure 20: Relative utility distributions for two values of RV_i (1 and 2), $q = 0.5$ (corresponding to a version of the qPL and BqPL models)

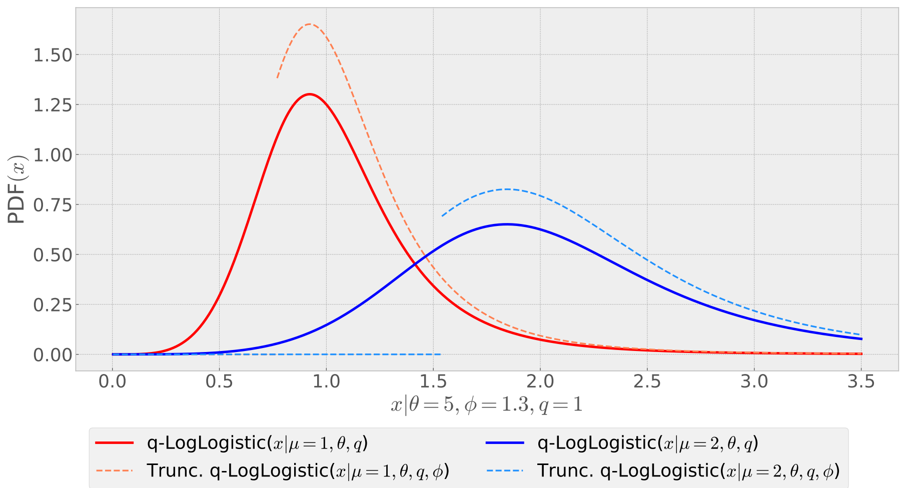


Figure 21: Relative utility distributions for two values of RV_i (1 and 2), $q = 1$ (corresponding to the MNW and BW models)

References

Abraham, I., D. Delling, A. V. Goldberg, and R. F. Werneck (Dec. 2013). "Alternative routes in road networks". en. In: *ACM Journal of Experimental Algorithmics* 18. DOI: 10.1145/2444016.2444019.

Bekhor, S., M. E. Ben-Akiva, and M. S. Ramming (Jan. 2002). "Adaptation of Logit Kernel to Route Choice Situation". en. In: *Transportation Research Record* 1805.1. Publisher: SAGE Publications Inc, pp. 78–85. DOI: 10.3141/1805-10.

Ben-Akiva, M. and M. Bierlaire (1999). "Discrete Choice Methods and their Applications to Short Term Travel Decisions". en. In: *Handbook of Transportation Science*. Ed. by R. W. Hall. Boston, MA: Springer US, pp. 5–33. DOI: 10.1007/978-1-4615-5203-1_2.

Bhat, C. R. (Dec. 1995). "A heteroscedastic extreme value model of intercity travel mode choice". In: *Transportation Research Part B: Methodological* 29.6, pp. 471–483. DOI: 10.1016/0191-2615(95)00015-6.

Borges, E. P. (Sept. 2004). "A possible deformed algebra and calculus inspired in nonextensive thermostatistics". In: *Physica A: Statistical Mechanics and its Applications*. News and Expectations in Thermostatistics 340.1, pp. 95–101. DOI: 10.1016/j.physa.2004.03.082.

Bovy, P. H. L. (Jan. 2009). "On Modelling Route Choice Sets in Transportation Networks: A Synthesis". In: *Transport Reviews* 29.1, pp. 43–68. DOI: 10.1080/01441640802078673.

Bovy, P. H. L. and S. Fiorenzo-Catalano (Jan. 2007). "Stochastic Route Choice Set Generation: Behavioral and Probabilistic Foundations". In: *Transportmetrica* 3.3. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/18128600708685672, pp. 173–189. DOI: 10.1080/18128600708685672.

Cascetta, E., A. Nuzzolo, F. Russo, and A. Vitetta (1996). "A modified logit route choice model overcoming path overlapping problems. Specification and some calibration results for interurban networks". en. In: ISSN: , pp. 697–711.

Cascetta, E. and A. Papola (Aug. 2001). "Random utility models with implicit availability/perception of choice alternatives for the simulation of travel demand". en. In: *Transportation Research Part C: Emerging Technologies* 9.4, pp. 249–263. DOI: 10.1016/S0968-090X(00)00036-X.

Castillo, E., J. M. Menéndez, P. Jiménez, and A. Rivas (May 2008). "Closed form expressions for choice probabilities in the Weibull case". In: *Transportation Research Part B: Methodological* 42.4, pp. 373–380. DOI: 10.1016/j.trb.2007.08.002.

Cazor, L., L. C. Duncan, D. P. Watling, O. A. Nielsen, and T. K. Rasmussen (2024a). "A smooth bounded choice model: formulation and application to three large-scale case studies". In: *Journal of Choice Modelling* (Under review).

Cazor, L., D. P. Watling, L. C. Duncan, O. A. Nielsen, and T. K. Rasmussen (Sept. 2024b). "A novel choice model combining utility maximization and the disjunctive decision rules, application to two case studies". In: *Journal of Choice Modelling* 52, p. 100510. DOI: 10.1016/j.jocm.2024.100510.

Chen, A., S. Pravinvongvuth, X. Xu, S. Ryu, and P. Chootinan (Oct. 2012). "Examining the scaling effect and overlapping problem in logit-based stochastic user equilibrium models". In: *Transportation Research Part A: Policy and Practice* 46.8, pp. 1343–1358. DOI: 10.1016/j.tra.2012.04.003.

Chikaraishi, M. and S. Nakayama (Nov. 2016). "Discrete choice models with q-product random utilities". In: *Transportation Research Part B: Methodological* 93, pp. 576–595. DOI: 10.1016/j.trb.2016.08.013.

Daly, A., S. Hess, and G. de Jong (Feb. 2012). "Calculating errors for measures derived from choice modelling estimates". In: *Transportation Research Part B: Methodological*. Emerging and Innovative Directions in Choice Modeling 46.2, pp. 333–341. DOI: 10.1016/j.trb.2011.10.008.

Danalet, A. and M. Bierlaire (2015). "Importance sampling for activity path choice". In: *15th Swiss Transport Research Conference*.

Duesenberry, J. S. (1949). *Income, saving and the theory of consumer behavior*. Harvard economic studies. Harvard University Press.

Duncan, L. C., D. P. Watling, R. D. Connors, T. K. Rasmussen, and O. A. Nielsen (May 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". en. In: *Transportation Research Part B: Methodological* 135, pp. 1–40. DOI: 10.1016/j.trb.2020.02.006.

— (Dec. 2022). "A bounded path size route choice model excluding unrealistic routes: formulation and estimation from a large-scale GPS study". In: *Transportmetrica A: Transport Science* 18.3. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/23249935.2021.1872730, pp. 435–493. DOI: 10.1080/23249935.2021.1872730.

Efron, B. and R. Tibshirani (1986). "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy". In: *Statistical science*. Publisher: JSTOR, pp. 54–75.

Fischer, S. M. (Nov. 2020). "Locally optimal routes for route choice sets". en. In: *Transportation Research Part B: Methodological* 141, pp. 240–266. DOI: 10.1016/j.trb.2020.09.007.

Fosgerau, M. and M. Bierlaire (June 2009). "Discrete choice models with multiplicative error terms". en. In: *Transportation Research Part B: Methodological* 43.5, pp. 494–505. DOI: 10.1016/j.trb.2008.10.004.

Frejinger, E., M. Bierlaire, and M. Ben-Akiva (Dec. 2009). "Sampling of alternatives for route choice modeling". In: *Transportation Research Part B: Methodological* 43.10, pp. 984–994. DOI: 10.1016/j.trb.2009.03.001.

Frejinger, E. and M. Bierlaire (Jan. 2010). "On Path Generation Algorithms for Route Choice Models". In: *Choice Modelling: The State-of-the-art and The State-of-practice*. Ed. by S. Hess and A. Daly. Emerald Group Publishing Limited, pp. 307–315. DOI: 10.1108/9781849507738-013.

Greene, W. H., D. A. Hensher, and J. Rose (Jan. 2006). "Accounting for heterogeneity in the variance of unobserved effects in mixed logit models". In: *Transportation Research Part B: Methodological* 40.1, pp. 75–92. DOI: 10.1016/j.trb.2005.01.005.

Gu, Y., A. Chen, and S. Kitthamkesorn (Sept. 2022). "Weibit choice models: Properties, mode choice application and graphical illustrations". In: *Journal of Choice Modelling* 44, p. 100373. DOI: 10.1016/j.jocm.2022.100373.

Hess, S. and K. Train (June 2017). "Correlation and scale in mixed logit models". In: *Journal of Choice Modelling* 23, pp. 1–8. DOI: 10.1016/j.jocm.2017.03.001.

Horowitz, J. L. and J. J. Louviere (May 1995). "What is the role of consideration sets in choice modeling?" In: *International Journal of Research in Marketing*. Consideration sets 12.1, pp. 39–54. DOI: 10.1016/0167-8116(95)00004-L.

Kitthamkesorn, S. and A. Chen (2013). "A path-size weibit stochastic user equilibrium model". en. In: *Transportation Research Part B: Methodological* 57.C. Publisher: Elsevier, pp. 378–397.

Koppelman, F. S. and C.-H. Wen (Feb. 2000). "The paired combinatorial logit model: properties, estimation and application". In: *Transportation Research Part B: Methodological* 34.2, pp. 75–89. DOI: 10.1016/S0191-2615(99)00012-0.

Łukawska, M., M. Paulsen, T. K. Rasmussen, A. F. Jensen, and O. A. Nielsen (Oct. 2023). "A joint bicycle route choice model for various cycling frequencies and trip distances based on a large crowdsourced GPS dataset". In: *Transportation Research Part A: Policy and Practice* 176, p. 103834. DOI: 10.1016/j.tra.2023.103834.

Martínez, F., F. Aguila, and R. Hurtubia (Mar. 2009). "The constrained multinomial logit: A semi-compensatory choice model". In: *Transportation Research Part B: Methodological* 43.3, pp. 365–377. DOI: 10.1016/j.trb.2008.06.006.

McFadden, D. (Nov. 1974). "The measurement of urban travel demand". en. In: *Journal of Public Economics* 3.4, pp. 303–328. DOI: 10.1016/0047-2727(74)90003-6.

Munizaga, M. A., B. G. Heydecker, and J. d. D. Ortúzar (Apr. 2000). "Representation of heteroskedasticity in discrete choice models". In: *Transportation Research Part B: Methodological* 34.3, pp. 219–240. DOI: 10.1016/S0191-2615(99)00022-3.

Nakayama, S. (June 2013). "q-Generalized Logit Route Choice an Network Equilibrium Model". In: *Procedia - Social and Behavioral Sciences*. 20th International Symposium on Transportation and Traffic Theory (ISTTT 2013) 80, pp. 753–763. DOI: 10.1016/j.sbspro.2013.05.040.

Nielsen, O. A. (2004). "A Large Scale Stochastic Multi-Class Schedule-Based Transit Model with Random Coefficients". en. In: *Schedule-Based Dynamic Transit Modeling: theory and applications*. Ed. by N. H. M. Wilson and A. Nuzzolo. Operations Research/Computer Science

Interfaces Series. Boston, MA: Springer US, pp. 53–77. DOI: 10.1007/978-1-4757-6467-3_4.

Petrin, A. and K. Train (Jan. 2003). *Omitted Product Attributes in Discrete Choice Models*. Working Paper. DOI: 10.3386/w9452.

Pougala, J., T. Hillel, and M. Bierlaire (Oct. 2023). "OASIS: Optimisation-based Activity Scheduling with Integrated Simultaneous choice dimensions". In: *Transportation Research Part C: Emerging Technologies* 155, p. 104291. DOI: 10.1016/j.trc.2023.104291.

Prato, C. G. (Jan. 2009). "Route choice modeling: past, present and future research directions". In: *Journal of Choice Modelling* 2.1, pp. 65–100. DOI: 10.1016/S1755-5345(13)70005-8.

Prato, C. G. and S. Bekhor (Jan. 2007). "Modeling Route Choice Behavior: How Relevant Is the Composition of Choice Set?" EN. In: *Transportation Research Record* 2003.1. Publisher: SAGE Publications Inc, pp. 64–73. DOI: 10.3141/2003-09.

Ramming, M. S. (2002). "Network knowledge and route choice". eng. Accepted: 2009-11-06T16:36:54Z. Thesis. Massachusetts Institute of Technology.

Rasmussen, T. K., L. C. Duncan, D. P. Watling, and O. A. Nielsen (Dec. 2024). "Local detouredness: A new phenomenon for modelling route choice and traffic assignment". In: *Transportation Research Part B: Methodological* 190, p. 103052. DOI: 10.1016/j.trb.2024.103052.

Simon, H. A. (1955). "A Behavioral Model of Rational Choice". In: *The Quarterly Journal of Economics* 69.1. Publisher: Oxford University Press, pp. 99–118. DOI: 10.2307/1884852.

Stadt, H. van de, A. Kapteyn, and S. van de Geer (1985). "The Relativity of Utility: Evidence from Panel Data". In: *The Review of Economics and Statistics* 67.2. Publisher: The MIT Press, pp. 179–187. DOI: 10.2307/1924716.

Tan, H., X. Xu, and A. Chen (May 2024). "On endogenously distinguishing inactive paths in stochastic user equilibrium: A convex programming approach with a truncated path choice model". In: *Transportation Research Part B: Methodological* 183, p. 102940. DOI: 10.1016/j.trb.2024.102940.

Tsallis, C. (1994). "What are the numbers that experiments provide?" In: *Quimica Nova* 17.6, pp. 468–471.

Vovsha, P. (Jan. 1997). "Application of Cross-Nested Logit Model to Mode Choice in Tel Aviv, Israel, Metropolitan Area". en. In: *Transportation Research Record* 1607.1. Publisher: SAGE Publications Inc, pp. 6–15. DOI: 10.3141/1607-02.

Watling, D. P., T. K. Rasmussen, C. G. Prato, and O. A. Nielsen (Aug. 2018). "Stochastic user equilibrium with a bounded choice model". en. In: *Transportation Research Part B: Methodological* 114, pp. 254–280. DOI: 10.1016/j.trb.2018.05.004.

Williams, H. C. W. L. and J. D. Ortuzar (June 1982). "Behavioural theories of dispersion and the mis-specification of travel demand models". In: *Transportation Research Part B: Methodological* 16.3, pp. 167–219. DOI: 10.1016/0191-2615(82)90024-8.

Xu, X., A. Chen, S. Kitthamkesorn, H. Yang, and H. K. Lo (Jan. 2015). "Modeling Absolute and Relative Cost Differences in Stochastic User Equilibrium Problem". In: *Transportation Research Procedia*. 21st International Symposium on Transportation and Traffic Theory Kobe, Japan, 5-7 August, 2015 7, pp. 75–95. DOI: 10.1016/j.trpro.2015.06.005.

Yai, T., S. Iwakura, and S. Morichi (June 1997). "Multinomial probit with structured covariance for route choice behavior". In: *Transportation Research Part B: Methodological* 31.3, pp. 195–207. DOI: 10.1016/S0191-2615(96)00025-2.

Yao, J. and A. Chen (Nov. 2014). "An analysis of logit and weibit route choices in stochastic assignment paradox". In: *Transportation Research Part B: Methodological* 69, pp. 31–49. DOI: 10.1016/j.trb.2014.07.006.

Zhang, J., H. Timmermans, A. Borgers, and D. Wang (2004). "Modeling traveler choice behavior using the concepts of relative utility and relative interest". en. In: *Transportation Research Part B: Methodological* 38.3. Publisher: Elsevier, pp. 215–234.