This is a repository copy of *Joint road toll pricing and capacity development in discrete transport network design problem*.

White Rose Research Online URL for this paper:
http://eprints.whiterose.ac.uk/93226/

Version: Accepted Version

Article:

Xu, M, Wang, G, Grant-Muller, S et al. (1 more author) (2017) Joint road toll pricing and capacity development in discrete transport network design problem. Transportation, 44 (4). pp. 731-752. ISSN 0049-4488

https://doi.org/10.1007/s11116-015-9674-2

Joint road toll pricing and capacity development in discrete transport network design problem

Meng Xu
State Key Laboratory of Rail Traffic Control and Safety
Beijing Jiaotong University
No. 3 of Shangyuan Residence, Haidian District, Beijing, 100044, China
Tel: +86-10-5168-7070
Fax: +86-10-5168-7127
Email: mengxu@bjtu.edu.cn

Guangmin Wang
School of Economics and Management
China University of Geosciences
Wuhan, China, 430074
E-mail: wgm97@163.com

Susan Grant-Muller
Institute for Transport Studies
University of Leeds
34-40 University Road, LS2 9JT Leeds, UK
E-mail: s.m.grant-muller@ leeds.ac.uk
Tel: +44 (0)113 343 6618

Ziyou Gao
School of Traffic and Transportation,
Beijing Jiaotong University
3 Shangyuancun, Haidian District
Beijing, China 100044
Tel: +86-10-51688193
E-mail: zygao@bjtu.edu.cn

Abstract:

The paper demonstrates a method to determine road network improvements that also involve the use of a road toll charge, taking the perspective of the government or authority. A general discrete network design problem with a road toll pricing scheme, to minimize the total travel time under a budget constraint, is proposed. This approach is taken in order to determine the appropriate level of road toll pricing whilst simultaneously addressing the need for capacity. The proposed approach is formulated as a bi-level programming problem. The optimal road capacity improvement and toll level scheme is investigated with respect to the available budget levels and toll revenues.

1. Introduction

Road transport plays a major role in driving both economic development and social activity. However, the presence of negative externalities such as congestion, pollution and accidents brings increasing social and environmental stress, which has emphasized the urgent need for an effective, efficient and socially feasible road transport system (Grant-Muller and Xu, 2014). The extraordinarily complex task of introducing developments to the road transport system is well recognised. Efforts to advance policies incrementally or by introducing a single policy measures often fail, leading to calls for 'integrative', 'holistic' policies that incorporate multiple policy instruments from sectors related to urban mobility. The need for more sustainable and integrative planning processes has been widely recognized, with research including that by (Givoni, 2013).

The future development and implementation of integrative policies in the road transport sector depends primarily on how well understood they are by decision makers. Road capacity investment and pricing issues have become crucial with an increased demand for the development of new transport infrastructure and the necessity to ensure the infrastructure can be both financed and utilized efficiently (Adler and Proost, 2010). Theoretical developments in both the road network design problem (NDP) and road toll pricing have received much attention over a period of time in published research. The NDP involves optimal allocation of budget to the expansion of existing links and/or to the addition of new candidate links so as to maximize the network performance (e.g. minimizing the total travel time or generalized cost, maximizing the social welfare) while accounting for the travellers' route choice behaviour (LeBlanc, 1975). The literature includes several comprehensive reviews of the NDP, examples being Yang and Bell (1998), and Farahani et al. (2013). The road toll pricing problem (as an effective means of both managing road traffic demand and raising additional revenue for road construction) has been studied extensively by both transportation practitioners and economists. Practical implementation has progressed rapidly and electronic road pricing schemes have been proposed and tested worldwide (Meng, et al. (2012); Yang and Huang (2005); Yang, et al., (2010)).

Considering the literature on both NDP and the road toll pricing problem, there have been few published studies concerned with the simultaneous determination of road toll pricing and capacity development, despite the strong connection between them and their importance for the transport system. In Section 2, a detailed summary of recent studies concerned with the problem of simultaneous determination of road toll pricing and capacity development is provided. In an era of tightly constrained budgets, the analysis of road tolls and capacity provides a useful framework to consider interactions between both decision variables. This is especially the case in developing counties where there may be long term financial constraints on road capacity improvement.

In this paper, we therefore investigate the joint decision making involved in the development of road infrastructure (capacity) alongside introduction of a road toll and the pricing problem. The revenue from a road toll charge is assumed to be a part of the infrastructure financing. In contrast to the case for a BOT (Build-Operate-Transfer) scheme, the government will consider road capacity development and road toll pricing together, with the purpose of minimising total travel time. However the tolled roads established by the government are not necessarily the newly built roads that require financing. In a departure from existing formulations for a continuous network design problem (CNDP), we propose a discrete network design problem (DNDP) approach. This is considered more appropriate for a transport network, as improvements such as lane expansions cannot be delivered in fractional quantities (Boyce and Janson, 1980).

We investigate toll pricing and capacity development within a discrete network design framework. Using this approach, different toll pricing and capacity expansion schemes can be compared explicitly, as can variances in total travel time with respect to budget levels and toll revenues. The research uses a new formulation for the joint road toll and capacity discrete network design problem, where the decision variables for toll and capacity are both integer variables. Whilst existing mixed-integer programming formulations generally stipulate a discrete variable for a particular link (of either the toll road or capacity expansion), and a continuous variable for the level of the toll (or capacity quantity), the formulation used in this research is different.

It defines discrete capacity variables with respect to the number of additional lanes for capacity expansion and an integer toll level, which lies within a pre-defined toll range. The discrete variable design (with toll pricing and capacity expansion) is more appropriate for use with a transport network, and results show that a road toll on the subset of links provides an approach to the optimal solution with respect to the total travel time.

For the capacity expansion problem with a discrete variable for capacity (i.e. number of additional lanes), Wang et al. (2013) provide a bi-level program formulation and a global optimization algorithm for the discrete network design problem. Hence, this paper can be viewed as extending the state of the art by viewing the road toll as a discrete variable for the joint road toll and capacity discrete network design problem. The issues addressed and problems solved in this paper are as follows. Firstly, the proposed modelling approach can support government decision-makers in identifying how, and under what circumstances, to set the budget levels and the toll pricing levels. It can also assist the public in understanding how the government measures will benefit transport development. Secondly, to solve the proposed bi-level programming problem (with an integer nonlinear program for the upper level problem, and a continuous nonlinear program for the lower level problem), we reformulate it as a single-level mixed-integer nonlinear program with nonlinear complementarity constraints. We then develop a relaxation approach by relaxing the complementarity constraints. The SBB solver in GAMS (2009) is employed to solve the relaxed mixed-integer nonlinear program at each iteration step.

The organization of this paper is as follows: in Section 2, we provide a brief literature review with respect to the combination/integration of urban road transport management and modelling approaches. Section 3 presents the bi-level programming model to describe the DNDP, simultaneously optimizing both road toll charging and road expansion under the fixed OD demand case. Section 4 presents a relaxation algorithm, using Solver in GAMS for convenience. In Section 5, numerical experiments are illustrated using a simple two-node network and the Sioux-Falls Network. Finally, conclusions are presented in Section 6.

2. Literature review

The combination or integration of measures for road transport tends to cover certain types of management policies, e.g., the development of sustainable urban transport strategies (Marsden, et al., 2010; May et al., 2005a and b), optimal urban transport strategies (May and Still, 2000; May, et al., 2006; Shepherd, et al., 2006; Kelly et al., 2008; Xu et al, 2015), and more recently, urban transport packages (Givoni, 2014). Modelling the combination of measures in the transport network framework is important. Although these management policies have been discussed, relatively few have been based on a transport network modelling process.

Studies into the simultaneous determination of a road toll level and road capacity have taken different approaches. The main studies have focused on investigating whether the revenue from socially optimal pricing on a road can cover the capital cost for constructing and maintaining the road. Mohring and Harwitz (1962) demonstrated that under certain conditions revenues from optimal pricing are just sufficient to cover the cost of optimal supply of road infrastructure. Yang and Meng (2000) considered the selection of toll and capacity for a private highway in a general road network under a BOT scheme. They investigated profitability and social welfare gain together with maximum profit and social welfare solutions in a toll-capacity two-dimensional space. Yang and Huang (2005) dealt with the relationship between pricing, capacity choice and financing in a general network with a single or multiple types of vehicles under the first-best and second-best conditions. Subprasom and Chen (2007) discussed the effects of regulation on highway pricing and capacity choice of a BOT scheme. Dimitriou et al. (2009) addressed the joint optimization of capacity investments and the toll charges problem imposed on multi-group users in monopolistic private highways within general road networks. They provided a game-theoretic formulation that leads to a non-convex bi-level program. With real-application results they demonstrated the importance of considering the spatial heterogeneity of prices, and the trade-off between investments and pricing strategies in regulated private highways. Verhoef et al. (2010) considered second-best cases where only a sub-set of links in a network is subject to tolling and/or capacity choice, and for cases with and without a

self-financing constraint. Using the 'long-run cost functions' for congested networks they demonstrated that, under certain assumptions, second-best long-run cost (or actually, generalized price) functions can be derived for most cases of interest. These can be used in an applied network model as a substitute for the conventional short-run user cost functions. Recently, building on the works by Tan et al. (2010) and Tan and Yang (2012), Niu and Zhang (2013) discussed the impact of demand uncertainty on the BOT contract design by optimizing a bi-objective problem via three decisions including toll, capacity and concession period.

Concerning development of a modelling approach to simultaneously determine the first and/or second best pricing and road capacity, Koh et al. (2009) investigated the first and second best jointly optimal toll and road capacity investment problems from both the policy and technical perspectives. They emphasized that, as with optimal pricing, optimal investment in capacity could not always be feasible, and therefore a second best optimum investment must be sought. Previous studies to minimize the total travel time or maximize the total social welfare of travellers in transport networks include Zhang and Wee (2012), as they proposed a new bi-level model for the implementation of congestion pricing, with a new objective function for maximizing the reserve capacity of networks. They further formulated this as a single-level optimization program with equilibrium constraints to circumvent the computational difficulty. Recently, Wang et al. (2014b) proposed a bi-level programming model for a joint optimal link-based tradable credit charging scheme and road capacity improvement. Their model integrates the improvement in the urban road network according to improvements in road capacity with a given budget constraint, and decreasing travel demand with the tradable credit scheme. They demonstrated that the tradable credit scheme offers a better combination of cost-effectiveness, administrative flexibility and distributional fairness compared to congestion pricing. To understand the trade-offs between conflicting objectives and to design a financially and environmentally sustainable transport system, Yin et al. (2014) addressed the simultaneous determination of road toll pricing and capacity investment in a congested road network in a multi-criteria decision-making framework. They presented

a goal programming approach to follow four major goals including cost recovery, service level, environmental and equity. The multi-objective road toll pricing and capacity investment problem was formulated as a bi-level goal programming model. The numerical results demonstrated that the priority structure of the goals can significantly affect road toll pricing and capacity investment decisions. Research that has focused on the development of algorithms includes work on a sensitivity analysis-based algorithm, a direct search algorithm (introduced in Yang and Huang (2005)), and a constraint cutting algorithm (given by Koh, et al. (2009)). Fan and Gurmu (2014) presented a bi-level genetic algorithm (GA)-based optimization solution methodology that can be used to determine the optimal solution for combined congestion pricing and capacity expansion problems.

Past research into the problem of simultaneous determination of road toll pricing and road capacity development has focused on investigating pricing and self-financing from modelling. Road capacity has been generally assumed to be theoretically continuous, whilst the toll is optimized under given constraints. These studies provide the theoretical background and an important framework for further study. The approach proposed in this paper, with a general DNDP and road toll pricing to minimize the total travel cost under a limited budget is, however, more appropriate for application in the practical transport planning and management contexts.

3. Problem formulation

3.1 Notation

In a given urban road network, we assume that the authority needs to improve the transport network in such a way that the total travel time is minimized. However, the authority faces a shortage of budget and therefore needs to receive revenue from road toll charges. Without loss of generality, we assume constant returns to scale in costs of capacity expansion (Mohring and Harwitz, 1962; Strotz, 1965; Keeler and Small, 1977), and fixed demand in the transport network. It is noted that network design and road pricing are both long-term decisions that impact on demand, and it would therefore be desirable to incorporate demand elasticity. In this study context, one can

simply use social welfare instead of total travel time as the design objective function by introducing elastic demand functions. However, incorporating elastic demand and changing the objective function would not fundamentally change the proposed model formulation and the method used to find the solution[1].

The following symbols are firstly defined:

Sets:

A Set of all links in the transport network

D_1 Subset of links in the transport network considered for additional lanes, $D_1 \subseteq A$

D_2 Subset of links in the transport network considered for the implementation of toll pricing, $D_2 \subseteq A$

LA_a Set of number of lanes to be added to link $a \in D_1$. For example, $LA_a = \{0, 1, 2\}$ means that the transport authority determines to either add zero (do-nothing), one, or two lanes to link $a \in D_1$

LT_a Set of values for toll pricing (j_a) imposed by the authority on link $a \in D_2$. For example, $LT_a = \{j_a | j_{min} \leq j_a \leq j_{max}, j_a \ is \ integer \ , a \in D_2\}$

N Set of nodes in the transport network

W Set of Origin-Destination (OD) pairs

R_w Set of paths between OD pair $w \in W$

Parameters/functions:

$\delta_{a,r}^w$ A binary coefficient which equals 1 if path $r \in R_w$ between OD pair $w \in W$ uses link $a \in A$, and 0 otherwise

B Total available budget

$c_a(i_a)$ Construction cost of adding i lanes to link $a \in D_1$, $i_a \in LA_a$; $c_a(i_a) = 0$ if $i_a = 0$

d_w Given travel demand for OD pair $w \in W$

$t_a(v_a, i_a, j_a)$ Travel time function on link a when i lanes are added and toll pricing j_a are set on link a, $i_a \in LA$, $j_a \in LT_a$, $a \in D_1 \cap D_2$

$t_a(v_a, i_a)$ Travel time function on link a when i lanes are added on link a, $i_a \in LA_a$, $a \in D_1$

$t_a(v_a, j_a)$ Travel time function on link a when toll pricing j_a are set on link a, $j_a \in LT_a$, $a \in D_2$

---

[1] We owe this point to an anonymous referee.

$t_a(v_a)$ Travel time function on link $a \in A - D_1 \cup D_2$

η A parameter setting for a more flexible formulation

Decision Variables:

$z1_a^{i_a}$ A binary decision variable which equals 1 if and only if i_a lanes are
added to link $a \in D_1, i_a \in LA_a$. $\mathbf{z1} = \{z1_a^{i_a}, i_a \in LA_a, a \in D_1\}$

$z2_a^{j_a}$ A binary decision variable which equals 1 if and only if road toll
pricing j_a is chosen for link $a \in D_2$, $j_a \in LT_a$. $\mathbf{z2} = \{z2_a^{j_a}, j_a \in LT_a, a \in D_2\}$

v_a Traffic flow on link $a \in A$, $\mathbf{v} = \{v_a, a \in A\}$

f_r^w Traffic flow on path $r \in R_w$ between OD pair $w \in W$, $\mathbf{f} = \{f_r^w, r \in R_w, w \in W\}$

3.2 Modelling approach

Consider that the authority plans to construct and operate a road under a given budget constraint B (i.e., the available public funding per unit time period). Therefore, the deficit per unit time period for the authority is no more than the budget constraint B, that is,

$$\sum_{a \in D_1} \sum_{i_a \in LA} c_a(i_a) z1_a^{i_a} - \eta \sum_{a \in D_2} \sum_{j_a \in LT} v_a z2_a^{j_a} j_a \leq B$$

where η is a parameter set to give a more flexible formulation, and the first term is the total construction cost of adding lanes to the network, the second term is the total revenue of toll pricing. Therefore, alongside congestion mitigation, road pricing could also be used for deficit reduction in the case of a limited budget.

The authority aims to minimize the social travel time in the DNDP with road toll pricing scheme $\mathbf{z2}$. Thus, the upper-level problem can be formulated as

[OP: original problem]:

[OP] $\qquad \underset{\mathbf{z1,z2}}{min} \sum_{a \in A} \sum_{i_a \in LA} \sum_{j_a \in LT} v_a t_a(v_a, i_a, j_a)$ $\qquad$ (1)

Subject to

$$\sum_{i_a \in LA} z1_a^{i_a} = 1, \forall a \in D_1 \tag{2}$$

$$\sum_{j_a \in LT} z2_a^{j_a} = 1, \forall a \in D_2 \tag{3}$$

$$\sum_{a \in D_1} \sum_{i_a \in LA} c_a(i_a) z1_a^{i_a} - \eta \sum_{a \in D_2} \sum_{j_a \in LT} v_a z2_a^{j_a} j_a \le B \tag{4}$$

$$z1_a^{i_a} = \{0, 1\}, \forall a \in A, i_a \in LA_a \tag{5}$$

$$z2_a^{j_a} = \{0, 1\}, \forall a \in A, j_a \in LT_a \tag{6}$$

The objective function (1) represents the total travel time where v_a is determined by the lower-level user equilibrium (UE) problem which will be presented in Eqs (7)-(10). Constraint (2) ensures that exactly one capacity level is chosen. Constraint (3) ensures that exactly one link toll level is used on each link. Constraint (4) guarantees that the deficit per unit time period of the authority does not exceed the total budget. Constraints (5) and (6) define that $z1_a^i$ and $z2_a^j$ are binary decision variables separately.

In model [OP], the relation between the capacity and construction cost for each link $a \in A$ can be any general function (Wang et al., 2013). For example, the cost of adding two lanes, c_a^2, is not necessarily equal to twice the cost of adding one lane, c_a^1. This is more reasonable than assuming the construction cost to be a linear function of capacity, as is usually done in the CNDP (Yang and Bell, 1998).

As the lower-level of the bilevel programming problem, it is assumed that travel demand is given and fixed, and users' route choice is characterized by the UE principle. The UE problem with fixed demand can be formulated below (Sheffi, 1985):

$$[UE] \; \min_{v} \; \sum_{a \in D_1} \sum_{i_a \in LA} z1_a^{i_a} \int_0^{v_a} t_a(\omega, i_a) d\omega +$$

$$\sum_{a \in D_2} \sum_{j_a \in LT} z2_a^{j_a} \int_0^{v_a} t_a(\omega, j_a) d\omega -$$

$$\sum_{a \in D_1 \cap D_2} \sum_{i_a \in LA} \sum_{j_a \in LT} z1_a^{i_a} z2_a^{j_a} \int_0^{v_a} t_a(\omega, i_a, j_a) d\omega +$$

$$\sum_{a \in A - D_1 \cup D_2} \int_0^{v_a} t_a(\omega) d\omega \tag{7}$$

Subject to

$$\sum_{r \in R_w} f_r^w = d_w, \forall w \in W \qquad (8)$$
$$v_a = \sum_{w \in W} \sum_{r \in R_w} f_r^w \delta_{a,r}^w, \forall a \in A \qquad (9)$$
$$f_r^w \geq 0, \forall r \in R_w, \forall w \in W \qquad (10)$$

Constraint (8) defines the demand conservation condition, Constraint (9) defines the relation between link flow and path flow and Constraint (10) requires nonnegative path flows.

In this study, before and after a capacity change and toll setting, it is assumed that link travel time functions $t_a(v_a)$, $t_a(v_a, i_a)$, $t_a(v_a, j_a)$, $t_a(v_a, i_a, j_a)$ are always strictly increasing and convex with respect to link flow v_a ($a \in A$) so that link flow solutions of both the UE and system optimization (SO) traffic assignment problems are unique. With this formulation, we could assume each of these new candidate links already exists by associating it with an infinite (or extremely large) free flow travel time and no toll setting in the initial transport network. If one (or more) lanes is added ($z1_a^{i_a}, i_a \in LA_a, a \in D_1$), then it is assumed that the actual free flow travel time is incorporated in the corresponding link travel time function and that a toll scheme will be considered. Compared with the conventional DNDP (that determines whether each link in a set of new candidate links should be added to the transport network), the DNDP formulation here embraces the conventional DNDP as a special case, and simultaneously, the road toll pricing design ($z2_a^{j_a}, j_a \in LT_a, a \in D_2$) is considered for a subset of the set A of all links in the transport network. The purpose of road pricing could be for both congestion mitigation and deficit reduction (in the case of a limited budget).

4. Solution method

The proposed bilevel programming model is flexible in determining extensions to road infrastructure and the design of road toll pricing from the perspective of government or high level policy makers. However, solving the model is challenging. Potential approaches to solve the model could use existing algorithms to solve the DNDP, for example, the support function method proposed by Gao et al.(2005), the active set technique presented by

Zhang, et al.(2009), the linear approach technique given by Farvaresh and Sepehri (2011), the branch and bound algorithm given by Farvaresh and Sepehri (2013), and the global optimization recently proposed by Wang et al. (2013). Beside these existing algorithms it would also be possible to reformulate the model then use software from the sector. Here we use GAMS (The General Algebraic Modeling System, www.gams.com). It is noted that the model [OP] is a mixed-integer nonlinear bilevel programming problem, in which the upper level decision variables are all integer and the lower level decision variables are all continuous. Existing approaches to solve a mixed-integer nonlinear bilevel programming problem include work by Moore and Bard (1990), Jan and Chern (1994), Sahin and Ciric (1998), Gümüş and Floudas (2005) and Mitsos (2010). These greatly expand the range of decision making problems that can be modelled and solved within a bilevel optimization framework. However, little attention is given to bilevel problems involving discrete decisions in the bilevel programming literature. This is primarily because these problems pose major algorithmic challenges in the development of effective solution strategies.

In this paper, a relaxation algorithm is proposed to solve the mixed-integer nonlinear bi-level programming problem. Because the objective and constraints of the lower level problem (UE) satisfy the convexity requirements of the Karush-Kuhn-Tucker (KKT) optimality conditions, the lower level problem can be replaced with its necessary and sufficient KKT optimality conditions. Thus, the bi-level programming model can be reformulated into the following single-level mixed-integer nonlinear optimization problem (MINLP):

[MINLP]

$$Z = \min_{z1,z2,v} \sum_{a \in A} \sum_{i_a \in LA} \sum_{j_a \in LT} v_a t_a(v_a, i_a, j_a) \tag{11}$$

Subject to

$$\sum_{i_a \in LA} z1_a^{i_a} = 1, \forall a \in D_1 \tag{12}$$

$$\sum_{j_a \in LT} z2_a^{j_a} = 1, \forall a \in D_2 \tag{13}$$

$$\sum_{a\in D_1}\sum_{i_a\in LA} c_a(i_a)z1_a^{i_a} - \eta\sum_{a\in D_2}\sum_{j_a\in LT} v_a z2_a^{j_a}j_a \leq B \tag{14}$$

$$z1_a^{i_a} = \{0,1\}, \forall a\in A\,, i_a\in LA_a \tag{15}$$

$$z2_a^{j_a} = \{0,1\}, \forall a\in A\,, j_a\in LT_a \tag{16}$$

$$\left[\begin{array}{c}\sum_{a\in D_1}\sum_{i_a\in LA_a} z1_a^{i_a} t_a(v_a,i_a)\delta_{a,r}^w + \sum_{a\in D_2}\sum_{j_a\in LT_a} z2_a^{j_a}t_a(v_a,j_a)\delta_{a,r}^w \\ -\sum_{a\in D_1\cap D_2}\sum_{i_a\in LA_a}\sum_{j_a\in LT_a} z1_a^{i_a} z2_a^{j_a} t_a(v_a,i_a,j_a)\delta_{a,r}^w + \\ \sum_{a\in A-D_1\cup D_2} t_a(v_a)\delta_{a,r}^w - \mu^w\end{array}\right] f_r^w = 0 \tag{17}$$

$$\sum_{a\in D_1}\sum_{i_a\in LA_a} z1_a^{i_a} t_a(v_a,i_a)\delta_{a,r}^w + \sum_{a\in D_2}\sum_{j_a\in LT_a} z2_a^{j_a}t_a(v_a,j_a)\delta_{a,r}^w$$
$$-\sum_{a\in D_1\cap D_2}\sum_{i_a\in LA_a}\sum_{j_a\in LT_a} z1_a^{i_a} z2_a^{j_a} t_a(v_a,i_a,j_a)\delta_{a,r}^w + \qquad \geq 0$$
$$\sum_{a\in A-D_1\cup D_2} t_a(v_a)\delta_{a,r}^w - \mu^w$$

$$\tag{18}$$

$$f_r^w \geq 0, \forall r\in R_w, \forall w\in W \tag{19}$$

Unfortunately, MINLP is non-convex because of the complementarity constraints (17)-(19) (Chen and Florian (1995); Luo et al. (1996); Scheel and Scholtes (2000)). To address this difficulty, some algorithmic approaches have been focused on avoiding this formulation (Ban et al. (2006); Bouza and Still (2007); Leyffer (2003)). The basic idea of these relaxation schemes is to get rid of the complicated complementarity constraints by replacing these conditions in a suitable way. There are different types of relaxation scheme, for example, Hoheisel et al. (2013) discussed five kinds of relaxation schemes including the global relaxation scheme by Scholtes (2001), the local relaxation scheme by Steffensen and Ulbrich (2010) and the relaxation scheme by Lin and Fukushima (2005), Kadrani et al. (2009), and Kanzow and Schwartz (2010).

In this paper, a relaxation algorithm will be adopted to iteratively tackle this MINLP. The main idea of this relaxation algorithm is as follows. The relaxation algorithm presented here introduces auxiliary parameters θ^w for each OD pair $\in W$, which can then be used to define relaxed complementarity slackness conditions, rather than the exact ones. We relax the complementarity slackness constraints (17) as follows:

$$
\begin{bmatrix}
\sum_{a \in D_1} \sum_{i_a \in LA_a} z1_a^{i_a} t_a(v_a, i_a)\delta_{a,r}^w + \sum_{a \in D_2} \sum_{j_a \in LT_a} z2_a^{j_a} t_a(v_a, j_a)\delta_{a,r}^w \\
- \sum_{a \in D_1 \cap D_2} \sum_{i_a \in LA_a} \sum_{j_a \in LT_a} z1_a^{i_a} z2_a^{j_a} t_a(v_a, i_a, j_a)\delta_{a,r}^w + \\
\sum_{a \in A - D_1 \cup D_2} t_a(v_a)\delta_{a,r}^w - \mu^w
\end{bmatrix} f_r^w \leq
$$
$$
\theta^w \tag{20}
$$

The MINLP can then become a relaxed MINLP problem (RMINLP, relaxed mixed-integer nonlinear optimization problem) by replacing equation (17) with equation (20). In this case, the RMINLP is non-convex and the MFCQ holds (Luo et al., 1996). Consequently, existing MINLP solution algorithms can be adopted to solve the RMINLP. The standard MINLP solver SBB is a GAMS solver for Mixed Integer Nonlinear Programming (MINLP) models, based on a combination of the standard Branch and Bound method from Mixed Integer Linear Programming and some of the standard NLP solvers already supported by GAMS (2009). It performs well on models that have difficult nonlinearities and possibly also on models that are fairly non-convex, more details about the SBB solver can refer to http://www.gams.com/dd/docs/solvers/sbb/index.html.

We use SBB to solve the RMINLP. The iterative algorithm can be summarized as follows:

Step 1. Initialization: Choose an initial auxiliary parameter θ_0^w for the complementarity slackness constraint. Set $\varepsilon \geq 0$, updating factor $0 < \lambda < 1$, and iteration number $= 0$.

Step 2. Major iteration: By setting θ_k^w as the auxiliary parameter for each complementarity condition in (20), solve the current RMINLP by using the SBB solver in GAMS (2009).

Step 3. Stop condition: If $\theta_k^w \geq \varepsilon$, stop and go to Step 4; otherwise set $\theta_{k+1}^w = \lambda \theta_k^w$, go to Step 2.

Step 4. Report solution: The optimal solution is achieved from the last run of Step 2.

The relaxation algorithm presented here is straightforward and easily implemented. In particular, many existing solution techniques for the MINLP can be used in Step 2 to tackle RMINLP. Suitable convergence results for the global relaxation scheme have already been known from the discussion given by Hoheisel et al. (2013). The fundamental approach of the relaxation scheme for mixed-integer nonlinear bi-level programming is to replace MINLP by a sequence of the parameterized RMINLP, which is similar to the relaxation method proposed in Scholtes (2001), but extends it to the mixed-integer case.

It is noted that, besides the relaxation methods, there are other ways to deal with the complementarity constraints, for example, dispensing with the complementarity constraints through the introduction of binary variables and a penalty method. All of these approaches have positive and negative features, and none is optimum for all problems. As an example, the binary variables can be used to remove the complementarity constraints as presented in Fortuny-Amat and McCarl (1981) (i.e., rewriting the complementary slackness condition by introducing binary variables and a large positive constant). This transforms the linear bi-level programming problem into a large mixed-integer programming problem. However, it is limited by computational speed due to the large size of the augmented problem (Wen and Hsu, 1991). The penalty method is another important class of algorithms to solve nonlinear bi-level programming problem, however they are generally limited to computing stationary points and local minima (Colson et al., 2005).

5. Numerical analysis

5.1 A two link network

We firstly consider a simple two-link network as shown in Fig 1, consisting of two nodes: node 1 and node 2 and two links: link a_1 and link a_2. The travel demand $d_1 = 100$. The travel time functions for both links are given by:

$$t_a(v_a) = t_a^0 \times \left(1 + 0.15 * \frac{v_a}{T_a}\right), a \in \{a_1, a_2\},$$
$$t_{a_1}^0 = 20, T_{a_1} = 1.5, t_{a_2}^0 = 70, T_{a_2} = 10.5,$$

where T_a is the current capacity of link $a \in \{a_1, a_2\}$ (i.e. $t_a(v_{a_1}) = 20 + 2 * v_{a_1}, t_a(v_{a_2}) = 70 + v_{a_2}$).

We assume the transport authority considers whether there is a need to add 0, 1 or 2 lanes to each link. The binary variables $z1_{a_1}^{i_{a_1}}$ and $z1_{a_2}^{i_{a_2}}$, $i_{a_1}, i_{a_2} \in \{0,1,2\}$ can then be used to indicate the multiple capacity level network design decisions. We assume that capacity per lane is 2, that is, $c_a^1 = 2$, the total budget $B = 7000$ and we set the parameter $\eta = 1$.

Concerning the construction cost function, we consider two cases:
  (i) The construction cost is a linear function of capacity, i.e. $c_{a_1}(i_{a_1}) = 2000i_{a_1}, c_{a_2}(i_{a_2}) = 3000i_{a_2}, i_{a_1}, i_{a_2} \in \{0,1,2\}$;
  (ii) The construction cost is a nonlinear function of capacity, i.e. $c_{a_1}(i_{a_1}) = 2000i_{a_1}^{1/2}, c_{a_2}(i_{a_2}) = 3000i_{a_2}^{1/3}, i_{a_1}, i_{a_2} \in \{0,1,2\}$.

We can also determine the toll level to make the two-link network reach SO.
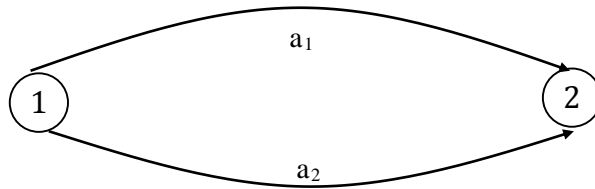

Figure 1 A two-link network

Table 1 Comparison of optimal solutions under different models.

Model	UE		SO		FBTP			DNDP			FBTP+DNDP			
	v_a	t_a	v_a	t_a	v_a	t_a	j_a	v_a	t_a	i_a	v_a	t_a	j_a	i_a
a_1	50	120	41.667	103.33	41.67	103.33	25	96.719	72.756	2	76.728	61.852	36	2
a_2	50	120	58.333	128.33	58.33	128.33	0	3.281	72.756	1	23.272	86.852	11	2
Z	12000		11791		11791			7275.590			6766.975			

Table 2 optimal solutions under a nonlinear construction cost

Model	DNDP			Toll+DNDP			
	v_a	t_a	i_a	v_a	t_a	j_a	i_a
a_1	97.059	72.941	2	76.728	61.852	25	2

a_2	2.941	72.941	0	23.272	86.852	0	2
Z	7275.590			6766.975			

We now present a comparison of optimal solutions under different models. For the convenience of comparison and demonstration, we firstly present the traffic pattern under different models including: the UE and SO solutions with a traffic assignment problem (TAP) (Sheffi, 1985), the first-best toll pricing problem (FBTP) (Hearn and Ramana, 1998; Yang and Huang, 2005), and DNDP (LeBlanc, 1975). The TAP under UE and SO conditions (which determines the UE and SO flow patterns in the transport network with fixed origin-destination demand), demonstrates firstly the traffic situation under the UE assumption that all users minimize their own individual travel costs on transport networks. Secondly it demonstrates the SO assumption that all users cooperate to minimize the total network cost. The UE and SO are two central concepts pertaining to the road pricing, the DNDP and combined DNDP and pricing problems. The FBTP[2] is solved to demonstrate the best performance under the road pricing scheme. The DNDP is solved to illustrate the optimal effect of increasing road infrastructure under given budget constraints. Finally, the results of our proposed model, which combines road toll pricing and capacity development from the perspective of government or high level decision makers, are presented. The comparisons are focused on the optimal traffic pattern and the objective function values. To avoid any bias during the process comparison, the objective function values of different models are obtained again by the relaxation algorithm given in Section 4. The results are given in Table 1.

Table 1 shows the link flow patterns under UE, SO, FBTP, DNDP and a combination of toll and DNDP. With the FBTP, the toll on link a_1 ($j_{a_1} = 25$) changes the traffic flow from UE to SO and the system total travel time decreases from $Z = 12000$ to $Z = 11791$. Under a budget constraint, $B = 7000$ with a linear construction cost (case i), the optimal road development pattern arising from the DNDP is to add two lanes for link a_1 and one lane

---

[2] Note that with the definition of first-best toll based on economic theory as the difference between marginal cost and average cost of traffic on a link; this is not integer by definition. Since we restrict the toll to an integer value in this paper, this will be a second best toll by definition. However, we still refer to it as the FBTP here as it has the same effect as a first-best toll.

for link a_2, in which case the system total travel time will decrease to $Z = 7275.59$.

We can further construct the combined first best toll pricing and capacity improvement model (FBTP + DNDP) with the first best toll pricing in the DNDP, as shown in Section 3.2. The upper level objective function still aims to minimize the total travel time under a budget constraint. With a policy of combined road toll pricing and capacity development, we can set the toll on link a_1 ($j_{a_1} = 36$) and on link a_2 ($j_{a_2} = 11$), and add two lanes for both link a_1 and link a_2. The system total travel time will decrease to $Z = 6766.975$. Under this case, the total construction cost is $c_{a_1}(i_{a_1}) + c_{a_2}(i_{a_2}) = 2000 * 2 + 3000 * 2 = 10000$, the toll revenue on both links is $v_{a_1} * j_{a_1} + v_{a_2} * j_{a_2} = 76.728 * 36 + 23.272 * 11 = 3018.2$, therefore the total available budget is $B + v_{a_1} * j_{a_1} = 10018.2$.

Under a budget constraint, $B = 5000$ with a nonlinear construction cost (case ii), the optimal road development pattern arising from the DNDP is to add two lanes on link a_1, and the system total travel time will decrease to $Z = 7275.59$. With a government policy of combined road toll pricing and capacity development, we can set the toll on link a_1 ($j_{a_1} = 25$) and add two lanes for both link a_1 and link a_2. In this case the system total travel time will decrease to 6766.975, the total construction cost is $c_{a_1}(i_{a_1}) + c_{a_2}(i_{a_2}) = 2000 * \sqrt[2]{2} + 3000 * \sqrt[3]{2} = 6608.2$, the toll revenue on link a_1 is $v_{a_1} * j_{a_1} = 76.728 * 25 = 1918.2$, therefore the total available budget is $B + v_{a_1} * j_{a_1} = 6918.2$.

Comparing combined road toll pricing and capacity development with other approaches (DNDP, FBTP, UE, and SO) we find that under the extension of both links, more travellers ($v_{a_1} = 76.728$) choose link a_1. With a fixed travel demand assumption, the new links added can reduce the total travel time and improve the level of transport service.

5.2 The Sioux-Falls network

The Sioux-Falls network, as shown in Figure 2, consists of 24 nodes, 76 links and 528 OD pairs, the parameters can be obtained from http://www.bgu.ac.il/~bargera/tntp/.
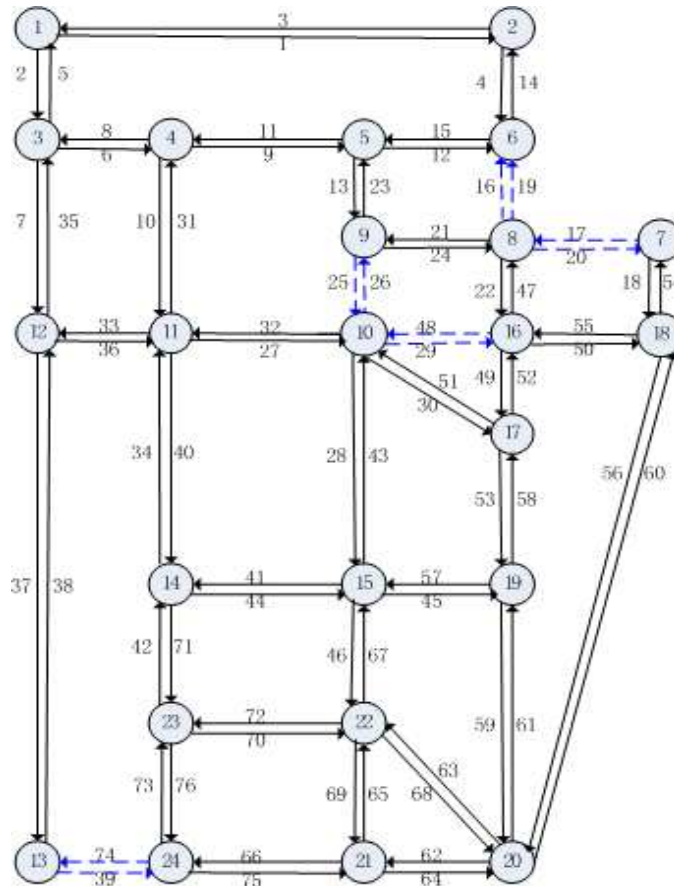

Figure 2 The Sioux-Falls network

Although it is possible to discuss the first-best toll (as shown for the simple network), it is not necessary for the purposes of the numerical demonstration here. To save space, we use a section of the links as the link toll set and the toll pricing problem therefore becomes a second-best road pricing problem. According to the UE assignment (Wang, et al., 2013; Wang et al., 2014a), we choose the following 10 links: (link 16, 17, 19, 20, 25, 26, 29, 39, 48, 74) as they are the most congested links in terms of the ratio of link flow to capacity. These 10 links are shown by dotted lines in Fig 2.

We assume the transport authority considers whether it needs to add 0, 1 or 2 lanes to these 10 links and that the capacity is 5 per lane. The budget is B

=5000. Without loss of generality, we assume a linear construction cost function, i.e., $c_a(i_a) = co_a * i_a$. The coefficients co_a for the 10 links are given in Table 3.

We use the SBB to solve the RMINLP in the relaxation algorithm. It performs well on models that have difficult nonlinearities and possibly also on models that are fairly non-convex. The computation time is less than 3 minutes.

Table 3 construction cost function coefficients

Node-Node	Link	co_a
6-8	16	260
7-8	17	400
8-6	19	260
8-7	20	400
9-10	25	250
10-9	26	250
10-16	29	480
13-24	39	340
16-10	48	480
24-13	74	340

(1) Comparison of optimal solutions under different models

Table 4 presents the link flow pattern (v_a), link cost (t_a) and toll j_a on different links, and link capacity extension (i_a) under different models (UE, SO, Second-best toll, DNDP and DNDP+toll). To avoid any bias in the comparisons, the objective function values of different models are obtained again using the relaxation algorithm proposed in this paper. As previously shown, the set of toll links and road capacity links in DNDP are not necessarily the same. It can be seen that in comparison to the UE case ($Z = 10062.743$), the second-best road pricing brings a reduction in total travel time ($Z = 10025.801$), however this is still not as low as the SO case ($Z = 9826.453$). Under the budget ($B = 5000$) and DNDP, Table 4 presents the optimal road expansion with a linear construction cost function given in Table 3 and illustrates the total travel time reducing to a level $Z = 7728.312$. A toll revenue 332.441 can be achieved with the

optical toll pricing scheme j_a (as shown in Table 4), and with a combination of toll pricing and DNDP the total travel time can reduce to 7029.808.

Table 4 Comparison of optimal solutions under different models.

Model		UE(TAP)		SO		Toll(Second-best)			DNDP			DNDP+Toll			
Node-Node	Link	v_a	t_a	v_a	t_a	v_a	t_a	j_a	v_a	t_a	i_a	v_a	t_a	j_a	i_a
6-8	16	14.017	22.114	13.786	20.816	13.944	21.694	2	18.190	7.238	2	19.228	2.832	0	2
7-8	17	14.053	7.641	14.631	8.453	14.127	7.740	0	13.018	6.418	0	11.662	5.201	0	0
8-6	19	14.017	22.114	13.786	20.816	13.944	21.694	2	16.500	11.817	1	17.600	2.584	5	2
8-7	20	14.053	7.641	14.631	8.453	14.127	7.740	0	13.862	7.394	0	12.301	5.725	3	0
9-10	25	24.013	6.990	23.901	6.916	24.240	7.143	0	25.740	6.078	1	25.702	3.600	2	2
10-9	26	24.013	6.990	23.901	6.916	24.240	7.143	0	26.331	5.100	2	27.088	3.741	0	2
10-16	29	11.926	26.848	11.769	25.718	11.936	26.917	1	15.370	11.015	2	18.440	6.653	3	2
13-24	39	11.728	20.896	11.640	20.395	10.715	15.769	9	15.733	9.381	2	16.668	4.893	4	2
16-10	48	11.926	26.848	11.769	25.718	11.936	26.917	1	15.909	11.903	2	20.971	7.765	0	2
24-13	74	11.728	20.896	11.640	20.395	10.715	15.769	9	16.053	9.833	2	17.072	4.983	2	2
Z		10062.743		9826.453		10025.801			7728.312			7029.808			
Toll revenue						272.518						332.441			

(2) Relationship between budget, toll pricing and total travel time

To continue the comparison of optimal solutions under different models, we examine the relationship between budget, toll pricing and total network travel time. For the case of the Sioux-Falls network we find that the maximum total investment (i.e., all links given in Table 3 with a two-lane expansion) is 6920, as shown in Table 5. That is, if the government budget is greater than 6920, there is no need for toll pricing and the total travel time will be $Z = 6881.759$. Table 5 also illustrates the toll revenue and corresponding total travel time (Z) under different budget levels. Generally, the lower the budget level, the higher the toll revenue and the lower the total travel time. However this is not the case when the budget is 5000, where the

optimal toll pricing scheme generates 332.441 total revenue. This is higher than the toll revenue of 145.344 with a budget level of 4000. Figure 3 illustrates the rate of reduction in total travel time with respect to the budget level, with the rate of reduction rate in total travel time declining as the budget level increases. Figure 4 indicates changes in toll revenue with respect to the budget and in contrast, a more erratic rate of decrease is seen as the budget level increases.

Table 5 Total travel time and toll revenue under different budget levels for the Sioux-Falls network

Budget (B)	1000	2000	3000	4000	5000	6000	≥ 6920
Toll Revenue	1168.888	676.197	657.551	145.344	332.441	133.318	91.143*
Total travel time (Z)**	7994.253	7613.050	7351.769	7152.049	7029.808	6891.008	6881.759

*Here the toll revenue is 91.143 not zero, as the toll is applied to reduce the total travel time.

**The maximum total investment cost, i.e., all links given in Table 3 with a two-lane expansion, is 6920. The effect is the same when total budget B is greater than 6920, given the subset D_1.
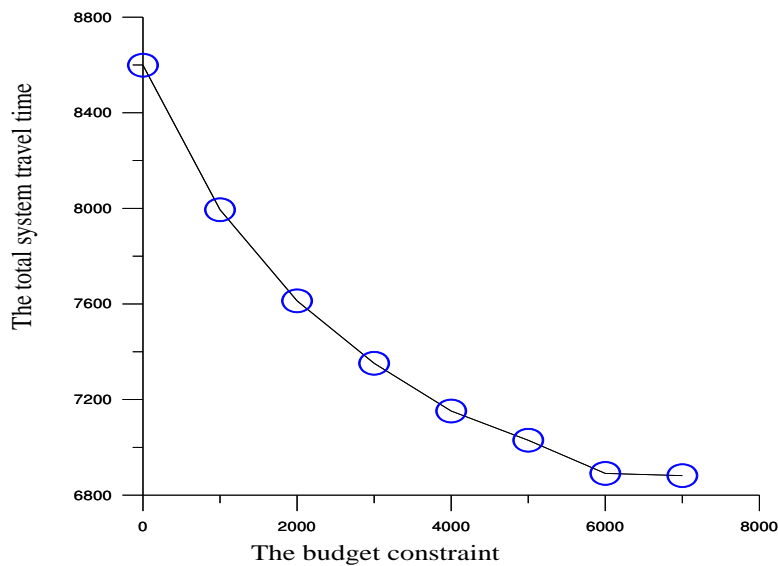

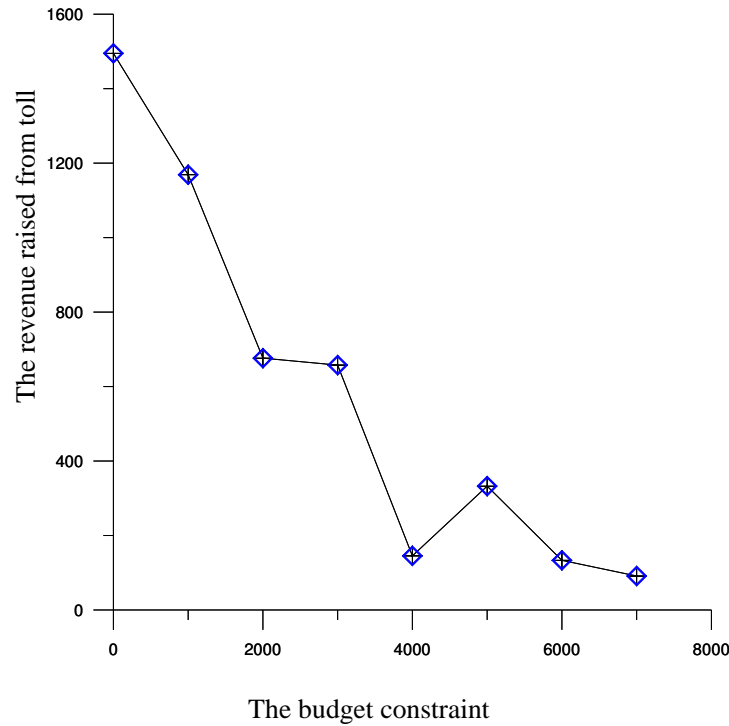Figure 3 Total travel time vs budget constraint for the Sioux-Falls network

Figure 4 Toll revenue vs budget level for the Sioux-Falls network

6. Conclusions

Research into (and applications of) packages of measures for transport development and management have received increasing recognition recently. In the context of the development and implementation of integrative measures in an urban road transport equilibrium network, this paper has presented research into the modelling of joint road toll pricing and capacity development, taking the perspective of government or highway authorities. The main contributions of this paper to the state of the art in the literature are as follows:

- The proposed formulation differs from existing mixed-integer programming formulations which apply where the model decision variables for selected links are discrete for either toll road or capacity expansion, and continuous for the level of toll or capacity. In contrast, the formulation proposed in this paper defines discrete variables for capacity (i.e. the number of additional lanes for capacity expansion) and for the toll level (i.e. an integer toll level within a pre-defined toll range). In the modelling approach described the two sets of links are, in fact, separated so that it is not necessary for D_1 and D_2 to coincide. This allows flexibility according to the policy being

designed, allowing, for example, for the revenues from tolling in one part of the network to be used to finance infrastructure improvements elsewhere in the network.

- To solve the bi-level program (with an integer nonlinear program for the upper level problem and a continuous nonlinear program for the lower level problem), we have first transformed it into a single-level mixed-integer nonlinear program with nonlinear complementarity constraints. A relaxation approach has then been developed to relax the complementarity constraints. At each iteration step, the SBB solver in GAMS was used to solve the relaxed mixed-integer nonlinear program. An existing business software (rather than bespoke software) was used to solve the complex mathematical programming problem, demonstrating that an application of the approach by transport planners/policymakers is practical.

This study further differs from existing analyses in two main aspects. Firstly, the toll pricing and capacity development in a discrete network design framework where different toll pricing and capacity expansion schemes can be compared explicitly. Secondly, the variances in total travel time with respect to the budget level and toll revenues are investigated explicitly. The discrete variable design with toll pricing and capacity expansion is more appropriate for a transport network. The proposed modelling approach can support government decision-makers to identify how, and under what circumstances, to set the budget level and the toll pricing level. It can also help in improving public understanding of how different transport policies can benefit the development of the transport network. Finally, the proposed model can be solved conveniently using existing optimal solvers, making it accessible in principle to the wider transport community and policy makers.

Some other key issues for further study are:

- Complex behaviour investigation. From the lower model (7-10), given the fixed travel demand, the link toll level and capacity level determined by the upper level, we use the UE principle to investigate the traveler's route choice behavior. However, network design and road pricing are long-term decisions and thus affect demand. They will also result in optimization of other ways, e.g., departure rates and individual travel mode choice. Further model development along these lines is a topic for further study.

- A multi-modal modelling approach for the additional complexities of the real world. With the proposed model framework, a further additional degree of complexity, i.e. public transport can be integrated. In a modern urban context, a policy that is only designed for cars, without taking public transport into account is largely unthinkable. Integrated transport planning and policy, especially in presence of pricing policies, requires multi-modal thinking.

References

Adler, N., Proost, S., 2010. Introduction to special issue of Transportation Research Part B Modelling non-urban transport investment and pricing, Transportation Research Part B, 44, 791-794.

Ban, J.X., Liu, H.X., Ferris, M.C., et al, 2006. A general MPCC model and its solution algorithm for continuous network design problem, Mathematical and Computer Modelling, 43 (5-6), 493-505.

Bazaraa, M.S., Sherali, H.D., Shetty, C.M., 2006. Nonlinear Programming- Theory and Algorithms. John Wiley & Sons, Hoboken, New Jersey.Boyce, D.E., Janson, B.N., 1980. A discrete transportation network design problem with combined trip distribution and assignment, Transportation Research Part B, 14(1), 147-154.

Chen, Y., Florian, M., 1995. The nonlinear bilevel programming problem: formulations, regularity and optimality conditions, Optimization, 32, 193-209.

Colson, B., Marcotte, P.and Savard, G., 2005. Bilevel programming: A survey, A Quarterly Journal of Operations Research, 3, 87-107.

D'Ouville, E.L., McDonald, J.F., 1990. Optimal road capacity with a suboptimal congestion toll, Journal of Urban Economics, 28(1), 24-49.

De Borger, B., Proost, S., 2001. Reforming transport pricing in the European Union. Edward Elgar, Cheltenham,UK.

Dimitriou, L., Tsekeris, T., Stathopoulos, A., 2009. Joint pricing and design of urban highways with spatial and user group heterogeneity, Netnomics, 10, 141-160.

Fan, W., Gurmu, Z., 2014. Combined decision making of congestion pricing and capacity expansion: Genetic algorithm approach, Journal of Transportation Engineering, 140(8), 04014031.

Farvaresh, H., Sepehri, M.M., 2011. A single-level mixed integer linear formulation for a bi-level discrete network design problem, 47(5), Transportation Research Part E, 623-640.

Farvaresh, H., Sepehri, M.M., 2013. A branch and bound algorithm for bilevel discrete network design problem, Networks and Spatial Economics, 13(1), 67-106.

Farahani, R.Z.; Miandoabchi, E.; Szeto, W.Y.; Rashidi, H., 2013. A review of urban transportation network design problems, European Journal of Operational Research, 229(2), 281-302.

Fortuny-Amat, J., McCarl B., 1981. A representation and economic interpretation of two-level programming problem, Journal of Operational Research Society, 32, 783-792.

GAMS., 2009. GAMS The solver manuals, GAMS Development Corporation, Washington, DC, USA.

Gao, Z.Y., Wu, J.J., Sun, H.J., 2005. Solution algorithm for the bilevel discrete network design problem, Transportation Research Part B, 39(6), 479-495.

Givoni, M.,2014. Addressing transport policy challenges through policy-packaging, Transportation Research Part A, 60, 1-8.

Givoni, M., Macmillen, J., Banister, D., Feitelson, E., 2013. From policy measures to policy packages, Transport Reviews 33(1), 1-20.

Grant-Muller, S., Xu, M., 2014. The role of tradable credit schemes in road traffic congestion management. Transport Reviews, 34(2), 128-149.

Gümüş, Z. H., Floudas, C.A., 2005. Global optimization of mixed-integer bilevel programming problems, Computational Management Science, 2(3), 181-212.

Hearn, D.W., Ramana, M.V., 1998. Solving congestion toll pricing models. In: Marcotte, P., Nguyen, S. (Eds.), Equilibrium and Advanced Transportation Modeling. Kluwer Academic Publishers, Boston, 109–124.

Hoheisel, T., Kanzow, C., Schwartz, A., 2013. Theoretical and numerical comparison of relaxation methods for mathematical programs with complementarity constraints. Mathematical Programming, 137, 257-288.

Jan, R.H., Chern, M.S., 1994. Nonlinear integer bilevel programing, European Journal of Operational Research, 72(3), 574-587.

Kadrani, A., Dussault, J.P., Benchakroun, A., 2009. A new regularization scheme for mathematical programs with complementarity constraints. SIAM Journal on Optimization, 20, 78-103.

Kanzow, C., Schwartz, A., 2010. A new regularization method for mathematical programs with complementarity constraints with strong convergence properties. Preprint 296. Institute of Mathematics, University of Wrzburg, Wrzburg.

Keeler, T.E., Small, K.A., 1977. Optimal peak-load pricing, investment, and service levels on urban expressways. Journal of Political Economy 85, 1-25.

Kelly, C., May, A., Jopson, A., 2008. The development of an option generation tool to identify potential transport policy packages. Transport Policy 15 (6), 361-371.

Koh, A., Shepherd, S.P., Sumalee, A. 2009. Second best toll and capacity optimisation in networks: solution algorithm and policy implications. Transportation, 36 (2), 147-165.

LeBlanc, L.J., 1975. An algorithm for the discrete network design problem, Transportation Science, 9(3), 183-199.

Leyffer, S., 2003. Mathematical programs with complementarity constraints. SIAG/OPT Views-and-News, 14(1), 15-18.

Lin, G.H., Fukushima, M., 2005. A modified relaxation scheme for mathematical programs with complementarity constraints. Annals of Operations Research, 133, 63-84.

Luo, Z.Q., Pang, J.S., Ralph, D., 1996. Mathematical programs with equilibrium constraints. Cambridge, UK: Cambridge University Press.

Magnanti, T.L., Wong, R.T., 1984. Network design and transportation planning: models and algorithms, Transportation Science, 18(1), 1-55.

Marsden, G., Frick, K.T., May, A.D., Deakin, E., 2010. Transfer of innovative policies between cities to promote sustainability: Case study evidence, Transportation Research Record, 89-96.

May, A.D., Karlstrom, A., Marler, N., Matthews, B., Minken, H., Monzon, A., Page, M., Pfaffenbichler, P., Shepherd, S., 2005b. Developing sustainable urban land use and transport strategies: a decision-makers' guidebook. Institute for Transport Studies, Leeds. Working paper.

May, A.D., Kelly, C., Shepherd, S.P., 2006. The principles of integration in urban transport strategies. Transport Policy 13 (4), 319-327.

May, A.D., Shepherd, S.P., Emberger, G., Ash, A., Zhang, X., Paulley, N., 2005a. Optimal land use and transport strategies: methodology and application to European cities. Transportation Research Record 1924, 129-138.

May, A.D., Still, B.J., 2000. The Instruments of Transport Policy. Working Paper WP545. Institute for Transport Studies, Leeds.

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

Mitsos, A., 2010. Global solution of nonlinear mixed-integer bilevel programs, Journal of Global Optimization, 47, 557-582.

Mohring, H., Harwitz, M., 1962. Highway Benefits. Northwestern University Press, Evanston, Illinois.

Moore, J.T., Bard, F.J., 1990. The mixed integer linear bilevel programming problem, Operations Research, 38(5), 911- 921.

Niu, B.Z., Zhang, J., 2013. Price, capacity and concession period decisions of Pareto-efficient BOT contracts with demand uncertainty, Transportation Research Part E, 53, 1-14.

Sahin, K.H., Ciric, A.R., 1998. A dual temperature simulated annealing approach for solving bilevel programming problems, Computers & Chemical Engineering, 23(1), 11-25.

Scheel, H., Scholtes, S., 2000. Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity, Mathematics of Operations Research, 22(1), 1-22.

Scholtes, S., 2001. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. SIAM Journal on Optimization, 11 (4),918-936.

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

Shepherd, S.P., Zhang, X., Emberger, G., May, A.D., Hudson, M., Paulley, N., 2006. Designing optimal urban transport strategies: the role of individual policy instruments and the impact of financial constraints. Transport Policy 13 (1), 49–65.

Steffensen, S., Ulbrich, M., 2010. A new relaxation scheme for mathematical programs with equilibrium constraints. SIAM Journal on Optimization, 20, 2504-2539.

Strotz, R.H., 1965. Urban transportation parables. In: Margolis, J. (Ed.), The Public Economy of Urban Communities. Resources for the Future, Washington, DC.

Subprasom, K., Chen, A., 2007. Effects of regulation on highway pricing and capacity choice of a build-operate-transfer scheme, Journal of Construction and Engineering Management, 133(1), 64-71.

Tan, Z.J., Yang, H., Guo, X.L., 2010. Properties of Pareto efficient contracts and regulations for road franchising. Transportation Research Part B, 44 (4), 415-433.

Tan, Z.J., Yang, H., 2012. Flexible build-operate-transfer contracts for road franchising under demand uncertainty. Transportation Research Part B, 46(10), 1419–1439.

Verhoef, E., Koh, A., Shepherd, S., 2010. Pricing, capacity and long-run cost functions for first-best and second-best network problems. Transportation Research Part B, 44, 870-885.

Wang, G.M., Gao, Z.Y., Xu, M., Sun, H.J., 2014a. Models and a relaxation algorithm for continuous network design problem with a tradable credit scheme and equity constraints, Computers & Operations Research, 41(1), 252-261.

Wang, G.M., Gao, Z.Y., Xu, M., Sun, H.J., 2014b. Joint link-based credit charging and road capacity improvement in continuous network design problem, Transportation Research Part A, 67, 1-14.

Wang, S.A., Meng, Q., Yang, H., 2013. Global optimization methods for the discrete network design problem, Transportation Research Part B, 50, 42-60.

Wen, U.P., and Hsu, S.T., 1991. Linear bi-level programming problem: a review, Journal of Operational Research Society, 42(2), 125-133.

Xu, M., Grant-Muller, S., Huang, H.J., Gao, Z.Y., 2015. Transport management measures in the post-olympic games period: Supporting sustainable urban mobility for Beijing? International Journal of Sustainable Development and World Ecology. 22(1), 50-63.

Yang, H., Bell, M.G.H., 1998. Models and algorithms for road network design: a review and some new developments. Transport Reviews, 18 (3), 257-278.

Yang, H., Huang, H.J., 2005.Mathematical and Economic Theory of Road Pricing. Elsevier Ltd. Oxford, UK.

Yang, H., Meng, Q., 2000. Highway pricing and capacity choice in a road network under a build-operate-transfer scheme. Transportation Research A, 34, 207-222.

Yang, H., Xu, W., Heydecker, B., 2010. Bounding the efficiency of road pricing, Transportation Research E, 46(1), 90-108.

Yin, Y., Li, Z.C., Lam, W.H.K., Choi, K., 2014. Sustainable toll pricing and capacity investment in a congested road network: A goal programming approach, Journal of Transportation Engineering, 04014062.

Zhang, L., Lawphongpanich, S. and Yin, Y. 2009., Reformulating and solving discrete network design problem via an active set technique. In Proceedings of 18th International Symposium on Transportation and Traffic Theory (Lam, W.H.K., Wong, S.C. and Lo, H.K. eds), Springer, 283-300.

Zhang, X.N., Wee, B., 2012. Enhancing transportation network capacity by congestion pricing with simultaneous toll location and toll level optimization, Engineering Optimization, 44(4), 477-488.