

Article

A Maximum Entropy Fixed-Point Route Choice Model for Route Correlation

Louis de Grange ¹, Sebastián Raveau ² and Felipe González ^{1,*}

¹ Escuela de Ingeniería Civil Industrial, Universidad Diego Portales, Santiago 8370191, Chile; E-Mail: louis.degrange@udp.cl

² Department of Transport Engineering and Logistics, Pontificia Universidad Católica de Chile, Santiago 7820436, Chile; E-Mail: sraveau@puc.cl

* Author to whom correspondence should be addressed; E-Mail: felipe.gonzalezr@udp.cl; Tel.: +56-2-26768119; Fax: +56-2-26768130.

Received: 19 March 2014; in revised form: 6 June 2014 / Accepted: 23 June 2014/

Published: 30 June 2014

Abstract: In this paper we present a stochastic route choice model for transit networks that explicitly addresses route correlation due to overlapping alternatives. The model is based on a multi-objective mathematical programming problem, the optimality conditions of which generate an extension to the Multinomial Logit models. The proposed model considers a fixed point problem for treating correlations between routes, which can be solved iteratively. We estimated the new model on the Santiago (Chile) Metro network and compared the results with other route choice models that can be found in the literature. The new model has better explanatory and predictive power than many other alternative models, correctly capturing the correlation factor. Our methodology can be extended to private transport networks.

Keywords: maximum entropy; route choice; fixed point; route overlapping; route correlation

1. Introduction

The present study formulates a new route choice model for public transport networks that features significant innovations compared to existing models. The main enhancement in the proposed model is the ability to simultaneously and explicitly integrate the traveler's lack of information (randomness or uncertainty) and the correlation between route alternatives (due to overlapping). The proposed model is

based on a multi-objective mathematical programming problem and its respective scalarized single-objective problem. The multi-objective problem considers the exogenous cost functions of the network, entropy of the route choice, and the covariance matrix for the route flows. An extension to networks with congestion (e.g., private transport with endogenous costs) is also proposed.

A traditional approach to defining the route choice process, and the subsequent traffic equilibrium, is to assume a deterministic behavior. This deterministic equilibrium usually states optimality conditions, such as minimizing transport costs or satisfying Wardrop's first principle of traffic equilibrium [1]. These models assume that travelers have perfect information and seek to unilaterally minimize their travel costs [2–5]. Typically, a mathematical programming model is formulated and solved by an iterative algorithm. If applied with care and understanding, a deterministic user-equilibrium model provides a simple but effective method of traffic assignment [6–12].

Alternatively, stochastic/probabilistic route choice models differ from deterministic formulations in that they incorporate the uncertainty, randomness and/or the heterogeneity of travelers and alternative routes, and passenger's imperfect knowledge. This is the approach followed in this study. Reviews of this class of models are found in Daganzo and Sheffi [13], Hazelton [14], Ramming [15], and Prashker and Bekhor [16]. Among these models, there are some that explicitly consider correlations between alternative routes, such as Cascetta *et al.* [17], Ben-Akiva and Bierlaire [18], Bekhor and Prashker [19,20], and Bovy *et al.* [21]. A complete review of these studies can be found in Prashker and Bekhor [16] and Prato [22]. Several of the models presented by the authors were considered for performing a comparison with our new model.

In the remainder of this paper, Section 2 contains a brief review of the literature that provides context for understanding the proposed model; Section 3 provides an analytic derivation of the new formulation; Section 4 applies the model to a medium-sized network (the Santiago Metro), comparing the results with existing models and proposing a version for private networks with congestion; and lastly, Section 5 summarizes the results and gives the main conclusions.

2. Literature Review

The formulation of the route choice or traffic assignment stage in transportation modeling has long followed an approach in which users minimize their generalized trip cost on the assumption of perfect knowledge of the transport network. Under this approach, travelers are considered to be homogeneous and each one is fully informed of the cost of each link on the network at any level of flow [1,2]. These assumptions are both rather strong even for a small network, and the results obtained are often not satisfactory. However, due to their simplicity and availability, many transport planners continue to apply such models, especially in large networks.

Because the perfect information assumption is usually not correct, there is a clear need for models that represent users who have incomplete or imperfect information on the transport system in regard to existing routes and their levels of congestion. Various route choice models in the specialized literature are based on system attributes perceived by travelers and their socioeconomic and demographic characteristics [13,15,16,23,24]. In these models, users behave in accordance with the costs they perceive. The socioeconomic and demographic variable data are usually obtained through user surveys or from network data records and are easily justified as an integral part of individuals' rational

decision-making processes. However, because the modeler lacks information on these processes, the modeled choices necessarily embody a degree of variability.

Regarding knowledge of routes, it is widely accepted that individual users do not know (or do not perceive/consider) all of the route possibilities between a given origin-destination pair [17,25–28]. This is a reality that should be incorporated into route choice models. Cascetta *et al.* [27] propose a Logit-type model of route perception and choice similar to those based on random utility theory. Ben-Akiva *et al.* [25] model interurban route choice as a two-stage process in which a set of routes is defined in the first stage, and the route choice is performed in the second stage.

Correlations between alternative routes can (to a certain extent) be indirectly addressed, according to Bovy and Hoogendoorn-Lanser [29], by hierarchical nested Logit and multi-nested GEV models in cases where it arises from overlapping segments and/or nodes. Prato and Bekhor [30] and Bekhor *et al.* [31] study the way in which individuals construct their set of route alternatives and the implications of route similarity for user behavior. Bliemer and Bovy [32] study the interdependence of these two aspects.

2.1. Multinomial Logit for Route Choice

The probability of choosing route p to travel between O-D pair w (P_w^p) can be estimated by discrete choice models in which the cost of a given route as perceived by a user (C_w^p) is assumed to be given by Equation (1):

$$C_w^p = c_w^p + \varepsilon_w^p \tag{1}$$

The cost is thus perceived as the sum of a deterministic component (c_w^p) and a random error component (ε_w^p). The latter can be interpreted in various ways, one of which is that the error reflects the user’s inaccurate perception of route cost due to, among other things, the scarcity of information. In this case, the first term (c_w^p) represents the real mean cost of the route. The expression for the choice probability (P_w^p) is a function of the assumed distribution of the error terms and whether or not they are independent. If we assume the errors are i.i.d. Gumbel-distributed with a scale parameter $\theta > 0$, we have a multinomial logit (MNL) choice model in which P_w^p is given by Equation (2) (see [5,23,33]):

$$P_w^p = \frac{\exp(-\theta c_w^p)}{\sum_{r \in p^w} \exp(-\theta c_w^r)} \tag{2}$$

where p^w is the set of routes uniting the O-D pair w .

An equivalent optimization problem that generates the stochastic assignment model (2) is (see [5,34,35]):

$$\begin{aligned} \min_{\{h_w^p\}} Z &= \sum_w \sum_{p \in p^w} c_w^p h_w^p + \frac{1}{\theta} \sum_w \sum_{p \in p^w} h_w^p (\ln h_w^p - 1) \\ \text{s.t.} & \\ \sum_{p \in p^w} h_w^p &= T_w \quad \forall w \end{aligned} \tag{3}$$

where T_w is the total number of trips (exogenous demand) between the O-D pair w and h_w^p is the flow along route p ($P_w^p = h_w^p / T_w$).

In problem (3), the set of routes p^w between each pair w must be predefined and maintained invariant during the assignment process. The optimality conditions lead to the assignment criterion given by Equation (2) in which users divide up among the alternative routes based on a Logit model.

It is worth noting that the objective function in Equation (3) considers two terms: the first term $\sum_w \sum_{p \in p^w} c_w^p h_w^p$ represents the total cost in the network, while the second term $\sum_w \sum_{p \in p^w} h_w^p (\ln h_w^p - 1)$ represents the negative value of the entropy. The latter is weighted by $1/\theta$ to scale both terms in the objective function. In this way, problem (3) can be interpreted as a bi-objective problem that simultaneously minimizes the total cost and maximizes the total entropy [34].

The principal limitation of model (3) is that it does not incorporate a structure for capturing correlations between routes. Possible direct extensions are either extremely simplistic or difficult to apply correctly to real-world-scale networks (for example, using a hierarchical Logit model) given that they cannot properly capture the various types of correlations between routes that share links with each other in different ways. In urban networks, the routes linking a given O-D pair will typically have many overlaps due to common links, with the result that the independent error assumption implicit in Logit-type models such as Equation (2) is unrealistic.

Different extensions of the multinomial logit model have been proposed to explicitly capture correlations between alternative routes, some of which are presented below. These models are used as a comparative basis for our proposed model.

2.2. C-Logit Model

In Cascetta *et al.* [17] and Cascetta *et al.* [27], the authors propose a joint implicit availability/perception (IAP) and route choice (C-Logit) model that also explicitly addresses the correlation issue (*i.e.*, the lack of route independence due to common links) and is analytically tractable even for large-scale networks.

The basic idea behind the model is to address route interdependence via a cost attribute called the “similarity factor”, which is added to the cost of the route in a conventional Logit model, instead of addressing it in terms of error non-independence as in the Probit model. The probability of choosing route p is then given by Equation (4):

$$P_w^p = \frac{\exp(-\theta c_w^p + CF_w^p)}{\sum_{r \in p^w} \exp(-\theta c_w^r + CF_w^r)} \tag{4}$$

where CF_w^p is the “similarity factor” for route p joining pair w and is constructed as follows:

$$CF_w^p = \beta \cdot \ln \sum_{a \in p} \left(\frac{l_a}{L_p} \cdot \sum_{r \in p^w} \delta_{ar} \right) \tag{5}$$

where β is a parameter to be calibrated that must be negative, l_a is the length of link a , L_p is the length of route p , and δ_{ar} is equal to 1 if link a belongs to some route r joining w but is 0 otherwise. Other specifications for CF_w^p may be found in Prato [22].

2.3. Path-Size Logit Model

Ben-Akiva and Ramming [36] develop the path-size logit (PSL) model, which also aims to correct for routes that have overlaps and are therefore correlated. It attempts to incorporate behavioral theory in Cascetta’s C-logit model. In this case, P_w^p is given by Equation (6):

$$P_w^p = \frac{\exp(-\theta c_w^p + \beta \cdot \ln PS_w^p)}{\sum_{r \in p^w} \exp(-\theta c_w^r + \beta \cdot \ln PS_w^r)} \tag{6}$$

where PS_w^p is the correction for the route size.

The correction principle applied in this model is as follows: a route with no links overlapping another route needs no correction and is therefore assigned a size of 1. At the other extreme, if there are J duplicate routes (*i.e.*, total overlap), each one has a size of $1/J$. Lastly, the length of routes with partial overlap is based on the sizes of the links, which are appropriately weighted on some criterion, such as the link’s contribution to the total length of the route. Thus, PS_w^p can take the following form:

$$PS_w^p = \sum_{a \in p} \left(\frac{l_a}{L_p} \cdot \frac{1}{\sum_{r \in p^w} \delta_{ar}} \right) \tag{7}$$

where the variables are the same as those employed in Equation (5) to define the similarity factor. Further specifications for PS_w^p may be determined in Bovy *et al.* [21].

2.4. Paired Combinatorial Logit Model

The paired combinatorial logit model (PCL) belongs to a family of models that derives from the generalized extreme value (GEV) model [33]. The general PCL model can be derived from the following generation function:

$$G(y_1, y_2, \dots, y_n) = \sum_{k=1}^{n-1} \sum_{j=k+1}^n (1 - \sigma_{kj}) (y_k^{1/1-\sigma_{kj}} + y_j^{1/1-\sigma_{kj}})^{1-\sigma_{kj}} \tag{8}$$

where:

- y_k characterizes each alternative (in our case, a route between an O-D pair),
- σ_{kj} is a similarity index between alternatives k , and
- n is the number of alternatives.

The model was adapted for route choice by Bekhor and Prashker [19], and, in our case, we use the expression used by Chen *et al.* [37]: $y_k = y_w^p = e^{-\theta c_w^p}$. Next, the probability of choosing route p when traveling between O-D pair w is Equation (9):

$$P_w^p = \frac{\sum_{q \in p^w, q \neq p} e^{\frac{-\theta c_w^p}{1-\sigma_w^{pq}}} (1 - \sigma_w^{pq}) \left(e^{\frac{-\theta c_w^p}{1-\sigma_w^{pq}}} + e^{\frac{-\theta c_w^q}{1-\sigma_w^{pq}}} \right)^{-\sigma_w^{pq}}}{\sum_{r=1}^{p^w-1} \sum_{m=r+1}^{p^w} (1 - \sigma_w^{rm}) \left(e^{\frac{-\theta c_w^r}{1-\sigma_w^{rm}}} + e^{\frac{-\theta c_w^m}{1-\sigma_w^{rm}}} \right)^{1-\sigma_w^{rm}}} \tag{9}$$

In this case, for a choice set of p^w alternative routes, there are $p^w(p^w - 1)/2$ pairs of alternatives. If σ_w^{pq} equals 0 for every pair (p,q) in w routes, then the PCL model reduces to a MNL model such as Equation (2). We use the similarity index proposed by Chen *et al.* [37]:

$$\sigma_w^{pq} = \frac{L_w^{pq}}{\sqrt{L_w^p \cdot L_w^q}} \tag{10}$$

where L_w^{pq} is the length of overlap between routes p and q and L_w^p and L_w^q are the respective lengths of routes p and q , respectively. It is possible (and convenient) to include an additional degree of freedom in this model by including a parameter β in Equation (10) as shown in Equation (11). This new parameter must be estimated together with the rest of the model’s parameters:

$$\sigma_w^{pq} = \beta \cdot \frac{L_w^{pq}}{\sqrt{L_w^p \cdot L_w^q}} \tag{11}$$

2.5. Cross-Nested Logit Model

The cross-nested model (CNL) also belongs to the GEV family of models and was adapted for the route choice context by Prashker and Bekhor [38] and Vovsha and Bekhor [39]. In that adaptation, the model considers a two-level hierarchical structure. The upper level includes all of the links in the network, and the lower level includes all of the routes that belong to p^w . Next, every route is assigned to the nests that represent the links that belong to it. In this case, the probability of choosing route p is given by Equation (12):

$$P_w^p = \frac{e^{-\theta c_w^p} \sum_{m \in M_w^p} \alpha_w^{mp} \left(\sum_{q \in p^w} \alpha_w^{mq} e^{-\theta c_w^q} \right)^{\mu-1}}{\sum_{m \in M_w^p} \left(\sum_{q \in p^w} \alpha_w^{mq} e^{-\theta c_w^q} \right)^\mu} \tag{12}$$

where:

- m characterizes the links and therefore the nests,
- M_w^p is the set of links that belongs to route p in the O-D pair w ,
- α_w^{mp} are parameters that represent the degree of inclusion of alternative p in nest m , and
- μ is the nesting coefficient. If $\mu = 1$, the model reduces to a MNL model.

The CNL model is adapted to a route choice context, defining the parameters α_w^{mp} depending on the topology of the network. Prashker and Bekhor [38] proposed the following expression:

$$\alpha_w^{mp} = \left(\frac{L_w^m}{L_w^p} \right)^\gamma \delta_w^{mp} \tag{13}$$

where L_w^m is the length of link m , L_w^p is the length of route p , δ_w^{mp} equals 1 if link m belongs to route p , and γ is a parameter that must be calibrated, which reflects the perception of the travelers regarding the similarity of the alternative routes. For the estimation of the CNL model, we use $\gamma = 1$ following Bekhor *et al.* [31].

3. Route Choice Model with Correlated Routes

In this section, we develop a new route choice model that simultaneously incorporates: (i) users with imperfect knowledge of the network; (ii) correlations between alternatives (in this case, routes); and (iii) the effect of demand or flow levels on network costs (congestion). As noted earlier, the proposed formulation is entropy-based with quadratic constraints.

3.1. Mathematical Formulation of Model

The proposed stochastic equilibrium assignment model that captures route correlation is based on the following multi-objective optimization problem (other applications of multi-objective models in transportation can be found in [34,35,40])

$$\begin{aligned}
 \min F_1 &= \sum_w \sum_{p \in p^w} c_w^p h_w^p \\
 \min F_2 &= \sum_w \sum_{p \in p^w} h_w^p (\ln h_w^p - 1) \\
 \min F_3 &= \sum_w \sum_{p \in p^w} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} (h_w^p - t_w)(h_w^q - t_w)] \\
 \text{s.t.} &: \sum_{p \in p^w} h_w^p = T_w, \quad \forall w \quad (\gamma_w)
 \end{aligned}
 \tag{14}$$

Objective F_1 relates to the total system cost. Objective F_2 attempts to maximize the entropy to determine the most likely routes; in probabilistic terms, it finds the most feasible route combinations for travelers in equilibrium. Combining F_1 and F_2 with the flow conservation constraints gives the stochastic assignment model expressed in Equation (3).

Objective F_3 , the novel element in the proposed model, explicitly incorporates correlations of flows between different routes, whether or not they join the same O-D pair w . The objective is constructed as a weighted sum of the divergences of the flows on the individual defined routes joining w from the average flow on those routes, where $t_w = \frac{T_w}{N_w}$ is the average flow, N_w is the number of defined routes (*i.e.*, the cardinality of p^w) and T_w is the (fixed) total number of trips between w . The parameters η_w^{pq} are exogenous and determine the degree of correlation (0 to 1) between the routes p and q of w . The values of the parameters can be defined in a number of ways (see [15,17,41]).

As with F_2 , objective F_3 is an information criterion [42]. If, for example, the flows on all traveled routes were uniform (that is, if $h_w^p = t_w, \forall p$), the value of F_3 would be 0 and thus contain no information. F_2 also takes its lowest possible value if $h_w^p = t_w, \forall p$.

An alternative optimization problem [34,43] for Equation (14) that generates the proposed stochastic equilibrium model is the following:

$$\begin{aligned}
 \min F &= \sum_w \sum_{p \in p^w} c_w^p h_w^p + \frac{1}{\theta} \sum_w \sum_{p \in p^w} h_w^p (\ln h_w^p - 1) + \frac{1}{\rho} \sum_w \sum_{p \in p^w} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} (h_w^p - t_w)(h_w^q - t_w)] \\
 \text{s.t.} &: \sum_{p \in p^w} h_w^p = T_w, \quad \forall w \quad (\gamma_w)
 \end{aligned}
 \tag{15}$$

The terms $1/\theta$ and $1/\rho$ are the respective relative weights of the two information criteria F_2 and F_3 with respect to the reference objective F_1 . θ and ρ are parameters to be estimated.

The first-order conditions for Equation (15) are:

$$c_w^p + \frac{1}{\theta} \ln h_w^p + \frac{1}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} (h_w^q - t_w)] + \gamma_w = 0 \tag{16}$$

where c_w^p could be, for a public transportation application, replaced by $L_w^p = \sum_k \beta_k X_{w,k}^p$. From Equation (16), we obtain:

$$h_w^p = \exp \left(-\theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} (h_w^q - t_w)] - \theta \gamma_w \right) \tag{17}$$

$$\sum_{p \in p^w} h_w^p = T_w = \exp(-\theta \gamma_w) \cdot \sum_{p \in p^w} \exp \left(-\theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} (h_w^q - t_w)] \right) \tag{18}$$

Dividing Equations (17) and (18) we have:

$$h_w^p = T_w \frac{\exp \left(-\theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} (h_w^q - t_w)] \right)}{\sum_{r \in p^w} \exp \left(-\theta c_w^r - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq r}} [\eta_w^{rq} (h_w^q - t_w)] \right)} \tag{19}$$

Because η_w^{pq} is an exogenous model parameter (defined by the modeler) and t_w is assumed to be constant (for calibration and modeling purposes), the term $\frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} t_w]$ is also constant. Letting

$\frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} t_w] = \alpha_w^p$, which is an intercept or specific constant, Equation (19) can then be rewritten as:

$$h_w^p = T_w \frac{\exp \left(\alpha_w^p - \theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} h_w^q] \right)}{\sum_{r \in p^w} \exp \left(\alpha_w^r - \theta c_w^r - \frac{\theta}{\rho} \sum_{\substack{q \in p^w \\ q \neq r}} [\eta_w^{rq} h_w^q] \right)} \tag{20}$$

This non-linear expression is similar in structure to the model specified by Ben-Akiva and Ramming [36], given here as Equation (6). The main difference is that the right-hand side of Equation (20) includes the endogenous variable h_w^q and is therefore a fixed-point function [44], whereas in Equation (6), the right-hand side contains only the model's exogenous variables.

An alternative can be specified for F_1 that can model service levels in public transportation networks. An example of such a replacement is $\tilde{F}_1 = \sum_w \sum_{p \in P^w} L_w^p h_w^p$, where $L_w^p = \sum_k \beta_k X_{w,k}^p$ is a generalized cost given by the weighted sum of attributes or explanatory variables $X_{w,k}^p$ that could represent, for example, the trip time, wait time, or the cost of route p between pair w .

We can define $V_w^p = \alpha_w^p - \theta c_w^p - \frac{\theta}{\rho} \sum_{\substack{q \in P^w \\ q \neq p}} [\eta_w^{pq} h_w^q]$ as the “utility level” of the proposed Logit-based model. By the route choice probability function in Equation (20), the marginal utility of the attribute $X_{w,k}^p$ can be written as:

$$\frac{\partial V_w^p}{\partial X_{w,k}^p} = \beta_k + \frac{\partial V_w^p}{\partial X_{w,k}^p} \cdot \frac{1}{T_w} \cdot \frac{\theta}{\rho} \sum_{\substack{q \in P^w \\ q \neq p}} \eta_w^{pq} \cdot h_w^p \cdot h_w^q \tag{21}$$

Clearing, we obtain:

$$\frac{\partial V_w^p}{\partial X_{w,k}^p} = \frac{\beta_k}{1 - \frac{1}{T_w} \cdot \frac{\theta}{\rho} \sum_{\substack{q \in P^w \\ q \neq p}} \eta_w^{pq} \cdot h_w^p \cdot h_w^q} \tag{22}$$

Because the denominator of Equation (21) is independent of attribute k , the marginal rates of substitution are generic and have the same functional form as traditional discrete choice models:

$$\frac{\partial V_r / \partial X_{w,k_1}^p}{\partial V_r / \partial X_{w,k_2}^p} = \frac{\beta_{k_1}}{\beta_{k_2}} \tag{23}$$

Thus, in the proposed fixed-point model with spatial correlation, the marginal rates of substitution between attributes can be obtained without any additional complexity.

To implement model (20), the parameters $(\alpha_w^p, \theta, \rho^*)$, where $\rho^* = \frac{\theta}{\rho}$ must first be estimated and the fixed-point Equation (20) and then solved given that the variable h_w^p appears on both sides of the equation.

However, because the endogenous variable (h_w^p) appears on the right-hand side of the model, the parameters $(\alpha_w^p, \theta, \rho^*)$ cannot be estimated using the maximum likelihood because the presence of the endogenous variable violates the assumption of independent marginal probability functions necessary for defining the likelihood function. To bypass this problem, we resort to the use of an instrumental variable to replace h_w^p as explained below.

3.2. Estimation of Model Parameters

An instrument or instrumental variable [45] is an exogenous variable that is highly correlated with an explanatory variable exhibiting endogeneity and can therefore be used as a replacement for the latter without a loss of asymptotic properties in the estimated parameters. In our case, a suitable instrument h_w^{p0} to replace h_w^p is:

$$h_w^{p0} = T_w \frac{\exp(-\lambda c_w^p)}{\sum_{r \in p^w} \exp(-\lambda c_w^r)} \tag{24}$$

This formula is a classic MNL model and can be easily estimated. Once this is done, the value of h_w^{p0} is substituted into the right-hand side of Equation (20), as follows:

$$h_w^p = T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} h_w^{q0}]\right)}{\sum_{r \in p^w} \exp\left(\alpha_w^r - \theta c_w^r - \rho^* \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{rq} h_w^{q0}]\right)} \tag{25}$$

Unlike Equation (20), the parameters $(\alpha_w^p, \theta, \rho^*)$ in Equation (25) can be directly estimated by the maximum likelihood [46–48] because h_w^{p0} does not exhibit endogeneity. Using these estimates, which we now denote $(\alpha_w^{p0}, \theta^0, \rho^{0*})$, we can estimate h_w^{p1} by Equation (26):

$$h_w^{p1} = T_w \frac{\exp\left(\alpha_w^{p0} - \theta^0 c_w^p - \rho^{0*} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} h_w^{q0}]\right)}{\sum_{r \in p^w} \exp\left(\alpha_w^{r0} - \theta^0 c_w^r - \rho^{0*} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{rq} h_w^{q0}]\right)} \tag{26}$$

The original model (26) is thus estimated iteratively from the following recursive relation (which represents the equilibrium of the fixed-point function):

$$h_w^{p(n+1)} = T_w \frac{\exp\left(\alpha_w^{p(n)} - \theta^{(n)} c_w^p - \rho^{(n)*} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{pq} h_w^{q(n)}]\right)}{\sum_{r \in p^w} \exp\left(\alpha_w^{r(n)} - \theta^{(n)} c_w^r - \rho^{(n)*} \sum_{\substack{q \in p^w \\ q \neq p}} [\eta_w^{rq} h_w^{q(n)}]\right)} \tag{27}$$

The iterative estimation process concludes when convergence is obtained; that is, $(\alpha_w^{p(n)}, \theta^{(n)}, \rho^{(n)*}) \approx (\alpha_w^{p(n-1)}, \theta^{(n-1)}, \rho^{(n-1)*})$. If the estimator of parameter $\rho^{(n)*}$ in Equation (27) is significantly different from 0, the null hypothesis of no correlation between route p and any of the other routes joining O-D pair w is rejected.

In an analogous fashion to the derivation of Equation (21), upon iteratively solving model (20), the marginal utility of attribute $X_{w,k}^p$ is then:

$$\frac{\partial V_w^{p(n+1)}}{\partial X_{w,k}^p} = \beta_k^{(n)} + \frac{\partial V_w^{p(n)}}{\partial X_{w,k}^p} \cdot \frac{1}{T_w} \cdot \rho^{(n)*} \sum_{\substack{q \in p^w \\ q \neq p}} \eta_w^{pq} \cdot h_w^{p(n)} \cdot h_w^{q(n)} \tag{28}$$

This expression generates the marginal utilities recursively in successive iterations. The model converges to the equilibrium of the fixed-point function as follows:

$$\frac{\partial V_w^{p(n+1)}}{\partial X_{w,k}^p} \approx \frac{\partial V_w^{p(n)}}{\partial X_{w,k}^p} \rightarrow \frac{\partial V_r / \partial X_{w,k_1}^{p(n+1)}}{\partial V_r / \partial X_{w,k_2}^{p(n+1)}} = \frac{\beta_{k_1}^{(n)}}{\beta_{k_2}^{(n)}} \tag{29}$$

The marginal rates of substitution thus continue to be generic and equal in their functional form to those of the traditional discrete choice models.

However, because the endogenous variable (h_w^p) appears on the right-hand side of the model, the parameters ($\alpha_w^p, \theta, \rho^*$) cannot be estimated using the maximum likelihood because the presence of the endogenous variable violates the assumption of independent marginal probability functions necessary for defining the likelihood function. To bypass this problem, we resort to the use of an instrumental variable to replace h_w^p as explained below.

3.3. Existence and Uniqueness of the Fixed-Point Problem Solution

To demonstrate the existence of a solution to the fixed-point problem expressed by Equation (27), we apply the Brouwer fixed-point theorem [49]. This theorem may be stated as follows:

Let H be a non-empty compact convex set of a finite-dimensional Euclidean space k , and let $f : H \rightarrow H$ be a continuous function. Then, f has a fixed point, *i.e.*, $\exists \mathbf{h} \in H : \mathbf{h} = f(\mathbf{h})$.

Below, it is shown that the system of equations describing the equilibrium in model (27) satisfies the theorem’s existence and completeness hypotheses.

Let $h = (h_w^p)_{p \in P^w, w \in W}$, $h \in H$, with $H = \otimes_{w \in W} T_w \Delta^{N_w}$ where $\Delta^n = \{x \in X_+^n : \mathbf{1}^T x = 1\}$ is the probability simplex. H is a nonempty convex compact set because it is a finite cross product of nonempty convex compact sets.

The function $f: H_w \rightarrow H_w$ is defined as:

$$f(\mathbf{h}_w) = \left(T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in P^w} [\eta_w^{pq} h_w^q]\right)}{\sum_{r \in P^w} \exp\left(\alpha_w^r - \theta c_w^r - \rho^* \sum_{q \in P^w} [\eta_w^{rq} h_w^q]\right)} \right)_p \tag{30}$$

This expression is a construction of two continuous functions and is therefore itself continuous. The parameters ($\alpha_w^p, \theta, \rho^*$) are estimated using the maximum likelihood given an h_w and thus are continuous functions with respect to the latter, thereby satisfying the Brouwer’s theorem hypotheses and proving the existence of at least one fixed point.

The function f is continuous because it is a composition of continuous functions. In this case, the parameters (α_w^p, θ, ρ) are continuous functions of h , which is given by the continuity and strict concavity (α_w^p, θ, ρ) of the log-likelihood function to be used in the multinomial logit model. The estimation of our model, which corresponds to the point where succession described in the previous section converges, belongs to the set H and will be a fixed point of the function f .

Using Brouwer’s Theorem, we can establish the existence of at least one fixed point, *i.e.*, $\exists h \in H : h = f(h)$, which is exactly the point of equilibrium that we want:

In formal terms:

$$\exists h^* \in H : f(h^*) = \left(T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} [\eta_w^{pq} h_w^{q*}]\right)}{\sum_{r \in p^w} \exp\left(\alpha_w^r - \theta c_w^r - \rho^* \sum_{q \in p^w} [\eta_w^{rq} h_w^{q*}]\right)} \right) \tag{31}$$

We now prove uniqueness for the case of $\rho^* \geq 0$. Assume that there are 2 different equilibrium points $h^*, g^* \in H$. It is therefore the case that:

$$\left(h^* \neq g^* \wedge \sum_p h_w^{p*} = \sum_p g_w^{p*} = T_w \right) \tag{32}$$

$$\Rightarrow \exists p^+, p^- \in \mathbb{N} : (p^+ \neq p^-; 1 \leq p^+, p^- \leq n; h_w^{p^{*+}} < g_w^{p^{*+}}; h_w^{p^{*-}} > g_w^{p^{*-}}) \tag{33}$$

$$\frac{h_w^{p^{*+}}}{g_w^{p^{*+}}} = \frac{\sum_{p \in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} [\eta_w^{pq} g_w^{q*}]\right)}{\sum_{p \in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} [\eta_w^{pq} h_w^{q*}]\right)} \times \exp\left\{ \rho^* \left(\sum_{q \in p^w} [\eta_w^{pq} h_w^{q*}] - \sum_{q \in p^w} [\eta_w^{pq} g_w^{q*}] \right) \right\} \tag{34}$$

Without a loss of generality, we can assume that:

$$\exists h^* \in H : f(h^*) = \left(T_w \frac{\exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} [\eta_w^{pq} h_w^{q*}]\right)}{\sum_{r \in p^w} \exp\left(\alpha_w^r - \theta c_w^r - \rho^* \sum_{q \in p^w} [\eta_w^{rq} h_w^{q*}]\right)} \right) \tag{35}$$

For the case of p^+ , we have:

$$h_w^{p^{*+}} > g_w^{p^{*+}} \Rightarrow \frac{h_w^{p^{*+}}}{g_w^{p^{*+}}} > 1 \tag{36}$$

$$\Rightarrow \underbrace{\frac{\sum_{p \in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} [\eta_w^{pq} g_w^{q*}]\right)}{\sum_{p \in p^w} \exp\left(\alpha_w^p - \theta c_w^p - \rho^* \sum_{q \in p^w} [\eta_w^{pq} h_w^{q*}]\right)}}_{\leq 1} \times e^{\underbrace{-\rho^* \left(\sum_{q \in p^w} [\eta_w^{pq} h_w^{q*}] - \sum_{q \in p^w} [\eta_w^{pq} g_w^{q*}] \right)}_{< 1}} < 1 \tag{37}$$

This is true given assumption (35) and provided that $\rho^* \geq 0$. From this, we deduce that $-\rho^* (h_w^{p^*} - g_w^{p^*}) < 0 \Rightarrow e^{-\rho^* (h_w^{p^*} - g_w^{p^*})} < 1$. There is thus a contradiction, and we conclude that a fixed point not only exists but is unique.

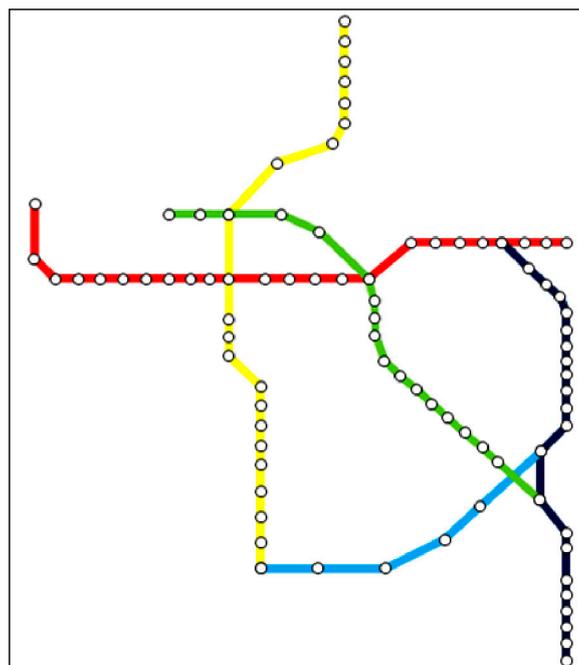
4. Numerical Results and Extensions

In Section 4.1, we present a route choice case study comparing the results of a real application of the proposed model to those generated by the classic MNL model in Equation (2), C-Logit model in Equation (4), path-size logit model in Equation (6), PCL model in Equation (9) and the CNL model in Equation (12). Next, in Section 4.2, we develop an extension of the proposed formulation with route correlations to traffic assignments on congested networks (private transportation).

4.1. Application to a Medium-Sized Network: The Santiago Metro

This case study applies the various models to route choice on the Metro rail transport system in the Chilean city of Santiago. Successive transfer points are chosen between origin and destination pairs that in many instances are joined by more than one feasible route (see Figure 1). The analysis focuses on the morning and evening peak periods (7 am to 9 am and 6 pm to 8 pm) when approximately 790,000 trips are taken across the system, 44% of which include transfers.

Figure 1. Santiago Metro in 2008.



Trip data were obtained through an O-D survey conducted on the Metro in which 92,800 system users, or approximately 12% of the total, participated. Because only those trips for which there existed more than one alternative route were retained in the data set, the number of individuals or observations eventually employed was 16,029, or approximately 40% of users who transferred.

When the survey was performed in October 2008, the Santiago Metro consisted of five lines and 85 stations, seven of which were transfer points. Of the 7140 O-D pairs in the network, 4985 (70%) of them required transfers. The reasonability criterion applied in including a route as a possible alternative for a given pair was that at least one surveyed traveler was observed to have used it. Data on the alternative routes for O-D pairs across the system that had more than one route on this criterion are given in Table 1. Although in the majority of cases there were only two alternatives, some pairs

had as many as four. Denser networks than this one would undoubtedly have a greater proportion of pairs with alternative routes.

Table 1. Origin-destination pairs and alternative routes, 2008.

No. of Observed Alternative Routes	% of All O-D Pairs	% of All Trips
2	97%	93%
3	3%	7%
4	<1%	<1%

For the proposed model (20), the generalized cost of traveling between pair w on route p is defined as $L_w^p = \beta_{time} X_{time,w}^p + \beta_{access} X_{access,w}^p + \beta_{trans} X_{trans,w}^p + \beta_{rr} D_{rr,w}^p + \beta_{old} D_{old,w}^p$, where $X_{time,w}^p$ is total trip time, $X_{access,w}^p$ is the access time (walking and waiting), $X_{trans,w}^p$ is the number of transfers along the route, $D_{rr,w}^p$ is a binary variable that takes the value 1 if route p is reasonable for travelling in w (based onDial [23]) and takes the value 0 if not, $D_{old,w}^p$ is a binary variable that takes the value 1 if the route is less than 10 years old (a proxy of how well known the route is) and takes the value 0 if it is older. The number of transfers on a given route alternative is considered as an indicator of the disutility of transferring. Because the Metro uses a flat fare system, the fare is the same regardless of the route and can therefore be omitted from the cost function.

The η_w^{pq} term, the spatial correlation factor due to route overlap between routes p and q joining pair w , is defined as suggested by Yai *et al.* [41]:

$$\eta_w^{pq} = \frac{D_w^{pq}}{\sqrt{D_w^p \cdot D_w^q}} \tag{38}$$

where D_w^{pq} is the length of overlap between routes p and q and D_w^p and D_w^q are the respective lengths of routes p and q . This expression stems from the traditional definition of the similarity factor in the C-logit model. As with the parameter β in that model (see Equation (5) above), the parameter $-\rho^*$ must be negative, ensuring that ρ^* is positive.

The proposed model was compared to a conventional MNL model that included only network service level variables. That is, ρ^* was set to 0. The comparison was performed using statistical tests. Thus, the six route choice models estimated for the Santiago Metro were the following:

- (a) Multinomial logit (MNL). This is the base model because it does not account for correlations and is also used to construct the route choice proxy variables.
- (b) C-logit.
- (c) Path-size logit (PSL).
- (d) PCL.
- (e) CNL.
- (f) Fixed-point model (FPM) with spatial correlations, which is our proposed model.

The estimation results are set out in Table 2. The proposed FPM converged in only four iterations at the 0.01% tolerance level. As can be seen, all explanatory variables had the correct sign and were statistically significant. In the C-logit, PCL, CNL and FPL models, the correlation parameter was significant; in the PSL model, the statistical significance was lower.

The models with better goodness-of-fit (log-likelihood value) were the CNL and our proposed FPL model, and both were very similar. In addition to the log-likelihood function, the following indicators were used to compare the MNL and fixed-point models based on their goodness of fit [50,51]:

- i. percent of correct predictions (*PCP*);
- ii. residual sum of squares (*RSS*): $S = \sum_i (Y_i - P_i)^2$, where Y_i is 1 if an alternative i is chosen and 0 otherwise and P_i is the probability predicted by the model of choosing alternative i ; and
- iii. weighted residual sum of squares (*WRSS*): $S' = \frac{\sum_i (Y_i - P_i)^2}{P_i(1 - P_i)}$.

The results of the three indicators for the estimated models are given in Table 3. The *PCP* indicator is practically the same for all models; therefore, it does not allow any analysis regarding forecasting capability. In regard to *RSS* and *WRSS*, once again the best models are the CNL and FPL models.

Table 3. Goodness-of-fit indicators for route choice models.

Indicator	MNL	C-Logit	PSL	PCL	CNL	FPM
<i>PCP</i>	81.4%	81.4%	81.4%	81.4%	81.6%	81.6%
<i>RSS</i>	2273.1	2272.0	2272.5	2253.9	2227.5	2236.5
<i>WRSS</i>	28,072.2	28,339.0	28,320.3	23,147.8	19,745.0	19,757.9

4.2. Extension of Proposed Fixed-Point Model to Traffic Assignment with Route Correlation

The proposed fixed-point model can be extended to the traffic assignment problem in any congested transportation networks. The multi-objective optimization problem would then take the following form:

$$\begin{aligned}
 \min F_1 &= \sum_a \int_0^{f_a} c_a(x) dx \\
 \min F_2 &= \sum_w \sum_{p \in P^w} h_w^p (\ln h_w^p - 1) \\
 \min F_3 &= \sum_w \sum_{p \in P^w} \sum_{\substack{q \in P^w \\ q \neq p}} [\eta_w^{pq} (h_w^p - t_w)(h_w^q - t_w)] \\
 \text{s.t.} &: \sum_{p \in P^w} h_w^p = T_w, \quad \forall w \quad (\gamma_w) \\
 & \sum_{\substack{p \in P^w \\ a \in p}} h_w^p = f_a, \quad \forall a
 \end{aligned}
 \tag{39}$$

Objective F_1 is the classic Beckmann transformation of the traffic assignment problem [2]. Objectives F_2 and F_3 were described above in Section 3.1 as information criteria. Combining F_2 and F_3 with flow conservation constraints gives the Fisk stochastic assignment model [52].

A substitute optimization problem for Equation (14) that defines a new stochastic traffic assignment model with route correlation is the following:

$$\begin{aligned}
 \min Z &= \sum_a \int_0^{f_a} c_a(x) dx + \frac{1}{\theta} \sum_{p \in P^w} \sum_w h_w^p (\ln h_w^p - 1) + \frac{1}{\rho} \sum_w \sum_{p \in P^w} \sum_{\substack{q \in P^w \\ q \neq p}} [\eta_w^{pq} (h_w^p - t_w)(h_w^q - t_w)] \\
 \text{s.t.} &: \sum_{p \in P^w} h_w^p = T_w, \quad \forall w \quad (\gamma_w) \\
 & \sum_{\substack{p \in P^w \\ a \in p}} h_w^p = f_a, \quad \forall a
 \end{aligned}
 \tag{40}$$

The first-order conditions for Equation (40) are:

$$h_w^p = T_w \frac{\exp\left(-\theta C_w^p - \frac{\theta}{\rho} \sum_{q \in P_w} [\eta_w^{pq} (h_w^q - t_w)]\right)}{\sum_{p \in P_w} \exp\left(-\theta C_w^p - \frac{\theta}{\rho} \sum_{q \in P_w} [\eta_w^{pq} (h_w^q - t_w)]\right)} \quad (41)$$

where $C_w^p = \sum_{\substack{p \in P_w \\ a \in p}} c_a(f_a^*)$.

5. Conclusions

A new transportation network route choice model was developed that explicitly incorporates the phenomenon of correlations between routes. The model is applicable to public transportation networks and extensible to the traffic assignment problem. The presence of correlations between route alternatives was contrasted empirically by means of classical econometric techniques.

A multi-objective problem was stated and a substitute problem then formulated whose optimality conditions yielded a logit specification with an endogenous variable constituting a fixed-point model that is estimated and solved iteratively. The estimation was performed by the maximum likelihood and by the use of instrumental variables due to the presence of endogeneity in the model's explanatory variables. The functional form of the proposed fixed-point model combined with the utilization of instrumental variables guaranteed both the existence and uniqueness of the solution.

The proposed model was compared with other route choice models reported in the literature in a case study of the Santiago Metro. The results obtained were both satisfactory and superior to many other existing formulations, and statistically equivalent to CNL model. Unlike the latter, the proposed fixed-point model was able to capture the correlation between routes and provided a better goodness of fit. A future research should be related to compare FPM with CNL in large networks. However, FPM model is able to include congestion in traffic equilibrium conditions; this is an advantage of our new model.

Although the proposed formulation could be more complex to estimate due to the iterative process that must be used to obtain the parameters, in practice the iterations converge rapidly. Furthermore, the additional estimation complexity does not complicate the derivation of project evaluation indicators such as the marginal rates of substitution, which retain the simple functional form of the MNL models.

Lastly, future research should address the apparent advantages of the proposed model on larger networks and traffic assignment problems.

Acknowledgments

This research was supported by the FONDECYT 11121199.

Author Contributions

Louis de Grange developed the theoretical model and analyzed the conditions of existence and uniqueness of the solution. Felipe González built the database and he also calibrated the new

model. Sebastián Raveau adapted the databases and then calibrate alternative models to compare our new model. All authors have read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Wardrop, J.G. Some theoretical aspects of road traffic research. In *ICE Proceedings: Engineering Divisions*; Thomas Telford: London, UK, 1952; pp. 325–362.
2. Beckmann, M.; McGuire, C.; Winsten, C.B. *Studies in the Economics of Transportation*; Yale University Press: New Haven, CT, USA, 1956.
3. Dijkstra, E.W. A note on two problems in connexion with graphs. *Numer. Math.* **1959**, *1*, 269–271.
4. De Cea, J.; Fernández, E. Transit assignment for congested public transport systems: An equilibrium model. *Transp. Sci.* **1993**, *27*, 133–147.
5. De Grange, L.; Muñoz, J.C. An equivalent optimization formulation for the traffic assignment problem with asymmetric linear costs. *Transp. Plan. Tech.* **2009**, *32*, 1–25.
6. Dafermos, S.C.; Sparrow, F.T. Traffic assignment problem for a general network. *J. Res. Natl. Bur. Stand. Sect. B-Math. Sci.* **1969**, *73*, 91–117.
7. Florian, M.A., Ed., *Traffic Equilibrium Methods*; Lecture Notes in Economics and Mathematical Systems; Springer-Verlag: Berlin, Germany, 1976; Volume 188.
8. Florian, M. A traffic equilibrium model of travel by car and public transit modes. *Trans. Sci.* **1977**, *11*, 166–179.
9. Dafermos, S. Traffic equilibrium and variational inequalities. *Transp. Sci.* **1980**, *14*, 42–54.
10. Patriksson, M. *Nonlinear Programming and Variational Inequality Problems: A Unified Approach*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1998; Volume 23.
11. Florian, M.; Hearn, D. Traffic assignment: Equilibrium models. *Proc. Inst. Mech. Eng. Part I J. Syst. Control Eng.* **2001**, *215*, 305–315.
12. Boyce, D.E.; Mahmassani, H.S.; Nagurney, A. A retrospective on beckmann, mcguire and winsten's studies in the economics of transportation. *Pap. Reg. Sci.* **2005**, *84*, 85–103.
13. Daganzo, C.F.; Sheffi, Y. On stochastic models of traffic assignment. *Transp. Sci.* **1977**, *11*, 253–274.
14. Hazelton, M.L. Some remarks on stochastic user equilibrium. *Transp. Res. Part B Methodol.* **1998**, *32*, 101–108.
15. Ramming, M.S. *Network Knowledge and Route Choice*; Massachusetts Institute of Technology: Cambridge, MA, USA, 2001.
16. Prashker, J.N.; Bekhor, S. Route choice models used in the stochastic user equilibrium problem: A review. *Transp. Rev.* **2004**, *24*, 437–463.
17. Cascetta, E.; Nuzzolo, A.; Russo, F.; Vitetta, A. A Modified Logit Route Choice Model Overcoming Path Overlapping Problems: Specification and Some Calibration Results for Interurban Networks. In *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, Lyon, France, 24–26 July 1996; Pergamon: Oxford, NY, USA 1996; pp. 697–711.

18. Ben-Akiva, M.; Bierlaire, M. Discrete choice methods and their applications to short term travel decisions. In *Handbook of Transportation Science*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1999; pp. 5–34.
19. Bekhor, S.; Prashker, J.N. Formulations of Extended Logit Stochastic User Equilibrium Assignments. In Proceedings of the 14th International Symposium on Transportation and Traffic Theory, Jerusalem, Israel, 20–23 July 1999.
20. Bekhor, S.; Prashker, J.N. Stochastic user equilibrium formulation for generalized nested logit model. *Transp. Res. Rec. J. Transp. Res. Board* **2001**, *1752*, 84–90.
21. Bovy, P.H.; Bekhor, S.; Prato, C.G. The factor of revisited path size: Alternative derivation. *Transp. Res. Rec. J. Transp. Res. Board* **2008**, *2076*, 132–140.
22. Prato, C.G. Route choice modeling: Past, present and future research directions. *J. Choice Model.* **2009**, *2*, 65–100.
23. Dial, R.B. A probabilistic multipath traffic assignment model which obviates path enumeration. *Transp. Res.* **1971**, *5*, 83–111.
24. Bell, M.G. Alternatives to dial's logit assignment algorithm. *Transp. Res. Part B: Methodol.* **1995**, *29*, 287–295.
25. Ben-Akiva, M.; Bergman, M.; Daly, A.J.; Ramaswamy, R. Modeling Inter-Urban Route Choice Behaviour. In Proceedings of the 9th International Symposium on Transportation and Traffic Theory, Delft, The Netherlands, 11–13 July 1984; VNU Press: Utrecht, The Netherlands, 1984; pp. 299–330.
26. Cascetta, E.; Papola, A.; Russo, F.; Vitetta, A. Implicit Availability/Perception Logit Models for Route Choice in Transportation Networks. In Proceedings of the 8th World Conference on Transport Research, Antwerp, Belgium, 12–17 July 1998.
27. Cascetta, E.; Russo, F.; Viola, F.; Vitetta, A. A Model of Route Perception in Urban Road Networks. *Transp. Res. Part B: Methodol.* **2002**, *36*, 577–592.
28. Ben-Akiva, M.; Bierlaire, M. Discrete choice models with applications to departure time and route choice. In *Handbook of Transportation Science*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; pp. 7–37.
29. Bovy, P.H.; Hoogendoorn-Lanser, S. Modelling route choice behaviour in multi-modal transport networks. *Transportation* **2005**, *32*, 341–368.
30. Prato, C.G.; Bekhor, S. Modeling route choice behavior: How relevant is the composition of choice set? *Transp. Res. Rec. J. Transp. Res. Board* **2007**, *2003*, 64–73.
31. Bekhor, S.; Ben-Akiva, M.E.; Ramming, M.S. Evaluation of choice set generation algorithms for route choice models. *Ann. Oper. Res.* **2006**, *144*, 235–247.
32. Bliemer, M.C.; Bovy, P.H. Impact of route choice set on route choice probabilities. *Transp. Res. Rec. J. Transp. Res. Board* **2008**, *2076*, 10–19.
33. McFadden, D. Modelling the Choice of Residential Location. In *Spatial Interaction Theory and Planning Models*; Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J., Eds.; North-Holland: Amsterdam, The Netherlands, **1978**, pp. 75–96.
34. De Cea, J.; Fernandez, J.E.; de Grange, L. Combined models with hierarchical demand choices: A multi-objective entropy optimization approach. *Transp. Rev.* **2008**, *28*, 415–438.

35. Donoso, P.; de Grange, L.; González, F. A maximum entropy estimator for the aggregate hierarchical logit model. *Entropy* **2011**, *13*, 1425–1445.
36. Ben-Akiva, M.; Ramming, S. Lecture notes: Discrete choice models of traveler behavior in networks. In Proceedings of the Advanced Methods for Planning and Management of Transportation Networks, Capri, Italy, 25–29 May 1998; Volume 25.
37. Chen, A.; Kasikitwiwat, P.; Ji, Z. Solving the overlapping problem in route choice with paired combinatorial logit model. *Transp. Res. Rec. J. Transp. Res. Board* **2003**, *1857*, 65–73.
38. Prashker, J.N.; Bekhor, S. Investigation of stochastic network loading procedures. *Transp. Res. Rec. J. Transp. Res. Board* **1998**, *1645*, 94–102.
39. Vovsha, P.; Bekhor, S. Link-nested logit model of route choice: Overcoming route overlapping problem. *Transp. Res. Rec. J. Transp. Res. Board* **1998**, *1645*, 133–142.
40. De Grange, L.; Fernández, E.; de Cea, J. A consolidated model of trip distribution. *Transp. Res. Part E: Logist. Transp. Rev.* **2010**, *46*, 61–75.
41. Yai, T.; Iwakura, S.; Morichi, S. Multinomial probit with structured covariance for route choice behavior. *Transp. Res. Part B: Methodol.* **1997**, *31*, 195–207.
42. Golan, A. Information and entropy econometrics—Editor’s view. *J. Econ.* **2002**, *107*, 1–15.
43. Marler, R.T.; Arora, J.S. Survey of multi-objective optimization methods for engineering. *Struct. Multidiscip. Optim.* **2004**, *26*, 369–395.
44. De Oliveira, J.A.; Papesso, E.R.; Leonel, E.D. Relaxation to fixed points in the logistic and cubic maps: Analytical and numerical investigation. *Entropy* **2013**, *15*, 4310–4318.
45. Greene, W.H. *Econometric Analysis*; Prentice Hall: Upper Saddle River, NJ, USA, 2011.
46. Donoso, P.; de Grange, L. A microeconomic interpretation of the maximum entropy estimator of multinomial logit models and its equivalence to the maximum likelihood estimator. *Entropy* **2010**, *12*, 2077–2084.
47. Gupta, M.; Srivastava, S. Parametric bayesian estimation of differential entropy and relative entropy. *Entropy* **2010**, *12*, 818–843.
48. Raveau, S.; Muñoz, J.C.; de Grange, L. A topological route choice model for metro. *Transp. Res. Part A: Policy Pract.* **2011**, *45*, 138–147.
49. Brouwer, L.E.J. Über abbildung von mannigfaltigkeiten. *Math. Ann.* **1911**, *71*, 97–115. (in German)
50. Copas, J. Unweighted sum of squares test for proportions. *J. R. Stat. Soc. Ser. C* **1989**, *38*, 71–80.
51. Wasserman, L. *All of Nonparametric Statistics*; Springer: New York, NY, USA, 2006.
52. Fisk, C. Some developments in equilibrium traffic assignment. *Transp. Res. Part B: Methodol.* **1980**, *14*, 243–255.