Article:

Wang, JYT and Ehrgott, M (2013) Modelling route choice behaviour in a tolled road network with a time surplus maximisation bi-objective user equilibrium model. Transportation Research Part B: Methodological, 57. 342 - 360. ISSN 0191-2615

https://doi.org/10.1016/j.trb.2013.05.011

Modelling route choice behaviour in a tolled road network with a time surplus maximisation bi-objective user equilibrium model

Judith Y.T. Wang[a,b,*], Matthias Ehrgott[a,c]

[a] *Department of Engineering Science, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand*

[b] *School of Civil Engineering & Institute for Transport Studies, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK*

[c] *Department of Management Science, Lancaster University Management School, Bailrigg, Lancaster LA1 4YX, UK*

Abstract

In this paper, we propose a novel approach to model route choice behaviour in a tolled road network with a bi-objective approach, assuming that all users have two objectives: (1) minimise travel time; and (2) minimise toll cost. We assume further that users have different preferences in the sense that for any given path with a specific toll, there is a limit on the time that an individual would be willing to spend. Different users can have different preferences represented by this indifference curve between toll and time. Time surplus is defined as the maximum time minus the actual time. Given a set of paths, the one with the highest (or least negative) time surplus will be the preferred path for the individual. This will result in a bi-objective equilibrium solution satisfying the time surplus maximisation bi-objective user equilibrium (TSmaxBUE) condition. That is, for each O-D pair, all individuals are travelling on the path with the highest time surplus value among all the efficient paths between this O-D pair.

*Corresponding author. Tel.: +44 113 343 3259 ; Fax: +44 113 343 3265.

Email addresses: j.y.t.wang@leeds.ac.uk (Judith Y.T. Wang), m.ehrgott@lancaster.ac.uk (Matthias Ehrgott)

We show that the TSmaxBUE condition is a proper generalisation of user equilibrium with generalised cost function, and that it is equivalent to bi-objective user equilibrium. We also present a multi-user class version of the TSmaxBUE condition and demonstrate our concepts with illustrative examples.
Keywords: Traffic assignment, route choice, equilibrium problem, multi-objective optimisation

1. Introduction

The last stage of a conventional four-stage transport planning model, traffic assignment, is essentially modelling the route choice behaviour of travellers and their interactions. Whether a traffic assignment model can realistically represent travel behaviour is, therefore, dependent on the behavioural assumptions behind the route choice model. In tolling analysis, there are basically two approaches in practice as described in Florian (2006): (1) models based on generalised cost path choice; and (2) models based on explicit choice of tolled facilities. These two approaches follow the principles of the two classic traffic assignment models in the literature, namely, the user equilibrium (UE) model and the stochastic user equilibrium (SUE) model.

Wardrop (1952) defined *user equilibrium* as:

> "No user can improve his travel time by unilaterally changing routes."

This is known as Wardrop's first principle which has two key assumptions: (1) all users have the same objective, i.e. to minimise travel time or generalised cost; and (2) users have perfect knowledge of the network, i.e. they know the travel times that would be encountered on all available routes between their origin and destination. The second assumption is considered to be a strong assumption. Dial

(1971) was the first to introduce a probabilistic assignment concept to address this problem. He proposed a probabilistic multipath traffic assignment model based on the following functional principles:

1. The model gives all efficient paths between a given origin and destination a non-zero probability of use, while all inefficient paths have a probability of zero.
2. All efficient paths of equal length have an equal probability of use.
3. When there are two or more efficient paths of unequal length, the shorter has the higher probability of use.

The meaning of ‘efficient’ paths in Dial’s model is defined as a path that does not backtrack, i.e. as it progresses from node to node, it always gets further from the origin and closer to the destination. Every link in an efficient path has its initial node closer to the origin than its final node and its final node closer to the destination than its initial node. In this manner, the set of ‘efficient’ paths can be considered as the reasonable choices. By introducing *diversion curves*, Dial (1971) incorporated the logit function into his model which enables the solution to be expressed in explicit form. However, congestion effects have not been considered in this model as link travel time is assumed to be constant.

SUE was developed by Daganzo and Sheffi (1977) based on variation of the first assumption of Wardrop’s first principle by considering the objective as minimising the *perceived* cost which is modelled as a stochastic function rather than the static generalised cost function. Daganzo and Sheffi (1977) defined *stochastic user equilibrium* as:

“No user can improve his perceived travel time by unilaterally changing routes”

In order to translate this SUE equilibrium condition into its mathematical definition, Daganzo and Sheffi introduced a user's perceived travel time function on route k, $\tilde{T}_k$, which has two components as follows:

$$\tilde{T}_k = T_k + \epsilon_k, \tag{1}$$

where T_k is the systematic component which is the measured travel time on route k; and ϵ_k is an error term representing the random component which varies from user to user.

Here ϵ is randomly distributed with a mean value of zero. Thus,

$$E\left(\tilde{T}_k\right) = T_k. \tag{2}$$

Every user then evaluates the travel time on all routes and selects the route k_{min} with the minimum perceived travel time, i.e.

$$\tilde{T}_{k_{min}} \leq \tilde{T}_k \text{ for all } k \neq k_{min}. \tag{3}$$

The mathematical conditions for SUE within this modelling framework are formally defined in Daganzo and Sheffi (1977). The assumption on the distribution of the error term, ϵ_k, varies. The most commonly used distributions are Gumbell and normal distributions, known as the logit and probit models, respectively. The assumption of the error term following Gumbell/normal distributions is the key linkage of Dial (1971)'s probabilistic model to Discrete Choice Models, which led to further development of SUE traffic assignment models that appeared later in the literature such as Fisk (1980)'s logit-based model and Sheffi and Powell (1982)'s probit model. It is important to note that in order to take congestion effects into consideration, travel time should be flow-dependent. Fisk (1980) was the first to consider the effect of congestion in a stochastic manner, as travel time is considered to be independent of traffic flow in the previous models (Daganzo and Sheffi, 1977; Dial, 1971).

A disadvantage of the probit model is well known as the intensive computational effort requiring Monte Carlo or other numerical techniques (Maher, 1992; Rosa and Maher, 2002). Logit models have their weaknesses but a very important advantage of having a closed form solution. Thus, the most commonly used stochastic traffic assignment model for toll analysis is a logit-based model as described in Florian (2006). The key weakness of the most commonly used logit-based model is the validity of the property of independence of irrelevant alternatives (IIA), which can be stated as:

"Where any two alternatives have a non-zero probability of being chosen, the ratio of one probability over the other is unaffected by the presence or absence of any additional alternative in the choice set (Luce and Suppes, 1965)."

When it comes to modelling path choice, the IIA property can be easily violated because of extensive overlapping of possible paths in a choice set for the same origin-destination (OD) pair. Over the last two decades, there were extensive developments in stochastic route choice models trying to address this weakness. Prashker and Bekhor (2004) provide a comprehensive review of the developments. Since the perceived cost function has two components as shown in Equation (1), this problem can be addressed by tackling either the systematic or the error component. In principle, the technique being used is to make adjustments to these two components such that the resulting solution reflects reality better. Prashker and Bekhor (2004) classified the techniques into three categories: (1) modifications of the basic multinomial logit (MNL) model, such as C-logit and path-size logit (PSL); (2) generalised extreme value (GEV) models, such as paired combinatorial logit (PCL) and cross-nested logit (CNL); and (3) logit kernel (LK) or mixed logit models. The first category adjusts the systematic component while the second and the third adjust the error component.

In this paper, we propose a novel approach to model route choice behaviour in a tolled network. We extend our work in Wang et al. (2010) on bi-objective traffic assignment to incorporate the capability to model the differences between individuals in terms of their willingness to pay. First of all, we assume that all users have two objectives: (1) to minimise travel time; and (2) to minimise toll cost. Users are all rational in the sense that given a choice set, they will only choose one of the *efficient* paths. Efficient paths are defined as the set of paths for each O-D pair for which neither time nor travel time can be improved without worsening the other (Wang et al., 2010). According to this definition, at equilibrium, all the used paths between a given O-D pair are *efficient*. We define *bi-objective user equilibrium* (BUE) as follows:

> "Under *bi-objective user equilibrium* conditions traffic arranges itself in such a way that no individual trip maker can improve either his/her toll or travel time or both without worsening the other objective by unilaterally switching routes."

Dial (1979) is one of the first to introduce multiple objectives in traffic assignment. According to BUE, when we consider time and toll cost separately, there is no need to add them up as generalised cost. However, in Dial's model (Dial, 1979, 1996, 1997), a simplification was made by adding time and toll cost in a linear choice function, which is essentially the same as the generalised cost function, but with a probabilistic component by assuming that the value-of-time (VOT) follows a certain probability density function. As discussed in Wang et al. (2010), Dial's approach might miss out some efficient paths. In Wang et al. (2010); Raith et al. (2013), we developed heuristics to find BUE solutions without missing efficient paths. It is clear that according to the BUE definition, there would be many possible equilibrium solutions rather than one as in conventional static UE. Given there are so many possible equilibrium solutions satisfying the BUE condition, we must further develop this model to incorporate the consideration of individual preferences in order to be able to replicate their route choice behaviour more realistically.

There is no doubt that route choice behaviour in a tolled road network is stochastic in nature since individuals might not choose the shortest path for all sorts of reasons and the willingness to pay would vary among individuals. As discussed above, probabilistic models such as Dial (1979)'s or the logit-based SUE traffic assignment models such as Fisk (1980)'s all possess some deficiencies. The philosophy behind the proposed model is to overcome these difficulties, including the possibility of missing efficient paths in Dial (1979)'s model and the limitations induced by the IIA property of the logit-based SUE traffic assignment model, by introducing an indifference function which can vary between individuals with no restrictions. As with any models, there are, however, some key assumptions to be made:

1. Users are all rational in the sense that they will only choose one of the *efficient* paths.
2. Users have different *preferences* which can be represented by an indifference function between toll and time. Users' behaviour as represented by this indifference function is rational, i.e. the maximum time that a user is willing to spend will always be shorter for higher toll.
3. Preferences among users vary in the sense that their preferred paths can be different, even though they are considering the same choice set.
4. Users have perfect knowledge of the network, as in standard user equilibrium models.

With this new approach, each individual will only choose from a reasonable choice set and choose according to his/her own preference.

This paper is organised as follows. In Section 2, we review standard user equilibrium for traffic assignment. In Section 3, we introduce bi- and multi-objective user equilibrium and investigate their relationship with single objective user equilibrium. Section 4 is devoted to the description of our new concept of time surplus maximisation bi-objective user equilibrium. We show that this generalises user equilibrium based on generalised cost functions and prove its equivalence to bi-objective user equilibrium. Section 5 provides an illustrative example of the idea, whereas Section 6 extends the idea to multiple user classes, which is then illustrated in Section 7. Finally, Section 8 discusses the importance of the findings in this paper and Section 9 concludes with an outlook for future research.

2. User Equilibrium

In this section we introduce equilibrium models of traffic assignment. Let $G = (N, A)$ denote a (transportation) network, where N is a set of $|N|$ nodes and $A \subset N \times N$ is a set of $|A|$ arcs or links. Moreover, let $Z \subset N \times N$ be a set of origin-destination pairs (O-D pairs) and for all $p \in Z$, let D_p denote the demand for travel between the origin and destination of O-D pair p. Equilibrium models attempt to determine the amount of traffic f_a on all links $a \in A$ under some assumptions on the behaviour of road users. One of these assumptions is that road users choose the path k^* between their origin and destination that minimises a cost function C_k:

$$k^* = \operatorname{argmin}\{C_k : k \in K_p\},$$

where K_p is the set of all simple paths from the origin of O-D pair p to its destination.

To formalise the idea of user equilibrium, let δ_a^k be an indicator with $\delta_a^k = 1$ if and only if link a is contained in path k and 0 otherwise. Then $f_a = \sum_{p \in Z} \sum_{k \in K_p} \delta_a^k F_k$, where F_k is the flow on path $k \in K_p$. The cost $C_k(\mathbf{F})$ of path k may depend on the entire vector $\mathbf{F} = (F_1, \ldots, F_{|K|})$ of flows on all paths $k \in K := \cup_{p \in Z} K_p$. The user equilibrium condition of Wardrop's first principle states that the cost of all used paths is equal and less than that which would be experienced by a single user on any unused route. It is well known that this principle assumes that all users are the same in that they want to minimise the cost C_k and that all users have perfect information about the cost function, see e.g. Sheffi (1985).

Let $U_p := \min_{k \in K_p} C_k(\mathbf{F})$ denote the minimum cost of any path for O-D pair $p \in Z$. Then, following e.g. Florian and Hearn (1995), the user equilibrium condition can be written mathematically as follows: Path flow vector $\mathbf{F}^*$ is an *equilibrium flow* if $\mathbf{F}^*$ satisfies conditions (4) – (8):

$$\begin{aligned}
F_k^*(C_k(\mathbf{F}^*) - U_p) &= 0 \quad \text{for all } k \in K_p \text{ and all } p \in Z, && (4)\\
C_k(\mathbf{F}^*) - U_p &\geq 0 \quad \text{for all } k \in K_p \text{ and all } p \in Z, && (5)\\
\sum_{k \in K_p} F_k^* - D_p &= 0 \quad \text{for all } p \in Z, && (6)\\
F_k^* &\geq 0 \quad \text{for all } k \in K, && (7)\\
U_p &\geq 0 \quad \text{for all } p \in Z. && (8)
\end{aligned}$$

Equation (4) states that if flow on path k is positive then the cost $C_k(\mathbf{F}^*)$ has to be minimal, whereas if $C_k(\mathbf{F}^*) > U_p$ then the flow on path k must be 0. Equation (5) says that all path costs are greater than or equal to the minimum. Equation (6) guarantees that demand is satisfied, whereas equations (7) and (8) postulate non-negativity of flow and cost. For future use, let us introduce

$$\Omega := \{\mathbf{F} : \mathbf{F} \text{ satisfies } (6) - (7)\} \tag{9}$$

to denote the set of all feasible path flow vectors $\mathbf{F}$.

Existence of a solution of the network equilibrium model (4) – (8) is guaranteed if the path cost functions $C_k(\mathbf{F})$ are all positive and continuous. In addition, for uniqueness of the solution, $C_k(\mathbf{F})$ must be strictly monotone (Florian and Hearn, 1995).

The most important cost function is travel time. In this paper we use the common Bureau of Public Roads (1964) function to model the relation between travel time and traffic flow on any link $a \in A$, i.e.

$$t_a\left(f_a\right) = t_a^0 \left[1 + \alpha \left(\frac{f_a}{C_a}\right)^{\beta}\right], \tag{10}$$

where t_a^0 is the free-flow travel time on link a, C_a is the practical capacity of link a in vehicles per time unit, and α, β are function parameters. If the cost function C_k considered in (4) – (8) is path travel time, then

$$C_k(\mathbf{F}) = T_k(\mathbf{F}) := \sum_{a \in k} t_a(f_a) \tag{11}$$

for all $k \in K$.

Conventional traffic assignment assumes that path cost functions $C_k(\mathbf{F})$ are additive and separable. Additivity means that $C_k(\mathbf{F}) = \sum_{a \in k} c_a(\mathbf{f})$ can be written as the sum of link cost functions $c_a(\mathbf{f})$, where $\mathbf{f} := (f_1, \ldots, f_{|A|})$ is the link flow vector. Separability means that the link cost functions $c_a(\mathbf{f})$ depend only on the flow f_a on link a, i.e. $c_a(\mathbf{f}) = c_a(f_a)$.

Under these assumptions it is well known (Beckmann et al., 1956) that the network equilibrium model (4) – (8) can be reformulated as a mathematical programme

$$\min \sum_{a \in A} \int_0^{f_a} c_a(x)dx, \tag{12}$$

$$\text{subject to} \sum_{k \in K_p} F_k = D_p \quad \text{for all } p \in Z, \tag{13}$$

$$F_k \geq 0 \quad \text{for all } k \in K, \tag{14}$$

$$f_a - \sum_{p \in Z} \sum_{k \in K_p} \delta_a^k F_k = 0 \quad \text{for all } a \in A. \tag{15}$$

Conventional traffic assignment based on travel time can, therefore, be solved by algorithms for optimising a convex function over a polyhedron. The first algorithm used for traffic assignment is the Frank-Wolfe algorithm (Frank and Wolfe, 1956) but many others such as path equilibration (Dafermos and Sparrow, 1969), gradient projection (Jayakrishnan et al., 1994) and projected gradient (Florian et al., 2009) methods have been proposed.

Many researchers have suggested more general cost functions than travel time, see e.g. Chen et al. (2010); Larsson et al. (2002). Most often $C_a(\mathbf{F})$ takes the form of a *generalised cost function* that incorporates a linear combination of travel time and a monetary component (Dial, 1996; Leurent, 1993). A generalised cost function is of the form

$$C_k(\mathbf{F}) = M_k(\mathbf{F}) + \alpha T_k(\mathbf{F}), \tag{16}$$

where $M_k(\mathbf{F})$ is the monetary cost associated with path k. This may be composed of different factors such as toll cost and vehicle operating costs. In addition, α is a value of time, i.e. it converts the travel time $T_k(\mathbf{F})$ into a monetary value. To solve traffic assignment problems with generalised cost function (16), one can apply the same algorithms as for conventional traffic assignment, depending on the properties of function $M_k(\mathbf{F})$. We note, however, that if $C_k(\mathbf{F})$ is not additive, it is necessary to calculate shortest paths based on non-additive costs (Gabriel and Bernstein, 1997), a research topic in its own right.

3. Bi-Objective User Equilibrium

The generalised cost function (16) combines a monetary component and travel time into a single function via value of time α. It is reasonable to assume that not all users will have the same value of time, so that a user will choose the route that minimises the generalised cost (16) with a user specific value of time. Dial (1979) realised this and interpreted the problem as bi-objective problem: Users would be wanting to minimise both travel time and monetary cost. He observed that at equilibrium, all used path will be efficient.

Definition 1. *Let $\mathbf{F} \in \Omega$ be a feasible flow and $M_k(\mathbf{F})$ and $T_k(\mathbf{F})$ be the monetary and time components of the cost of path k for all $k \in K_p$.*

1. *Path k is* efficient*, if there is no path $k' \in K_p$ such that $M_{k'}(\mathbf{F}) \leq M_k(\mathbf{F})$ and $T_{k'}(\mathbf{F}) \leq T_k(\mathbf{F})$ with at least one inequality being strict.*
2. *If $M_{k'}(\mathbf{F}) \leq M_k(\mathbf{F})$ and $T_{k'}(\mathbf{F}) \leq T_k(\mathbf{F})$ with at least one strict inequality then path k'* dominates *path k and cost vector $(T_{k'}(\mathbf{F}), M_{k'}(\mathbf{F}))$ dominates $(T_k(\mathbf{F}, M_k(\mathbf{F}))$.*

Dial (1979) describes this idea and an algorithm to find the efficient paths which makes use of the generalised cost function (16) with flow independent objectives. Leurent (1993) applies the idea in traffic assignment and designs an algorithm to compute the equilibrium in a tolled road network with toll cost and time as the objectives, where only time is flow dependent. As in Dial (1979), Leurent (1993) assumes that users make their route choice decisions based on a generalised cost function and a continuous value of time distribution is considered. Dial (1996, 1997) further develops his idea of 1979 into more efficient algorithms to find the efficient paths and to solve the bi-objective equilibrium problem in which both criteria can be flow dependent.

As we have demonstrated in Wang et al. (2010), and as the example in Section 5 shows, the procedures of Dial (1997) and Leurent (1993) will only compute equilibrium flows that allow positive flows on a subset of all efficient paths, namely those that are shortest path with respect to the generalised cost function (16) for some positive value of α. Since all efficient paths can be rational route choices, the work of Leurent and Dial appears to be limited by the use of the functional form (16) and its underlying assumption of an additive utility function. Removing this form and allowing for more than two objectives one arrives at the definition of multi-objective user equilibrium.

Definition 2. *Let $G = (N, A)$ be a network, $Z \subset N \times N$ be a set of O-D pairs and for all $p \in Z$, let D_p be the demand of O-D pair p. Let $C_k^{(i)}(\mathbf{F}), i = 1, \ldots, r$ be r cost functions of path k and let $\mathbf{C}_k(\mathbf{F})$ denote the cost vector of path k. Feasible flow $\mathbf{F}^* \in \Omega$ is a* multi-objective equilibrium flow*, if whenever $\mathbf{C}_k(\mathbf{F}^*)$ dominates $\mathbf{C}_{k'}(\mathbf{F}^*)$ for $k, k' \in K_p$ for any $p \in Z$ then $F_{k'} = 0$.*

Definition 2 is the multi-objective generalisation of the equilibrium conditions (4) – (8). Only efficient paths can carry positive flow, whereas dominated paths have zero flow. If $r = 2$, we talk about bi-objective equilibrium flow. In the case of $r = 1$, Definition 2 reduces to the standard equilibrium condition. In Wang et al. (2010) we have shown that even for the case $r = 2$ and even if both objectives are separable and additive, and one of the objectives does not depend on flow, the multi-objective user equilibrium condition is not equivalent to a multi-objective version of Beckmann's formulation (12) – (15). A discussion of the similarities and differences of multi-objective equilibrium, optimisation, and vector inequality problems is provided in Raith and Ehrgott (2011).

Moreover, there are usually infinitely many flow vectors $\mathbf{F} \in \Omega$ that satisfy the condition of Definition 2.

The concept of multi-objective user equilibrium therefore provides a general framework for investigating equilibrium flows in the presence of multiple objectives. Under the assumption that the objectives considered are those relevant for users' route choice, one of these multi-objective equilibrium solutions will be realised in practice. Which one that is will depend on user preferences and trade-offs between the objectives. The simplest model of user preferences is the additive form as shown in (16). In this paper, we develop a more general model. But before we proceed to this model, we formally show that the multi-objective user equilibrium is a proper generalisation of the single objective model with generalised cost (16).

Proposition 1. *Let $G = (N, A)$ be a network, $Z \subset N \times N$ be a set of O-D pairs and for all $p \in Z$, let D_p be the demand of O-D pair p. Let $r = 2$. Let $\mathbf{F}^*$ be an equilibrium flow with respect to generalised cost function $C(\mathbf{F}) := C^{(1)}(\mathbf{F}) + \alpha C^{(2)}(\mathbf{F})$ for some positive number α. Then $\mathbf{F}^*$ is also a bi-objective equilibrium flow for objective functions $\mathbf{C}^{(1)}(\mathbf{F})$ and $\mathbf{C}^{(2)}(\mathbf{F})$.*

PROOF. Assume the contrary. Then there exists a $p \in Z$ and two paths $k, k' \in K_p$ with $F_k, F_{k'} > 0$ such that $\mathbf{C}_{k'}(\mathbf{F})$ dominates $\mathbf{C}_k(\mathbf{F})$. Then, because $\alpha > 0$, it holds that

$$C_{k'}^{(1)}(\mathbf{F}) + \alpha C_{k'}^{(2)}(\mathbf{F}) < C_k^{(1)}(\mathbf{F}) + \alpha C_k^{(2)}(\mathbf{F}),$$

contradicting the equilibrium condition for the generalised cost function.

In fact, with the same argument, it is possible to show a more general result.

Theorem 1. *Let $G = (N, A)$ be a network, $Z \subset N \times N$ be a set of O-D pairs and for all $p \in Z$, let D_p be the demand of O-D pair p. Let $g : \mathbb{R}^r \to \mathbb{R}$ be a strictly increasing function in all r arguments. Let $\mathbf{F}^*$ be an equilibrium flow with respect to generalised cost function $C(\mathbf{F}) := g(\mathbf{C}(\mathbf{F}))$. Then $\mathbf{F}^*$ is also a multi-objective equilibrium flow for objective functions $C^{(1)}(\mathbf{F}), \ldots, C^{(r)}(\mathbf{F})$.*

In Section 4, we will address the case of equilibrium problems with $r = 2$ objectives, where $C_k^{(1)}(\mathbf{F}) = T_k(\mathbf{F})$ and $C_k^{(2)}(\mathbf{F}) = M_k(\mathbf{F})$ and the monetary objective consists of exogenously defined tolls. We investigate bi-objective user equilibrium for these functions and a specific nonlinear function to combine them.

4. The Time Surplus Maximisation Model

In this section, we develop the time surplus maximisation equilibrium model as a new model for route choice behaviour in tolled road networks. From now on, we will consider two objective functions, namely, travel time $C_k^{(1)}(\mathbf{F}) = T_k(\mathbf{f}) = \sum_{a \in k} t_a(x_a)$, where $t_a(x_a)$ is the travel time function (10) and toll $C_k^{(2)}(\mathbf{F}) = M_k(\mathbf{f}) = \tau_k = \sum_{a \in k} \tau_a$, with exogenously defined link tolls τ_a. Hence, both path objectives (travel time and toll) are additive, link travel time and link toll are separable, and link toll does not depend on flow.

4.1. The Indifference Function

To start with, we assume that given an O-D pair p, each user has an indifference function between toll and time. For any given path k with a specific toll, there is a limit on the time that a user would be willing to spend. We model this indifference function as a function $T_p^{max} : \mathbb{R} \to \mathbb{R}$ that is strictly decreasing, i.e. $T_p^{max}(\tau_k^1) < T_p^{max}(\tau_k^2)$ if $\tau_k^1 > \tau_k^2$. This takes into account that users would expect to spend less time in traffic if they need to pay a higher toll. An example of an indifference curve is shown in Figure 1.

Time surplus is defined as the time that the user would be willing to spend minus the actual travel time. The time surplus for a path can be positive or negative. Given a choice set of paths, the one with the highest time surplus will be the preferred path for the individual.

A *positive* time surplus value can be viewed as virtually the pleasure for an individual obtained from choosing this path, whereas a *negative* time surplus value can represent an unfavourable choice and the magnitude of this path being disliked. One would expect that given a set of efficient paths with both positive and negative time surplus values, only the one with positive time surplus values will be considered. For example, an individual with an indifference curve as shown in Figure 1

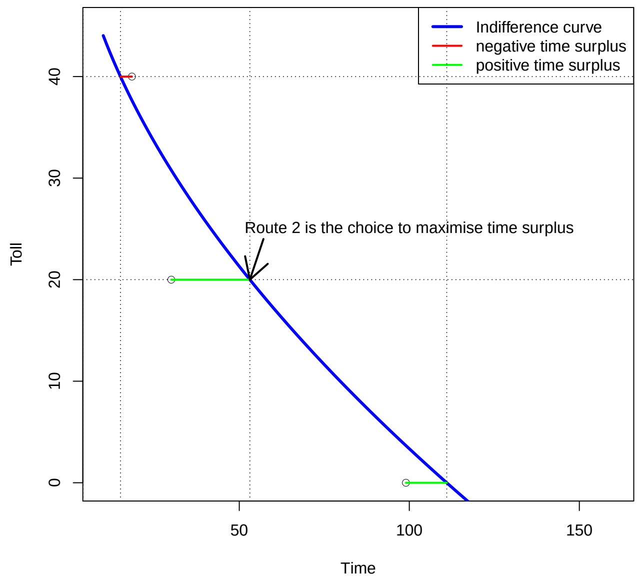

Figure 1: An indifference curve between time and toll.

will only consider the two paths that have positive time surplus, i.e the ones with $\tau_k = 20$ and the one with $\tau_k = 0$. Among these two, the one with $\tau_k = 20$ is considered more attractive as the time surplus value is higher.

There is, however, the possibility that all the efficient paths have negative time surplus values for a user who is both unwilling to pay and to spend time. In that case, we will have to assume either this user would not travel at all or will have to make a choice based on the negative values. In this paper, we assume that the total demand is inelastic and hence the user will choose the path with the least negative time surplus value.

4.2. The Time Surplus Maximisation BUE Condition

Given the indifference curves T_p^{max} for all $p \in Z$, we define time surplus for path $k \in K_p$ as

$$TS_k(\mathbf{F}) := T_p^{max}(\tau_k) - T_k(\mathbf{f}) = T_p^{max}\left(\sum_{a \in k} \tau_a\right) - \sum_{a \in k} t_a(f_a). \tag{17}$$

.

We note that function $TS_k(\mathbf{F})$ is not additive because $T_p^{max}(\tau)$ is only defined for OD pair p but neither for paths nor for links and, therefore, cannot be written as the sum of link indifference functions. Moreover, T_P^{max} may be non-linear. Hence, all equilibrium models using this function will be path based. Assuming that users choose the path k^* with maximum time surplus, i.e.

$$k^* = \mathrm{argmin}\{TS_k(\mathbf{F}) : k \in K_p\}, \tag{18}$$

we can now formulate the Time Surplus Maximisation Bi-objective User Equilibrium (TSmaxBUE) condition.

Definition 3. *Path flow vector* $\mathbf{F}^*$ *is called a* time surplus maximisation bi-objective user equilibrium *flow if* $F_k > 0 \Rightarrow TS_k^{max}(\mathbf{F}^*) \geq TS_{k'}^{max}(\mathbf{F}^*)$ *for all* $k, k' \in K_p$*, or equivalently, if* $T_k^{max}(\mathbf{F}) > TS_{k'}^{max}(\mathbf{F}) \Rightarrow F_{k'} = 0$.

In words, the TSmaxBUE condition states that

"Under the *Time Surplus Maximisation equilibrium* condition traffic arranges itself in such a way that no individual trip maker can improve his/her time surplus by unilaterally switching routes,"

or alternatively

"Under the *Time Surplus Maximisation equilibrium condition* all individuals are travelling on the path with the highest time surplus value among all the efficient paths between each O-D pair."

Next, we show that the TSmaxBUE model is a special case of the general multi-objective user equilibrium model of Definition 2, but that it is more general than the single objective user equilibrium with generalised cost function (16) based on value of time.

Theorem 2. *Let $G = (N, A)$ be a network, $Z \subset N \times N$ be a set of O-D pairs with demand $D_p > 0$ for all $p \in Z$. Let τ_a denote the toll of link a and $t_a(f_a)$ be the travel time function of link a. Assume that $\mathbf{F}^*$ is a TSmaxBUE flow. Then $\mathbf{F}^*$ is also a bi-objective equilibrium flow with respect to the objectives $C^{(1)}(\mathbf{F}) = T_k(\mathbf{f})$ and $C^{(2)}(\mathbf{F}) = \tau_k$.*

PROOF. We have to show that all paths k with $\mathbf{F}_k^* > 0$ are efficient paths with respect to $C^{(1)}$ and $C^{(2)}$. So assume that $\mathbf{F}^*$ is such that there is some $p \in Z$ and $k, k' \in K_p$ such that $\mathbf{C}(F_k)$ dominates $\mathbf{C}(F)^*_{k'}$. That is, $T_k(F_k^*) \leq T_{k'}(F_{k'}^*)$ and $\tau_k \leq \tau_{k'}$ with one strict inequality.

Then we have

$$TS_k(F_k^*) = T_p^{max}(\tau_k) - T_k(F_k^*) > T_{k'}^{max}(\tau_{k'}) - T_{k'}(F_{k'}^*) = TS_{k'}(F_{k'}) \quad (19)$$

because of the dominance and because TS^{max} is a strictly decreasing function. Clearly (19) contradicts the assumption that $\mathbf{F}^*$ is a TSmaxBUE flow.

It is even possible to prove the converse of Theorem 2.

Theorem 3. *Let $G = (N, A)$ be a network, $Z \subset N \times N$ be a set of O-D pairs with demand $D_p > 0$ for all $p \in Z$. Let τ_a denote the toll of link a and $t_a(f_a)$ be the travel time function of link a. Assume that $\mathbf{F}^*$ is a bi-objective equilibrium flow, with respect to objectives $C^{(1)}(\mathbf{F})$ and $C^{(2)}(\mathbf{F})$ as in Theorem 2. Then there exists an indifference function T^{max} such that $\mathbf{F}^*$ is also a TSmaxBUE flow.*

PROOF. Let $\mathbf{F}^*$ be a bi-objective equilibrium flow. According to Definition 2, all paths with positive flow are efficient. Let K_p^* be the set of all efficient paths for O-D pair $p \in Z$. Then for paths $k, k' \in K_p^*$ we have that $\tau_k > \tau_{k'}$ implies $T_k(F_k) < T_{k'}(F_{k'})$ and can therefore order the paths in $K_p^* = \{1, \ldots |K_p^*|\}$ in such a way that $\tau_k > \tau_{k'}$ and $T_k < T_{k'}$ if and only if $k > k'$.

In case there is no efficient path with $\tau_k = 0$ or $\tau_k = \max\{\tau_k : k \in K\}$ we add (one of) the points $(\tau_0 = 0, T_0 = \max\{T_k(D_p) : k \in K, p \in Z\})$ and $(\tau_{|K_p^*|+1} = \max\{\tau_k : k \in K\}, T_{|K_p^*|+1} = 0)$ to the sequence (τ_k, T_k). We define $T^{max}(\tau)$ as the uniquely determined piecewise linear function through the points $(\tau_k, T_k), k = 0, \ldots |K_p^*| + 1$. Clearly $T^{max}(\tau)$ is strictly decreasing and non-negative.

Now observe that for $\mathbf{F}^*$ we have that $TS_k(\mathbf{F}^*) = 0$ for all efficient paths $k \in K_p^*$. It remains to show that there does not exist a path with positive time surplus. To see this, assume that $l \in K_p$ is such a path. $TS_l(\mathbf{F}^*) > 0$ implies that $T_l(F_l^*) < TS^{max}(\tau_l)$. Then either $(\tau_l, T_l(F_l^*)) < (\tau_k, T_k(F_k^*))$ for some $k \in K_p^*$, contradicting the definition of K_p^* or there are $k_1, k_2 \in K_p^*$ such that $\tau_{k_1} < \tau_l < \tau_{k2}$ and $T_{k_1}(F_{k_1}^*) > T_l(F_l^*) > T_{k_2}(F_{k_2}^*)$. In this case, path l does not dominate nor is it dominated by any path k in K_p^*. Hence path l is itself efficient, therefore used in the definition of T^{max}, which implies $TS_l = 0$.

Theorems 2 and 3 imply that the time surplus maximisation equilibrium concept is equivalent to the bi-objective user equilibrium, although, of course, the function TS^{max} is in general not known beforehand. We notice that this function is piecewise linear, non-negative and continuous, but in general neither convex nor concave. Concavitiy/convexity of the indifference curve TS^{max} indicates willingness/reluctance to pay, so that TSmaxBUE equilibrium flows with concave/convex indifference curves will form a subset of all bi-objective equilibrium flows that is more realistic than arbitrary decreasing indifference curves.

The next result shows that every equilibrium flow with respect to generalised cost function $C(\mathbf{F}) = \tau_k + \alpha T_k(F_k)$, where $\alpha > 0$ is a positive constant, is also a TSmaxBUE flow.

Theorem 4. *Let $G = (N, A)$ be a network, $Z \subset N \times N$ be a set of O-D pairs with demand $D_p > 0$ for all $p \in Z$. Let τ_a denote the toll of link a and $t_a(f_a)$ be the travel time function of link a. Assume that $\mathbf{F}^*$ is an equilibrium flow with respect to the generalised cost objective $C(\mathbf{F}) = \tau_k + \alpha T_k(\mathbf{f})$. Then there exists an indifference curve T^{max} such that $\mathbf{F}^*$ is also a TSmaxBUE flow.*

PROOF. Let $\mathbf{F}^*$ be an equilibrium flow with respect to C and for all $p \in Z$ define $T_p^{max}(\tau) := a_0 - \frac{1}{\alpha}\tau$ for some $a_0 > 0$, e.g $a_0 = \max\{T_k(D_p) : k \in K_p\}$. We need to show that for any pair of paths k and k', with time surplus $TS_k(\mathbf{F})$ defined using the just defined functions $T_p^{max}(\tau)$, $TS_k(F_k) > TS_{k'}(F_{k'})$, implies that $F_{k'} = 0$.

$$\begin{aligned}
TS_k(F_k) &> TS_{k'}(F_{k'}) \quad \Leftrightarrow \\
a_0 - \frac{1}{\alpha}\tau_k - T_k(F_k) &> a_0 - \frac{1}{\alpha}\tau_{k'} - T_{k'}(F_{k'}) \quad \Leftrightarrow \\
\frac{1}{\alpha}\tau_{k'} + T_{k'}(F_{k'}) &> \frac{1}{\alpha}\tau_k + T_k(F_k) \quad \Leftrightarrow \\
\tau_{k'} + \alpha T_{k'}(F_{k'}) &> \tau_k + \alpha T_k(F_k) \quad \Leftrightarrow \\
C(F_{k'}) &> C(F_k)
\end{aligned}$$

Hence, the equilibrium condition for generalised cost function C implies that $F_{k'} = 0$.

The proof of Theorem 4 reveals that any generalised cost equilibrium flow is a special case of a TSmaxBUE flow, with the choice of a linear indifference curve T^{max}. Notice that only the slope $1/\alpha$ of this curve is important, but not its axis intercept a_0. In the example of Section 5, we will see that the converse of Theorem 4 does not hold. We can therefore summarise the relationships between generalised cost equilibrium, time surplus maximisation equilibrium and bi-objective equilibrium in Figure 2.

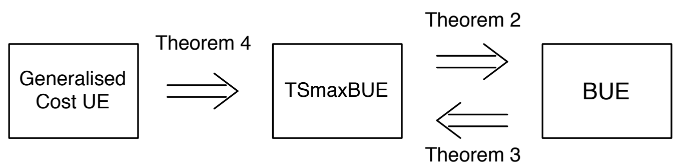

Figure 2: The relationship between equilibrium concepts discussed in this paper.

The proof of Theorem 2 shows that the time surplus of a dominated path is never better than that of any efficient path dominating it, we only include *efficient* paths in the choice set which gives us a reasonable choice set. We also note that the time surplus maximisation BUE model basically follows similar functional principles as outlined in Dial (1971).

1. Traffic will only be assigned to *efficient* paths. Note that we define *efficient* paths differently but basically the meaning of our definition also identifies the set of reasonable choices.
2. All dominated (inefficient) paths will have zero probability of use.
3. If there are two or more efficient paths, the one with the highest time surplus will be chosen.

We believe that single objective equilibrium models based on generalised cost functions of the form (16) are restrictive, because they essentially imply, as Theorem 4 shows, a linear indifference curve between toll and time. Moreover, (Dial, 1997) and (Leurent, 1993) in fact violate the first functional principle above, because some efficient paths in the sense of Definition 1 will always have zero flow.

It is more realistic to assume that there will be users who are willing to pay to ensure short travel times, whereas others may be reluctant to pay any tolls, and would accept high travel times in order to avoid tolls. The latter would have convex indifference curves, while the former users' indifference curves will be concave. Hence, the variability between individuals in terms of willingness to pay is modelled by the indifference function which leads to their differences in behaviour. We can now classify the various types of equilibrium flow as in Figure 3. Generalised user equilibrium flows with cost function (16) are TSmaxBUE equilibrium flows with linear indifference curves. More general TSmaxBUE equilibrium flows are generated by convex or concave indifference curves, whereas all bi-objective user equilibrium flows are TSmaxBUE equilibrium flows with arbitrary strictly decreasing indifference curves. The proof of Theorem 4 shows that such a curve may be neither convex nor concave.

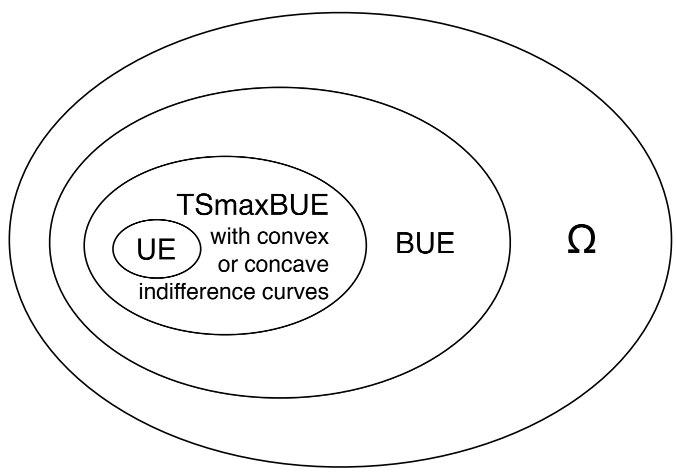

Figure 3: Several classes of equilibrium flows.

We introduce the TSmaxBUE concept with multiple user classes in Section 6, but first, we briefly address solving the TSmaxBUE traffic assignment problem and present a small illustrative example.

In order to be able to solve the TSmaxBUE problem, we use the framework of a *generalised time function* as introduced in Larsson et al. (2002). Larsson et al. (2002) consider time based traffic equilibrium, where users minimise travel time T_k and monetary cost τ_k via a a generalised time function

$$\theta_k = T_k + g(\tau_k), \tag{20}$$

where $g : \mathbb{R} \to \mathbb{R}$ is a nonlinear function, called the *time equivalent of money*. Larsson et al. (2002) showed that the equilibrium problem with generalised time (20) is equivalent to an optimisation problem.

We introduce the following function $g : \mathbb{R} \to \mathbb{R}$:

$$g(x) = h(0) - h(x), \tag{21}$$

where $h : \mathbb{R} \to \mathbb{R}$ is a strictly decreasing function on $\mathbb{R}_0^+$. Clearly, g is a strictly increasing function of x on $\mathbb{R}_0^+$. We substitute T^{max} for h and define the path cost function

$$C_k(\mathbf{F}) := \sum_{a \in k} t_a(f_a) + g(\tau_k) = \sum_{a \in k} t_a(f_a) + T^{max}(0) - T^{max}(\tau_k). \tag{22}$$

We observe that because T^{max} is a strictly decreasing function of τ_k, maximising time surplus is equivalent to minimising C_k. Moreover, $C_k(\mathbf{F})$ is positive because $T^{max}(\tau) > 0$ for any $\tau \geq 0$ and because travel times are positive. Path cost function $C^k(F_k)$ in equation (22) is therefore a generalised time function of form (20), and we can apply the results of Larsson et al. (2002) and formulate the time surplus maximisation equilibrium problem as a single objective equilibrium problem with generalised time function (22). Applying the results of Larsson et al. (2002), it follows that this equilibrium problem is equivalent to the optimisation problem (23) – (26), which under our assumptions satisfies the conditions for unique link flow solutions in Larsson et al. (2002).

$$\min \sum_{a\in A} \int_0^{f_a} t_a(x)dx + \sum_{p\in Z} \sum_{k\in K_p} F_k g(\tau_k) \tag{23}$$

$$\begin{aligned} \sum_{k\in K_p} F_k &= D_p \quad \text{for all } p \in Z, & (24)\\ F_k &\geq 0 \quad \text{for all } k \in K, & (25)\\ f_a - \sum_{p\in Z} \sum_{k\in K_p} \delta_a^k F_k &= 0 \quad \text{for all } a \in A. & (26) \end{aligned}$$

In Section 5, we provide an example illustrating the time surplus maximisation BUE concept.

5. A Four Node Example

5.1. Network Specification

Now we consider a four node network as shown in Figure 4 with link characteristics as shown in Table 1.

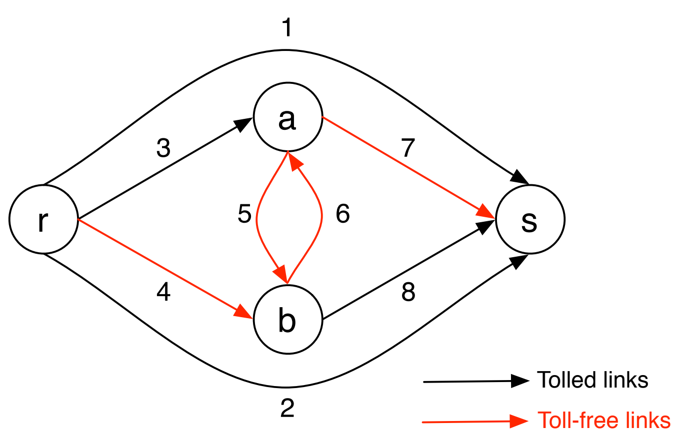

Figure 4: A four node network.

Table 1: Link characteristics of the four node network.

Link	Type	Distance (km)	Free-flow travel time (mins)	Toll ($)	Capacity (veh/hr)
1	Expressway	30	18.0	20	3600
2	Highway	30	22.5	15	3600
3	Arterial	10	12.0	1	1800
4	Arterial	20	24.0	0	1800
5	Arterial	2	2.4	0	1800
6	Arterial	5	6.0	0	1800
7	Arterial	20	24.0	0	1800
8	Arterial	10	12.0	1	1800

The single O-D pair is (r, s) and there are only six feasible routes in this network. The routes and their characteristics are listed in Table 2. Note that Route 1 and Route 2 are the direct routes, with Route 1 being the fastest with the highest toll while Route 6 is the only toll-free route and the slowest. The total demand from r to s is fixed at 10,000 vehicles per hour which is just a little bit lower than the network corridor capacity of 10,800 vehicles per hour. In order to define the indifference curve, we only need to specify the values of $TS^{max}(\tau_k)$ for $\tau_k = 0, 1, 2, 15, 20$. These values are shown in the last column of Table 2.

The solution $\mathbf{F}^*$ shown in Table 3 is a TSmaxBUE solution. The values of travel time and toll for the four routes with nonzero flow are illustrated in Figure 5. As Theorem 2 states, all routes with positive flow are efficient. Toll-free Route 6 is also efficient, but has zero flow because its time surplus, even at free-flow travel time is negative and less than the equilibrium value. Note that Routes 3 and 4 have

Table 2: Route characteristics of the four node network.

Route	Path	Length	Free-flow Travel Time	Toll	Max Time
1	1	30	18.0	20	25
2	2	30	22.5	15	40
3	$3-7$	30	36.0	1	50
4	$4-8$	30	36.0	1	50
5	$3-5-8$	22	26.4	2	49
6	$4-6-7$	45	54.0	0	51

identical toll, travel time, and flow and, therefore, show as a single dot in Figure 5.

Table 3: Time surplus maximisation BUE solution.

Route	Flow	Travel time	Time Surplus
1	2384.6	18.52	6.48
2	4839.2	33.52	6.48
3	202.9	43.52	6.48
4	202.9	43.52	6.48
5	2370.4	42.52	6.48
6	0.0	54.00	-3.00

We also notice that two of the efficient routes are not optimal for generalised cost (16) for any positive value of time. Hence, this example demonstrates that there are TSmaxBUE flows that are not equilibrium flows for generalised cost functions (16), even if a continuous distribution of value of time such as suggested in Dial (1997) is considered. Together with Theorem 4, this means that the time surplus maximisation bi-objective user equilibrium is indeed more general than generalised cost user equilibrium.

Figure 6 shows in addition to time and toll values the time surplus for each route. This is of course equal for each used route and larger than the (negative) time surplus of unused Route 6.

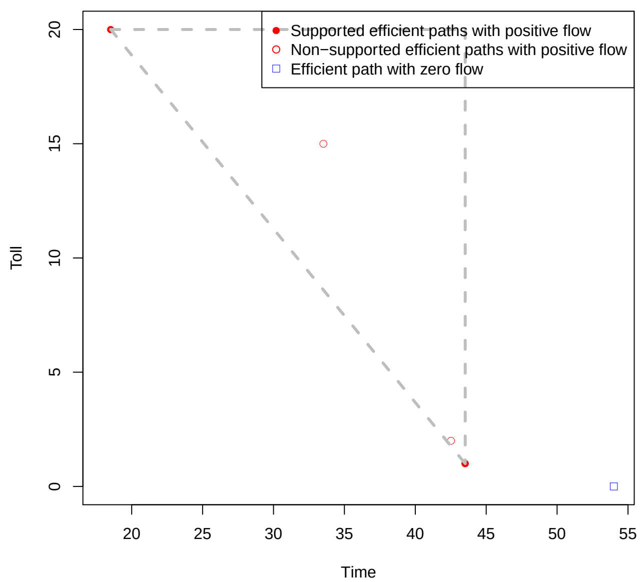

Figure 5: Efficient paths do not all optimise generalised cost.

6. Time Surplus Maximisation User Equilibrium with Multiple User Classes

In Section 4.2, we indicated that the concept of indifference curve that underlies the time surplus maximisation bi-objective user equilibrium lends itself to multi-user class traffic assignment. The shape of the indifference curve models users' attitude towards tolls in terms of willingness to pay. Users who are unwilling to

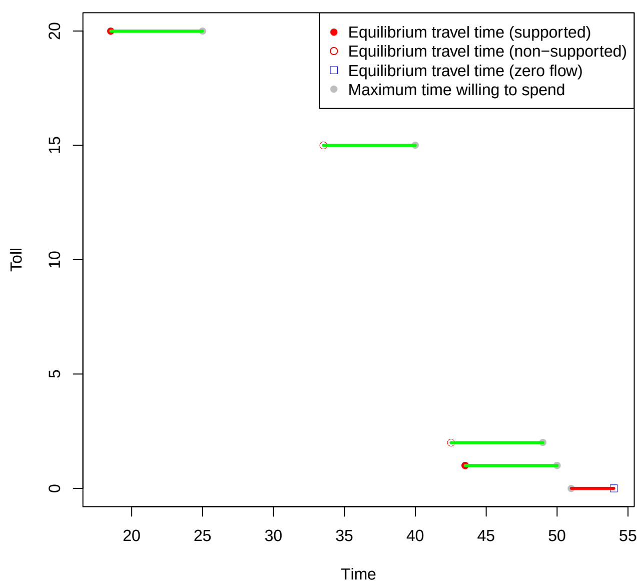

Figure 6: Time surplus at equilibrium.

pay tolls would accept higher maximum travel times at zero tolls to avoid the tolls. The shape of their indifference curve would be convex, as in Figure 7, whereas users with a strong preference for short travel time would accept any toll in order to ensure short travel times. Their indifference curve would be concave as in Figure 8.

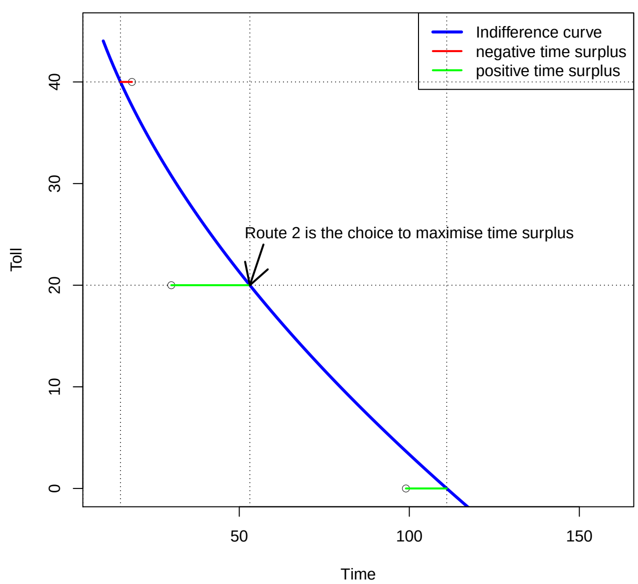

Figure 7: A convex indifference curve.

The limiting case for users who are reluctant to pay (whose indifference curve is convex) is a quasi-convex function defined as in Equation (27)

$$\underline{T_p^{max}}(\tau) := \begin{cases} 0 & \text{if } 0 < \tau \leq \max\{\tau_k : k \in K_p\} \\ \max\{T_k(D_p) : k \in K_p\} & \text{if } \tau = 0, \end{cases} \tag{27}$$

whereas the limiting case for a user insensitive to paying tolls (with a concave

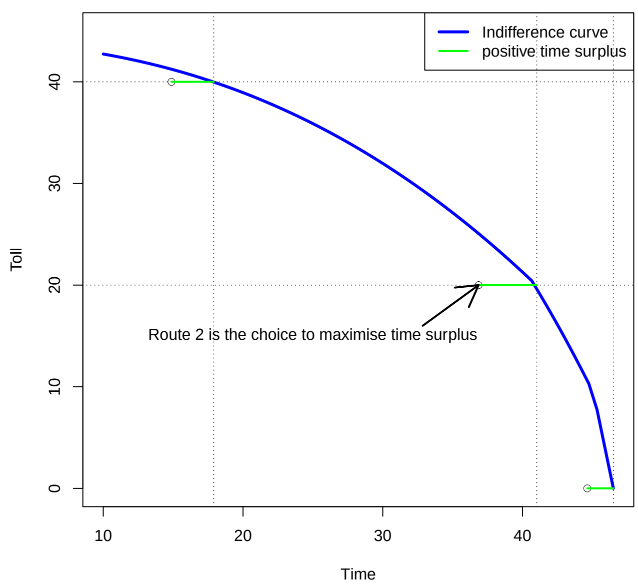

Figure 8: A concave indifference curve.

indifference curve) would be the quasi-concave function as in Equation (28)

$$\overline{T_p^{max}}(\tau) := \begin{cases} \max\{T_k(\mathbf{D_p}) : k \in K_p\} & \text{if } 0 \leq \tau < \max\{\tau_k : k \in K_p\} \\ 0 & \text{if } \tau = \max\{\tau_k : k \in K_p\}. \end{cases} \tag{28}$$

Notice that neither $\overline{T_p^{max}}$ nor $\underline{T_p^{max}}$ are strictly decreasing and are, therefore, excluded from being used in the definition of indifference curves in Section 4.1 and the time surplus function (17).

To extend Definition 3 to the case of multiple user classes, we let M be the finite set of user classes and denote by D_{pm} the demand for travel for O-D pair $p \in Z$ and user class m, for all $m \in M$ and $p \in Z$. $T_{pm}^{max}(\tau)$ and $TS_{km}(\mathbf{F})$ are, respectively, the indifference curve for user class m on origin-destination pair p and the time surplus function of user class m on path $k \in K_p$. Moreover, we index path flows by user class, i.e. F_{km} denotes the flow on path k for user class m. The set of feasible flows for traffic assignment with multiple user classes is defined as

$$\Omega^M := \left\{ \mathbf{F} \in \mathbb{R}^{|K|\cdot|M|} : \sum_{k \in K_p} F_{km} = D_{pm} \text{ for all } p \in Z \text{ and } m \in M \right\}. \tag{29}$$

Definition 4. *Path flow vector* $\mathbf{F}^* \in \Omega^M$ *is called a* TSmaxBUE flow with multiple user classes *if for all* $m \in M$ *it holds that* $F_{km} > 0 \Rightarrow TS_{km}(\mathbf{F}^*) \geq TS_{k'm}(\mathbf{F}^*)$ *for all* $k, k' \in K_p$*, or equivalently, if* $T_{km}(\mathbf{F}^*) > TS_{k'm}(\mathbf{F}^*) \Rightarrow F_{k'm} = 0$.

To find a solution of the time surplus maximisation bi-objective user equilibrium model with multiple user classes, we do not extend the method proposed in Larsson et al. (2002) as shown in Section 4.2, Equations (23) – (26). Instead, because the functions C_{km} defined analogously to Equation (22) are positive and demand is fixed and positive, we can formulate the problem as a nonlinear complementarity problem (Aashtiani, 1979; Chen et al., 2010) as shown in (30) – (35).

Let U_{pm} be a variable that denotes the minimal value of C_{km} for O-D pair p and user class m

$$
\begin{aligned}
\left(C_{km}(\mathbf{F}) - U_{pm}\right) F_{km} &= 0 \quad \text{for all } k \in K_p, p \in Z \text{ and } m \in M && (30)\\
\sum_{k \in K_p} F_{km} - D_{pm} &= 0 \quad \text{for all } p \in Z \text{ and } m \in M && (31)\\
C_{km}(\mathbf{F}) - U_{pm} &\geq 0 \quad \text{for all } k \in K \text{ and } m \in M && (32)\\
\sum_{k \in K_p} F_{km} - D_{pm} &\geq 0 \quad \text{for all } p \in Z \text{ and } m \in M && (33)\\
F_{km} &\geq 0 \quad \text{for all } k \in K \text{ and } m \in M && (34)\\
U_{pm} &\geq 0 \quad \text{for all } p \in Z \text{ and } m \in M && (35)
\end{aligned}
$$

Following (Lo and Chen, 2000) this problem can be solved by optimising the gap function

$$\phi(a,b) = \frac{1}{2}\left(\sqrt{a^2+b^2} - (a+b)\right)^2 \tag{36}$$

applied to the NCP (30) – (35). This leads to the optmisation problem

$$
\begin{aligned}
\min \sum_{m \in M} \sum_{p \in Z} \sum_{k \in K_p} \frac{1}{2}\left[\sqrt{F_{km}^2 + \left(C_{km}(F_{km}) - U_{pm}\right)^2} - \left(F_{km} + C_{km}(F_{km}) - U_{pm}\right)\right]^2 \\
+ \sum_{m \in M} \sum_{p \in Z} \frac{1}{2}\left[\sqrt{U_{pm}^2 + \left(\sum_{k \in K_p} F_{km} - D_{pm}\right)^2} - \left(U_{pm} + \sum_{k \in K_p} F_{km} - D_{pm}\right)\right]^2 \quad (37)
\end{aligned}
$$

We notice that, as is common with traffic assignment problems with multiple user classes, there is no uniqueness of link or path flows by user class. In Section 7, we provide an example to illustrate the time surplus maximisation user equilibrium with multiple user classes. We compare this to both user equilibrium based on linear generalised cost (16) and stochastic user equilibrium, both with multiple user classes defined by different values of time.

7. A Three Link Example

7.1. Network Specification

Now we consider a three link example as shown in Figure 9 with route characteristics as shown in Table 4. Note that Route 1 is the fastest with the highest toll while Route 3 is toll free and the slowest. The total demand from r to s is fixed at 15,000 vehicles per hour. The link travel time is assumed to be a function of traffic flow following the Bureau of Public Roads (1964) function as shown in Equation (10). There are three user classes with different levels of willingness to pay. Their respective indifference curves to the toll values are shown in Table 5. Since there is only one O-D pair, we omit the index p hereafter.

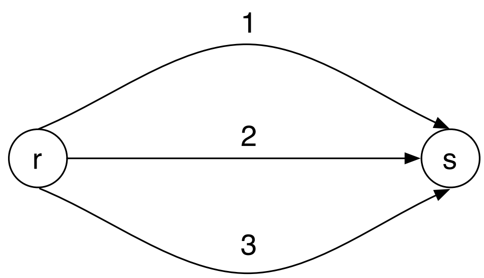

Figure 9: A three link example network.

Table 4: Route characteristics of the three link network.

Route	Type	Distance (km)	Free flow travel time (mins)	Toll ($)	Capacity (veh/hr)
1	Expressway	20	12	40	4000
2	Highway	50	30	20	5400
3	Arterial	40	40	0	4800

Table 5: Maximum time willing to spend.

Route	Class 1	Class 2	Class 3
k	T_1^{max}	T_2^{max}	T_3^{max}
1	12.5	17.5	22.5
2	32.5	37.5	42.5
3	65.0	75.0	85.0

7.2. The Conventional Solutions (UE, SUE and Social Optimum) with a Single User Class

Assuming demand is inelastic, i.e. all users must travel, the solution space for this three link network can be represented two-dimensionally as shown in Figure 10, with contours of the total travel time. We first identified the following solutions, as shown in Figure 10, in the conventional way:

1. the UE solution without tolls;
2. the UE solution with tolls, assuming VOT being \$1 per minute;
3. the SUE solution based on a multinomial logit formulation as shown in Equation (38)

$$P_k = \frac{e^{\theta U_k}}{\sum_{a \in A} e^{\theta U_a}}, \tag{38}$$

   where P_k is the probability of path k to be chosen; U_k is the utility of choosing path k; U_k is a function of the travel time t_k and toll τ_k, i.e. $U_k = -t_a(x_a) \times VOT - \tau_k$; and θ is the model parameter for calibration (assuming $\theta = 0.05$); and
4. the Social Optimum (SO) solution, by minimising total travel time, i.e. replacing Equation (12) in the optimisation problem of Equations (12) – (15)

with Equation (39)

$$\min Z(\mathbf{f}) = \sum_{a \in A} f_a t_a. \tag{39}$$

7.3. The BUE Solution Space

In order to illustrate the BUE solution space in this three link example, we first identify the BUE solution space where the BUE equilibrium condition applies. Because tolls are independent of flow and $\tau_1 > \tau_2 > \tau_3$, the BUE condition is satisfied whenever

$$t_1(f_1) < t_2(f_2) < t_3(f_3). \tag{40}$$

It is, therefore, enough to draw the curves defined by $t_1(f_1) = t_2(f_2)$ and $t_2(f_2) = t_3(15,000 - f_1 - f_2)$. The BUE solution space is illustrated three-dimensionally with total travel time as the third dimension in Figure 11 and two-dimensionally in Figure 12. We then examine the distribution of link flow and link travel time in this discretised BUE solution space. The boxplots of the link flow and link travel time are illustrated in Figures 13 and 14, respectively. The link travel time on the toll-free route has a range of 40 minutes to 612 minutes corresponding to a flow range of 1,000 to 15,000 vehicles per hour. The latter case corresponds to the case of putting all the demand on Route 3; the resulting solution will have a link travel time of 612 minutes on Route 3 while the link travel times on Route 1 and 2 are free-flow at 12 minutes and 30 minutes. This solution satisfies the BUE definition but obviously we would expect that someone would want to pay if the travel time is 612 minutes on the toll-free route. Observations made from this three link example strongly support the urgent need for further specification of the equilibrium conditions to represent route choice behaviour more realistically.

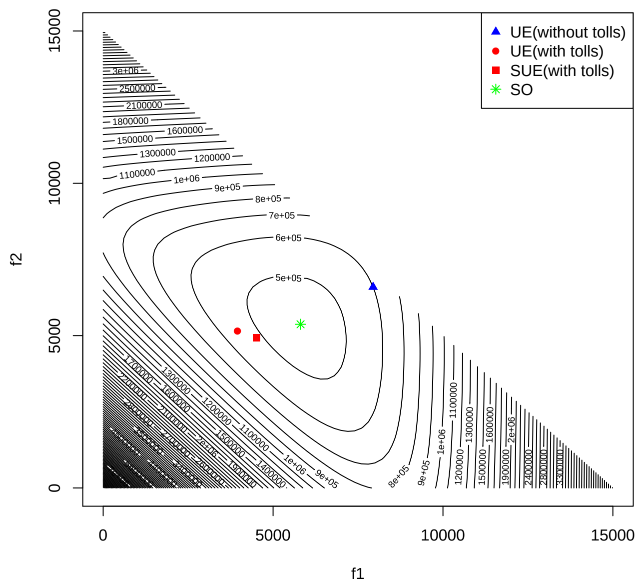

Figure 10: Total travel time contours in the solution space of the three link example.

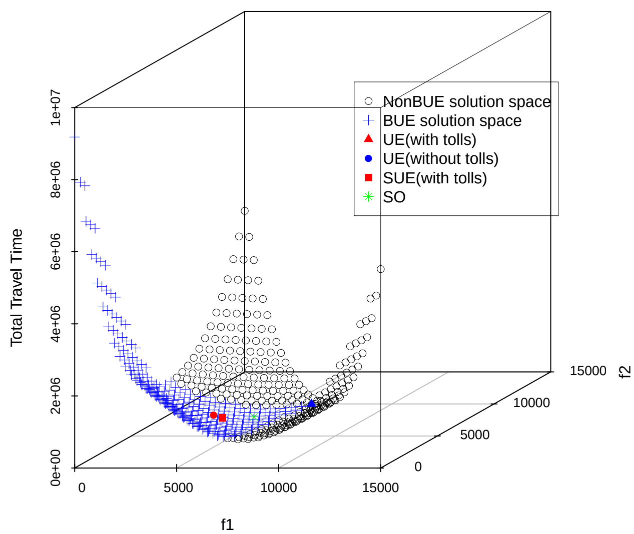

Figure 11: Three-dimensional plot of the BUE solution space.

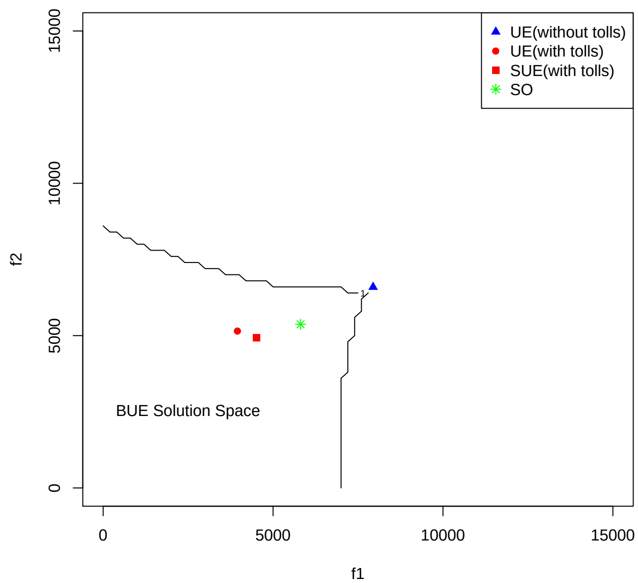

Figure 12: The BUE solution space boundary.

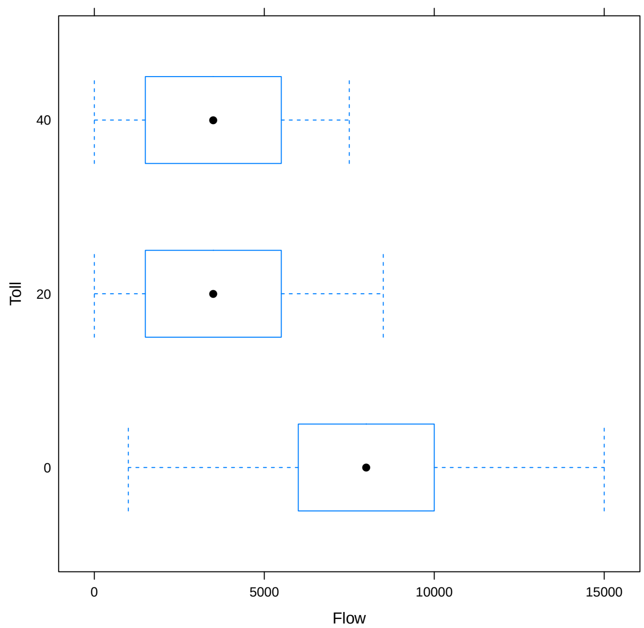

Figure 13: Boxplot of link flow.

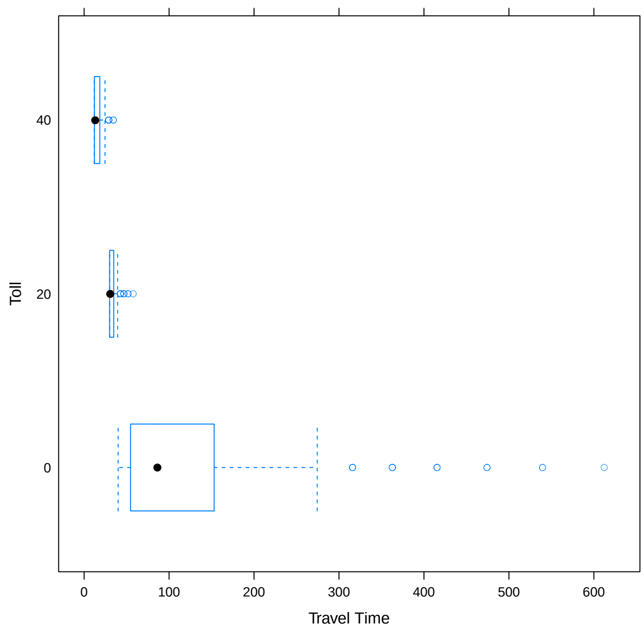

Figure 14: Boxplot of link travel time.

7.4. *Time Surplus Maximisation BUE Solution versus UE and SUE Solutions with Multiple User Classes*

Now we examine the case of multiple user classes for the TSmaxBUE model and the conventional UE and SUE models. The three user classes for the TSmaxBUE are as defined in Table 5, while those for UE and SUE are defined in Table 6. Note that the VOT values are assigned such that Class 1 has the highest VOT value representing the group that is most willing to pay while Class 3 has the lowest representing those most unwilling to pay. The $\theta-$value for the SUE cases is fixed at 0.1 representing a relatively low sensitivity case for illustration purpose.

Table 6: Multiple user class test parameters for UE & SUE.

Parameter	Class 1	Class 2	Class 3
VOT in UE & SUE	\$3	\$2	\$1
θ in SUE	0.1	0.1	0.1
Demand	5000 veh/h	5000 veh/h	5000 veh/h

We solved the TSmaxBUE case with the NCP formulation as shown in equations (30) – (35), the UE multiple user class case with the mathematical formulation in Yang and Huang (2004), Equations (3)–(7), and the SUE multiple user class case with the heuristics in Florian (2006). The solutions are as shown in Figure 15. The following observations are made:

1. The behaviour as modelled by the UE model is the most extreme. All users in Class 1 will choose the most expensive tolled Route 1 while all users in Class 3 will choose the toll-free Route 3. Class 2 will choose only Routes 1 and 2 with a higher proportion on Route 2.

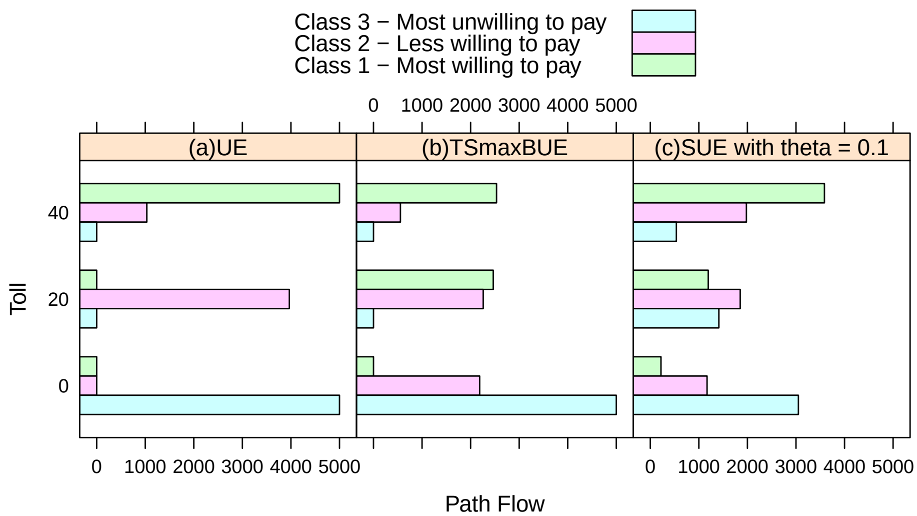

Figure 15: Comparisons of equilibrium flows in the UE, SUE, and TSmaxBUE solutions of the three link example.

2. In the TSmaxBUE solution shown in the middle of Figure 15, the users in each class choose their routes based on their respective indifference curves. Class 1 users will choose Routes 1 and 2; Class 2 users will use all three routes; and Class 3 users, who are most unwilling to pay, will all choose Route 3. Note, however, that this is only one possible TSmaxBUE solution, due to the solution with multiple user classes not being unique.
3. To illustrate the other extreme to the UE model, we chose a low $\theta-$value of 0.1 for the SUE case, the users are relatively less sensitive to the differences in utility values on each route. As a result, all classes will use all three routes with the proportions influenced by their VOT values. That is, more users from Class 1 will use Route 1 while more users from Class 3 will use Route 3.

8. Discussion

Modelling route choice behaviour is not an easy task, as clearly there are many factors influencing the decision. In fact, it is well known from empirical studies that the three most important factors influencing route choice behaviour are travel time, travel time reliability and monetary cost (e.g. Abdel-Aty et al., 1995; Brownstone and Small, 2005; Lam and Small, 2001; Liu et al., 2004). In this paper, we consider two of the three most important factors, i.e. time and toll. In a tolled network, this task is even more challenging as one would expect that users might have a strong opinion on whether they want to pay at all or not; and for those who are willing to pay, the time they are willing spend might vary a lot for the same toll value. The conventional modelling approaches, namely, UE and SUE, have relied on the specification of a value of time by an individual, which is assumed to be a constant. Such model structure is very restrictive as it implies that individuals will trade off time and money in the same way for any duration of the trip. It is natural to think that the longer the duration, the more stressful the trip would be and one would be more willing to pay. The indifference curve between toll and time is very likely to be non-linear, which is also supported by empirical evidence (Hensher and Truong, 1985). Although Hensher and Truong (1985) recommend that value of travel time savings should be specified as constant for planning purposes, the results from their experiments have indicated the presence of non-linear effects of time on travel behaviour. In other words, modelling route choice behaviour with a constant value of time might not be adequate. By considering the trade off between toll and time in a two-dimensional space, we can model variability among users with no restrictions.

By modelling the equilibrium with a bi-objective approach, only efficient paths will be included in the choice set, which creates a reasonable choice set for each individual naturally.

In terms of modelling the sensitivity of individuals to the differences in toll and time between alternatives, the use of indifference curves is also more flexible than the use of sensitivity parameters in the logit model, since the indifference curves can be of any form, convex or concave, as long as they are strictly decreasing.

9. Conclusion and Outlook

In this paper we have introduced a new model for route choice in tolled road networks. The model is based on the idea of bi-objective user equilibrium, which refers to the condition that traffic will arrange itself in such a way that no user can decrease travel time, or toll, or both without worsening the other. Since bi-objective user equilibrium allows many possible solutions, not all of which are meaningful in practice, we have augmented the concept with the idea of time surplus maximisation. This idea assumes that a user has an indifference function defining for any value of toll the maximum time he/she is willing to spend for travel between an origin and a destination. The preference of a user can be determined by the time surplus defined as maximum time willing to spend minus actual travel time. Users are rational and will choose a route with maximum time surplus among all efficient paths. We demonstrated that this model overcomes drawbacks of earlier UE models based on generalised cost. We have demonstrated that our model is more general than traditional models using a (linear) generalised cost function, and is therefore more versatile in modelling route choice behaviour. We also discussed the time surplus maximisation concept for the case of multiple user classes which have different indifference curves, i.e. the indifference function can model variability among users. To construct the indifference curves, we can conduct surveys to determine the maximum time one would be willing to spend for a given toll for each O-D pair. User classes can then be formed by grouping users with similar indifference curves as one user class.

Our research opens up many avenues for future work. It will be interesting to compare the routes identified as efficient in our model to empirical approaches for choice set generation, which are based on behavioural principles, see e.g. Bekhor et al. (2006) and Bovy and Fiorenzo-Catalano (2007). Further investigation on uniqueness of path and link flows based on network topology, as discussed in Milchtaich (2005) and Richman and Shimkin (2007) also deserves attention.

In this paper we have only considered the case of inelastic demand, i.e. even users with only routes with negative time surplus in their choice sets will have to choose a route (the one with least negative time surplus) and travel. It is natural to extend the model to the elastic case, where users may not travel if their time surplus is negative on all efficient paths. In this case, we would look at replacing the function of Equation (17) with $TS_k(\mathbf{F}) = \left[T_p^{max}(\tau_k) - T_k(\mathbf{F})\right]_+$, i.e. either positive time surplus or zero if time surplus is negative, as the route choice function. This complicates analysis considerably and is a topic of current research.

In the future, we will also investigate other combinations of two objectives that are relevant for route chice. We will also look at the inclusion of travel time reliability in a multi-objective extension of our model. Travel time reliability has empirically been shown to be one of the three main factors influencing route choice along with travel time and monetary cost.

Finally, we note that we have only considered two objectives in user equlibrium models. Naturally, it is also of interest to consider bi-objective system optimum models. Guo and Yang (2009) and Chen and Yang (2012) propose such models. We have proposed a first idea for integrating bi-objective user equilibrium with bi-objective system optimum models in Wang and Ehrgott (2013).

10. Acknowledgments

The authors would like to acknowledge Dr Jon Pearce for his assistance with MATLAB coding. This research was partially supported by the Marsden Fund, grant number 9075 362506.

References

Aashtiani, H., 1979. The Multi-Modal Traffic Assignment Problem. Ph.D. thesis. Operations Research Center, MIT, Cambridge, MA.

Abdel-Aty, M.A., Kitamura, R., Jovanis, P.P., 1995. Investigating effect of travel time variability on route choice using repeated-measurement stated preference data. Transportation Research Record 1493, 39–45.

Beckmann, M.J., McGuire, C.B., Winsten, C.B., 1956. Studies in the Economics of Transportation. Yale University Press, New Haven, CT.

Bekhor, S., Ben-Akiva, M., Ramming, M., 2006. Evaluation of choice set generation algorithms for route choice models. Annals of Operations Research 144 (1), 235–247.

Bovy, P., Fiorenzo-Catalano, S., 2007. Stochastic route choice set generation: Behavioral and probabilistic foundations. Transportmetrica 3 (3), 173–189.

Brownstone, D., Small, K., 2005. Valuing time and reliability: Assessing the evidence from road pricing demonstrations. Transportation Research Part A 39 (4), 279–293.

Bureau of Public Roads, 1964. Traffic Assignment Manual. U.S. Department of Commerce, Urban Planning Division, Washington D.C.

Chen, A., Oh, J., Park, D., Recker, W., 2010. Solving the bicriteria traffic equilibrium problem with variable demand and nonlinear path costs. Applied Mathematics and Computation 217 (7), 3020–3031.

Chen, L., Yang, H., 2012. Managing congestion and emissions in road networks with tolls and rebates. Transportation Research Part B 46 (8), 933–948.

Dafermos, S., Sparrow, F., 1969. The traffic assignment problem for a general network. Journal of Research of the National Bureau of Standards Series B 73 (2), 91–118.

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

Dial, R., 1971. A probabilistic multipath traffic assignment model which obviates path enumeration. Transportation Research 5 (2), 83–113.

Dial, R., 1979. A model and algorithm for multicriteria route-mode choice. Transportation Research Part B 13 (4), 311–316.

Dial, R., 1996. Bicriterion traffic assignment: Basic theory and elementary algorithms. Transportation Science 30 (2), 93–111.

Dial, R., 1997. Bicriterion traffic assignment: Efficient algorithms plus examples. Transportation Research Part B 31 (5), 357–379.

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

Florian, M., 2006. Network equilibrium models for analyzing toll highways, in: Lawphongpanich, S., Hearn, D.W., Smith, M.J. (Eds.), Mathematical and Computational Models for Congestion Charging. Springer, New York, pp. 105–115.

Florian, M., Constantin, I., Florian, D., 2009. A new look at projected gradient method for equilibrium assignment. Transportation Research Record 2090, 10–16.

Florian, M., Hearn, D., 1995. Network equilibrium models and algorithms, in: Ball, M. (Ed.), Network Routing, Elsevier, Amsterdam. pp. 485–550.

Frank, F., Wolfe, P., 1956. An algorithm for quadratic programming. Naval Research Logistics Quarterly 3 (1-2), 95–110.

Gabriel, S., Bernstein, D., 1997. The traffic equilibrium problem with nonadditive path costs. Transportation Science 31 (4), 337–348.

Guo, X., Yang, H., 2009. User heterogeneity and bi-criteria system optimum. Transportation Research Part B 43 (4), 379–390.

Hensher, D., Truong, P., 1985. Valuation of travel time savings. Journal of Transport Economics and Policy 19 (3), 237–261.

Jayakrishnan, A., Tsai, W.K., Rajadhyaksha, S., 1994. A faster path-based algorithm for traffc assignment. Transportation Research Record 1443, 75–83.

Lam, T., Small, K., 2001. The value of time and reliability: Measurement from a value pricing experiment. Transportation Research Part E 37 (2-3), 231–251.

Larsson, T., Lindberg, P.O., Patriksson, M., Rydergren, C., 2002. On traffic equilibrium models with a nonlinear time/money relation, in: Patriksson, M., Labbé, M. (Eds.), Transportation Planning: State of the Art, Kluwer Academic Publishers, Dordrecht. pp. 19–31.

Leurent, F., 1993. Cost versus time equilibrium over a network. European Journal of Operational Research 71 (2), 205–221.

Liu, H., Recker, W., Chen, A., 2004. Uncovering the contribution of travel time reliability to dynamic route choice using real-time loop data. Transportation Research Part A 38 (6), 435–453.

Lo, H.K., Chen, A., 2000. Traffic equilibrium problem with route-specific costs: Formulation and algorithms. Transportation Research Part B 34, 493–513.

Luce, R.D., Suppes, P., 1965. Preference, utility and subjective probability, in: Luce, R.D., Bush, R.R., Galanter, E. (Eds.), Handbook of Mathematical Psychology, John Wiley & Sons. pp. 249–410.

Maher, M., 1992. SAM – A stochastic assignment model, in: Griffiths, J. (Ed.), Mathematics in Transport Planning and Control, Oxford University Press, Oxford. pp. 121–132.

Milchtaich, I., 2005. Topological conditions for uniqueness of equilibrium in networks. Mathematics of Operations Research 30 (1), 225–244.

Prashker, J.N., Bekhor, S., 2004. A review on route choice models used in the stochastic user equilibrium problem. Transport Reviews 24 (4), 437–463.

Raith, A., Ehrgott, M., 2011. On vector equilibria, vector optimization and vector variational inequalities. Journal of Multi-Criteria Decision Analysis 18 (1-2), 39–54.

Raith, A., Wang, J.Y.T., Ehrgott, M., Mitchell, S., 2013. Solving multi-objective traffic assignment. Annals of Operations Research DOI 10.1007/s10479-012-12841.

Richman, O., Shimkin, N., 2007. Topological uniqueness of nash equilibrium for selfish routing with atomic users. Mathematics of Operations Research 32 (1), 215–232.

Rosa, A., Maher, M., 2002. Algorithms for solving the probit path-based stochastic user equilibrium traffic assignment problem with one or more user classes, in: Taylor, M. (Ed.), Transportation and Traffic Theory in the 21st Century – Proceedings of the 15th International Symposium on Transportation and Traffic Theory, Adelaide, Australia, 16-18 July 2002, Pergamon Press, Oxford. pp. 371–392.

Sheffi, Y., 1985. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall.

Sheffi, Y., Powell, W.B., 1982. An algorithm for the equilibrium assignment problem with random link times. Networks 12 (2), 191–207.

Wang, J.Y.T., Ehrgott, M., 2013. Bi-objective network equilibrium, traffic assignment and road pricing, in: Helber, S. (Ed.), Operations Research Proceedings 2012, Springer Verlag, Berlin. In press.

Wang, J.Y.T., Raith, A., Ehrgott, M., 2010. Tolling analysis with bi-objective traffic assignment, in: Ehrgott, M., Naujoks, B., Stewart, T., Wallenius, J. (Eds.), Multiple Criteria Decision Making for Sustainable Energy and Transportation Systems. Springer Verlag, Berlin, pp. 117–129.

Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers: Engineering Divisions 1, 325–362.

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