Koster, Paul; Verhoef, Erik T.; Shepherd, Simon; Watling, David

Working Paper

Probabilistic Choice and Congestion Pricing with Heterogeneous Travellers and Price-Sensitive Demand

Tinbergen Institute Discussion Paper, No. 14-078/VIII

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Koster, Paul; Verhoef, Erik T.; Shepherd, Simon; Watling, David (2014) : Probabilistic Choice and Congestion Pricing with Heterogeneous Travellers and Price-Sensitive Demand, Tinbergen Institute Discussion Paper, No. 14-078/VIII, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/107791

TI 2014-078/VIII
Tinbergen Institute Discussion Paper

Probabilistic Choice and Congestion Pricing with Heterogeneous Travellers and Price-Sensitive Demand

Paul Koster[1]
Erik T. Verhoef[1]
Simon Shepherd[2]
David Watling[2]

[1] *VU University Amsterdam, and Tinbergen Institute, the Netherlands;*
[2] *University of Leeds, United Kingdom.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Probabilistic Choice and Congestion Pricing with Heterogeneous Travellers and Price-Sensitive Demand[☆]

Paul Koster[a,b,*], Erik Verhoef[a,b], Simon Shepherd[c], David Watling[c]

[a] *Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands*
[b] *Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam, The Netherlands*
[c] *Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, UK*

Abstract

This paper deals with first-best and second-best congestion pricing of a stylised two-link network with probabilistic route choice of travellers. Travellers may have heterogeneous values of travel times and may differ in their idiosyncratic route preferences. We derive first-best and second-best tolls taking into account how the overall network demand responds to generalized costs including the tolls that are levied. We show that with homogeneous values of times the welfare losses of second-best pricing, of one link only, may be smaller if route choice is probabilistic. Furthermore, we show that with heterogeneous values of times, common second-best tolls and group-differentiated tolls can be very close when route choice is governed by random utility maximisation, leading to low welfare losses from the inability to differentiate tolls.

Keywords: Stochastic User Equilibrium, Second-best Congestion Pricing, Preference Heterogeneity, Scale Heterogeneity, Probabilistic Choice

[☆]Erik Verhoef and Paul Koster gratefully acknowledge the financial support of the Advanced ERC Grant OPTION # 246969. David Watling and Simon Shepherd acknowledge the support of the EPSRC in funding the research as part of the project "Competitive Cities : The network and long-term impacts of fiscal management of transport demand" grant number EP/H021345/1."

[*]Corresponding author. Fax +31 20 5986004, phone +31 20 5982369.
Email addresses: p.r.koster@vu.nl (Paul Koster), e.t.verhoef@vu.nl (Erik Verhoef), S.P.Shepherd@its.leeds.ac.uk (Simon Shepherd), D.P.Watling@its.leeds.ac.uk (David Watling)

1. Introduction

Probabilistic route choice is a common assumption in transportation science. Instead of considering deterministic trade-offs, route utility is then considered as a random variable, depending on a deterministic part and a random part. This paper is concerned with the implications of probabilistic choice for congestion pricing. It presents analytical results for a stylised two-link road network with price-sensitive demand and heterogeneous travellers. These results will enhance our understanding of the economic properties of first-best and second-best congestion pricing in transport networks when route choice is governed by random utility.

Since the early seventies of the previous century, economists have tried to fit in psychological theories of probabilistic choice in economic choice theory. McFadden (1974) and Manski (1977) interpreted the random part of utility as the result of the limited ability of the analyst to observe individuals' preferences, resulting in a choice model where the utility function is random from the perspective of the researcher, but the choices of individuals conditional on their unobserved idiosyncratic preferences are deterministic. Additional assumptions on the distribution of the idiosyncratic preferences then lead to the well-known and widely applied logit model. In the absence of income effects, consumer surplus for this model can then be expressed by the logsum measure, which exhibits returns to variety. This means that, all else equal, more choice alternatives lead to a higher expected consumer surplus (Williams, 1977; Small and Rosen, 1981; de Jong et al., 2007).

Earlier studies have derived congestion tolls for stochastic user equilibrium network models. For example, Yang (1999) and Maher et al. (2005) derive first-best congestion tolls in a road network with price-insensitive demand and homogeneous travellers. They assume that the number of alternative routes and road capacities are fixed, and show that, -as with deterministic user equilibrium (DUE)-, first-best tolls internalise the marginal external costs. Others have included heterogeneous preferences in stochastic user equilibrium (SUE) network models. For example, Yang and Huang (2004) have analysed network models with travellers that have heterogeneous values of times (VOTs). Heterogeneity in the relative size of the random part of utility (scale heterogeneity), has been less frequently analysed. Huang and Li (2007) propose a network model with price-insensitive demand where some drivers are equipped with Advanced Travellers Information Systems (ATIS). In their model the VOT follows a log-normal distribution and group-specific idiosyncratic preferences are assumed. Simulation is used to obtain the benefits of ATIS.

Many earlier studies have addressed user heterogeneity in the presence of second-best congestion pricing in static and dynamic congestion models (see for example: Verhoef et al. (1995); Small and Yan (2001); Verhoef and Small (2004); Mahmassani et al. (2005); Lu et al. (2006); Zhang et al. (2008); Clark et al. (2009); Jiang et al. (2011); van den Berg and Verhoef (2011a,b, 2013)). When second-best pricing is applied and road users have heterogeneous VOTs, travellers will self-select their best route and therefore it is valuable from a welfare perspective to offer toll differentiated roads (Arnott et al., 1992; Verhoef et al., 1995; Small and Yan, 2001; Light, 2009). For example, a welfare optimal strategy may be to have one route with a lower toll and higher congestion, and one with the opposite. The latter

will then be used by travellers with a higher VOT.

The present paper contributes to the literature on congestion pricing by presenting a stylised model for a two route network when route choice is probabilistic and preferences are heterogeneous. We show how the SUE welfare function and congestion tolls are related to the DUE welfare function and congestion tolls, and develop an economic meaningful approach to analyse SUE when choice probabilities of routes are logits. For DUE, welfare is given by the area under the demand curve up to the equilibrium number of travellers minus the total user costs. For logit SUE an additional positive network entropy term is added that captures the total benefits related to individual returns to variety (Fisk, 1980). The welfare function has this structure for the SUE model with homogeneous and heterogeneous preferences. As with the DUE model of Verhoef et al. (1996), the SUE model can then be solved using Lagrangian techniques, leading to analytical marginal expressions for first-best and second-best tolls.

First, we will derive analytical expressions for first-best and second-best congestion tolls with homogeneous VOTs and returns to variety. We show that probabilistic choice does not affect the marginal first-best tolls. These are equal to the standard Pigouvian tolls of the DUE model. However, for asymmetric route costs, the levels of these first-best tolls differ for SUE and DUE, despite the equality of the toll rule, because equilibrium flows are different. Second, we derive a probabilistic second-best toll with homogeneous VOTs and returns to variety that has the deterministic second-best toll of Verhoef et al. (1996) as a limiting case. For second-best pricing, we find that the toll rules of the DUE model and the SUE model diverge. This is because when a toll is levied, the substitution effect to the untolled route depends on the relative size of the random part of utility in the total utility. More deterministic route choices naturally will lead to a stronger response to the toll in terms of diversion onto the tolled route. Because drivers are less responsive to tolls in the SUE model, these second-best tolls will be higher. This reflects that in a deterministic setting, the toll is further below marginal external cost on the tolled route when there is a stronger effect on congestion costs on the untolled route.

Third, we derive first-best and second-best toll rules for the two-route case with heterogeneous preferences. Here, we assume a finite number of groups, with each group having a different valuation of travel time, and a different degree of returns to variety. Our model thus allows for scale heterogeneity, meaning that the returns of variety may differ between groups. Compared to the homogeneous case, a discrete distribution of VOTs and scale parameters increases the empirical plausibility of the model. For example, the VOTs of individuals may be different because of variations in income (see Small (2012) for a review on heterogeneity in VOTs).

We show that if congestion tolls are group-specific, the first-best SUE tolls are isomorphic to the DUE tolls. The marginal expressions do depend on the group-specific valuations of travel times, but are independent of the idiosyncratic part of utility. But again, the SUE toll levels may be different if route costs are asymmetric because equilibrium aggregate usage levels are. We also show that the common first-best toll is equal to the group-specific first-best toll. This is because we assume that travellers only differ in their travel time valuation and therefore the marginal change in external costs due to an additional driver is independent

of the type of traveller. Group-specific first-best tolls are therefore equal across groups for a given link.
For group-specific second-best tolls, we find an analytical solution for the group-specific toll. The level of this second-best toll depends on the valuation of route variety of the different groups. For non-differentiated (common) second-best tolls, we are not able to derive an analytical closed-form expression, and therefore this case will be analysed only numerically. Our numerical results confirm the analytical expressions for first and second-best tolls and give additional insights on the role of probabilistic choice. We assume that there are two groups: one with high value of time, and one with low value of time. With (almost) deterministic route choice there will be a toll differentiated equilibrium, where the high value of time group uses the tolled road and the low value of time group uses the untolled route. However, when the randomness in utility increases, this separation disappears due to idiosyncratic route preferences becoming more important in route choice, and a pooled equilibrium is optimal. The extent to which the roads are differentiated in order to accommodate the needs of distinct groups therefore crucially depends on the size of the random part of utility and the heterogeneity in the value of travel times. More deterministic route choice and larger heterogeneity in travel time valuations increase the likelihood of having a toll differentiated optimal equilibrium.
The paper proceeds as follows. Section 2 introduces the probabilistic route choice model. Section 3 discusses first-best congestion pricing policies with homogeneous and heterogeneous VOTs. Section 4 introduces the model for second-best congestion pricing with homogeneous and heterogeneous VOTs. Section 5 discusses the numerical results and section 6 concludes.

2. The random utility framework

This section introduces the probabilistic route choice model. Consider the following random utility function of choice t of individual n_k belonging to group k, choosing route r from a finite set of R alternative routes:

$$U_{ktr} = G_{kr} + \epsilon_{ktr}. \tag{1}$$

Random utility U_{ktr} therefore depends on a deterministic part G_{kr} and a stochastic idiosyncratic route preference ϵ_{ktr}. The individual n_k therefore makes repeated route choices where the route preferences are random. The cause of these random preference may be that there are unobserved route characteristics that affect route choice. For example, one route may have a petrol station or a route may be closer to the child care. This leads to unobserved preferences for a certain route, which may differ from one choice occasion to another. If idiosyncratic preferences are identically and independently distributed with a Gumbel distribution, McFadden (1974) showed that the probability that individual n_k chooses route x from the set of R routes is given by:

$$P_{kx} = \frac{\exp(V_{kx})}{\sum_{r=1}^{R} \exp(V_{kr})}, \tag{2}$$

where $V_{kr} = \theta_k G_{kr}, \forall r = 1...R$. The scale parameter θ_k governs the relative importance of the unobserved idiosyncratic part of the utility in the total utility. There are two extreme cases to be considered. First, the random part of utility may be very large ($\theta_k \to 0$), resulting in route choices that are independent of the deterministic part of utility. Choice probabilities then converge in the limit to $1/R$. Second, the random part of the utility may become very small ($\theta_k \to \infty$), resulting in a deterministic route choice model for individual n_k. Because the random term of the utility is interpreted as individuals' unobserved preferences, there are returns to variety for individual travellers, meaning that adding an additional alternative will lead to higher maximum expected utility. This becomes more evident if we derive the expected generalised costs for individual n_k, which is given by the negative of the logsum (Williams (1977); Small and Rosen (1981); de Jong et al. (2007)):

$$GC_k = -\frac{1}{\theta_k} \int P_{kr} dV_{kr} = -\frac{1}{\theta_k} \ln \left[\sum_{r=1}^{R} \exp(V_{kr}) \right], \tag{3}$$

where the minus sign is due to the fact that we transform the logsum to expected generalised route costs (including tolls). Furthermore, it is assumed that the marginal utility of income is equal to 1. There is a variety discount, meaning that an increase in the number of routes will always have a positive effect on the generalised user costs as long as $\theta_k < \infty$, and $GC_k < \infty$ resulting in non-zero choice probabilities for all alternatives. This is an appealing feature of the logsum measure, since it implies that having more routes is valued positively by travellers. The stochastic model has the deterministic model as a limiting case. When $\theta_k \to \infty$ in equation 3, the utility differences between the alternatives must become 0 in equilibrium, because otherwise all travellers will travel on one route. Therefore we have $V_{kr} \to V_k$ and $GC_k \to -V_k$, which is equal to the negative part of the systematic utility.

3. First-best congestion pricing, two route case

3.1. Homogeneous preferences

We start our analysis of pricing with first-best congestion pricing in a stylised two-route setting. This model can be viewed as a probabilistic version of the model in Verhoef et al. (1996). We assume that there is only one group, meaning that all travellers are identical in systematic and random utilities. Therefore we can write $n_k \equiv n$, $V_{kr} \equiv V_r$ and $\theta_k \equiv \theta$. Traveller n has a willingness to pay to enter the road network reflected by an inverse demand $D(n)$. A traveller enters the road network if this willingness to pay is higher than the expected generalised costs of equation 3. Our modelling framework therefore assumes a two stage decision, where drivers first make their decision to enter the road on the basis of the expected generalised costs, and then make their route choice on the basis of a individual 'draw' of their idiosyncratic route preference. This is a plausible assumption as long as travellers do not consider the decision to travel for every choice occasion t, but rather make this decision for a series of choices.

If we assume that tolls and congestion costs enter systematic utility in an additive separable

way, the deterministic utilities for routes U and T are given by:

$$V_r = -\theta \left(f_r + c_r(N_r) \right), r \in \{U, T\}, \tag{4}$$

where f_r is the toll on route r, and $c_r(N_r)$ is the travel cost for route r, which is increasing in the route flow N_r. Because the cost coefficient is normalised to -1, systematic utility is expressed in monetary units. For the two route case, equilibrium is implicitly defined by:

$$P_r = \frac{\exp(V_r)}{\exp(V_T) + \exp(V_U)} = \frac{N_r}{N_U + N_T} = \frac{N_r}{N}. \tag{5}$$

This is an implicit condition because probabilities in turn determine $c_r(N_r)$ through congestion effects. These conditions show that equilibrium probabilities can always be expressed by the number of drivers on the two routes. Using equation 3, the expected generalised costs for the two route case is given by:

$$GC_r = -\frac{1}{\theta} \ln \left[\exp(V_T) + \exp(V_U) \right]. \tag{6}$$

These generalised costs both capture the congestion costs, the tolls and the variety discount. This becomes more evident if we rewrite equation 6 using equations 5 and 4:

$$\begin{aligned} GC_r = -\frac{1}{\theta} \ln \left[\exp(V_T) + \exp(V_U) \right] = \ln \left[\frac{\exp(V_r)}{P_r} \right] = -\frac{1}{\theta} \left(V_r - \ln \left[\frac{N_r}{N} \right] \right) = \\ f_r + c_r(N_r) + \frac{1}{\theta} \ln \left[\frac{N_r}{N} \right]. \end{aligned} \tag{7}$$

This is a convenient way of rewriting the expected generalised costs because it becomes immediately clear that the deterministic model results as a limiting case when $\theta \to \infty$ (Akamatsu, 1997). For given route costs $f_r + c_r(N_r)$, the variety discount $\frac{1}{\theta} \ln \left[\frac{N_r}{N} \right]$ always decreases expected generalised costs, because choice probabilities N_r/N are always smaller than 1, leading to a negative ln-term. For given equilibrium route flows and tolls, the stochastic choice model therefore always has lower generalised costs then the deterministic model.

The social surplus S is given by the social benefits (the integral under the inverse demand curve, the toll revenues and the variety discounts) minus the sum of total maximum expected costs. We can therefore multiply equation 7 with the number of travellers on each route to obtain total costs. The toll payments of the travellers are a money transfer to the regulator, and they drop out of the social surplus function (we assume a zero shadow price of public funds). To stay as close as possible to the two-route model of Verhoef et al. (1996), we assumed that inverse demand is given by $D(n)$. In the deterministic model, this reflects that the routes are pure substitutes. Here it shows that the imperfect substitutability is fully captured by the logit model. Social surplus is then given by

$$S = \int_0^N D(n)dn - N_T c_T(N_T) - N_U c_U(N_U) - \frac{1}{\theta} \left(N_T \ln \left[\frac{N_T}{N} \right] + N_U \ln \left[\frac{N_U}{N} \right] \right). \tag{8}$$

The first part of this equation captures the consumer surplus and the deterministic total user costs. The second part is always positive and captures the total variety discounts for given route flows N_T and N_U. For given N_T and N_U, a smaller θ, and hence a higher randomness of route utility, will lead to higher variety discounts. The total variety discounts are equal to the negative of the Shannon entropy multiplied by the number of travellers, and the inverse of the scale parameter (Shannon, 1948). For the multinomial logit model, this relationship between Shannon entropy and variety discounts has long been recognised (Erlander, 1977; Fisk, 1980; Miyagi, 1986).

Entropy is higher when route probabilities are more alike. Therefore, if route probabilities are more similar, variety discounts are higher. This is intuitive: routes add more to the variety discounts if they are used more equally in equilibrium. The total variety discounts can fully be expressed by the choice probabilities and the total number of travellers. Any change in the congestion costs or the toll on a route will only have an effect on the variety discounts via the equilibrium choice probabilities. We consider first-best congestion pricing with a welfare-maximizing regulator, setting a toll on route U and route T. The Lagrangian is given by

$$\begin{aligned}\mathcal{L} = \int_0^N D(n)dn - N_T c_T(N_T) - N_U c_U(N_U) - \frac{1}{\theta}\left(N_T \ln\left[\frac{N_T}{N}\right] + N_U \ln\left[\frac{N_U}{N}\right]\right) \\ +\lambda_T\left(D(N) - f_T - c_T(N_T) - \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right]\right) + \lambda_U\left(D(N) - f_U - c_U(N_U) - \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right]\right).\end{aligned} \tag{9}$$

The constraints govern equilibrium on both routes because travellers keep on entering the road up to the point where the marginal benefits $D(N)$ are equal to the expected generalised costs of Equation 7. This condition holds for both routes and therefore the expected generalised costs of the two routes are also equal in equilibrium. The setup of Equations 8 and 9 separates the overall demand response from the substitution between routes and has a clear advantage over a setup with an outside alternative, because then θ governs both the elasticity of demand and the substitution between routes. Then the limiting case of $\theta \to \infty$ is not easily interpretable, since this results in a deterministic model with perfect elastic demand. In order to find the first-best congestion tolls, the following first-order conditions need to be solved jointly:[1]

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial N_T} = D(N) - c_T(N_T) - N_T c_T'(N_T) - \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right] + \lambda_T\left(D'(N) - c_T'(N_T) - \frac{1}{\theta}\frac{N_T}{N_U N}\right) \\ +\lambda_U\left(D'(N) + \frac{1}{\theta}\frac{1}{N}\right) = 0.\end{aligned} \tag{10}$$

[1]Here we use: $\frac{\partial}{\partial N_T}\left(N_T \ln\left[\frac{N_T}{N}\right] + N_U \ln\left[\frac{N_U}{N}\right]\right) = \ln\left[\frac{N_T}{N}\right]$, $\frac{\partial}{\partial N_U}\left(N_T \ln\left[\frac{N_T}{N}\right] + N_U \ln\left[\frac{N_U}{N}\right]\right) = \ln\left(\frac{N_U}{N}\right)$, $\frac{\partial}{\partial N_T}\ln\left[\frac{N_T}{N}\right] = \frac{N_T}{N_U N}$, $\frac{\partial}{\partial N_U}\ln\left[\frac{N_U}{N}\right] = \frac{N_U}{N_T N}$ and $\frac{\partial}{\partial N_U}\ln\left[\frac{N_T}{N}\right] = \frac{\partial}{\partial N_T}\ln\left[\frac{N_U}{N}\right] = \frac{1}{N}$

$$\frac{\partial \mathcal{L}}{\partial N_U} = D(N) - c_U(N_U) - N_U c'_U(N_U) - \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right] + \lambda_T\left(D'(N) + \frac{1}{\theta}\frac{1}{N}\right) + \lambda_U\left(D'(N) - c'_U(N_U) - \frac{1}{\theta}\frac{N_U}{N_T N}\right) = 0. \tag{11}$$

$$\frac{\partial \mathcal{L}}{\partial f_T} = -\lambda_T = 0. \tag{12}$$

$$\frac{\partial \mathcal{L}}{\partial f_U} = -\lambda_U = 0. \tag{13}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_T} = D(N) - f_T - c_T(N_T) - \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right] = 0. \tag{14}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_U} = D(N) - f_U - c_U(N_U) - \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right] = 0. \tag{15}$$

Equations 12 and 13 show that the Lagrangian multipliers of both routes are 0 in the socially optimal equilibrium. This is intuitive, because these multipliers reflect the marginal change in social surplus for a marginal change in the congestion toll on a route. In equilibrium, this marginal change should be 0, otherwise the tolls would non-optimal by definition. This matches insights from deterministic models (Verhoef, 2002). Substituting Equations 12-15 in Equations 10 and 11 we obtain:

$$\begin{aligned} f_T &= N_T c'_T(N_T), \\ f_U &= N_U c'_U(N_U). \end{aligned} \tag{16}$$

These first-best toll rules are therefore isomorphic to the standard Pigouvian toll rules of the DUE model. Tolls internalise marginal external cost to make people behave according to the social optimum, when acting in their own self-interest. In the probabilistic model we may not fully observe all the individual benefit components, but through the first-best tolls travellers are correctly taking into account all relevant aspects (their own costs and benefits, be it observable to the regulator or not, and the impact on other travellers), so they behave so as to maximise welfare. Even though the toll rules are the same for DUE and SUE, absolute toll levels of the SUE and DUE model may still diverge when route costs are asymmetric. This is because in the DUE model, the route costs will be equal in equilibrium resulting in equal route flows N_T and N_U. The SUE model can have unequal equilibrium route costs and route flows. This asymmetric case is analysed in more detail in section 5.

3.2. Group-differentiated and common first-best tolls with heterogeneous preferences

Next, we proceed with the analysis of first-best congestion pricing with heterogeneous preferences. Assume that there are K distinct groups in the population. The inverse demand for travelling, the valuation of travel time, and the variety discount are assumed to be group-specific. The heterogeneity in travel time valuation enters the model via the deterministic route costs $c_{rk}(N_r)$ for group k, with N_r being the total number of travellers using route r. Scale heterogeneity is captured by having a group-specific scale parameter θ_k. Finally,

heterogeneity in overall demand response is captured by having a group-specific inverse demand $D_k(n_k)$. Let N_{Tk} be the number of travellers of group k that use route T, and N_{Uk} the number of travellers of group k that use route U. The total number of travellers in a group is $N_k = N_{Tk} + N_{Uk}$. Here we have $N_T = \sum_{k=1}^{K} N_{Tk}$, $N_U = \sum_{k=1}^{K} N_{Uk}$ and $N_T + N_U = \sum_{k=1}^{K} N_k = N$. Because the number of groups can be chosen, our model can approximate any continuous distribution of preferences arbitrarily closely. Deterministic route costs are determined by the total number of travellers on each route. To simplify matters, these travel costs are assumed to be equal up to a group-specific multiplicative term, implying that $c_{Tk}(N_T) = \alpha_k c_T(N_T)$ and $c_{Uk}(N_U) = \alpha_k c_U(N_U), \forall k = 1...K$. If $c_r(N_r)$ is interpreted as the travel time on route r, this model can be viewed as a model with travellers having different valuations of travel time α_k. This group-specific valuation of travel time then converts the route travel time $c_r(N_r)$ to monetary units. To save notation we define $\bar{N_T}^{\alpha} = \sum_{k=1}^{K} \alpha_k N_{Tk}$ as the preference weighted number of travellers at route T, and $\bar{N_U}^{\alpha} = \sum_{k=1}^{K} \alpha_k N_{Uk}$ as the preference weighted number of travellers for route U. Each group has an inverse demand function $D_k(n_k)$. Equilibrium choice probabilities for group k for the two routes are then given by $P_{Tk} = \frac{N_{Tk}}{N_k}$ and $P_{Uk} = \frac{N_{Uk}}{N_k}$. The generalised costs for route r for group k are given by the negative of the group-specific logsum, which again can be rewritten as:

$$GC_{rk} = f_{rk} + c_{rk}(N_r) + \frac{1}{\theta_k} \ln \left[\frac{N_{rk}}{N_k} \right]. \tag{17}$$

Deterministic costs are governed by the total number of travellers on route r, whereas the total returns to variety depend on the group-specific route probabilities and the inverse of the group-specific scale parameter θ_k. Returns to variety for group k are thus fully determined by the equilibrium number of travellers for route k at both routes. The interaction of the groups in the network is captured in the deterministic route costs. The total returns to variety for the SUE network model are then given by the sum of the group-specific Shannon entropies, multiplied with the group-specific number of travellers and the inverse of the group-specific scale parameters. The Lagrangian is given by a straightforward extension of Equation 9 to K groups:

$$\begin{aligned} \mathcal{L} = \sum_{k=1}^{K} \int_0^{N_k} D_k(n_k) dn_k - \sum_{k=1}^{K} \alpha_k N_{Tk} c_T(N_T) - \sum_{k=1}^{K} \alpha_k N_{Uk} c_U(N_U) \\ - \sum_{k=1}^{K} \frac{1}{\theta_k} \left(N_{Tk} \ln \left[\frac{N_{Tk}}{N_k} \right] + N_{Uk} \ln \left[\frac{N_{Uk}}{N_k} \right] \right) + \\ \sum_{k=1}^{K} \lambda_{Tk} \left(D_k(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln \left[\frac{N_{Tk}}{N_k} \right] \right) + \\ \sum_{k=1}^{K} \lambda_{Uk} \left(D_k(N_k) - f_{Uk} - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln \left[\frac{N_{Uk}}{N_k} \right] \right). \end{aligned} \tag{18}$$

For all groups, the marginal willingness to pay should be equal to the expected generalised costs in equilibrium, resulting in $2K$ equilibrium constraints and corresponding Lagrangian

multipliers. The system can be solved using the first-order conditions with respect to N_{Tl}, N_{Ul}, the Lagrange multipliers and the tolls. In Appendix A we show that the group-specific first-best tolls with heterogeneous preferences are given by

$$\begin{aligned} f_{Tk} &= \bar{N_T}^{\alpha} c'_T(N_T), \\ f_{Uk} &= \bar{N_U}^{\alpha} c'_U(N_U). \end{aligned} \tag{19}$$

Marginal first-best tolls on the routes are therefore equal to the deterministic case with differentiated tolls. As with first-best tolling with homogeneous preferences, probabilistic choice only has an effect on the tolls via the equilibrium number of travellers on both routes. Furthermore, equation 19 shows that the first-best tolls are equal for all groups. This is because the change in external costs for an additional traveller is assumed to be the same for all groups. For external costs it does not matter to which group the traveller belongs, since travel time losses increase with the same amount independent of the type of traveller. This could be different if different groups had different impacts on travel times, as would be likely with trucks versus passenger cars (see for example Palma et al. (2008) and Parry (2008) on trucks). Differentiated and common tolls will not coincide for second-best congestion pricing, as we will show in the next section.

4. Second-best congestion pricing

4.1. Homogeneous preferences

In many cases first-best pricing is not feasible and often not accepted because travellers do not have the opportunity to travel on a untolled route. Tolling one of the two routes (a form of second-best congestion pricing), may then be a viable alternative. In this section we analyse congestion pricing with probabilistic choice in the presence of an untolled alternative. The SUE model developed in this section is a probabilistic version of the DUE model of Verhoef et al. (1996) which has its roots in the earlier DUE analysis of Marchand (1968) and Lévy-Lambert (1968). For this deterministic model, the substitution between the routes plays an important role for determining the second-best toll. Because in the SUE model this substitution is governed by θ, we expect that second-best tolls will depend on θ too. The systematic route utility for the tolled route is given by $V_T = -\theta\left(f_T + c_T(N_T)\right)$ whereas for the untolled route it is given by $V_U = -\theta c_U(N_U)$. Equilibrium is implicitly defined by equation 5, but the equilibrium conditions are different compared to the first-best case, because no toll is levied on route U. This has an effect on the expected generalised costs (the negative of the logsum), and on overall entry, because overall demand is responsive to generalised cost levels. The expected generalised costs for route T are equivalent to 7 and given by $GC_T = f_T + c_T(N_T) + \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right]$, whereas the generalised route costs of the untolled route are $GC_U = c_U(N_U) + \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right]$. Because tolls are a cost for the travellers and a benefit for the government, the toll revenues will not enter the total social surplus. Therefore the expression for the total social surplus (equation 8) will not change. Because we have price-sensitive demand, travellers enter the road up to the point where the marginal willingness to pay is equal to to the generalised route costs. The Lagrangian is then given by:

$$\mathcal{L} = \int_0^N D(n)dn - N_T c_T(N_T) - N_U c_U(N_U) - \frac{1}{\theta}\left(N_T \ln\left[\frac{N_T}{N}\right] + N_U \ln\left[\frac{N_U}{N}\right]\right) + \lambda_T\left(D(N) - f_T - c_T(N_T) - \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right]\right) + \lambda_U\left(D(N) - c_U(N_U) - \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right]\right). \tag{20}$$

Not surprisingly, this Lagrangian reduces to the DUE model of Verhoef et al. (1996) for $\theta \to \infty$. The second-best toll can be found by solving the following system of first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial N_T} = D(N) - c_T(N_T) - N_T c_T'(N_T) - \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right] + \lambda_T\left(D'(N) - c_T'(N_T) - \frac{1}{\theta}\frac{N_T}{N_U N}\right) + \lambda_U\left(D'(N) + \frac{1}{\theta}\frac{1}{N}\right) = 0. \tag{21}$$

$$\frac{\partial \mathcal{L}}{\partial N_U} = D(N) - c_U(N_U) - N_U c_U'(N_U) - \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right] + \lambda_T\left(D'(N) + \frac{1}{\theta}\frac{1}{N}\right) + \lambda_U\left(D'(N) - c_U'(N_U) - \frac{1}{\theta}\frac{N_U}{N_T N}\right) = 0. \tag{22}$$

$$\frac{\partial \mathcal{L}}{\partial f_T} = -\lambda_T = 0. \tag{23}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_T} = D(N) - f_T - c_T(N_T) - \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right] = 0. \tag{24}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_U} = D(N) - c_U(N_U) - \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right] = 0. \tag{25}$$

Using Equations 21, 23 and 24 we obtain:

$$f_T = N_T c_T'(N_T) - \lambda_U\left(D'(N) + \frac{1}{\theta}\frac{1}{N}\right). \tag{26}$$

Using Equations 25 and 23, we can solve Equation 22 for λ_U:

$$\lambda_U = \frac{N_U c_U'(N_U)}{D'(N) - c_U'(N_U) - \frac{1}{\theta}\frac{N_T}{N_U N}}. \tag{27}$$

This Lagrangian multiplier is non-positive implying that if it was possible to raise a positive toll on route U this would result in a welfare increase because tolls enter the constraints in Equation 20 negatively. This is expected because raising a toll on route U will bring the equilibrium closer to the first-best optimum. Substituting Equation 27 in Equation 26 gives:

$$f_T = N_T c_T'(N_T) - N_U c_U'(N_U)\frac{-D'(N) - \frac{1}{\theta}\frac{1}{N}}{c_U'(N_U) - D'(N) - \frac{1}{\theta}\frac{N_T}{N_U N}}. \tag{28}$$

The first term in Equation 28 is equal to the marginal external costs in the second-best equilibrium. The second term is more complicated and corrects for the marginal costs of congestion caused by substitution to the untolled route when a toll is levied on route T. More precisely: the marginal external costs on route U is multiplied by a fraction, which is for sufficiently high $c'_U(N_U)$ in between 0 and 1. This fraction gives the weight of this correction term, relative to the marginal external costs on route T. This weight depends on the relative size of the random part of utility in the total utility, the slope of the inverse demand, the ratio of the equilibrium number of travellers on both routes, the total number of travellers, and the slope of the congestion cost function of the untolled route. It shows that the second-best toll depends in a complicated way on θ, since θ has a direct positive effect on the numerator and the denominator of the correction term, but also has an indirect effect on Equation 28 via the equilibrium number of travellers. This last effect is the result of additional entry if there are higher variety discounts.

A more detailed look at Equation 28 shows that despite the additional terms due to stochastic route choice, it has a similar analytical structure as the toll rule for deterministic route choice and can be written as $f_T = MEC_T + MEC_U \frac{\partial N_U}{\partial N_T}$, where MEC_r is the marginal external cost on route r. The marginal external costs on route U are weighted with a term $\frac{\partial N_U}{\partial N_T}$. More specifically, Equation 20 shows that the term $D'(N) + \frac{1}{\theta}\frac{1}{N}$ is the change in the constraint for route U due to a marginal change in N_T, whereas $-c'_U(N_U) + D'(N) + \frac{1}{\theta}\frac{N_T}{N_U N}$ is the change in the constraint for route U due to a marginal change in N_U. The ratio therefore gives $\frac{\partial N_U}{\partial N_T}$. As opposed to the first-best toll rules of Equation 16, the toll rules of the DUE model of Verhoef et al. (1996) and our SUE model differ even for the case with symmetric route costs. Several limiting cases can be considered.

First, the second-best toll rule of the deterministic model is a limiting case of the stochastic model when its random component vanishes:

$$\lim_{\theta \to \infty} f_T = N_T c'_T(N_T) - N_U c'_U(N_U) \frac{-D'(N)}{c'_U(N_U) - D'(N)}. \tag{29}$$

This toll is isomorphic to the toll rule for the DUE model developed by Verhoef et al. (1996). The SUE model therefore generalises the DUE model because it has the DUE model as a special case.

Second, for perfectly overall inelastic demand, $D'(N) \to -\infty$, and the toll rule becomes equal to the difference in marginal external costs on the two routes:

$$\lim_{D'(N) \to -\infty} f_T = N_T c'_T(N_T) - N_U c'_U(N_U). \tag{30}$$

This toll rule is isomorphic to the toll rule of the DUE model with price-insensitive demand of Verhoef et al. (1996). Because there is no effect of tolling on the overall demand, the regulator only seeks to find the optimal route split. This produces the first-best outcome, so in itself it is no surprise that as with first-best tolls, the toll rules for the DUE and the SUE model become identical again. The level of the toll in Equation 30 may well be different for DUE and SUE for asymmetric route costs, because θ has an effect on the optimal route

split. Furthermore, Equation 30 may be negative if in equilibrium the marginal external costs on route U are higher than the marginal external costs on route T. This means that travellers on route T would receive a subsidy instead of paying a toll.
Third, with perfectly elastic overall demand the toll rule becomes

$$\lim_{D'(N)\to -0} f_T = N_T c'_T(N_T) - N_U c'_U(N_U) \frac{-\frac{1}{\theta}\frac{1}{N}}{c'_U(N_U) - \frac{1}{\theta}\frac{N_T}{N_U N}}. \tag{31}$$

This is clearly different from the corresponding toll 29, where the second term vanishes as $-D'(N)$ becomes 0. For perfectly elastic overall demand the marginal toll rule depends on θ, because the substitution between the routes depends on the variety discount, whereas the use of route U would be fully independent of f_T with deterministic route choice. The reason is that for deterministic route costs and perfectly elastic demand, the toll on route T cannot affect the use of route U, so there is no benefit from taking route U into account in the toll rule. In the stochastic model, there remains an effect of the toll on the use of route U, and this is accounted for in the toll rule.
Fourth, if route U is uncongested, $c'_U(N_U) \to 0$, and the toll rule 28 reduces to:

$$\lim_{c'_U(N_U)\to 0} f_T = N_T c'_T(N_T). \tag{32}$$

which is again isomorphic to the toll rule in the deterministic model. The absence of congestion on route U then means that this route is optimally priced when it is not tolled. The regulator may therefore ignore route U, and needs only to consider the unconstrained optimal regulation of route T.
Finally, it turns out that both the deterministic model and the stochastic model are part of a broader class of models with user benefits $B(N_T, N_U)$. The first derivative of this benefit function is the inverse demand and should be equal to the congestion costs plus the toll in equilibrium. The second-best congestion toll is then given by (Small and Verhoef (2007), equation 4.47):

$$f_T = N_T c'_T(N_T) - N_U c'_U(N_U) \frac{-\frac{\partial^2 B}{\partial N_U \partial N_T}}{c'_U(N_U) - \frac{\partial^2 B}{\partial N_U^2}}. \tag{33}$$

For deterministic route choice we have $B(N_T, N_U) = \int_0^N D(n)dn$, whereas for stochastic route choice we observe from Equation 8 that there are additional returns to variety resulting in: $B(N_T, N_U) = \int_0^N D(n)dn - \frac{1}{\theta}\left(N_T \ln\left[\frac{N_T}{N}\right] + N_U \ln\left[\frac{N_U}{N}\right]\right)$. If we substitute the second order derivatives of this benefit function in Equation 33 we arrive at Equation 28.

4.2. Group-specific second-best tolling with heterogeneous preferences

This section generalises the SUE model of the previous section by deriving group-specific second-best congestion tolls with heterogeneous travellers. We use a similar setup as in section 3.2 where K distinct groups have different preferences for congestion costs, returns

to variety and the inverse demand curve. The Lagrangian is given by:

$$\begin{aligned}\mathcal{L} = \sum_{k=1}^{K} \int_0^{N_k} D_k(n_k) dn_k - \sum_{k=1}^{K} \alpha_k N_{Tk} c_T(N_T) - \sum_{k=1}^{K} \alpha_k N_{Uk} c_U(N_U) \\ - \sum_{k=1}^{K} \frac{1}{\theta_k} \left(N_{Tk} \ln \left[\frac{N_{Tk}}{N_k} \right] + N_{Uk} \ln \left[\frac{N_{Uk}}{N_k} \right] \right) + \\ \sum_{k=1}^{K} \lambda_{Tk} \left(D'_k(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln \left[\frac{N_{Tk}}{N_k} \right] \right) + \\ \sum_{k=1}^{K} \lambda_{Uk} \left(D'_k(N_k) - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln \left[\frac{N_{Uk}}{N_k} \right] \right)\end{aligned} \tag{34}$$

In Appendix B we first show that the group-specific Lagrangian multipliers of route U are non-positive, implying that the possibility to raise a positive toll on route U for any of the groups would increase welfare. Furthermore, we show that the second-best group-specific toll for group k is given by:

$$f_{Tk} = \bar{N_T}^{\alpha} c'_T(N_T) - \bar{N_U}^{\alpha} c'_U(N_U) \frac{-D'(N_k) - \frac{1}{\theta_k}\frac{1}{N_k}}{\alpha_k c'_U(N_U) - D'_k(N_k) + \frac{1}{\theta_k}\frac{N_{Tk}}{N_{Uk}N} + \phi_k}, \tag{35}$$

where

$$\phi_k = \sum_{\substack{l=1 \\ l \neq k}}^{K} \alpha_l c'_U(N_U) \frac{D'_k(N_k) - \frac{1}{\theta_k}\frac{N_{Tk}}{N_{Uk}N_k}}{D'_l(N_l) - \frac{1}{\theta_l}\frac{N_{Tl}}{N_{Ul}N_l}} > 0. \tag{36}$$

The first part in equation 35 is related to the external costs on the tolled route and is isomorphic to the first-best toll with heterogeneous preferences (19). The second part in Equation 35 takes into account the substitution effect to the other route which is different for each group. Several limiting cases can be considered. First, when there is only one group, $\phi_k \to 0$, and Equation 35 reduces to Equation 28. Second, the DUE group-specific toll is a special case for which $\theta_k \to \infty, \forall k = 1...K$. This results in:

$$f_{Tk} = \bar{N_T}^{\alpha} c'_T(N_T) - \bar{N_U}^{\alpha} c'_U(N_U) \frac{-D'(N_k)}{\alpha_k c'_U(N_U) - D'_k(N_k) - c'_U(N_U) \sum_{\substack{l=1 \\ l \neq k}}^{K} \alpha_k \frac{D'_k(N_k)}{D'_l(N_l)}}. \tag{37}$$

If the slopes of the demand curves of all groups are equal we have $D'_l(N_l) \equiv D'_k(N_k) \equiv D'$ this reduces to:

$$f_{Tk} = \bar{N_T}^{\alpha} c'_T(N_T) - \bar{N_U}^{\alpha} c'_U(N_U) \frac{-D'}{c'_U(N_U) \sum_{l=1}^{K} \alpha_l - D'}. \tag{38}$$

This implies that the DUE model with equal slopes of the demand curves lead to common second-best tolls for all groups because $c'_U(N_U) \sum_{l=1}^{K} \alpha_k$ has the same value for all groups.

For the SUE model with equal slopes of the inverse demand curves, tolls are still differentiated between groups, because the substitution effect to the untolled route does depend on the equilibrium number of travellers of each group on each route and the group-specific scale parameters θ_k.
We were not able to derive analytical solutions for the common second-best toll case (undifferentiated between groups). The welfare for common second-best tolls will be lower than for the group-specific second-best tolls, because the inability to differentiate the tolls between user groups imposes an additional constraint. The Lagrangian problem is equivalent to 34 with $f_{Tk} \equiv f_T$. However, the next section will include numerical results for this case.

5. Numerical results

5.1. Introduction and calibration

Our numerical results build on the DUE model of Verhoef et al. (1996), who assumed linear inverse demand and linear cost functions. We shall use the DUE case as a benchmark case to which we judge the implications of moving from a DUE to SUE framework, considering sensitivity of the results and toll rules to variations in the variety discounts. The DUE model of Verhoef et al. (1996) assumes a linear inverse demand:

$$D(N) = \delta_1 - \delta_2(N_T + N_U), \tag{39}$$

and linear congestion costs for route r defined as:

$$c_r(N_r) = \kappa_r + \beta_r N_r, \tag{40}$$

where the base case assumed parameter values $\delta_1 = 50; \delta_2 = 0.02; \kappa_1 = \kappa_2 = 20$ and $\beta_1 = \beta_2 = 0.02$ for both routes. This implies that both routes are assumed to be identical in the base case resulting in non-intervention equilibrium flows of $N_T = N_U = 750$. Substituting these values in Equation 40 gives equilibrium average costs of $20 + 0.02 \times 750 = 35$ and marginal social costs of $20 + 0.04 \times 750 = 50$. Applying optimal first-best tolling results in average costs of 30 and marginal social costs of 40, whereas the toll is given by $40 - 30 = 10$. The socially optimal number of travellers is given by 500 for both routes.
In what follows we consider the various toll rules and welfare implications under the homogeneous cases with the variety discount captured by θ, heterogeneity in values of time between groups (α_k), and heterogeneity in preferences between groups (θ_k) for the symmetric case. In order to make a comparison between such cases, we calibrate the initial link flows at the non-intervention equilibrium to be equal to the UE non-intervention flows, and adjust the inverse demand function(s) accordingly. We also impose a constraint on the flow weighted average value of time so that this is equal to the value used in the homogeneous case. This calibration of the model to observed flows and average value of time not only ensures that the flows are consistent at the non-intervention case, but also that the initial aggregate welfare levels are maintained across models. We are though aware that symmetric examples may hide some impacts, in particular the change in route flows even in the non-intervention case θ is adjusted. For this reason we also develop an asymmetric example (again based on

Verhoef et al. (1996)), where we adjust the systematic route utility with the introduction of a specific constant on one route, to maintain the equilibrium route flows. This asymmetric example is used to illustrate the fact that θ can affect the tolls and resulting flows even in the first-best homogeneous case. Calibration of the asymmetric case is directed to Appendix C.

When adapting the model to a SUE with heterogeneous preferences, we purposely calibrate the demand function and initial group-specific demands such that the DUE non-intervention route flows (and demands) are retrieved as we move between cases. As we are dealing with linear inverse demand and cost functions these adjustments are rather straightforward and additive in nature. Users are willing to travel until their willingness to pay is equal to the logsum and therefore we have to shift the inverse demand curve by an amount equal to the difference between the DUE average costs (35 at the no toll equilibrium in our example) and the no toll stochastic average costs. Hence the new inverse demand function can be written as:

$$D(N) = \delta_1 + \nu + \delta_2(N_T + N_U), \tag{41}$$

where the shift term $\nu < 0$ is equal to the logsum evaluated at the DUE demand level (1500 in this case) minus the DUE systematic average costs. Because the variety discount decreases average costs, the inverse demand curve needs to be shifted downwards to maintain the same equilibrium non-intervention flows.

When we move to the case where groups are characterised by different values of time, we impose the condition that the flow-weighted average value of time remains equal 1, which is the assumed value with homogeneous travellers, reflecting that in the base calibration no explicit distinction between a travel time function and the average user cost function is made. In addition we maintain the initial demand, so that we have conditions as follows:

$$\sum_{k=1}^{K} \alpha_k \frac{N_k^0}{N^0} = 1; N^0 = \sum_{k=1}^{K} N_k^0, \tag{42}$$

where superscript N_k^0 refers to the non-intervention values of total group-specific demand and N^0 is total non-intervention demand. For the case with two groups we have:

$$N_1^0 = N^0 \frac{1-\alpha_2}{\alpha_1 - \alpha_2}; N_2^0 = N^0 - N_1^0. \tag{43}$$

For the DUE case with two groups, we have to adjust the group-specific demand functions to account for the change in value of time in order to maintain the initial systematic average costs as for the homogeneous case. This is achieved by adjusting the intercept and slope of the group-specific inverse demand curves as follows:

$$D_k(N_k) = \delta_1 \alpha_k - (\delta_1 - \nu_k)\alpha_k \frac{N^0}{N_k^0}. \tag{44}$$

This procedure is best demonstrated by an example. Let the values of time be $\alpha_1 = 0.8$ and $\alpha_2 = 1.3$. This gives initial flows of $N_1^0 = 1500\frac{1-1.3}{0.8-1.3} = 900$ and $N_2^0 = 1500 - 900$ and

intercepts of $\delta_1\alpha_1 = 50 \times 0.8 = 40$ and $\delta_1\alpha_2 = 50 \times 1.3 = 65$. Assuming $\nu_1 = \nu_2 = 0$ for the deterministic model, the slopes of the inverse demand curves for both groups are given by $-\frac{\alpha_1(\delta_1-\nu_1)}{N_1^0} = -\frac{15\times0.8}{900} = -\frac{1}{75}$, and $-\frac{\alpha_2(\delta_1-\nu_1)}{N_2^0} = -\frac{15\times1.3}{600} = -\frac{13}{400}$ respectively. Substituting these values back in the inverse demand function gives equilibrium non-intervention average costs of $40 - \frac{1}{75} \times 900 = 28$ and $65 - \frac{13}{400} \times 600 = 45\frac{1}{2}$. Note that their ratio corresponds to $\frac{1.3}{0.8}$, so that the equilibrium travel time will be equal for both groups.
For the no toll DUE equilibrium, group-specific and total welfare can be calculated using these equilibrium average user costs. Because of linear inverse demand, welfare is given by the triangular area above the average cost curve. For group k the welfare is denoted by $W_{k,0}$. The group-specific welfare is given by $W_{1,0} = \frac{1}{2} \times (\delta_1\alpha_1 - 28) \times N_1^0 = \frac{1}{2} \times 12 \times 900 = 5400$ and $W_{2,0} = \frac{1}{2} \times (\delta_1\alpha_2 - 45\frac{1}{2}) \times N_2^0 = \frac{1}{2} \times 19\frac{1}{2} \times 600 = 5850$. Total initial welfare is then given by $\hat{W}_0 = W_{1,0} + W_{2,0} = 5400 + 5850 = 11250$, which is the required welfare of the homogeneous DUE case of Verhoef et al. (1996). This procedure for DUE can be extended to SUE by shifting the group-specific demand curves with the group-specific correction terms ν_1 and ν_2 in such a way that non-intervention average user costs and welfare levels are maintained.

5.2. First-best congestion pricing with homogeneous travellers and asymmetric route costs

We start with first-best tolling in the homogeneous value of time SUE model with symmetric route costs. When route costs are symmetric, the first-best tolls from equation 16 were found to give the same optimal tolls and flows as for the DUE case: tolls of 10 and flows of 500 on each link for all chosen values of θ. However, this is a special case, because with asymmetric route costs θ has an effect on the toll via its impact on the equilibrium numbers of travellers on both routes, which in turn directly enters the first-best toll rules. To illustrate this, assume that the routes have different free-flow travel times κ_r, with $\kappa_T = 20$ and $\kappa_U = 10$. This changes the non-intervention flows on both routes and from Verhoef et al. (1996) these are $N_T^0 = 625$ and $N_U^0 = 1125$ for the DUE case. If we now introduce a preference for variety in the SUE case then these non-intervention flows would be different. As discussed in the previous section, we seek to maintain the observed route flows in the non-intervention case. Therefore we introduce a route specific constant for route U, which represents a route-specific preference not related to travel time and toll. In Appendix C we show that the required calibrated constant for any chosen value of θ amounts to:

$$ASC_U = \frac{1}{\theta} \ln \left[\frac{N_T^0}{N_U^0} \right], \tag{45}$$

where the flows are from the DUE non-intervention case. This results in a negative constant being added to the shorter route U, which attracts more users to compensate for the returns to variety term. Table 1 shows the results for different values of θ. The flows for the $\theta = 10$ case are close to the UE solution of $458\frac{1}{3}$ and $708\frac{1}{3}$ of Verhoef et al. (1996). This implies the corresponding tolls are also close to the first-best tolls for DUE. As opposed to the symmetric case, an increase in returns to variety changes equilibrium route flows which in turn impact the optimal toll level. The equilibrium flows on the shorter route U, increase with decreasing θ as the alternative specific constant ASC_U increases with decreasing θ. The

Table 1: Tolls, flows and welfare for FB tolling with asymmetric route costs.

θ	f_T	f_U	N_T	N_U	N	$\hat{W}_{FB}$
10	9.2	14.2	458.1	708.5	1166.6	21041.4
1	9.1	14.2	456.2	709.9	1166.1	21038.6
0.5	9.2	14.2	454.3	711.2	1165.5	21035.8
0.1	8.9	14.4	443.1	719.0	1162.1	21020.0

flows on the longer route decrease when θ decreases. The tolls follow the flows as implied by the marginal first-best toll rules of 16. Compared to DUE, the overall demand and the corresponding welfare slightly reduce as returns to variety increase.

5.3. First-best congestion tolling with heterogeneous travellers

For first-best tolling with heterogeneous values of time and stochastic route choice we use group-specific values of time of $\alpha_1 = 0.8$ and $\alpha_2 = 1.3$. As described in section 5.1 we maintain the average value of time and welfare at the no toll equilibrium using initial flows of 900 and 600 respectively. It was confirmed numerically that the first-best tolls from equations 19 were optimal and the resulting tolls, flows and welfare are shown in Table 2. For low values of θ, there was only one solution with common first-best tolls which are higher than in the homogeneous case. As with the homogeneous case, the first-best toll solutions are independent of θ due to the symmetry in average route costs for low values of θ. In DUE, almost all the low value of time group were priced off route T, with around 500 remaining in the higher value of time group. The total welfare is larger than for the homogeneous case despite the total demand being only 960.8 users. This is due to the new average value of time being 1.06 at the first-best equilibrium, because more high value of time users enter the road. For $\theta = 10$ the model is close to the UE case. There we obtain several solutions

Table 2: Tolls, flows and welfare for FB tolling with heterogeneous value of time

θ	f_{T1}	f_{T2}	f_{U1}	f_{U2}	N_{T1}	N_{T2}	N_{U1}	N_{U2}	$\hat{W}_{FB1}$	$\hat{W}_{FB2}$	$\hat{W}_{FB}$
10*	10.2	10.2	10.2	10.2	229.4	251.0	229.4	251.0	6081.7	9212.5	15294.1
10	11.4	11.4	9.1	9.1	0.3	437.5	471.8	60.4	5791.9	9558.4	15350.3
1	10.2	10.2	10.2	10.2	229.4	251.0	229.4	251.0	6081.7	9212.5	15294.1
0.5	10.2	10.2	10.2	10.2	229.4	251.0	229.4	251.0	6081.7	9212.5	15294.1

Note: the * optimum is a local minimum that satisfies the marginal toll expression. For $\theta < 0.5$ the tolls, flows and welfare levels are similar.

that satisfy the first-best toll expressions of Equation 19 with heterogeneity.[2] The common toll solution in the first row of Table 2, turns out to be local minimum that satisfies the

[2]Due to symmetry we leave out two solutions because it is always possible to swap the route flows. we leave this out

analytical toll expressions, but not the second order condition for a welfare maximum. For different values of θ there exist two possible solutions with group flows tending towards a differentiated toll equilibrium with a high number of high value of time users on the link with a high toll, and a high number of low value of time users at the other route. This result occurs also in the heterogeneous DUE case of Arnott et al. (1992) and Verhoef and Small (2004), but eventually disappears in the SUE case when θ becomes sufficiently low. Route preferences of individuals then become so stochastic that toll differentiation is not beneficial in welfare terms. The toll differentiated equilibrium then dissipates due to the lower sensitivity to deterministic costs, and the toll differentiated solution is "smoothed" out, by randomness in route choice. Toll differentiated equilibria in our model become more likely for two reasons. First, the likelihood of these equilibria to occur increases when the route choice model becomes more deterministic, so for higher values of θ. Second, when values of times are more heterogeneous, a toll differentiated equilibrium is more likely to occur because it is more beneficial to offer differentiated roads (see Small and Yan (2001) and Verhoef and Small (2004)). For our model this implies that when we increase the difference between α_1 and α_2, while keeping the average value of time constant, a toll differentiated equilibrium starts occurring for lower values of θ.

5.4. Second-best tolling, homogeneous values of time

When first-best tolling is not feasible, second-best tolling with a toll on route T might be a realistic and viable alternative. The numerical results in this section confirm the optimal toll rule of Equation 28. Figure 1 shows how the welfare improvement varies with the second-best toll on route T for different values of θ. Quite intuitively, as θ increases, the solution of the second-best toll tends towards the UE solution of 5.45 of Verhoef et al. (1996). The general tendency in Figure 1 is that the optimal second-best toll increases when route preferences become more stochastic. The reason is that travellers are less responsive to the deterministic part of utility, and therefore the behavioural response to the toll onto route T is less strong. This allows the regulator to more fully internalize the marginal external costs on route T, without spillovers upon route U mitigating the gains, and therefore SB tolls can be higher when the variety discount increases. Randomness in utility thus mitigates the central inefficiency under second-best tolling. Table 3 shows the optimal second-best tolls, route flows, and relative efficiencies $\omega = \frac{\hat{W}_{SB}-\hat{W}_0}{\hat{W}_{FB}-\hat{W}_0}$. The latter is defined as the welfare gain due to second-best regulation divided by the welfare gain due to first-best regulation, where non-intervention is the benchmark (see Verhoef et al. (1995)). As expected, the optimal toll with SUE increases when θ decreases, because road users become less sensitive to the deterministic part of average costs. The total demand decreases as users become less cost sensitive because the toll increases. The relative efficiency increases with decreasing θ, as the induced welfare losses on route U become smaller. This implies that the welfare losses due to second-best congestion pricing are lower when a stochastic route choice model is used.

5.5. Second-best tolling, heterogeneous returns to variety

Next, we allow for heterogeneity in the scale of utility, and thus in the importance of unobserved preferences, between two groups of equal size and with their value of time set to

Figure 1: Welfare gains $\hat{W}_{SB} - \hat{W}_0$ against toll level f_T for second-best tolling with homogeneous values of times for varying values of θ.

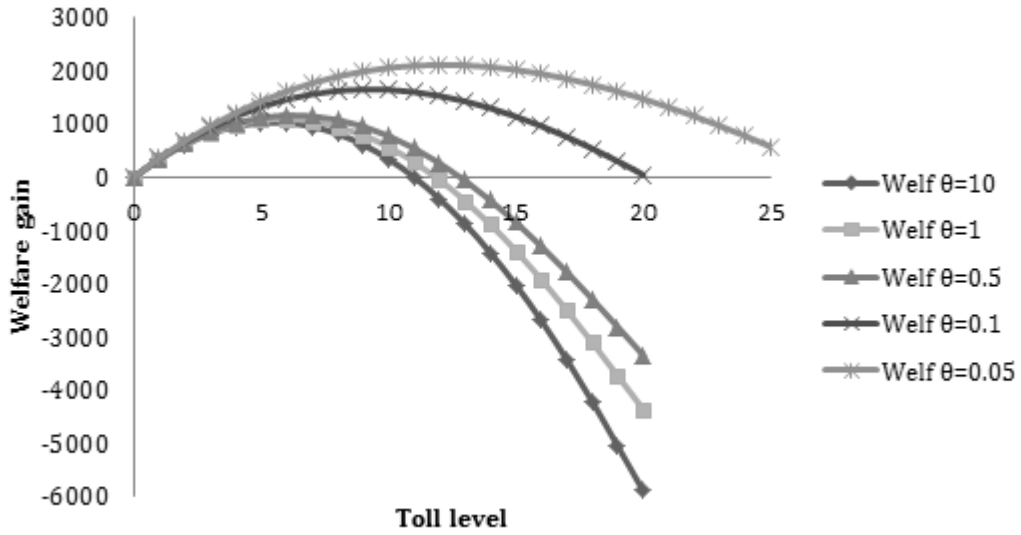

Table 3: Tolls, route flows and welfare for second-best tolling with symmetric route costs and homogeneous values of times.

θ	f_T	N_T	N_U	N	$\hat{W}_{SB} - \hat{W}_0$	ω
UE	5.45	545.00	818.00	1363.00	1023.7	0.273
10	5.50	544.95	817.74	1362.69	1029.9	0.275
1	5.87	540.63	813.67	1354.30	1093.9	0.292
0.5	6.38	532.72	809.92	1342.63	1163.0	0.310
0.1	9.36	507.89	768.73	1276.62	1653.5	0.441
0.05	11.73	501.07	722.20	1223.27	2111.3	0.563

1. We will compare the results with the homogeneous second-best toll case of the previous section. The example follows the symmetric case, where θ_2 is varied for group 2 holding θ_1 constant at 10. This allows us to study the effect of heterogeneous returns to variety. Demand is calibrated as discussed in section 5.1. As we deal with symmetric route costs, the initial group flows are split equally between the links in the no toll case. Table 4 shows the results for the second-best group-specific tolls from Equation 35, which were confirmed numerically to give the optimal tolls. The first row of Table 4 shows the result for homogeneous returns to variety and has the same toll as the toll in the second row of Table 3. Table 4 shows the total welfare so that we can examine the differences between groups. The base welfare is 11250 so the total welfare gain corresponds to the reported value in Table 3. An increase in θ_2 results in a decrease in the optimal toll for group 1 and an increase in the optimal toll for group 2. Since group 2 has higher returns to variety, this group is less responsive to the toll, a higher toll can be charged, and more of the congestion externalities of route T can be internalized. The tolls for group 2 are consistently higher than those

Table 4: Tolls, flows and welfare for SB tolling with heterogeneous returns to variety

θ_1	θ_2	f_{T1}	f_{T2}	N_{T1}	N_{T2}	N_{U1}	N_{U2}	N_1	N_2	$\hat{W}_{SB1}$	$\hat{W}_{SB2}$	$\hat{W}_{SB}$
10	10	5.50	5.50	272.5	272.5	408.9	408.9	545.0	817.7	6140.0	6140.0	12279.9
10	1	5.44	5.98	300.5	243.4	384.5	429.9	543.9	814.4	6325.3	5989.5	12314.7
10	0.5	5.37	6.51	304.1	238.9	384.8	425.8	543.0	810.6	6380.1	5972.3	12352.4
10	0.1	4.94	10.07	311.4	226.0	404.8	378.4	537.4	783.2	6668.4	5928.5	12596.9

for the same value of θ in the homogeneous case in Table 3, because the said mechanism prevails whenever $\theta_2 < 10$. Consistent with the toll levels, the equilibrium flows on the tolled link for group 1(2) increases (decreases) as returns to variety increase for group 2. The group-specific welfare levels show that group 1 benefits from the increase in returns to variety of group 2. The result that the toll is higher for the second group as θ decreases can be inferred from the toll rule of Equation 35, where the second term representing the group-specific route substitution and demand effects. This term decreases with decreasing own returns to variety and increases for returns to variety of other groups.

5.6. Second-best congestion tolling, heterogeneous values of time, group-specific tolls

This section present the results for group-specific second-best tolls with heterogeneous values of times of $\alpha_1 = 0.8$ and $\alpha_1 = 1.3$. Table 5 presents the numerical results for different values of the scale parameter θ. The SB toll for both groups first decreases in θ and then increases for lower values of θ. A lower θ means that more low value of time travellers and fewer higher value of time travellers will use the tolled route. This leads to a downward adjustment of the first direct term in Equation 35 which captures the marginal external costs of route T. But a further decrease in θ also means that spillovers become less and less important, and that means that the second term in Equation 35 decreases. This effect raises the value of the second-best toll. The U-shaped pattern in Table 5 is the combined result of these two opposing forces.

Table 5: Tolls, route flows and welfare for differentiated second-best tolls with heterogeneous values of time

θ	f_{T1}	f_{T2}	N_{T1}	N_{T2}	N_{U1}	N_{U2}	N_1	N_2	$\hat{W}_{SB1}$	$\hat{W}_{SB2}$	$\hat{W}_{SB}$
10	8.28	8.55	0.00	496.2	793.1	41.7	793.1	537.9	4193.2	8947.0	13140.2
1	7.37	7.72	89.1	404.2	675.2	152.3	764.3	556.5	4550.7	8151.3	12702.0
0.5	7.31	7.25	175.5	331.2	583.6	230.1	759.1	561.3	5125.3	7521.3	12646.5
0.25	8.08	7.33	219.7	288.8	521.1	276.2	740.8	564.9	5433.0	7303.4	12736.5
0.1	11.0	8.45	230.1	263.0	450.2	307.6	680.3	570.6	5604.9	7512.7	13117.6
0.05	14.7	10.4	218.8	254.8	381.2	319.3	600.0	574.0	5624.6	8008.2	13632.8

5.7. Second-best congestion tolling, heterogeneous values of time, common tolls

It may well be that the regulator is not able to distinguish the value of times for different groups and that only common tolls can be applied. Table 6 shows that the numerically determined common second-best tolls are between the group-differentiated second-best tolls of Table 5. Because tolls cannot be differentiated between groups, welfare levels are by definition always lower or equal than in the previous section. When we compare differentiated and common second-best tolls for higher values of θ, the higher value of time users may benefit further from a common toll. This is because the common second-best toll prices off more low value of time users, and therefore route T is more catered towards the high value of time group. Common tolls therefore appear to benefit the high value of time users when route choice is more deterministic. Table 5 shows that for $\theta = 0.5$, the second-best tolls of both groups are almost equal, meaning that differentiation of tolls between groups is hardly beneficial. If we compare this result with Table 6 we find that the welfare level for $\theta = 0.5$ is (almost) equal. The benefits of toll differentiation are therefore strongly influenced by the randomness of route preferences and are highest for very low and very high values of θ.

Table 6: Tolls, flows and welfare for common second-best tolls with heterogeneous values of time

θ	f_T	N_{T1}	N_{T2}	N_{U1}	N_{U2}	N_1	N_2	$\hat{W}_{SB1}$	$\hat{W}_{SB2}$	$\hat{W}_{SB}$
1	7.60	73.2	420.0	688.9	139.1	493.2	828.0	4427.3	8271.3	12698.8
0.5	7.28	177.9	328.9	581.8	231.9	506.8	813.7	5142.1	7504.4	12646.5
0.1	9.94	255.9	235.0	448.3	313.9	490.9	762.2	5851.7	7234.1	13085.7

6. Conclusions

We extended the two route model of Verhoef et al. (1996) along two dimensions. First, we assumed that route choice is governed by random utility maximization. Second, we included heterogeneous preferences for travel time savings and derived some useful analytical expressions for first-best and second-best congestion tolls. Further investigation is needed to see if our analytical approach can be applied to more general networks.

Our main results show that when values of time are homogeneous, welfare losses due to second-best pricing are lower for stochastic route choice than for deterministic route choice. When preferences for time savings are heterogeneous, the picture is less clear cut because the benefits of product differentiation (or value-pricing) of roads first decreases and then increases when route choice becomes more stochastic. In line with this, we find that stochastic route choice may result in common second-best congestion tolls that are close to the group-specific (differentiated) tolls. As we showed in our numerical analysis there are cases where the welfare loss due to the inability to differentiate tolls is negligible. If such cases are realistic is an empirical question, and therefore further empirical investigation of stochastic route preferences may help to provide a more detailed estimate of the benefits of value pricing.

Bibliography

Akamatsu, T. (1997). Decomposition of path choice entropy in general transport networks. *Transportation Science*, 31(4):349–362.

Arnott, R., de Palma, A., and Lindsey, R. (1992). Route choice with heterogeneous drivers and group-specific congestion costs. *Regional Science and Urban Economics*, 22(1):71–102.

Clark, A., Sumalee, A., Shepherd, S., and Connors, R. (2009). On the Existence and Uniqueness of First Best Tolls in Networks with Multiple User Classes and Elastic Demand. *Transportmetrica*, 5(2):141–157.

de Jong, G., Daly, A., Pieters, M., and van der Hoorn, T. (2007). The logsum as an evaluation measure: Review of the literature and new results. *Transportation Research Part A: Policy and Practice*, 41:874–889.

Erlander, S. (1977). Accessibility, entropy and the distribution and assignment of traffic. *Transportation Research*, 11(3):149–153.

Fisk, C. (1980). Some developments in equilibrium traffic assignment. *Transportation Research Part B: Methodological*, 14(3):243–255.

Huang, H. and Li, Z. (2007). A multiclass, multicriteria logit-based traffic equilibrium assignment model under ATIS. *European Journal of Operational Research*, 176(3):1464–1477.

Jiang, L., Mahmassani, H., and Zhang, K. (2011). Congestion Pricing, Heterogeneous Users, and Travel Time Reliability. *Transportation Research Record: Journal of the Transportation Research Board*, 2254:58–67.

Lévy-Lambert, H. (1968). Tarification des Services à Qualité Variable–Application aux Péages de Circulation. *Econometrica: Journal of the Econometric Society*, 36:564–574.

Light, T. (2009). Optimal highway design and user welfare under value pricing. *Journal of Urban Economics*, 66(2):116–124.

Lu, C., Zhou, X., and Mahmassani, H. (2006). Variable toll pricing and heterogeneous users: Model and solution algorithm for bicriterion dynamic traffic assignment problem. *Transportation Research Record: Journal of the Transportation Research Board*, 1964:19–26.

Maher, M., Stewart, K., and Rosa, A. (2005). Stochastic social optimum traffic assignment. *Transportation Research Part B: Methodological*, 39:753–767.

Mahmassani, H., Zhou, X., and Lu, C.-C. (2005). Toll Pricing and Heterogeneous Users: Approximation Algorithms for Finding Bicriterion Time-Dependent Efficient Paths in Large-Scale Traffic Networks. *Transportation Research Record: Journal of the Transportation Research Board*, 1923:28–36.

Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, 8:229–254.

Marchand, M. (1968). A note on optimal tolls in an imperfect environment. *Econometrica: Journal of the Econometric Society*, 36:575–581.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Frontiers in econometrics, Academic Press, New York. P. Zarembka (ed.)*, volume In P. Zare, pages 105–142.

Miyagi, T. (1986). On the Formulation of a Stochastic User Equilibrium Model Consistent with the Random Utility Theory; A Conjugate Dual Approach. *Research for Tomorrow's Transport Requirements. Proceedings of the Fourth World Conference on Transport Research*, 2:1619–1635.

Palma, A. D., Kilani, M., and Lindsey, R. (2008). The merits of separating cars and trucks. *Journal of Urban Economics*, 64(2):340–361.

Parry, I. (2008). How should heavy-duty trucks be taxed? *Journal of Urban Economics*, 63(2):651–668.

Shannon, C. (1948). A note on the concept of entropy. *Bell System Technological Journal*, 27:379–423.

Small, K. and Verhoef, E. (2007). *The economics of urban transportation*. Routledge.

Small, K. A. (2012). Valuation of travel time. *Economics of Transportation*, 1:2–14.

Small, K. A. and Rosen, H. S. (1981). Applied Welfare Economics with Discrete Choice Models. *Econometrica*, 49:105–130.

Small, K. A. and Yan, J. (2001). The Value of Value Pricing of Roads: Second-Best Pricing and Product Differentiation. *Journal of Urban Economics*, 49:310–336.

van den Berg, V. and Verhoef, E. (2013). Congestion pricing in a road and rail network with heterogeneous values of time and schedule delay. *Transportmetrica A: Transport Science*, 10(5):377–400.

van den Berg, V. and Verhoef, E. T. (2011a). Congestion tolling in the bottleneck model with heterogeneous values of time. *Transportation Research Part B: Methodological*, 45:60–78.

van den Berg, V. and Verhoef, E. T. (2011b). Winning or losing from dynamic bottleneck congestion pricing?. The distributional effects of road pricing with heterogeneity in values of time and schedule delay. *Journal of Public Economics*, 95:983–992.

Verhoef, E., Nijkamp, P., and Rietveld, P. (1995). Second-Best Regulation of Road Transport Externalities. *Journal of Transport Economics and Policy*, 29:147–167.

Verhoef, E., Nijkamp, P., and Rietveld, P. (1996). Second-best congestion pricing: the case of an untolled alternative. *Journal of Urban Economics*, 40(3):279–302.

Verhoef, E. T. (2002). Second-best congestion pricing in general networks. Heuristic algorithms for finding second-best optimal toll levels and toll points. *Transportation Research Part B: Methodological*, 36:707–729.

Verhoef, E. T. and Small, K. A. (2004). Product Differentiation on Roads: Constrained Congestion Pricing with Heterogeneous Users. *Journal of Transport Economics and Policy*, 38:127–156.

Williams, H. C. W. L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning A*, 9:285–344.

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

Yang, H. and Huang, H. H. J. (2004). The multi-class, multi-criteria traffic network equilibrium and systems optimum problem. *Transportation Research Part B: Methodological*, 38:1–15.

Zhang, L., Levinson, D. M., and Zhu, S. (2008). Agent-based model of price competition, capacity choice, and product differentiation on congested networks. *Journal of Transport Economics and Policy*, 42:435–461.

Appendix A. Derivation of the first-best toll with heterogeneous preferences

Define $\bar{N_T}^{\alpha} = \sum_{k=1}^{K} \alpha_k N_{Tk}$ as the preference weighted average number of travellers at route T, and $\bar{N_U}^{\alpha} = \sum_{k=1}^{K} \alpha_k N_{Uk}$ as the preference weighted number of travellers for route U. Then 18 is given by:

$$\begin{aligned}\mathcal{L} = &\sum_{k=1}^{K} \int_0^{N_k} D_k(n_k) dn_k - \bar{N_T}^{\alpha} c_T(N_T) - \bar{N_U}^{\alpha} c_U(N_U) \\ &- \sum_{k=1}^{K} \frac{1}{\theta_k} \left(N_{Tk} \ln \left[\frac{N_{Tk}}{N_k} \right] + N_{Uk} \ln \left[\frac{N_{Uk}}{N_k} \right] \right) + \\ &\sum_{k=1}^{K} \lambda_{Tk} \left(D_k'(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln \left[\frac{N_{Tk}}{N_k} \right] \right) + \\ &\sum_{k=1}^{K} \lambda_{Uk} \left(D_k'(N_k) - f_{Uk} - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln \left[\frac{N_{Uk}}{N_k} \right] \right)\end{aligned} \tag{A.1}$$

The first-order conditions are given by:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial N_{Tl}} = &D_l(N_l) - \bar{N_T}^{\alpha} c_T'(N_T) - \alpha_l c_T(N_T) - \frac{1}{\theta_l} \ln \left[\frac{N_{T_l}}{N_l} \right] - \sum_{k=1}^{K} \lambda_{Tk} \alpha_k c_T'(N_T) + \\ &\lambda_{Tl} \left(D_l'(N_l) - \frac{1}{\theta_l} \frac{N_{Tl}}{N_{Ul} N_l} \right) + \lambda_{Ul} \left(D_l'(N_l) + \frac{1}{\theta_l} \frac{1}{N_l} \right) = 0, \forall l = 1...K.\end{aligned} \tag{A.2}$$

$$\frac{\partial \mathcal{L}}{\partial N_{Ul}} = D_l(N_l) - \bar{N_U}^{\alpha} c'_U(N_U) - \alpha_l c_U(N_U) - \frac{1}{\theta_l} \ln\left[\frac{N_{U_l}}{N_l}\right] - \sum_{k=1}^{K} \lambda_{Uk}\alpha_k c'_U(N_U) + \lambda_{Ul}\left(D'_l(N_l) - \frac{1}{\theta_l}\frac{N_{Ul}}{N_{Tl}N_l}\right) + \lambda_{Tl}\left(D'_l(N_l) + \frac{1}{\theta_l}\frac{1}{N_l}\right) = 0, \forall l = 1...K. \tag{A.3}$$

$$\frac{\partial \mathcal{L}}{\partial f_{Tl}} = -\lambda_{Tl} = 0, \forall l = 1...K. \tag{A.4}$$

$$\frac{\partial \mathcal{L}}{\partial f_{Ul}} = -\lambda_{Ul} = 0, \forall l = 1...K. \tag{A.5}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_{Tl}} = D_l(N_l) - f_{Tl} - \alpha_l c_T(N_T) - \frac{1}{\theta_l} \ln\left[\frac{N_{Tl}}{N_l}\right] = 0, \forall l = 1...K. \tag{A.6}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_{Ul}} = D_l(N_l) - f_{Ul} - \alpha_l c_U(N_U) - \frac{1}{\theta_l} \ln\left[\frac{N_{Ul}}{N_l}\right] = 0, \forall l = 1...K. \tag{A.7}$$

Equations A.4 and A.5 show that the Lagrangian multipliers are 0. Using A.4, A.5 and A.6 in A.2 we obtain:

$$f_{Tl} = \bar{N_T}^{\alpha} c'_T(N_T), \forall l = 1...K. \tag{A.8}$$

For the toll of route U we use A.4, A.5 and A.7 in A.3:

$$f_{Ul} = \bar{N_U}^{\alpha} c'_U(N_U), \forall l = 1...K. \tag{A.9}$$

Because $\bar{N_T}^{\alpha}$ and $\bar{N_U}^{\alpha}$ are equal for all groups, the tolls are equal for all groups.

Appendix B. Second-best congestion pricing with heterogeneous preferences and group-specific tolls

The Lagrangian is given by:

$$\begin{aligned}\mathcal{L} = &\sum_{k=1}^{K} \int_0^{N_k} D_k(n_k)dn_k - \bar{N_T}^{\alpha} c_T(N_T) - \bar{N_U}^{\alpha} c_U(N_U) \\ &- \sum_{k=1}^{K} \frac{1}{\theta_k}\left(N_{Tk} \ln\left[\frac{N_{Tk}}{N_k}\right] + N_{Uk} \ln\left[\frac{N_{Uk}}{N_k}\right]\right) + \\ &\sum_{k=1}^{K} \lambda_{Tk}\left(D'_k(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln\left[\frac{N_{Tk}}{N_k}\right]\right) + \\ &\sum_{k=1}^{K} \lambda_{Uk}\left(D'_k(N_k) - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln\left[\frac{N_{Uk}}{N_k}\right]\right)\end{aligned} \tag{B.1}$$

The first-order conditions are given by:

$$\frac{\partial \mathcal{L}}{\partial N_{Tl}} = D_l(N_l) - \bar{N_T}^{\alpha} c'_T(N_T) - \alpha_l c_T(N_T) - \frac{1}{\theta_l} \ln\left[\frac{N_{T_l}}{N_l}\right] - \sum_{k=1}^{K} \lambda_{Tk}\alpha_k c'_T(N_T) + \lambda_{Tl}\left(D'_l(N_l) - \frac{1}{\theta_l}\frac{N_{Tl}}{N_{Ul}N_l}\right) + \lambda_{Ul}\left(D'_l(N_l) + \frac{1}{\theta_l}\frac{1}{N_l}\right) = 0, \forall l = 1...K. \tag{B.2}$$

$$\frac{\partial \mathcal{L}}{\partial N_{Ul}} = D_l(N_l) - \bar{N_U}^{\alpha} c'_U(N_U) - \alpha_l c_U(N_U) - \frac{1}{\theta_l} \ln\left[\frac{N_{U_l}}{N_l}\right] - \sum_{k=1}^{K} \lambda_{Uk}\alpha_k c'_U(N_U) + \lambda_{Ul}\left(D'_l(N_l) - \frac{1}{\theta_l}\frac{N_{Ul}}{N_{Tl}N_l}\right) + \lambda_{Tl}\left(D'_l(N_l) + \frac{1}{\theta_l}\frac{1}{N_l}\right) = 0, \forall l = 1...K. \tag{B.3}$$

$$\frac{\partial \mathcal{L}}{\partial f_{Tl}} = -\lambda_{Tl} = 0, \forall l = 1...K. \tag{B.4}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_{Tl}} = D_l(N_l) - f_{Tl} - \alpha_l c_T(N_T) - \frac{1}{\theta_l} \ln\left[\frac{N_{Tl}}{N_l}\right] = 0, \forall l = 1...K. \tag{B.5}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_{Ul}} = D_l(N_l) - \alpha_l c_U(N_U) - \frac{1}{\theta_l} \ln\left[\frac{N_{Ul}}{N_l}\right] = 0, \forall l = 1...K. \tag{B.6}$$

From B.4 we find that the group-specific Lagrangian multipliers for route T are all 0. Substituting B.4 and B.6 in B.3 gives:

$$\frac{\partial \mathcal{L}}{\partial N_{Ul}} = -\bar{N_U}^{\alpha} c'_U(N_U) - \sum_{k=1}^{K} \lambda_{Uk}\alpha_k c'_U(N_U) + \lambda_{Ul}\left(D'_l(N_l) - \frac{1}{\theta_l}\frac{N_{Ul}}{N_{Tl}N_l}\right) = 0, \forall l = 1...K. \tag{B.7}$$

From B.4, B.5 and B.2 we have:

$$f_{Tl} = \bar{N_T}^{\alpha} c'_T(N_T) - \lambda_{Ul}\left(D'_l(N_l) + \frac{1}{\theta_l}\frac{1}{N_l}\right), \forall l = 1...K. \tag{B.8}$$

The solution for the group-specific Lagrangian multipliers for route U can be obtained using the system of K equations of B.7. This system can be written in matrix notation:

$$A\lambda_U = b, \tag{B.9}$$

where λ_U is the $K \times 1$ vector with unknown multipliers, b is the $K \times 1$ vector with each element equal to $\bar{N_U}^{\alpha} c'_U(N_U)$, and A is the following $K \times K$ matrix:

$$A = \begin{pmatrix} D'_1(N_1) - \alpha_1 c'_U(N_U) - \frac{1}{\theta_1}\frac{N_{T1}}{N_{U1}N_1} & -\alpha_2 c'_U(N_U) & \dots & -\alpha_K c'_U(N_U) \\ -\alpha_1 c'_U(N_U) & D'_2(N_2) - \alpha_2 c'_U(N_U) - \frac{1}{\theta_2}\frac{N_{T2}}{N_{U2}N_2} & \dots & -\alpha_K c'_U(N_U) \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_1 c'_U(N_U) & -\alpha_1 c'_U(N_U) & \dots & D'_K(N_K) - \alpha_K c'_U(N_U) - \frac{1}{\theta_1}\frac{N_{TK}}{N_{UK}N_K} \end{pmatrix}$$

The solution for the vector λ_U can be found by Cramers' rule. Let $A_l(b)$ be the matrix A with column l replaced by the vector b. The solution for the lth Lagrangian multiplier is given by a ratio of determinants:

$$\lambda^*_{Ul} = \frac{\det(A_l(b))}{\det(A)}, \tag{B.10}$$

and therefore we need $\det(A) \neq 0$ to have a unique solution. Equation B.10 can be made more explicit using analytical expressions for the determinants. Because the matrix A has many common elements, its determinant can be written in a tractable closed-form:

$$\det(A) = \prod_{k=1}^{K} \left[D'_k(N_k) - \frac{1}{\theta_k} \frac{N_{Tk}}{N_{Uk} N_k} \right] - \sum_{k=1}^{K} \alpha_k c'_U(N_U) \prod_{\substack{m=1 \\ m \neq k}}^{K} \left[D'_m(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m} \right]. \tag{B.11}$$

Using $\prod_{\substack{m=1 \\ m \neq k}}^{K} \left[D'_m(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m} \right] = \frac{\prod_{k=1}^{K} \left[D'_k(N_k) - \frac{1}{\theta_k} \frac{N_{Tk}}{N_{Uk} N_k} \right]}{D'_m(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m}}$, we can divide out the first product term in B.11:

$$\det(A) = \prod_{k=1}^{K} \left[D'_k(N_k) - \frac{1}{\theta_k} \frac{N_{Tk}}{N_{Uk} N_k} \right] \left(1 - \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m}} \right). \tag{B.12}$$

Because $D'_k(N_k) - \frac{1}{\theta_k} \frac{N_{Tk}}{N_{Uk} N_k} < 0, \forall k = 1...K$, the first part in B.12 will be a product of negative numbers resulting in a number that is unequal to 0. Because $\alpha_k c'_U(N_U) > 0$, and $D'_m(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m} < 0$, the summation is over negative numbers resulting in a negative number for the part between large brackets. Therefore B.12 is unequal to 0 and a unique solution for the Lagrangian multipliers exists. The solution B.10 can be further investigated by using the following analytical expression for the determinant $\det(A_l(b))$:

$$\det(A_l(b)) = \bar{N_U}^{\alpha} c'_U(N_U) \prod_{\substack{l=1 \\ l \neq k}}^{K} \left[D'_l(N_l) - \frac{1}{\theta_l} \frac{N_{Tl}}{N_{Ul} N_l} \right] \tag{B.13}$$

We have $\det(A_l(b)) \neq 0$ implying that all the Lagrangian multipliers for route U have a unique non-zero value. Substituting B.12 and B.13 in B.10 gives:

$$\lambda^*_{Ul} = \frac{\bar{N_U}^{\alpha} c'_U(N_U) \prod_{\substack{l=1 \\ l \neq k}}^{K} \left[D'_l(N_l) - \frac{1}{\theta_l} \frac{N_{Tl}}{N_{Ul} N_l} \right]}{\prod_{k=1}^{K} \left[D'_k(N_k) - \frac{1}{\theta_k} \frac{N_{Tk}}{N_{Uk} N_k} \right] \left(1 - \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m}} \right)}, \tag{B.14}$$

which can be rewritten as:

$$\lambda^*_{Ul} = \bar{N_U}^{\alpha} c'_U(N_U) \frac{1}{\left(D'_l(N_l) - \frac{1}{\theta_l} \frac{N_{Tl}}{N_{Ul} N_l} \right) \left(1 - \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m}} \right)}. \tag{B.15}$$

Taking the lth term out of the summation this reduces to:

$$\lambda^*_{Ul} = \bar{N_U}^{\alpha} c'_U(N_U) \frac{1}{D'_l(N_l) - \frac{1}{\theta_l} \frac{N_{Tl}}{N_{Ul} N_l} - \alpha_l c'_U(N_U) - c'_U(N_U) \sum_{\substack{k=1 \\ k \neq l}}^{K} \alpha_k \frac{D'_l(N_l) - \frac{1}{\theta_l} \frac{N_{Tl}}{N_{Ul} N_l}}{D'_k(N_k) - \frac{1}{\theta_k} \frac{N_{Tk}}{N_{Uk} N_k}}}. \tag{B.16}$$

This shows that the Lagrangian multipliers for each group are non-positive. Because tolls enter the constraints negatively, an increase in the group-specific toll f_{Ul} from 0 (which is the SB case under consideration) to a positive value will lead to higher welfare. Substituting B.16 in B.8 gives:

$$f_{Tl} = \bar{N_T}^{\alpha} c'_T(N_T) - \bar{N_U}^{\alpha} c'_U(N_U) \frac{D'_l(N_l) + \frac{1}{\theta_l}\frac{1}{N_l}}{D'_l(N_l) - \frac{1}{\theta_l}\frac{N_{Tl}}{N_{Ul}N_l} - \alpha_l c'_U(N_U) - c'_U(N_U)\sum_{\substack{k=1\\k\neq l}}^{K} \alpha_k \frac{D'_l(N_l) - \frac{1}{\theta_l}\frac{N_{Tl}}{N_{Ul}N_l}}{D'_k(N_k) - \frac{1}{\theta_k}\frac{N_{Tk}}{N_{Uk}N_k}}}. \tag{B.17}$$

Multiplying the nominator and the denominator of the fractional part by -1 results in 35.

Appendix C. Calibration of the asymmetric route flows

If we want to calibrate the model in the no-toll case for given values of θ, we have observed number of travellers for both routes and the corresponding total number of travellers. We therefore also have the observed route probabilities which are functions of these. The inverse demand is assumed to be linear and is given by 39. In the no-toll equilibrium we have two conditions that need to be satisfied, since the inverse demand should be equal to the generalised costs. Using 6 and 40 and assuming $\beta_T = \beta_U = \beta$ results in:

$$\begin{aligned} \delta_1 - \delta_2(N_T + N_U) &= \kappa_T + \beta N_T + \frac{1}{\theta}\ln\left[\frac{N_T}{N}\right], \\ \delta_1 - \delta_2(N_T + N_U) &= ASC_U + \kappa_U + \beta N_T + \frac{1}{\theta}\ln\left[\frac{N_U}{N}\right], \end{aligned} \tag{C.1}$$

where ASC_U is the alternative specific constant for route U. Solving C.1 for ASC_U gives:

$$ASC_U = \kappa_T - \kappa_U + \beta(N_T - N_U) + \frac{1}{\theta}\ln\left[\frac{N_T}{N_U}\right]. \tag{C.2}$$

We want to have N_T and N_U as the flows in deterministic user equilibrium, implying $\kappa_T + \beta N_T^0 = \kappa_U + \beta N_U^0 \implies \kappa_T - \kappa_U + \beta(N_T^0 - N_U^0) = 0$. Substituting in C.2 gives:

$$ASC_U = \frac{1}{\theta}\ln\left[\frac{N_T^0}{N_U^0}\right]. \tag{C.3}$$

The symmetric case $N_T^0 = N_U^0$ is a special case and gives $ASC_U = 0$. This completes the calibration for asymmetric route flows.