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 reduced 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 of non-differentiated tolls.

Keywords: Stochastic User Equilibrium, System Optimum, Second best Congestion Pricing, Scale Heterogeneity, Probabilistic Choice

1. Introduction

Probabilistic route choice is widely applied 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 effect of probabilistic choice on 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. In the absence of income effects, consumer surplus 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 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 present of second-best congestion pricing in static and dynamic congestion models (see for example: Verhoef et al. (1995); 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 (2011); Van den Berg and Verhoef (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 integral under the demand curve 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 the 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. A more deterministic route choice 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 (e.g. business and non-business travellers).

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. 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. 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 analytical expressions, 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 and a pooled equilibrium is obtained. The extend towards the roads are differentiated in order to accommodate the needs of distinct groups therefore crucially does depend 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 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 $k$. The individual chooses out of a finite set of $R$ alternative routes:

$$U_{ktr} = G_{kr} + \epsilon_{ktr}. \quad (1)$$

Random utility $U_{ktr}$ therefore depends on a deterministic part $G_{kr}$ and a stochastic idiosyncratic route preference $\epsilon_{ktr}$. The individual $n_k$ therefore makes repeated route choices where his preferences for routes 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:

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

where $V_{kr} = \theta_k G_{kr}, \forall r = 1...R$. The scale parameter $\theta_k$ governs the relative importance of the unobserved idiosyncratic part of the utility in the total utility. There are two extreme case 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, meaning that adding an additional alternative will lead to higher maximum expected utility. This becomes more evident if we derive the maximum expected utility for individual $n_k$, which is given by the negative of the logsum (Williams (1977); Small and Rosen (1981); de Jong et al. (2007)):

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

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 $\theta_k < \infty$. 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 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 function $D(n)$. A traveller enters the road network if this willingness to pay is higher than the expected generalised costs of equation 3. This 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.

If we assume that tolls and congestion costs are additive separable, the deterministic utilities for routes $U$ and $T$ are given by:

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

(4)

where $f_r$ is the toll on route $r$, and $c_r(N_r)$ is the travel cost for route $r$ that increases 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:

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

(5)

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:

$$GC_r = -\frac{1}{\theta} \ln (\exp(V_T) + \exp(V_U)) .$$

(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:

$$GC_r = -\frac{1}{\theta} \ln (\exp(V_T) + \exp(V_U)) = \ln \left( \frac{\exp(V_r)}{P_r} \right) = -\frac{1}{\theta} \left( V_r - \ln \left( \frac{N_r}{N} \right) \right) = f_r + c_r(N_r) + \frac{1}{\theta} \ln \left( \frac{N_r}{N} \right) .$$

(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 decrease 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 therefore they drop out of the social surplus function. To stay as close as possible to the two-route model of Verhoef et al. (1996), it is assumed that overall demand can be reflected by a single inverse demand function \( D(n) \). 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 \( \theta \) 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 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

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

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 in then \( \theta \) 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 of 9 need to be solved jointly:

\[
\frac{\partial L}{\partial N_T} = D(N) - c_T(N_T) - N_T c'_T(N_T) - \frac{1}{\theta} \ln \left( \frac{N_T}{N} \right) + \lambda_T \left( D'(N) - c'_T(N_T) - \frac{1}{\theta} \frac{N_T}{N_U N} \right) \\
\quad + \lambda_U \left( D'(N) + \frac{1}{\theta} \frac{1}{N} \right) = 0. 
\] (10)

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

\[
\frac{\partial L}{\partial f_T} = -\lambda_T = 0. 
\] (12)

\[
\frac{\partial L}{\partial f_U} = -\lambda_U = 0. 
\] (13)

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

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

\[
f_T = D(N) - \frac{1}{\theta} \ln \left( \frac{N_T}{N} \right) = N_T c'_T(N_T), 
\] (16)

\[
f_U = D(N) - \frac{1}{\theta} \ln \left( \frac{N_U}{N} \right) = N_U c'_U(N_U).
\]
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, even though they acting acting in their own 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. The absolute toll levels of the SUE and DUE model may 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. The 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 curve 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 $k$. Scale heterogeneity is captured by having a group-specific scale parameter $\theta_k$. Finally, heterogeneity in overall demand response is captured by having a group-specific inverse demand curve $D_k(n_k)$. 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 $N_k = N_{Tk} + N_{Uk}$. 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 travel 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 $c_{Tk}(N_{Tk}) = \alpha_k c_T(N_T)$ and $c_{Uk}(N_{Uk}) = \alpha_k c_U(N_U)$, $\forall k = 1...K$. 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 $\alpha_k$.

This group-specific valuation of travel time converts the route travel time $c_{rk}(N_r)$ to monetary units. To save notation we define $\bar{N}_T^{\alpha} = \sum_{k=1}^K \alpha_k N_{Tk}$ as the preference weighted 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 $U$. 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:

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

\[
\mathcal{L} = \sum_{k=1}^{K} \int_{0}^{N_k} D_k(n_k) \, dn_k - \sum_{k=1}^{K} \alpha_k N_{Tk} c_T(N_T) - \sum_{k=1}^{K} \alpha_k N_{Uk} c_U(N_U) \\
- \sum_{k=1}^{K} \frac{1}{\theta_k} \left( N_{Tk} \ln \left( \frac{N_{Tk}}{N_k} \right) + N_{Uk} \ln \left( \frac{N_{Uk}}{N_k} \right) \right) + \\
\sum_{k=1}^{K} \lambda_{Tk} \left( D_k(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln \left( \frac{N_{Tk}}{N_k} \right) \right) + \\
\sum_{k=1}^{K} \lambda_{Uk} \left( D_k(N_k) - f_{Uk} - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln \left( \frac{N_{Uk}}{N_k} \right) \right).
\] (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

\[
f_{Tk} = \bar{N}_{Tk} c_T'(N_T), \\
f_{Uk} = \bar{N}_{Uk} c_U'(N_U).
\] (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 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. Therefore the group-differentiated first-best tolls coincide with the common first-best toll for all groups. This is not the case 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 an untolled route. Therefore tolling one of the two routes (a form of second-best congestion pricing), may be a viable alternative. In this section we analyse congestion pricing with probabilistic choice in the presence of an untolled
The systematic route utility for the tolled route is given by 

\[ V = \text{willingness to pay} \]

is equal to the generalised route costs. The Lagrangian is given by:

\[ \text{expression for the total social surplus (equation 8) therefore will not change. Because we} \]

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. The expression for the total social surplus (equation 8) therefore 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 the generalised route costs. The Lagrangian is given by:

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

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:

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

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

\[
\frac{\partial \mathcal{L}}{\partial f_T} = -\lambda_T = 0.
\]

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

\[
\frac{\partial \mathcal{L}}{\partial \lambda_U} = D(N) - c_U(N_U) - \frac{1}{\theta} \ln \left( \frac{N_U}{N} \right) = 0.
\]
Using 21, 23 and 24 we obtain:

\[ f_T = D(N) - c_T(N_T) - \frac{1}{\theta} \ln \left( \frac{N_T}{N} \right) = N_T c'_T(N_T) - \lambda_U \left( D'(N) + \frac{1}{\theta} \frac{N_T}{N} \right). \]  

(26)

Using 25 and 23, we can solve 22 for \( \lambda_U \):

\[ \lambda_U = \frac{N_U c'_U(N_U)}{D'(N) - c'_U(N_U) - \frac{1}{\theta} \frac{N_T}{N T N}}. \]  

(27)

Substituting 27 in 26 gives:

\[ f_T = N_T c'_T(N_T) - N_U c'_U(N_U) \frac{-D'(N) - \frac{1}{\theta} \frac{N_T}{N T N}}{c'_U(N_U) - D'(N) - \frac{1}{\theta} \frac{N_T}{N T N}}. \]  

(28)

The first term in 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 \( T \). This correction term depends on the relative size of the random part of utility in the total utility, the slope of the inverse demand curve, 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 \( \theta \), since \( \theta \) has a direct positive effect on the numerator and the denominator of the correction term, but also has an indirect effect on 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 28 shows that 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, 20 shows that the term \( D'(N) + \frac{1}{\theta} \frac{N_T}{N T N} \) is the change in the constraint for route \( U \) due to a marginal change in \( N_T \), whereas \( -c'_U(N_U) + D'(N) + \frac{1}{\theta} \frac{N_T}{N T N} \) is the change in the constraint for route \( U \) due to a marginal change in \( N_U \).

As opposed to the first-best toll rules of 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:

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

(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:

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

(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. The level of the toll in $30$ may well be different for DUE and SUE for asymmetric route costs, because $\theta$ has an effect on the optimal route split. Furthermore, $30$ 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

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

$$\lim_{c'_U(N_U) \to 0} f_T = N_T c'_T(N_T).$$  \hspace{1cm} (32)

which is again isomorphic to the toll rule in the deterministic model. The absence of a toll on route $U$ then means that the route is optimally priced. The regulator may therefore ignore route $U$ and needs only to consider the unconstrained optimal regulation of route $T$.

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:

$$\mathcal{L} = \sum_{k=1}^{K} \int_0^{N_k} D_k(n_k)dn_k - \sum_{k=1}^{K} \alpha_k N_{Tk} c_T(N_T) - \sum_{k=1}^{K} \alpha_k N_{Uk} c_U(N_U)$$
$$- \sum_{k=1}^{K} \frac{1}{\theta_k} \left( N_{Tk} \ln \left( \frac{N_{Tk}}{N_k} \right) + N_{Uk} \ln \left( \frac{N_{Uk}}{N_k} \right) \right) +$$
$$\sum_{k=1}^{K} \lambda_{Tk} \left( D'_k(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln \left( \frac{N_{Tk}}{N_k} \right) \right) +$$
$$\sum_{k=1}^{K} \lambda_{Uk} \left( D'_k(N_k) - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln \left( \frac{N_{Uk}}{N_k} \right) \right)$$  \hspace{1cm} (33)
In Appendix B we show that the second-best group-specific toll for group $k$ is given by:

$$f_{Tk} = \bar{N}_T \alpha c_T(N_T) - \bar{N}_U \alpha c_U(N_U) \frac{-D'(N_k) - \frac{1}{\theta_k} \frac{N_k}{N_k} \sum_{l=1}^{K} \alpha_l c'_U(N_U)}{\alpha_k c'_U(N_U) - D'_k(N_k) + \frac{1}{\theta_k} \frac{N_k}{N_k} + \phi_k},$$

(34)

where

$$\phi_k = \sum_{l=1}^{K} \alpha_l c'_U(N_U) \frac{D'_k(N_k) - \frac{1}{\theta_k} \frac{N_k}{N_k}}{D'_l(N_l) - \frac{1}{\theta_l} \frac{N_l}{N_l}} > 0.$$  

(35)

The first part in equation 34 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 34 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 34 reduces to 28. Second, the DUE group-specific toll is a special case for which $\theta_k \to \infty$, $\forall k = 1...K$. This results in:

$$f_{Tk} = \bar{N}_T \alpha c_T(N_T) - \bar{N}_U \alpha c_U(N_U) \frac{-D'(N_k)}{\alpha_k c'_U(N_U) - \frac{1}{\theta_k} \frac{N_k}{N_k} \sum_{l=1}^{K} \alpha_k D'_l(N_l)}. \quad (36)$$

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:

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

(37)

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 $\theta_k$. We were not able to derive analytical solutions for the common second-best toll case. 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 33 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 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) assumed linear demand functions:

\[ D(N) = \delta_1 - \delta_2 (N_T + N_U), \]  
(38)

and congestion costs on route \( r \) defined as:

\[ c_r(N_r) = \kappa_r + \beta_r N_r, \]  
(39)

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 39 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 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 \( \theta \), heterogeneity in values of time between groups \( (\alpha_k) \), and heterogeneity in preferences between groups \( (\theta_k) \) 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 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. The calibration process is simplified by the use of the symmetric case and which allows us to investigate impacts of singular changes in model assumptions.

We are though aware that symmetric examples may hide some impacts, in particular the change in route flows even in the non-intervention case \( \theta \) is adjusted. For this reason we develop an asymmetric example (again based on Verhoef et al. (1996)) where we adjust the choice mechanism 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 \( \theta \) 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 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 correct the 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 demand function may be written as:

\[ D(N) = \delta_1 + \delta_2 (N_T + N_U) + \nu, \]  
(40)
where the correction 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 remain constant and are equal to 1, meaning that 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:

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

where superscript $N^0_k$ 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:

$$
N^0_1 = N^0 \frac{1 - \alpha_2}{\alpha_1 - \alpha_2}; N^0_2 = N^0 - N^0_1.
$$

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:

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

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 $N^0_1 = 1500 \times \frac{1.3 - 1.3}{0.8 - 1.3} = 900$ and $N^0_2 = 1500 - 900$ 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 $-\frac{\alpha_1(\delta_1 - \nu_1)}{N^0_1} = -\frac{15 \times 0.8}{900} = -\frac{1}{75}$, and $-\frac{\alpha_2(\delta_1 - \nu_1)}{N^0_2} = -\frac{15 \times 1.3}{600} = -\frac{13}{400}$ 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 $\frac{13}{60}$, so that the equilibrium travel time will be equal for both groups.

For the no toll 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 $W_{1,0} = \frac{1}{2} \times (\delta_1 \alpha_1 - 28) \times N^0_1 = \frac{1}{2} \times 12 \times 900 = 5400$ and $W_{2,0} = \frac{1}{2} \times (\delta_1 \alpha_2 - 45\frac{1}{2}) \times N^0_2 = \frac{1}{2} \times 19\frac{1}{2} \times 600 = 5850$. 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 $\nu_1$ and $\nu_2$ in such a way that non-intervention average user costs and welfare levels are maintained.

15
5.2. Numerical results, 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 \( \theta \). However, this is a special case because with asymmetric route costs \( \theta \) has an effect on the toll via the equilibrium number of travellers. To illustrate this, assume that the routes have different free-flow travel times \( \kappa_r \), 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 \( N^0_T = 625 \) and \( N^0_U = 1125 \) 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 calibrated constant for any chosen value of \( \theta \) is given by:

\[
ASC_U = \frac{1}{\theta} \frac{N^0_T}{N^0_U},
\]

(44)

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 \( \theta \). The flows for the \( \theta = 10 \) case are close to the UE solution of 458.13 and 708.51 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 \( \theta \) as the alternative specific constant \( ASC_U \) increases with decreasing \( \theta \). The flows on the longer route decrease when \( \theta \) 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.

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( f_T )</th>
<th>( f_U )</th>
<th>( N_T )</th>
<th>( N_U )</th>
<th>( N )</th>
<th>( \hat{W}_{FB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>9.2</td>
<td>14.2</td>
<td>458.1</td>
<td>708.5</td>
<td>1166.6</td>
<td>21041.4</td>
</tr>
<tr>
<td>1</td>
<td>9.1</td>
<td>14.2</td>
<td>456.2</td>
<td>709.9</td>
<td>1166.1</td>
<td>21038.6</td>
</tr>
<tr>
<td>0.5</td>
<td>9.2</td>
<td>14.2</td>
<td>454.3</td>
<td>711.2</td>
<td>1165.5</td>
<td>21035.8</td>
</tr>
<tr>
<td>0.1</td>
<td>8.9</td>
<td>14.4</td>
<td>443.1</td>
<td>719.0</td>
<td>1162.1</td>
<td>21020.0</td>
</tr>
</tbody>
</table>

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 $\theta$, 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 $\theta$ due to the symmetry in average route costs for low values of $\theta$. 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 that satisfy the first-best toll expressions 19 with heterogeneity. Due to symmetry we leave out two solutions because it is always possible to swap the route flows. The common toll solution in the first row of Table 2, turns out to be local minimum and 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 UE case Arnott et al. (1992) but eventually disappears in the SUE case when $\theta$ becomes sufficiently low. Route preferences of individuals then become so stochastic that toll differentiation is not beneficial in welfare terms.

Figure 1 illustrates the toll differentiated and common toll equilibria by showing the welfare as a function of the toll on route $T$ while keeping the toll on route $U$ at the optimal level. The toll differentiated equilibrium dissipates due to the lower sensitivity to deterministic costs and the multiple solutions are “smoothed” out. Toll differentiated equilibria in our model occur for two reasons. First, the likelihood of these equilibria to occur increases when the route choice model becomes more deterministic, so for higher value of $\theta$. 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 $\alpha_1$ and $\alpha_2$, while keeping the average value of time constant, a toll differentiated equilibrium occurs for lower values of $\theta$ than for a lower difference between $\alpha_1$ and $\alpha_2$.

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$f_{T1}$</th>
<th>$f_{T2}$</th>
<th>$f_{U1}$</th>
<th>$f_{U2}$</th>
<th>$N_{T1}$</th>
<th>$N_{T2}$</th>
<th>$N_{U1}$</th>
<th>$N_{U2}$</th>
<th>$\hat{W}_{FB1}$</th>
<th>$\hat{W}_{FB2}$</th>
<th>$\hat{W}_{FB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10*</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>229.4</td>
<td>251.0</td>
<td>229.4</td>
<td>251.0</td>
<td>6081.7</td>
<td>9212.5</td>
<td>15294.1</td>
</tr>
<tr>
<td>10</td>
<td>11.4</td>
<td>11.4</td>
<td>9.1</td>
<td>9.1</td>
<td>0.3</td>
<td>437.5</td>
<td>471.8</td>
<td>60.4</td>
<td>5791.9</td>
<td>9558.4</td>
<td>15350.3</td>
</tr>
<tr>
<td>1</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>229.4</td>
<td>251.0</td>
<td>229.4</td>
<td>251.0</td>
<td>6081.7</td>
<td>9212.5</td>
<td>15294.1</td>
</tr>
<tr>
<td>0.5</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>10.2</td>
<td>229.4</td>
<td>251.0</td>
<td>229.4</td>
<td>251.0</td>
<td>6081.7</td>
<td>9212.5</td>
<td>15294.1</td>
</tr>
</tbody>
</table>

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.
5.4. Second-best tolling, homogeneous values of time

Because first-best tolling is often not feasible, second-best tolling with a toll on route $T$ is a realistic and viable alternative. The numerical results in this section confirm the optimal toll rule of 28. Figure 2 shows how the welfare improvement varies with the second-best toll on route $T$ for different values of $\theta$. As $\theta$ 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 2 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 on 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. Table 3 shows the optimal second-best tolls, route flows and relative efficiency $\omega = \frac{W_{SB} - W_0}{W_{FB} - W_0}$, which is 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 $\theta$ 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 $\theta$, 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 2: 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 $\theta$.

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

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$f_T$</th>
<th>$N_T$</th>
<th>$N_U$</th>
<th>$N$</th>
<th>$\hat{W}_{SB} - \hat{W}_0$</th>
<th>$\omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>UE</td>
<td>5.45</td>
<td>545.00</td>
<td>818.00</td>
<td>1363.00</td>
<td>1023.7</td>
<td>0.273</td>
</tr>
<tr>
<td>10</td>
<td>5.50</td>
<td>544.95</td>
<td>817.74</td>
<td>1362.69</td>
<td>1029.9</td>
<td>0.275</td>
</tr>
<tr>
<td>1</td>
<td>5.87</td>
<td>540.63</td>
<td>813.67</td>
<td>1354.30</td>
<td>1093.9</td>
<td>0.292</td>
</tr>
<tr>
<td>0.5</td>
<td>6.38</td>
<td>532.72</td>
<td>809.92</td>
<td>1342.63</td>
<td>1163.0</td>
<td>0.310</td>
</tr>
<tr>
<td>0.1</td>
<td>9.36</td>
<td>507.89</td>
<td>768.73</td>
<td>1276.62</td>
<td>1653.5</td>
<td>0.441</td>
</tr>
<tr>
<td>0.05</td>
<td>11.73</td>
<td>501.07</td>
<td>722.20</td>
<td>1223.27</td>
<td>2111.3</td>
<td>0.563</td>
</tr>
</tbody>
</table>
1. We will compare the results with the homogeneous second-best toll case of the previous section. The example follows the symmetric case, where $\theta_2$ is varied for group 2 holding $\theta_1$ 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 34, 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 $\theta_2$ 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 and a higher toll is needed to arrive at the optimum. The tolls for group 2 are consistently higher than those for the same value of $\theta$ in the homogeneous case in Table 3. 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 $\theta$ decreases can be inferred from the toll rule 34, 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.

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

<table>
<thead>
<tr>
<th>$\theta_1$</th>
<th>$\theta_2$</th>
<th>$f_{T1}$</th>
<th>$f_{T2}$</th>
<th>$N_{T1}$</th>
<th>$N_{T2}$</th>
<th>$N_{U1}$</th>
<th>$N_{U2}$</th>
<th>$N_1$</th>
<th>$N_2$</th>
<th>$\hat{W}_{SB1}$</th>
<th>$\hat{W}_{SB2}$</th>
<th>$\hat{W}_{SB}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>10</td>
<td>5.50</td>
<td>5.50</td>
<td>272.5</td>
<td>272.5</td>
<td>408.9</td>
<td>408.9</td>
<td>545.0</td>
<td>817.7</td>
<td>6140.0</td>
<td>6140.0</td>
<td>12279.9</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>5.44</td>
<td>5.98</td>
<td>300.5</td>
<td>243.4</td>
<td>384.5</td>
<td>429.9</td>
<td>543.9</td>
<td>814.4</td>
<td>6325.3</td>
<td>5989.5</td>
<td>12314.7</td>
</tr>
<tr>
<td>10</td>
<td>0.5</td>
<td>5.37</td>
<td>6.51</td>
<td>304.1</td>
<td>238.9</td>
<td>384.8</td>
<td>425.8</td>
<td>543.0</td>
<td>810.6</td>
<td>6380.1</td>
<td>5972.3</td>
<td>12352.4</td>
</tr>
<tr>
<td>10</td>
<td>0.1</td>
<td>4.94</td>
<td>10.07</td>
<td>311.4</td>
<td>226.0</td>
<td>404.8</td>
<td>378.4</td>
<td>537.4</td>
<td>783.2</td>
<td>6668.4</td>
<td>5928.5</td>
<td>12596.9</td>
</tr>
</tbody>
</table>

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

This section presents the results for group-specific second-best tolls with heterogeneous values of times. Table 5 presents the numerical results for different values of the scale parameter $\theta$. The SB toll for both groups first decreases in $\theta$ and then increases for lower values of $\theta$. A lower $\theta$ 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 34 which captures the marginal external costs of route $T$. But a further decrease in $\theta$ also means that spillovers become less and less important, and that means that the second term in 34 decreases. This effect raises the value of the second-best toll. The U-shaped pattern in Table 5 is the result.
Table 5: Tolls, route flows and welfare for differentiated second-best tolls with heterogeneous values of time

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

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 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 lower than in the previous section. When we compare differentiated and common second-best tolls for higher values of \( \theta \), 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 differentiated towards the high value of time group. Differentiated tolls therefore appear to benefit the high value of time users when route choice is more deterministic. Table 6 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 5 we find that the welfare level for \( \theta = 0.5 \) is (almost) equal. The benefits of toll differentiation are therefore strongly influenced by the returns to variety and are highest for very low and very high returns to variety.

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

<table>
<thead>
<tr>
<th>( \theta )</th>
<th>( f_T )</th>
<th>( N_{T1} )</th>
<th>( N_{T2} )</th>
<th>( N_{U1} )</th>
<th>( N_{U2} )</th>
<th>( N_1 )</th>
<th>( N_2 )</th>
<th>( \hat{W}_{SB1} )</th>
<th>( \hat{W}_{SB2} )</th>
<th>( \hat{W}_{SB} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.60</td>
<td>73.2</td>
<td>420.0</td>
<td>688.9</td>
<td>139.1</td>
<td>493.2</td>
<td>828.0</td>
<td>4427.3</td>
<td>8271.3</td>
<td>12698.8</td>
</tr>
<tr>
<td>0.5</td>
<td>7.28</td>
<td>177.9</td>
<td>328.9</td>
<td>581.8</td>
<td>231.9</td>
<td>506.8</td>
<td>813.7</td>
<td>5142.1</td>
<td>7504.4</td>
<td>12646.5</td>
</tr>
<tr>
<td>0.1</td>
<td>9.94</td>
<td>255.9</td>
<td>235.0</td>
<td>448.3</td>
<td>313.9</td>
<td>490.9</td>
<td>762.2</td>
<td>5851.7</td>
<td>7234.1</td>
<td>13085.7</td>
</tr>
</tbody>
</table>

6. Conclusions

We extended the classical 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 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. We also 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 these cases are realistic is an empirical question, and therefore further empirical investigation of stochastic route preferences may therefore help to provide a more detailed estimate of the benefits of value pricing.
Bibliography


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

Define \( \bar{N}^\alpha_T = \sum_{k=1}^K \alpha_k N_{Tk} \) as the preference weighted average number of travellers at route \( T \), and \( \bar{N}^\alpha_U = \sum_{k=1}^K \alpha_k N_{uk} \) as the preference weighted number of travellers for route \( U \). Then (A.4) is given by:

\[
\mathcal{L} = \sum_{k=1}^K \int_0^{N_k} D_k(n_k) dn_k - \bar{N}^\alpha_T c_T(N_T) - \bar{N}^\alpha_U c_U(N_U) \\
- \sum_{k=1}^K \frac{1}{\theta_k} \left( N_{Tk} \ln \left( \frac{N_{Tk}}{N_k} \right) + N_{uk} \ln \left( \frac{N_{uk}}{N_k} \right) \right) + \\
\sum_{k=1}^K \lambda_{Tk} \left( D_k'(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln \left( \frac{N_{Tk}}{N_k} \right) \right) + \\
\sum_{k=1}^K \lambda_{Uk} \left( D_k'(N_k) - f_{Uk} - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln \left( \frac{N_{uk}}{N_k} \right) \right)
\]  
(A.1)

The first-order conditions are given by:

\[
\frac{\partial \mathcal{L}}{\partial N_{Tl}} = D_t(N_t) - \bar{N}^\alpha_T c_T(N_T) - \alpha_l c_T(N_T) - \frac{1}{\theta_l} \ln \left( \frac{N_{Tl}}{N_t} \right) - \sum_{k=1}^K \lambda_{Tk} \alpha_k c_T'(N_T) + \\
\lambda_{Tl} \left( D_t'(N_t) - \frac{1}{\theta_l} \frac{N_{Tl}}{N_{Ul} N_t} \right) + \lambda_{Ul} \left( D_l'(N_l) + \frac{1}{\theta_l} \frac{1}{N_l} \right) = 0, \forall l = 1\ldots K.
\]  
(A.2)

\[
\frac{\partial \mathcal{L}}{\partial N_{Ul}} = D_t(N_t) - \bar{N}^\alpha_U c_U(N_U) - \alpha_l c_U(N_U) - \frac{1}{\theta_l} \ln \left( \frac{N_{Ul}}{N_t} \right) - \sum_{k=1}^K \lambda_{Uk} \alpha_k c_U'(N_U) + \\
\lambda_{Ul} \left( D_t'(N_t) - \frac{1}{\theta_l} \frac{N_{Ul}}{N_{Tl} N_t} \right) + \lambda_{Tl} \left( D_l'(N_l) + \frac{1}{\theta_l} \frac{1}{N_l} \right) = 0, \forall l = 1\ldots K.
\]  
(A.3)

\[
\frac{\partial \mathcal{L}}{\partial f_{Tl}} = -\lambda_{Tl} = 0, \forall l = 1\ldots K.
\]  
(A.4)
The first-order conditions are given by:

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

\[
\frac{\partial L}{\partial \lambda_{Ti}} = D_i(N_i) - f_{Ti} - \alpha_i c_T(N_T) - \frac{1}{\theta_i} \ln \left( \frac{N_{Ti}}{N_i} \right) = 0, \forall l = 1...K. \tag{A.6}
\]

\[
\frac{\partial L}{\partial \lambda_{Ul}} = D_i(N_i) - f_{Ul} - \alpha_i c_U(N_U) - \frac{1}{\theta_i} \ln \left( \frac{N_{Ul}}{N_i} \right) = 0, \forall l = 1...K. \tag{A.7}
\]

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:

\[
f_{Ti} = \tilde{N}_T^\alpha c_T(N_T), \forall l = 1...K. \tag{A.8}
\]

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

\[
f_{Ul} = \tilde{N}_U^\alpha c_U(N_U), \forall l = 1...K. \tag{A.9}
\]

Because \( \tilde{N}_T^\alpha \) and \( \tilde{N}_U^\alpha \) 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:

\[
\mathcal{L} = \sum_{k=1}^{K} \int_{N_k}^{N} D_k(n_k)dn_k - \tilde{N}_T^\alpha c_T(N_T) - \tilde{N}_U^\alpha c_U(N_U)
- \sum_{k=1}^{K} \frac{1}{\theta_k} \left( N_{Tk} \ln \left( \frac{N_{Tk}}{N_k} \right) + N_{Uk} \ln \left( \frac{N_{Uk}}{N_k} \right) \right) + \\
\sum_{k=1}^{K} \lambda_{Tk} \left( D_k'(N_k) - f_{Tk} - \alpha_k c_T(N_T) - \frac{1}{\theta_k} \ln \left( \frac{N_{Tk}}{N_k} \right) \right) + \\
\sum_{k=1}^{K} \lambda_{Uk} \left( D_k'(N_k) - \alpha_k c_U(N_U) - \frac{1}{\theta_k} \ln \left( \frac{N_{Uk}}{N_k} \right) \right) \tag{B.1}
\]

The first-order conditions are given by:

\[
\frac{\partial \mathcal{L}}{\partial N_{Ti}} = D_i(N_i) - \tilde{N}_T^\alpha c_T(N_T) - \alpha_i c_T(N_T) - \frac{1}{\theta_i} \ln \left( \frac{N_{Ti}}{N_i} \right) - \sum_{k=1}^{K} \lambda_{Tk} \alpha_k c_T(N_T) + \\
\lambda_{Ti} \left( D_i'(N_i) - \frac{1}{\theta_i} \frac{N_{Ti}}{N_{Ul} N_i} \right) + \lambda_{Ul} \left( D_i'(N_i) + \frac{1}{\theta_i} \frac{1}{N_i} \right) = 0, \forall l = 1...K. \tag{B.2}
\]

\[
\frac{\partial \mathcal{L}}{\partial N_{Ul}} = D_i(N_i) - \tilde{N}_U^\alpha c_U(N_U) - \alpha_i c_U(N_U) - \frac{1}{\theta_i} \ln \left( \frac{N_{Ul}}{N_i} \right) - \sum_{k=1}^{K} \lambda_{Uk} \alpha_k c_U(N_U) + \\
\lambda_{Ul} \left( D_i'(N_i) - \frac{1}{\theta_i} \frac{N_{Ul}}{N_{Ul} N_i} \right) + \lambda_{Ti} \left( D_i'(N_i) + \frac{1}{\theta_i} \frac{1}{N_i} \right) = 0, \forall l = 1...K. \tag{B.3}
\]
\[\frac{\partial L}{\partial f_{Tl}} = -\lambda_{Tl} = 0, \forall l = 1...K. \quad (B.4)\]

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

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

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:

\[\frac{\partial L}{\partial N_{Ul}} = -\hat{N}_U^T \alpha^t c_U(N_U) - \sum_{k=1}^{K} \lambda_{Uk} \alpha_k c_U(N_U) + \lambda_{Ul} \left( D_l'(N_l) - \frac{1}{\theta_l} \frac{N_{Ul}}{N_{Tl} N_l} \right) = 0, \forall l = 1...K. \quad (B.7)\]

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

\[f_{Tl} = \hat{N}_T^T \alpha^t c_T(N_T) - \lambda_{Ul} \left( D_l'(N_l) + \frac{1}{\theta_l} \frac{N_{Ul}}{N_l} \right) \quad (B.8)\]

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:

\[A \lambda_U = b, \quad (B.9)\]

where \(\lambda_U\) is the \(K\times 1\) vector with unknown multipliers, \(b\) is the \(K\times 1\) vector with each element equal to \(\hat{N}_U^T \alpha^t c_U(N_U)\), and \(A\) is the following \(K\times K\) matrix:

\[A = \begin{bmatrix} D_1'(N_1) - \alpha_1 c_U'(N_U) - \frac{1}{\theta_1} \frac{N_{T1}}{N_1 N_U} & -\alpha_2 c_U'(N_U) & \cdots & -\alpha_K c_U'(N_U) \\ -\alpha_1 c_U'(N_U) & D_2'(N_2) - \alpha_2 c_U'(N_U) - \frac{1}{\theta_2} \frac{N_{T2}}{N_2 N_U} & \cdots & -\alpha_K c_U'(N_U) \\ \vdots & \vdots & \ddots & \vdots \\ -\alpha_1 c_U'(N_U) & -\alpha_1 c_U'(N_U) & \cdots & D_K'(N_K) - \alpha_K c_U'(N_U) - \frac{1}{\theta_K} \frac{N_{TK}}{N_K N_U} \end{bmatrix} \]

The solution for the vector \(\lambda_U\) can be found by Cramer’s rule. Let \(A_l(b)\) be the matrix \(A\) with column \(l\) replaced by the vector \(b\). The solution for the \(l\)th Lagrangian multiplier is given by a ratio of determinants:

\[\lambda_{Ul}^* = \frac{\text{det}(A_l(b))}{\text{det}(A)}, \quad (B.10)\]

and therefore we need \(\text{det}(A) \neq 0\) 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, it can be written in a tractable closed-form:

\[
\text{det}(A) = \prod_{k=1}^{K} \left[ D_k'(N_k) - \frac{1}{\theta_k} \frac{N_{Tk}}{N_{Uk} N_k} \right] - \sum_{k=1}^{K} \alpha_k c_U'(N_U) \prod_{m=1}^{K} \left[ D_m'(N_m) - \frac{1}{\theta_m} \frac{N_{Tm}}{N_{Um} N_m} \right].
\]

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

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

Because $D'_k(N_k) - \frac{1}{\theta_k N_U k N_k} < 0, \forall k = 1..K$, 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 $D'_m(N_m) - \frac{1}{\theta_m N_U m N_m} < 0$, 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 $\det(A_t(b))$:

$$\det(A_t(b)) = \tilde{N}_U^* c'_U(N_U) \prod_{i=1}^{K} \left[D'_i(N_i) - \frac{1}{\theta_i N_U i N_i}\right]$$  \hspace{1cm} (B.13)

We have $\det(A_t(b)) \neq 0$ 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:

$$\lambda^*_U = \frac{\tilde{N}_U^* c'_U(N_U) \prod_{k \neq k}^{K} \left[D'_k(N_k) - \frac{1}{\theta_k N_U k N_k}\right] \left(1 - \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m N_U m N_m}}\right)}{\prod_{k=1}^{K} \left[D'_k(N_k) - \frac{1}{\theta_k N_U k N_k}\right] \left(1 - \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m N_U m N_m}}\right)},$$  \hspace{1cm} (B.14)

which can be rewritten as:

$$\lambda^*_U = \tilde{N}_U^* c'_U(N_U) \frac{1}{\left(D'_i(N_i) - \frac{1}{\theta_i N_U i N_i}\right) \left(1 - \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m N_U m N_m}}\right)}.$$  \hspace{1cm} (B.15)

Taking the $l$th term out of the summation this reduces to:

$$\lambda^*_U = \tilde{N}_U^* c'_U(N_U) \frac{1}{D'_i(N_i) - \frac{1}{\theta_i N_U i N_i} \alpha_l c'_U(N_U) - c'_U(N_U) \sum_{k \neq l}^{K} \frac{D'_i(N_i) - \frac{1}{\theta_i N_U i N_i}}{D'_k(N_k) - \frac{1}{\theta_k N_U k N_k}} \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m N_U m N_m}}}.$$  \hspace{1cm} (B.16)

Substituting B.16 in B.8 gives:

$$f_{Tl} = \tilde{N}_T^* c'_T(N_T) - \tilde{N}_U^* c'_U(N_U) \frac{D'_i(N_i) + \frac{1}{\theta_i N_i}}{D'_i(N_i) - \frac{1}{\theta_i N_U i N_i} \alpha_l c'_U(N_U) - c'_U(N_U) \sum_{k \neq l}^{K} \frac{D'_i(N_i) - \frac{1}{\theta_i N_U i N_i}}{D'_k(N_k) - \frac{1}{\theta_k N_U k N_k}} \sum_{k=1}^{K} \frac{\alpha_k c'_U(N_U)}{D'_m(N_m) - \frac{1}{\theta_m N_U m N_m}}}.$$  \hspace{1cm} (B.17)

Multiplying the nominator and the denominator of the fractional part by $-1$ results in 34.
Appendix C. Calibration of the asymmetric route flows

If we want to calibrate the model in the no-toll case for given values of $\theta$, 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 38. 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 39 and assuming $\beta_T = \beta_U = \beta$ results in:

\[
\begin{align*}
\delta_1 - \delta_2(N_T + N_U) &= \kappa_T + \beta N_T + \frac{1}{\theta} \frac{N_T}{N}, \\
\delta_1 - \delta_2(N_T + N_U) &= ASC_U + \kappa_U + \beta N_T + \frac{1}{\theta} \frac{N_U}{N},
\end{align*}
\]

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

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

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:

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

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