van der Weijde, Adriaan Hendrik; van den Berg, Vincent A. C.; Verhoef, Erik T.

Working Paper

Stochastic User Equilibrium Traffic Assignment with Price-sensitive Demand: Do Methods matter (much)?

Tinbergen Institute Discussion Paper, No. 13-209/VIII

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

TI 2013-209/VIII
Tinbergen Institute Discussion Paper

Stochastic User Equilibrium Traffic Assignment with Price-sensitive Demand: Do Methods matter (much)?

Adriaan Hendrik van der Weijde
Vincent A. C. van den Berg
Erik T. Verhoef

Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen Institute, The Netherlands.

Stochastic user equilibrium traffic assignment with price-sensitive demand: Do methods matter (much)?*

Adriaan Hendrik van der Weijde† Vincent A.C. van den Berg
Erik T. Verhoef

Department of Spatial Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081HV Amsterdam, The Netherlands
Tinbergen Institute, Gustav Mahlerplein 117, 1082MS Amsterdam, The Netherlands

Abstract

We compare three stochastic user equilibrium traffic assignment models (multinomial probit, nested logit, and generalized nested logit), using a congestible transport network. We test the models in two situations: one in which they have theoretically equivalent coefficients, and one in which they are calibrated to have similar traffic flows. In each case, we examine the differences in traffic flows between the SUE models, and use them to evaluate policy decisions, such as profit-maximizing tolling or second-best socially optimal tolling. We then investigate how the optimal tolls, and their performance, depend on the model choice, and hence, how important the differences between models are. We show that the differences between models are small, as a result of the congestibility of the network, and that a better calibration does not always lead to better traffic flow predictions. As the outcomes are so similar, it may be better to use computationally more efficient logit models instead of probit models, in at least some applications, even if the latter is preferable from a conceptual viewpoint.

1 Introduction

Stochastic user equilibrium (SUE) traffic assignment models are used in a wide range of applications. Starting with Dial (1971), who proposed a simple logit

*Financial support from ERC Advanced Grant #246969 (OPTION) is gratefully acknowledged. The authors thank Paul Koster, Andrew Koh, and the participants of the 2013 INFORMS Annual Meeting for helpful comments.

†Corresponding author. Tel.: +31 20 59 86106. E-mail: h.vander.weijde@vu.nl

model to analyze route choice in a network, many different formulations have been developed. Choosing a SUE model for a specific application is therefore not straightforward, especially since there are only a few studies that compare different formulations systematically. The present paper attempts to remedy this.

In a deterministic Wardropian equilibrium (Wardrop, 1952), the generalized prices (i.e. costs plus tolls) of all used routes are equal, and lower than those of all unused routes. This implicitly assumes that users have perfect knowledge about the costs of all routes, and that all relevant user attributes can be perfectly observed. SUE models instead use random utility discrete choice theory, which assumes that the utility users derive from a given route has a stochastic part, which cannot be explained by an observer, but which follows a known distribution. The resulting equilibrium differs from the Wardropian equilibrium in that the systematic (or, deterministic) generalized prices of all routes are typically not equal. Moreover, although the probability that a route will be used can approach 0 arbitrarily closely, it never reaches it. Besides being more realistic in many settings, this is also often computationally convenient.

In this paper, we first consider two 'workhorse' user equilibrium models: a multinomial probit, and a logit model. Both models are used regularly in the recent literature (see e.g. for logit: Meng et al., 2004; Yang, 1999; Yang et al., 2001, and for probit: Connors et al., 2007; Uchida et al., 2007; Meng et al., 2012). Probit models can account for partially overlapping routes and routes with significantly different lengths, while simple logit models can not properly handle overlap, and assume that all route costs are subject to the same level of stochasticity. However, logit models have closed form solutions for choice probabilities, while probit equilibria can only be determined using sampling techniques or numerical integration, and are therefore highly computationally intensive. It is therefore very useful to know how important the differences between these models are likely to be in real-world situations.

Although many studies investigate the differences between alternative discrete choice models, these almost exclusively focus on estimation, rather than simulation. For estimation, it is important that models fit existing data well. For simulation, the representation of one particular flow pattern is not of primary concern, as most models can be calibrated so as to achieve that. Instead, it is important to see what happens if we move away from the calibrated state, through, for instance, tolling. The different models imply different substitution patterns, so the effects of these changes could be different. This could have important implications for operational decisions and policies.

The few studies that do discuss the differences between logit and probit models for simulation purposes mostly date from the 70s, when these models were first developed (e.g. Florian and Fox, 1976; Daganzo and Sheffi, 1977). These studies compare simple versions of both models in small 'toy' networks; their results do not necessarily carry over to more realistic situations. We therefore compare the different SUE models in a setting that has:

1. A network that is more realistic than the simplest 'toy' networks that

have been used in the past. Specifically, we model a network with several pairs of origins and destinations (ODs) and multiple overlapping routes between each OD-pair, which share links with routes that are used between other OD-pairs. Using this network, we can not only look at the effects of different SUE models on an OD-level, but also examine how these effects interact in a network.

2. Congestion; specifically, we use the realistic congestion function proposed by US Bureau of Public Roads (1964). As we will show in section 2.2, congestion can have a large effect on the differences between SUE models.

3. Price-sensitive demand, through the addition of an alternative 'virtual', uncongested route between each OD-pair, that does not overlap with any others.

4. Parameters that are calibrated correctly. If all assignment models are calibrated to the same traffic flows, as would happen in reality, this might give different results than a situation where the assignment models are theoretically equivalent (i.e. where possible, have comparable parameters). Therefore, we consider two cases: one in which the assignment models have comparable parameters, and one in which the logit models are calibrated to the flows resulting from the probit model. We also investigate how the introduction of additional (alternative-specific) parameters, which can be used for calibration, affects the differences between the models.

Although the standard logit model is still widely used, several generalizations and extensions have been proposed. As an example of such an advanced logit model, we also include the generalized nested logit (GNL) model in our analysis. A relatively novel model, the GNL (Wen and Koppelman, 2001) can partially account for overlapping routes (though it does so in a different way than the probit model); for this reason, it has been proposed as a possible SUE assignment model (Bekhor and Prashker, 2002). Unlike the probit model, it cannot properly account for the fact that routes have different lengths; however, it does have closed-form probabilities.

To examine the differences between the assignment models, we test their performance in a 12-node network. In particular, we evaluate how the different assignment models affect policy decisions, such as profit-maximizing and welfare-maximizing tolls, and the effects of such policies. As we will show, the differences between models are, in many instances, small. However, improper calibration can lead to large differences, and significant reductions in welfare.

The next section defines the assignment models, and explores the theoretical differences between them, using a very simple network model with only one origin and one destination. Section 3 gives an overview of the methodology we use for the more realistic simulations, the results of which are presented in section 4. Section 5 concludes.

2 Theory

2.1 SUE assignment models

All the stochastic user equilibrium assignment models we examine are based on random utility theory. In these models, the utility users derive from an alternative (which, in traffic assignment, can be a route through a network, the use of a specific link, or the option to stay home and not travel) consists of a deterministic, and a stochastic part:

$$U_{ir} = V\left(-c_r\right) + \varepsilon_{ir} \tag{1}$$

where U_{ij} is the utility user i derives from alternative r, which is made up of a deterministic value function V and a stochastic term ε_{ir}. For simplicity, we define the value as a function of the generalized price c_r of an alternative only; we assume that benefits are the same for all alternatives. Since we have only one user class, V is also independent of personal characteristics. Although, in much of the recent literature, this random term is assumed to capture the uncertainty that transport *users* face, it was originally intended to capture measurement errors made by the *observer* (reflecting, for instance, the fact that users differ in unobservable characteristics, which influence the utility they derive from the use of an alternative). Different assignment models assume different joint distributions for the random terms.

Multinomial probit (MNP)

The first discrete choice model we consider is one of the most flexible. In the multinomial probit (MNP) model, the random terms ε follow a multivariate normal distribution; $\varepsilon \sim \mathcal{N}(0, \mathbf{\Sigma})$, where the index m refers to markets, or OD-pairs. Importantly, if $\mathbf{\Sigma} \neq \mathbf{I}$, random terms are allowed to be correlated across all alternatives, and the variances can differ across alternatives. We can therefore set $\mathbf{\Sigma}$ such that the covariances between different routes, $0 \leq \sigma_{rr'} < 1$, indicate to what extent routes overlap, and variances σ_r^2 that vary with the length of each route.

As already stated, we will model the price-sensitivity of demand through the inclusion of a no-travel alternative, which has a given cost (or disutility). This is not the only way to make travel demand price-sensitive. It is also possible to make the total travel demand between an origin and a destination a direct function of the expected utility of the discrete choice. However, particularly in probit models, this is considerably more complicated, and for our purposes, it yields few advantages. Our approach still results in a downward-sloping demand curve, which is non-linear, and whose slope reflects by the elasticity of the no-travel alternative choice probability.

There are several ways to derive the covariances $\sigma_{rr'}$ and variances σ_r^2. The most straightforward way, which we will use, is to formulate link-based random terms ε_{il}. The total random utility faced by a user who takes a given route r, ε_{ir} is then the sum of all link-based random terms, over all links that constitute the route. Since the sum of two or more independent normal distributions is also a normal distribution, the covariance between two routes is then the sum of the variances of the links they share. We set link-based variances $\sigma_l = sK_l$, where K_l is the length of link l, and s a parameter. For simplicity, we set the variance of the no-travel alternative to 1. The variance of route r is then simply $\sum_l a_{lr}\sigma_l^2$, where $a_{lr} = 1$ if link l is part of route r, and zero otherwise. The covariance between routes r and $r\prime$ is $\sum_l a_{lr}a_{lr\prime}\sigma_l^2$; the covariance between any route and the no-travel alternative is zero. We use s to control the average route variance $\tilde{\sigma}^2$, which we normalize to 1. Finally, we set $V(-c_r) = -c_r/10$, where the coefficient is chosen such that, in equilibrium, a significant number of routes is used. However, note that, in this model, we cannot differentiate between the average variance of the random terms and this coefficient. Setting $\tilde{\sigma}^2 = 10$ and $V(-c_r) = -c_r$ would yield the same results.

MNP models models do not have closed-form probabilities; instead, numerical methods are necessary to determine them. We use a Monte Carlo simulation, where random terms ε are sampled from the multivariate normal distribution described above. Based on these sampled random terms and the deterministic route costs, probabilities can be estimated. Naturally, since these probabilities depend on traffic volumes, we need an iterative procedure to find the user equilibrium assignment. We use the Method of Successive Averages (MSA), which consists of the following steps:

1. Determine initial link flows v_{l0} (e.g. estimate probabilities based on free-flow travel costs) and obtain link-based travel costsχ_l. Set $n = 1$.

2. Calculate auxiliary link flows v_l^A based on costs χ_l.

3. Set $v_{ln+1} = v_{ln} + n^{-1}\left(v_l^A - v_{ln}\right)$ and recalculate costs χ_l. Increase n by 1.

4. Repeat steps 2–3 until $|v_{ln+1} - v_{ln}|$ becomes sufficiently small.

Since step 2 requires a Monte-Carlo simulation, and the number of iterations the MSA needs to converge can be high, this model can be computationally intensive. It can, however, account for both overlapping routes, and routes of different lengths, which the other models we will examine can not fully take into account.

Logit

The simplest logit model assumes that all random terms ε_{ir} are independent. However, the introduction of elastic demand through the addition of a no-travel alternative makes this assumption particularly unrealistic; it is logical to assume that users first decide whether to travel or not, and choose a route only after they have decided to travel. We therefore use a somewhat more advanced nested logit (NL) model (Ben-Akiva, 1974), in which one nest contains all the physical routes, and the other the no-travel alternative. In the NL model, the random terms follow a multi-variate extreme value distribution, and there is correlation within, but not between, nests. Hence, if we only consider the users that have already chosen to travel, the traffic assignment model is still an standard logit model; the NL formulation just allows us to model the demand elasticity correctly. To avoid confusion, however, we will refer to this model as an NL model in what follows.

The Generalized Extreme Value (GEV) function of this model (see McFadden, 1978) for a given OD-pair (or market) m is

$$G_m\left(e^{V(-C)}, e^{V(-c_1)}, ..., e^{V(-c_R)}\right) = e^{V(-C)} + \left(\sum_r a_{rm}\left(e^{V(-c_r)}\right)^{\frac{1}{\mu}}\right)^{\mu} \tag{2}$$

where $a_{rm} = 1$ if route r serves market m and zero otherwise, c_r is the cost of route r and C the cost incurred by a user who decides not to travel. $0 < \mu \leq 1$ is a parameter, indicating the dissimilarity between the option to not travel on the one side, and all possible routes on the other (more precisely, this implies a correlation of $(1-\mu)^2$ between the random terms attributed to different routes (Daganzo and Kusnic, 1993)). We use linear value functions $V(-C) = -\beta C$ and $V(-c_r) = -\beta c_r$.

As stated before, we will examine two versions of this model. In the first version, we set $\beta = \pi/\sqrt{10 \times 6}$ to achieve the same variance as MNP model: since logit models assume that the random terms follow a Gumbell distribution, which has a variance of $\pi^2/6$, and the MNP model defined above has an average variance $\tilde{\sigma}^2 = 1$ and $\beta = 1/10$, we have to correct for this in the NL model. We also set μ such that $(1-\mu)^2$ is equal to the average overlap between routes. In the second version, we instead estimate β and μ, using Maximum Likelihood Estimation (MLE) on the route flows resulting from the MNP model; the resulting parameters are different than those in the first model.

Whatever the value of β, the resulting route choice probabilities have closed forms:

$$p_{mr} = \frac{\left(e^{-\beta c_r}\right)^{\frac{1}{\mu}}}{\sum_{r'} a_{r'm}\left(e^{-\beta c_{r'}}\right)^{\frac{1}{\mu}}} \frac{\left(\sum_{r'} a_{r'm}\left(e^{-\beta c_{r'}}\right)^{\frac{1}{\mu}}\right)^{\mu}}{\left(\sum_{r'} a_{r'm}\left(e^{-\beta c_{r'}}\right)^{\frac{1}{\mu}}\right)^{\mu} + e^{-\beta C}} \tag{3}$$

and consist of two parts. The first gives the probability that a randomly selected user chooses route r, given that that particular traveler has already decided to travel. The second part gives the probability that this user travels at all. Note that, in this case, $\sum_r p_{mr} < 1$, since some users will choose not to travel.

Although this NL model allows for correlation between nests, and can therefore control for the difference between traveling and not traveling, it still assumes that alternatives within a nest are independent. Hence, contrary to the MNP, overlapping routes are not properly accounted for. Moreover, the variance is the same for all alternatives, so it is also impossible to properly account for the fact that routes have different lengths.

Nested logit with alternative-specific constants (NL-ASC)

To see how better calibration affects the difference between models, we also examine a version of the NL model that has additional parameters, which can be used for calibration. We do this by defining an alternative-specific constant A_r for every route (not including the no-travel option), such that the value functions become:

$$V(-c_r) = A_r - \beta c_r \tag{4}$$

Using MLE, these constants are then estimated, together with the β and μ. There is no constant for the no-travel alternative, to avoid overspecification.

Naturally, the flow pattern resulting from this model will be closer to the MNP flows in the calibrated state, as there are many additional parameters to use for calibration. However, as the MNP model does not have alternative-specific constants, the NL-ASC model may perform worse when we move away from the calibrated state, for instance, when a link in the network is tolled. The direction and magnitude of this effect depends on how the introduction of these ASCs changes the calibrated values of β and μ.

Generalized nested logit (GNL)

Like the MNP model, the GNL model (Wen and Koppelman, 2001; see Daly and Bierlaire, 2006 for an alternative formulation), also sometimes called the cross-nested logit model, can control for overlap in routes. In the GNL, alternatives can belong to several nests; inclusion parameters α_{rl} indicate which share of alternative r belongs to nest l. In a traffic assignment model, each link would be a nest (with an additional nest containing the option not to travel), and the share of a route that uses a specific link can be approximated using free-flow travel speeds (Bekhor and Prashker, 2002). The GEV function for market m is then

$$G_m\left(e^{V(-C)}, e^{V(-c_1)}, ..., e^{V(-c_R)}\right) = e^{V(-C)} + \sum_l \left(\sum_r a_{rm}\left(\alpha_{rl} e^{V(-c_r)}\right)^{\frac{1}{\mu_l}}\right)^{\mu_l} \tag{5}$$

where

$$\alpha_{rl} = \left.\frac{\chi_l}{c_r}\right|_{n_l\ \forall l} \tag{6}$$

and χ_l if the cost of using link l. Again, we define $V(-C) = -\beta C$ and $V(-c_r) = -\beta c_r$. Probabilities still have a closed form, and are given by

$$p_{mr} = \sum_l \left(\frac{\left(\alpha_{rl} e^{-\beta c_r}\right)^{\frac{1}{\mu_l}}}{\sum_{r\prime} a_{r\prime m}\left(\alpha_{r\prime l} e^{-\beta c_r}\right)^{\frac{1}{\mu_l}}} \frac{\left(\sum_{r\prime} a_{r\prime m}\left(\alpha_{r\prime l} e^{-\beta c_r}\right)^{\frac{1}{\mu_l}}\right)^{\mu_l}}{\sum_l \left(\sum_{r\prime} a_{r\prime m}\left(\alpha_{r\prime l} e^{-\beta c_r}\right)^{\frac{1}{\mu_l}}\right)^{\mu_l} + e^{-\beta C}} \right) \tag{7}$$

Note that, although this expression is very similar to Eq. 3, the crucial difference is that there are now no longer two nests per market (travel and no travel); instead the number of nests is equal to the number of routes plus one. Again, Eq. 7 only gives the route choice probabilities; there is also the option to not travel, so $\sum_r p_{mr} < 1$. Bekhor and Prashker (2002) propose to directly relate μ_l to the network topology, by making it an inverse function of the average inclusion coefficient of routes using link l. In our setting, all links share the exact same characteristics, so we also average over all links to get a single $\mu = \sum_l (1/L)(1 - (1/R_l)\sum_r \alpha_{rl})$, where L is the total number of links, and R_l the number of routes passing through link. We then set $\mu_l = \mu \, \forall l$, while β has the same value as in the theoretically equivalent NL model. As with the NL model, we also examine a calibrated version, in which we use MLE to find a β and μ that provides the best fit with the probit route flows.

Although GNL models have closed-form probabilities, which is very convenient for a traffic assignment model, the implied covariances between alternatives do not, and are not easy to calculate. They are given by

$$Cov\left(\varepsilon_r, \varepsilon_{r\prime}\right) = \iint_{\mathbb{R}} \left(F\left(\varepsilon_r, \varepsilon_{r\prime}\right) - F\left(\varepsilon_r\right) F\left(\varepsilon_{r\prime}\right)\right) d\varepsilon_r d\varepsilon_{r\prime} \tag{8}$$

where

$$F\left(\varepsilon_r\right) = \exp\left(-\exp\left(-\beta\varepsilon_r\right)\right) \tag{9}$$

and

$$F\left(\varepsilon_r, \varepsilon_{r\prime}\right) = \exp\left(-\sum_l \left(\left(\alpha_{rl} e^{-\beta\varepsilon_r}\right)^{\frac{1}{\mu_l}} + \left(\alpha_{r\prime l} e^{-\beta\varepsilon_r}\right)^{\frac{1}{\mu_l}}\right)^{\mu_l}\right) \tag{10}$$

(Marzano and Papola, 2008; Lemp et al., 2010). These covariances are different than those of the MNP model, but do capture some of the overlap between routes.

2.2 Model differences

Before considering a more complex setting, it is useful to look at the simplest possible network in which the differences between the models above can be illustrated. This simple network, which is often used in the early literature on discrete choice models (e.g. Florian and Fox, 1976; Daganzo and Sheffi, 1977) has three routes, of which two partially overlap. Fig. 1 gives a graphical representation of such a network, where all routes between A and C have a length of 1, and the two routes that pass through B share a length of $1 - x$. We will have an independent route 1; the others routes are denoted 2 and 3. For simplicity we assume, for a moment, that the no-travel alternative is not available, such that the NL model collapses to the simplest possible multinomial logit model.

If there is no congestion, and hence, all link costs are constant and equal, the logit model will assign 1/3 of the total flow to each of the routes, which,

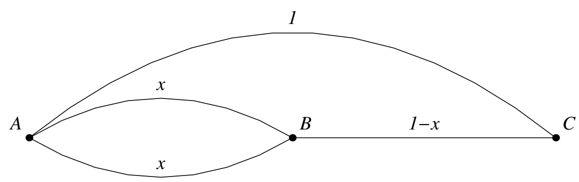

Figure 1: Simple network

arguably, is unrealistic. A probit model can account for the overlap, by setting

$$\sum = \begin{bmatrix} \sigma^2 & 0 & 0 \\ 0 & \sigma^2 & (1-x)\sigma^2 \\ 0 & (1-x)\sigma^2 & \sigma^2 \end{bmatrix} \tag{11}$$

The probabilities of the two overlapping routes will then increase in x, and the probability of the independent route will decrease, as shown in the left-hand panel of Fig. 2. Daganzo and Sheffi (1977) show a similar figure, and assuming that the MNP model is the correct one, argue that the logit model is often highly biased, because it fails to take the correlations between routes into account.

However, in traffic assignment models, links are usually congestible, so the costs of each route depend on the fraction of travelers that uses it. In our simple model, we can introduce congestion by, for instance, making link costs per unit of distance a linear function of link flows:

$$\mathbf{c} = \begin{bmatrix} 1 + 3p_1 \\ 1 + 3\left(xp_2 + (1-x)\left(p_2 + p\right)\right) \\ 1 + 3\left(xp_3 + (1-x)\left(p_2 + p_3\right)\right) \end{bmatrix} \tag{12}$$

where p_l is the probability, or fraction of the total flow assigned to route l, and the gradient of the congestion cost is set at 3 to generate a realistic fraction of congestion costs to total costs. The right-hand panel of Fig. 2 shows the fraction of the total flow assigned to route 1 that results from these route costs.

As Fig. 2 shows, the introduction of congestion significantly reduces the difference between the logit and probit models. This happens because, when the links are congestible, users of the two overlapping routes impose a congestion externality on each other, since they both use link BC. This makes routes 2 and 3 less attractive, especially if x is small. The random terms in the logit model are still independent, but the systematic utilities $V(-c_r)$ now share a common term. Hence, the difference between the logit and probit models becomes less important. It does not completely disappear, however, and is still significant if the amount of overlap between routes 2 and 3 is big.

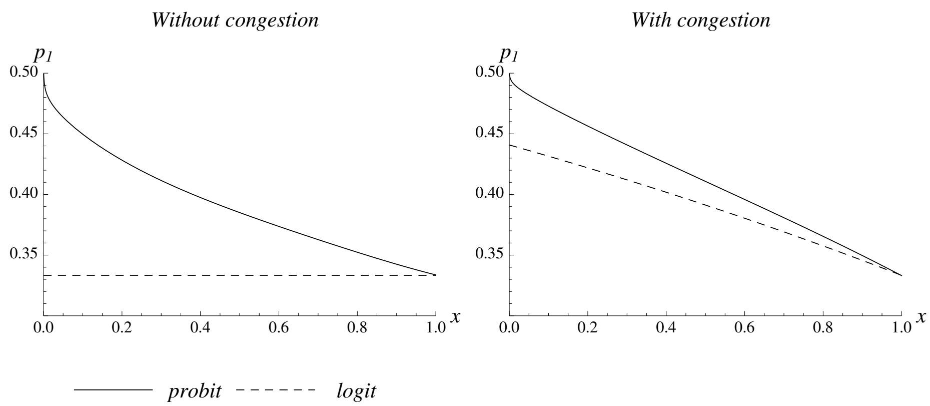

Figure 2: Logit vs. probit

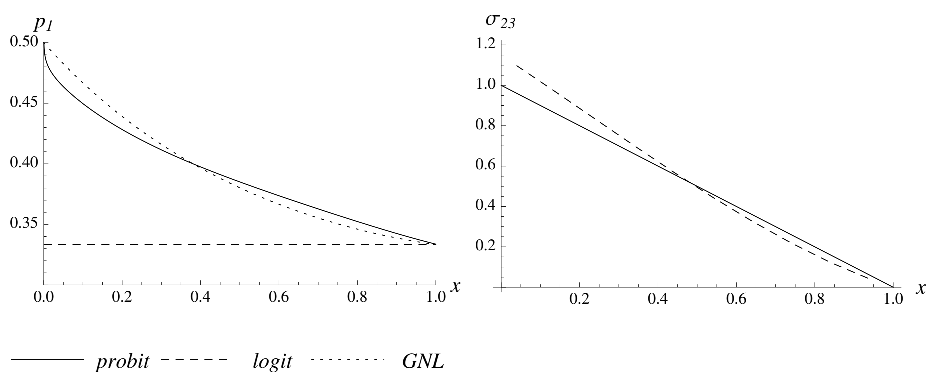

Figure 3: GNL

So far, we have only compared the simplest possible logit model with the probit model. The left-hand panel in Fig. 3 shows the results of a similar exercise, without congestion, this time including the GNL model, with

$$\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & x & 0 & 1-x \\ 0 & 0 & x & 1-x \end{bmatrix} \tag{13}$$

and $\mu = 2x/3$, which is chosen to generate route choice probabilities close to the MNP model. As Fig. 3 shows, the MNP and GNL fractions are very close, even in the model without congestion. The right-hand panel of Fig. 3 shows the covariance between the two overlapping routes in both models, which are calculated using Eqs. 8 and 11; they are also very close. Hence, the GNL model seems to approximate the MNP model well, at least in this simple example. Moreover, there is no clear relation between the amount of overlap and the difference between the two models.

The simple network in Fig. 1, has only one market; travel from A to C. In more realistic models, there are more complex network effects; links are used by travelers in multiple markets. Depending on the network structure, this can further reduce the difference between the assignment models. If, for instance, the top link in Fig. 1 is also part of route that serves another market, and that route partially overlaps another, the two differences between logit and probit probabilities may cancel each other out. Moreover, note that, for there to be a difference between the models, all three routes (two overlapping, and one separate) must not only exist, but also be used by a significant fraction of travelers. In a larger congested network, this is not necessarily the case; a large amount of congestion on a major link between two large nodes could, for instance, make that link so unattractive to travelers between other nodes, that all routes using it would be assigned very low choice probabilities.

The introduction of an additional alternative, representing the option not to travel, affects the models in different ways. On the one hand, the addition of a non-overlapping alternative in every market makes it more likely that a situation such as the one in Fig. 1 exists, and hence, that there are significant differences between assignment models. On the other hand, however, if that alternative is added as a separate nest in a nested logit model, as we will do, this gives an extra parameter (the dissimilarity between traveling and not traveling) which can be used to calibrate the logit model, and hence, reduce the difference with the probit model.

3 Simulation methodology

3.1 Network

To compare our three assignment models, we will apply them to the simple network shown in Fig. 4, where all links are bidirectional (and congestion levels are direction-specific), and the size of each node indicates the potential demand for travel. Nodes indicated with asterisks are connection nodes only. All links are 10 km long, and we consider all possible non-circular routes. This network layout allows us to examine partially overlapping routes, and interactions between markets. Importantly, there is not only overlap of potential but unused routes; in the base case equilibrium, which we will present below, there is still a significant amount of overlap if we only consider routes that are used by significant fractions (e.g., $> 5\%$) of travelers.

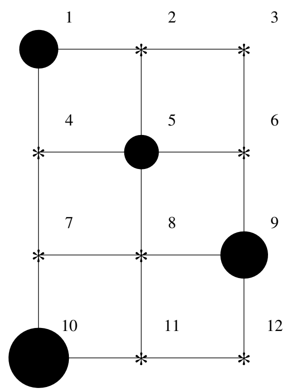

Figure 4: Network

All links in the network are congestible; we model congestion using the well-known Bureau of Public Roads (BPR) function (US Bureau of Public Roads, 1964), a widely used approximation of the congestion costs of highway travel:

$$\chi_l = vot \cdot \frac{K_l}{S_f} \left(1 + 0.15\,(v_l/V_l)^4\right) + f_l \tag{14}$$

where vot is the value of time, K_l the length of link l, S_f the free-flow travel speed (which we assume to be the same on all links), f_l a potential toll (or fare), and v_l/V_l the volume-capacity ratio on link l. We set the latter set such that the ratio of congestion costs to total costs remains within realistic limits; specifically, such that the term multiplying the value of time in Eq. 14 is between 2 and 12. The traffic volume on a link is the sum of the realized demand of every route using the link:

$$v_l = \sum_m \sum_r a_{rm} a_{lr} p_{mr} D_m \tag{15}$$

In the logit models, route choice probabilities p_{mr} and route flows $p_{mr}D_m$ can be calculated directly. In the probit model, this is more involved. Each iteration of the MSA uses new multivariate draws, which are independent of the previous one; hence, there is simulation noise in every iteration. Link flows are calculated as a weighted average of all auxiliary flows in previous iterations, which averages out the simulation noise. Route flows can be calculated in every iteration, but as they are not averaged, they will contain a much larger amount of noise.

The most straightforward way to accurately estimate route flows is to run one last iteration with a large sample size, after the link flows have converged, and calculate route flows in that iteration. This can be highly computationally intensive, especially in larger networks. Another solution, which we will use, is to calculate a successive average of route flows in all iterations; a similar procedure as the one followed to obtain link flows (except for the fact that averaged link flows are used in the following iterations, whereas route flows are just stored). This method uses information that is already available, and gives consistent results, provided that the number of iterations is large enough, such that the remaining simulation noise in the resulting average route flow is sufficiently small.

Finally, the cost of taking a given route is simply the sum of the the link-based costs over all links that make up the route

$$c_r = \sum_l a_{lr} \chi_l \tag{16}$$

3.2 Demand

The potential demand (or demand function intercept) in market m is calculated with a simple gravity model, such that it increases in the size N_i of each of the two nodes that form the market, and decreases in the distance K_m between them:

$$D_m = \frac{\prod_{i \in m} N_i}{\delta K_m} \tag{17}$$

where δ is a parameter. The realized demand depends on the generalized price of travel, trough the SUE models.

3.3 Welfare and profit

It is difficult to define a consistent welfare measure across all models. Logsums are a natural choice for the logit models, but the lack of a closed form in the probit model makes calculating an equivalent measure there more complicated. We therefore use the Rule of Half, which is often used for policy analyses. It approximates the welfare gains from a certain policy policy change (in our case, a change in toll) by

$$\Delta W = \sum_r \frac{1}{2} \left(q_r^1 + q_r^2 \right) \left(c_r^1 - c_r^2 \right) + \left(\pi^2 - \pi^1 \right) \tag{18}$$

where $\left(\pi^2 - \pi^1 \right)$ is the profit change resulting from the change in toll, c_r^1 and c_r^2 the route costs before and after the change respectively, and q_r^1 and q_r^2 the numbers of users choosing each route (this includes the no-travel alternative) before and after the change. Hence, it approximates the demand curve between q_r^1 and q_r^2 with a linear function. To calculate the welfare gains of a specific toll f, we use the cumulative benefits of all changes, up to a toll of f. This is

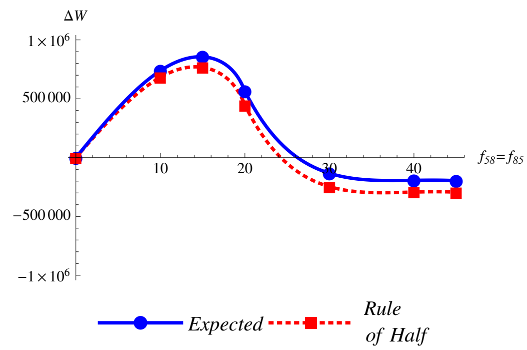

Figure 5: Expected welfare and Rule of Half approximation

a significantly better approximation than simply using the Rule of Half once, where the situation before the policy change would have a zero toll (see Nellthorp and Hyman, 2003). In the latter case, the whole demand function between a zero toll and a toll of f would be approximated with a linear function, which is particularly inappropriate in a discrete choice setting, where demand functions are usually highly convex. Our approach uses a piecewise linear approximation. The step size is, to some extent, arbitrary, but the decision to use just a single step would be too; moreover, increasing the number of steps has a quickly decreasing impact on the results.

Although this piecewise linearization is obviously a simplification, it performs well. Figure 5 shows, for the calibrated NL model, how the approximation compares to the expected welfare when we vary the toll one of the links in the network; the difference between the two is only substantial if the toll is so high that the flow on the tolled link approaches zero, which is generally not a situation of interest. Since the expected welfare is much more difficult to calculate in a probit model, we use the approximation in what follows.

3.4 Parameters

Initial parameters are shown in Table 1. We initially also set all $f_l = 0$, and estimate the parameters in the calibrated versions of the logit models using the flows resulting from the probit model in that situation.

4 Simulation results

4.1 Theoretically equivalent models

We first compare link flows resulting from the three theoretically equivalent models: the MNP, NL with comparable coefficients, and GNL with comparable coefficients. Starting from a situation where no links are tolled, we increase the toll on the link between nodes 5 and 8, in both directions. Fig. 6 shows the

Parameter	Value
c	35
vot	10
V_l	15000
S_f	120
N_i	$\begin{bmatrix} 225000 & 0 & 0 \\ 0 & 200000 & 0 \\ 0 & 0 & 275000 \\ 350000 & 0 & 0 \end{bmatrix}$
δ	50000

Table 1: Initial parameters

resulting flows v, as a function of these tolls $f_{58} = f_{85}$, for four representative links. Flow patterns on the other links either do not vary significantly with this toll, or display similar patterns as those in Fig. 6.

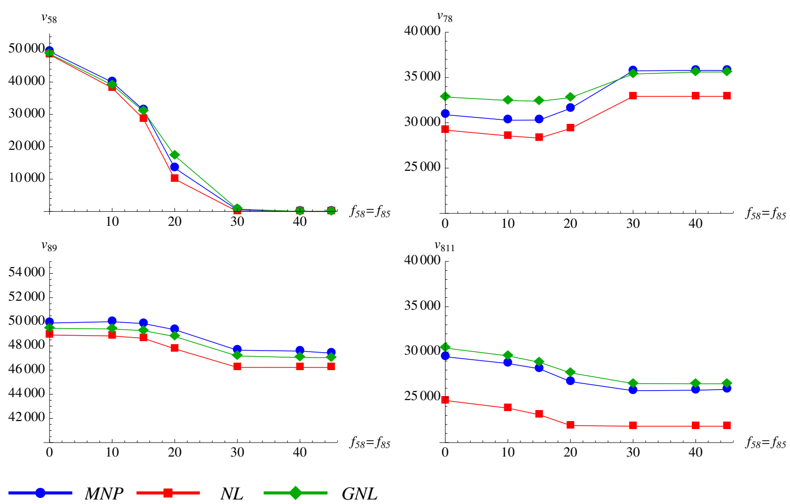

Figure 6: Link flows in theoretically equivalent models

Since these three models imply different substitution patterns, their resulting flows are, of course, not exactly the same. However, the differences between the models are relatively small, even on the link that is tolled (the top left panel in Fig. 6). As we have shown in section 2.2, this is, to a large extent, the result of congestion. Because all links are congested, the utilities users derive from routes always correlate if routes overlap, even in the NL model. The presence or absence of correlation between the stochastic parts of the utilities still has an impact, but not nearly as big as one might expect. Moreover, since most links are used in several markets, OD-level differences may partially cancel each other out on the network level.

Fig. 6 also shows that the GNL model, which allows for some correlation between random terms of routes, often results in flows closer to the Probit model, and it is never significantly further away.

4.2 Calibrated models

Although the difference between the theoretically equivalent models is a good benchmark, models are usually calibrated, which might increase or decrease the differences. Calibration obviously brings the traffic flows closer together in the calibrated point (in our case, where no links are tolled), but this could reduce the predictive power of the models.

Since we are only interested in the differences between models, and not in determining which mode best fits a particular dataset, we examine these effects by calibrating the NL and GNL models to the Probit route flows. Fig. 7 shows the flows on four representative links, resulting from the Probit, calibrated NL and calibrated GNL models.

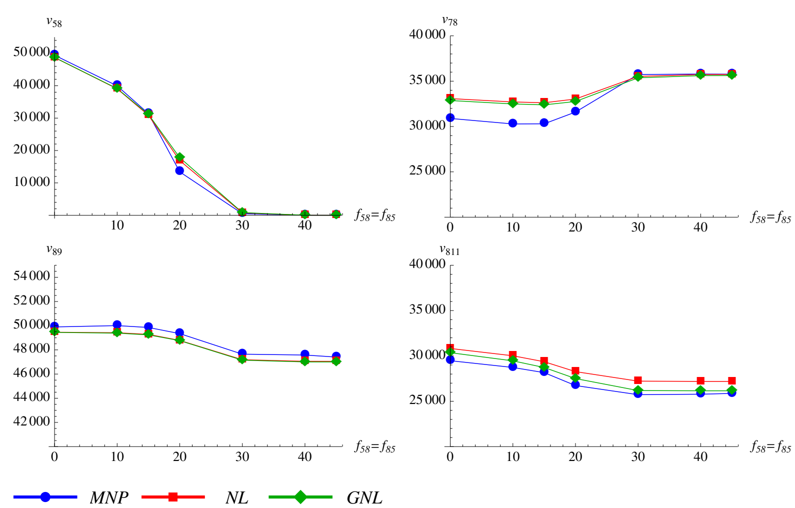

Figure 7: Link flows in calibrated models

As before, the differences between the models are small; in fact, the calibration has decreased them almost everywhere. It seems that, in this particular setting, a good choice of model parameters can compensate for the fundamental differences in the variance-covariance structure of the models. Moreover, the GNL model is now always closer to the Probit flows than the NL model is.

4.3 Overcalibration

The logit models used above are simple, and only have a few parameters that can be calibrated. Since, as we have seen above, calibration can bring the SUE models closer together, it might be advantageous to use a more flexible logit model with alternative-specific constants, as defined in Eq. 4. Fig. 8 shows the same Probit and NL flows as Fig. 7, but in addition, the flows resulting from a calibrated NL model with alternative-specific constants.

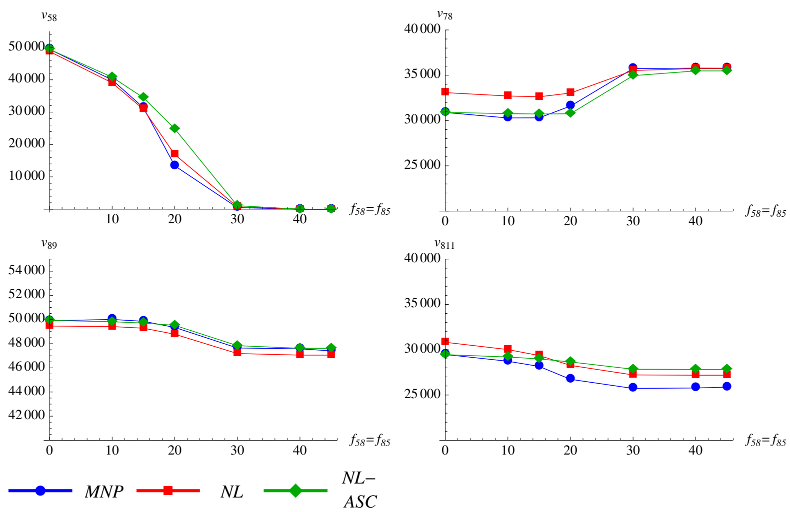

Figure 8: Link flows: NL vs NL-ASC

Naturally, the introduction of more calibration parameters allows the NL-ASC model to be closer to the Probit model when tolls are zero. If we move away from this calibrated state, however, the NL-ASC model is significantly further away from the Probit model than any other model we have examined, at least for the tolled link. This happens because the introduction of ASCs also changes the other calibrated parameters, β and μ. A decrease in μ makes routes less similar, and thus poorer substitutes. This decreases the price elasticity of a demand for each route. Conversely, a decrease in β leads to larger differences in route costs, which increases elasticities. The introduction of ASCs may change both parameters in each direction. In this case, both β and μ are lower, but the effect of the latter parameter is larger, which leads to an decrease in the price elasticity of demand for link 58. Hence, better calibration of the initial equilibrium is not always good for out-of-equilibrium predictions, especially if there is no theoretical foundation (e.g., a correction for the number of left-hand turns in a route) for the addition of more parameters.

4.4 Implications for tolling

Although the differences in link flows between the various SUE models are small, it is difficult to say whether these small differences are important without specifying where these flows are going to be used for. We therefore examine two situations: one in which link 58 is tolled (in both directions) by a private operator, which maximizes its profits, and one in which the toll on this link is set by a social planner, which maximizes social welfare. Other policies, such as capacity extensions, are likely to give similar results, as they also affect the cost of using specific links (although indirectly, through a higher of lower congestibility, rather than directly).

Fig. 9 gives the private operator's profit, as a function of $f_{58} = f_{85}$, for three models: MNP, calibrated NL, and calibrated NL-ASC. Naturally, the differences between the models are only important if they lead to different optimal tolls, or affect the choice for tolling as such. The Probit and NL models result in optimal tolls that are very similar; the profit levels are also comparable. If the Probit model is the correct model, using an NL model to obtain tolls would only reduce profits by 0.3%. Using an NL model with alternative-specific constants, however, would result in a significantly higher toll, and a profit loss of 8.6%, a direct result of the fact that this model uses a lower β and μ, which lowers the demand elasticity on link 58.

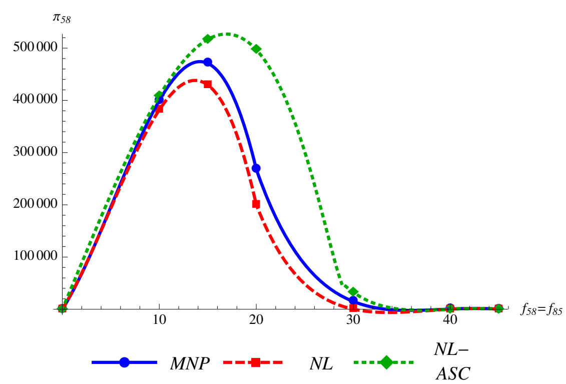

Figure 9: Profits as a function of tolls

Fig. 10 shows the result of a similar exercise, in which a social planner maximizes the social welfare improvement resulting from tolling, as defined in Eq. 18. This figure is very similar to Fig. 9; again, there is only a minimal difference in optimal tolls between the MNP and NL models. If the Probit model is correct, using an NL model to obtain optimal tolls results in only a 1.2% welfare loss. Using the NL-ASC model, however, results in a significantly different toll, and a welfare loss of 5.4% if. This implies that, although simple logit-based SUE models often give very similar results to more realistic probit models, overcalibration can negatively affect policy effectiveness.

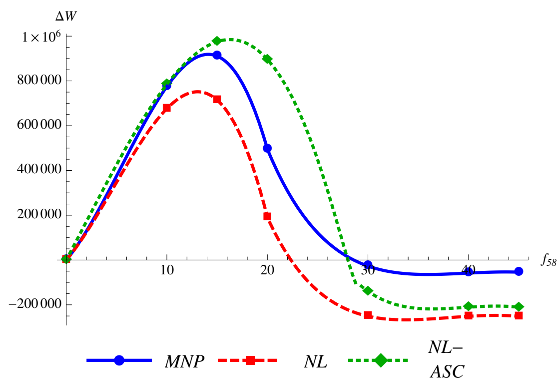

Figure 10: Changes in welfare as a function of tolls

4.5 Sensitivity

In Figs. 9 and 10, optimal tolls are higher in the MNP model than in the NL model, and even higher than in the ASC model. This is, however, not systematic. If, instead of link 58, another link is tolled, the results are different, as illustrated in Fig. 11, where we toll the link between nodes 10 and 11. Although this figure shows profits only, welfare follows a very similar pattern.

Because this link is much less central, it is used by fewer markets. Hence, although tolling still has a local effect, it has a much smaller impact on the other links in the model. As a result, the differences between the models are negligible. Although, here, both logit models result in a higher toll than the MNP model, this difference is so minor that the use of even the toll from an

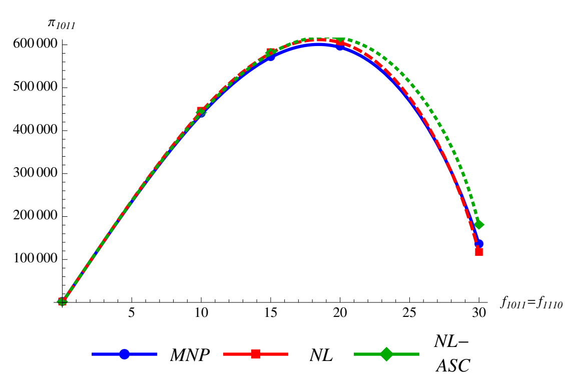

Figure 11: Tolling link 1011

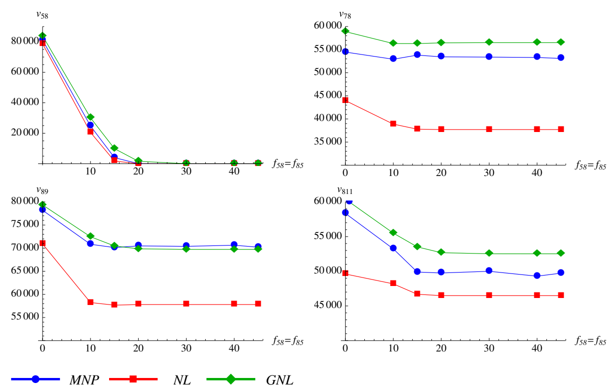

Figure 12: Lower congestibility – flows

NL-ASC model results in less than a 0.04% profit loss if the MNP model is correct.

Figs. 12–14 show link flows and welfare for the various SUE models, in a situation where the congestibility of the links is lower (specifically, $\chi_l = vot \cdot \frac{K_l}{S_f}\left(1 + 0.15\,(v_l/V_l)^2\right) + f_l$, where the square replaces the fourth power of the original BPR function). Note that this does not only decrease the congestibility of all links, but through that, also affects demand levels, and hence all demand elasticities. As Fig. 12 shows, the differences between models are larger than in the base case parameterization. This confirms that it is congestion that brings the models close together. As before, the GNL flows are much closer to the MNP flows than the NL flows. Moreover, as Fig. 13 shows, NL-ASC flows are now even further away from flows in the other models; in particular, the price elasticity of demand is much higher on the tolled link. As a result, the socially optimal NL-ASC toll is much lower, as Fig. 14 shows. If NL-ASC tolls are used while the MNP tolls are correct, this results in a 78% welfare loss, while using NL tolls would only reduce welfare by 28%.

5 Conclusions

We have shown that, in a small but realistic congestible transport network, simple logit SUE models can give results that are very similar to more general probit models. This result stems mostly from the fact that transport networks are congestible, which implies that the systematic utilities that users derive from overlapping routes are correlated, even if the random utilities corresponding to the routes are not. Moreover, in networks, OD-level differences can potentially cancel each other out. The differences in link flows between models are not systematic, i.e., one model does not always result in higher profit-maximizing or socially optimal tolls than another; this depends on the characteristics of the

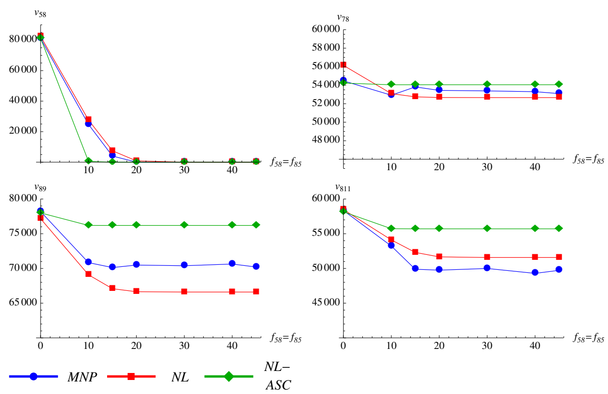

Figure 13: Lower congestibility – overcalibration

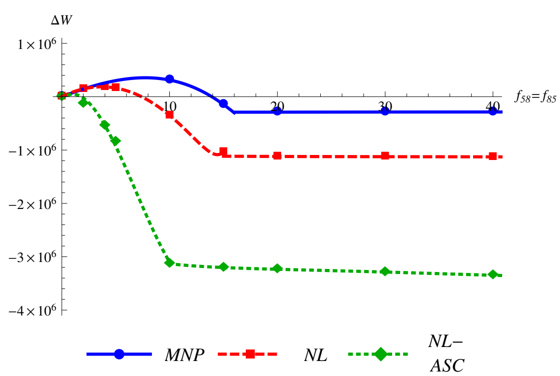

Figure 14: Lower congestibility – welfare

network and the links on which particular policies are enacted.

Since logit models are much less computationally intensive, this indicates that they might, at least in heavily congested settings, be a better choice. As logit models need no simulation, they can lead to more accurate results and allow for studying of more and more complex policy instruments and games (e.g., tax competition, networks with multiple operators, etc.). However, we also find that models can be overcalibrated, especially when parameters are introduced that have no theoretical justification. This is, for instance, the case if alternative-specific constants are introduced in a logit model when there is no theoretical basis for their inclusion. In that case, significant differences between models can arise, and careful evaluation of the various possibilities is necessary.

We have focused on a few representative SUE models, and have disregarded others; in particular, link-based route choice models, as proposed by, e.g. Fosgerau et al. (2013). Although these models are very useful for estimation, their complexity makes them less suitable for many simulation application. Further research has to determine how large different these models are from existing logit and probit models in practical situations.

Naturally, our results were obtained for a very specific situation. We have chosen our network such as to maximize the possible differences between models, and have chosen realistic parameters. It is unlikely that other networks or parameterizations would give significantly different results, and our sensitivity analyses confirm this. The network we have examined, though small, features a large number of overlapping routes between all OD-pairs; moreover, many of these routes are used in equilibrium. We therefore do not expect different results, with stronger contrasts, for larger networks.

References

Bekhor, S., Prashker, J. N., 2002. Stochastic user equilibrium formulation for generalized nested logit model. Transportation Research Record 1752, 84–90.

Ben-Akiva, M., 1974. Structure of passenger travel demand models. Transportation Research Record 526, 26–42.

Connors, R. D., Sumalee, A., Watling, D., 2007. Sensitivity analysis of the variable demand probit stochastic user equilibrium with multiple user classes. Transportation Research Part B 41 (6), 495–615.

Daganzo, C. F., Kusnic, M., 1993. Two properties of the nested logit model. Transportation Science 27, 395–400.

Daganzo, C. F., Sheffi, Y., 1977. On stochastic models of traffic assignment. Transportation Science 11 (3), 253–274.

Daly, A., Bierlaire, M., 2006. A general and operational representation of generalised extreme value models. Transportation Research Part B 40 (4), 285–305.

Dial, R. B., 1971. A probabilistic multipath assignment model which obviates path enumeration. Transportation Research 5, 83–111.

Florian, M., Fox, B., 1976. On the probabilistic origin of Dial's multipath traffic assignment model. Transportation Research 10, 339–341.

Fosgerau, M., Frejinger, E., Karlstrom, A., 2013. A link based network route choice model with unrestricted choice set. MPRA Paper 48707.

Lemp, J. D., Kockelman, K. M., Damien, P., 2010. The continuous cross-nested logit model: formulation and application for departure time choice. Transportation Research Part B 44 (5), 646–661.

Marzano, V., Papola, A., 2008. On the covariance structure of the cross-nested logit model. Transportation Research Part B 42 (2), 83–98.

McFadden, D., 1978. Modelling the choice of residential location. In: Karlquist, A. (Ed.), Spatial interaction theory and residential location. North-Holland, pp. 75–96.

Meng, Q., Lee, D.-H., Cheu, R. L., Yang, H., 2004. Logit-based stochastic user equilibrium problem for entry-exit toll schemes. Journal of Transportation Engineering 130 (6), 805–813.

Meng, Q., Liu, Z., Wang, S., 2012. Optimal distance tolls under congestion pricing and continuously distributed value of time. Transportation Research Part E 48 (5), 937–957.

Nellthorp, J., Hyman, G., 2003. Alternatives to the Rule of Half in matrix-based appraisal. In: Laird, J. J., Nellthorp, J., Mackie, P. J., Ahmed, F. (Eds.), Toolkit for the Economic Evaluation of World Bank transport Projects. World Bank, Washington, DC.

Uchida, K., Sumalee, A., Watling, D., Connors, R., 2007. A study on network design problems for multi-modal networks by probit-based stochastic user equilibrium. Networks and Spatial Economics 7, 213–240.

US Bureau of Public Roads, 1964. Traffic Assignment Manual. US Bureau of Public Roads, Washington, DC.

Wardrop, J. G., 1952. Some theoretical aspects of road traffic research. Proceedings of the Institute of Civil Engineers 1, 325–78.

Wen, C. H., Koppelman, F. S., 2001. The generalized nested logit model. Transportation Research Part B 35 (7), 627–641.

Yang, H., 1999. System optimum, stochastic user equilibrium, and optimal link tolls. Transportation Science 33 (4), 354–360.

Yang, H., Meng, Q., Bell, M. G. H., 2001. Simultaneous estimation of the origin-destination matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. Transportation Science 35 (2), 107–123.