Transportation Research Part B 76 (2015) 93–116

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Cordon toll competition in a network of two cities: Formulation and sensitivity to traveller route and demand responses

D.P. Watling, S.P. Shepherd, A. Koh *

Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, United Kingdom

A R T I C L E I N F O

Article history:
Received 31 July 2013
Received in revised form 17 February 2015
Accepted 18 February 2015
Available online 30 March 2015

Keywords:
Route choice
Stochastic User Equilibrium
Equilibrium problems with equilibrium constraints
Multiple Nash Equilibria
Complementarity problem

A B S T R A C T

While there exists extensive literature on the first- and second-best tolling of congested transportation networks, much of it presumes the existence of a single agent responsible for toll-setting. The present paper extends the small but growing body of work studying the impact of several agents independently regulating tolls on different parts of a network. Specifically we consider the problem of a network consisting of two 'cities', each city independently regulated by a city 'authority' able to set a single cordon toll for entry to the city. It is supposed that each authority aims to maximise the social welfare of its own residents, anticipating the impact of its toll on travellers' route and demand decisions, while reacting to the toll level levied by the other authority. In addition, we model the possibility of the cities entering into a 'tax-exporting agreement', in which city A agrees to share with city B the toll revenues it collects from city B residents using city A's network. It is assumed that the sensitivity of travellers, in terms of their route and demand responses, is captured by an elastic demand, Stochastic User Equilibrium (SUE) model. Conditions for a Nash Equilibrium (NE) between cities are set out as an Equilibrium Problem with Equilibrium Constraints (EPEC). It is shown that weaker, 'local' solutions to the EPEC (which we term LNE for local NE) satisfy a *single* variational inequality, using the smooth implicit function of the SUE map. Standard variational algorithms may then be used to identify such LNE solutions, allowing NE solutions to be identified from this candidate set; we test the use of a Sequential Linear Complementarity Problem algorithm. Numerical results are reported in which we see that the sensitivity of travellers may affect many factors, including: the number of LNE solutions, the initial conditions for which algorithms might determine such solutions, the gap between LNE and a global regulator solution, and the incentive for cities to cooperate in terms of tax-exporting.

1. Introduction

There are many ways in which transport authorities may influence travel behaviour in some socially-desirable manner through pricing mechanisms. One such instrument which has attracted considerable research attention is the use of road tolls—levied on drivers when using a transport facility or crossing a city cordon, for example—so as to both manage the demand for private car travel and to influence route choice. Much of the literature on tolling has focused on a benevolent regulator who sets tolls on the whole or some part of a congested transportation network so as to maximise some measure

* Corresponding author.
E-mail address: a.koh@its.leeds.ac.uk (A. Koh).

http://dx.doi.org/10.1016/j.trb.2015.02.007
0191-2615/© 2015 The Authors. Published by Elsevier Ltd.

Nomenclature

Symbol	Interpretation
A, B	Authorities A and B
$\lvert\cdot\rvert$	cardinality operator
$y = \mathrm{SOL}(\cdot)$	y is obtained as the solution of the functions in parenthesis
#	matrix transpose operator
θ	logit dispersion parameter
$\mathcal{L}$	index set of links in the traffic network $\mathcal{L} = \{1, 2, \ldots, \lvert\mathcal{L}\rvert\}$
τ_i	a single toll level chosen for links within authority $i, i \in \{A, B\}$ such that $\boldsymbol{\tau} = (\tau_{\mathrm{A}}, \tau_{\mathrm{B}})^{\#}$
v_l	link flows for link $l, l \in \mathcal{L}$ collected in vector form as $\mathbf{v} = (v_1, v_2, \ldots, v_{\lvert\mathcal{L}\rvert})^{\#}$
$t_l(v_l)$	generalised travel time as a separable, differentiable and monotonically increasing function of link flow on link l when the flow on link l is $v_l, l \in \{1, 2, \ldots, \lvert L\rvert\}$ and collected in vector form as $\mathbf{t}(\mathbf{v}) = (t_1(v_1), t_2(v_2), \ldots, t_{\lvert\mathcal{L}\rvert}(v_{\lvert\mathcal{L}\rvert}))^{\#}$
K	index set of Origin Destination (OD) pairs in the network $K = \{1, 2, \ldots, \lvert K\rvert\}$
q_k	the travel demand (level) for OD movement k. $K = \{1, 2, \ldots, \lvert K\rvert\}$ and collected in vector form as $\mathbf{q} = (q_1, q_2, \ldots, q_{\lvert K\rvert})^{\#}$
$\boldsymbol{\Lambda}$	$\lvert\mathcal{L}\rvert \times 2$ Toll link-authority incidence matrix $\boldsymbol{\Lambda}$, with elements defined as follows: $\Lambda_{li} = \begin{cases} 1, \text{ if link } l, l \in \mathcal{L} \text{ may be tolled by authority } i, i \in \{A, B\} \\ 0, \text{otherwise} \end{cases}$
$\mathcal{R}$	the index set of all acyclic routes, $\mathcal{R} = \{1, 2, \ldots, \lvert\mathcal{R}\rvert\}$ and $\mathcal{R}_k \subseteq \mathcal{R}$, is the subset serving OD pair $k, k \in \{1, 2, \ldots, \lvert K\rvert\}$.
$\boldsymbol{\Delta}$	$\lvert\mathcal{L}\rvert \times \lvert\mathcal{R}\rvert$ link-route incidence matrix, with elements defined as follows: $\Delta_{lr} = \begin{cases} 1, \text{ if link } l, l \in \mathcal{L} \text{ is part of route } r, r \in \{1, 2, \ldots, \lvert\mathcal{R}\rvert\} \\ 0, \text{otherwise} \end{cases}$
$\boldsymbol{\Gamma}$	$\lvert K\rvert \times \lvert R\rvert$ OD-route incidence matrix with elements defined as follows: $\Gamma_{kr} = \begin{cases} 1, \text{ if } r \in \mathcal{R}_k, k \in \{1, 2, \ldots, \lvert K\rvert\}; r \in \{1, 2, \ldots, \lvert\mathcal{R}\rvert\} \\ 0, \text{otherwise} \end{cases}$
D	(convex) demand-feasible set
$\mathbf{f}$	vector of path flows
$\mathbf{c}_{[k]}$	the vector of route cost for OD movement k. $k = \{1, 2, \ldots, \lvert K\rvert\}$
$p_r(\mathbf{c}_{[k]}; \theta)$	for OD pair k, $k \in \{1, 2, \ldots, \lvert K\rvert\}$, probability of using route r, $r \in \mathcal{R}_k$ when the perceived distribution of route utilities is parameterized in a random utility model by $\mathbf{c}_{[k]}$ and θ. This is collected in vector form as $\mathbf{p}_{[k]}(\mathbf{c}_{[k]}; \theta)$
$s_k(\mathbf{c}_{[k]}; \theta)$	composite generalised travel time function for OD pair k, $k \in \{1, 2, \ldots, \lvert K\rvert\}$ and collected in vector form $\mathbf{s}(\mathbf{c}; \theta)$
$d_k(s_k)$	separable, bounded, continuously differentiable and monotonically decreasing demand function for OD pair k, $k \in \{1, 2, \ldots, \lvert K\rvert\}$ such that $q_k = d_k(s_k)$ is the demand level for OD movement k when that OD movement's composite generalised travel time is s_k. These functions are collected in vector form as $\mathbf{d}(\mathbf{s}) = (d_1(s_1), d_2(s_2), \ldots, d_{\lvert K\rvert}(s_{\lvert K\rvert}))^{\#}$
$g_k(q_k)$	the inverse demand function for OD pair k, $k \in \{1, 2, \ldots, \lvert K\rvert\}$ such that $s_k = g_k(q_k)$ is the OD composite generalised travel time with demand level q_k
t_l^0	free flow travel time for link $l, l \in \mathcal{L}$
$\Psi_{GLOBAL}(\mathbf{v}, \mathbf{q}, \boldsymbol{\tau})$	global welfare measure
$h(\boldsymbol{\tau})$	global welfare measure as a function of tolls given that link flows and demands are implicitly defined by a SUE model
$\mathbf{v}^{(\mathrm{A})}, \mathbf{v}^{(\mathrm{B})}$	disaggregation of link flows by resident authority origin of trips such that $\mathbf{v} = \mathbf{v}^{(\mathrm{A})} + \mathbf{v}^{(\mathrm{B})}$
$\mathbf{q}^{(\mathrm{A})}, \mathbf{q}^{(\mathrm{B})}$	disaggregation of demands by resident authority origin of trips such that $\mathbf{q} = \mathbf{q}^{(\mathrm{A})} + \mathbf{q}^{(\mathrm{B})}$
$\Psi_{\mathrm{A}}^0(\mathbf{v}^{(\mathrm{A})}, \mathbf{v}^{(\mathrm{B})}, \mathbf{q}^{(\mathrm{A})}, \mathbf{q}^{(\mathrm{B})}, \tau_{\mathrm{A}} \vert \tau_{\mathrm{B}})$	welfare measure for Authority A computed at equilibrium link flows and demands dependent on the tolls set by opposing authority without tax exporting ($\alpha = 0$) given τ_{B}
$\Psi_{\mathrm{A}}^1(\mathbf{v}^{(\mathrm{A})}, \mathbf{v}^{(\mathrm{B})}, \mathbf{q}^{(\mathrm{A})}, \mathbf{q}^{(\mathrm{B})}, \tau_{\mathrm{A}} \vert \tau_{\mathrm{B}})$	welfare measure for Authority A computed at equilibrium link flows and demands dependent on the tolls set by opposing authority with tax exporting ($\alpha = 1$) given τ_{B}
$\boldsymbol{\Omega}^{(\mathrm{A})}$	$\lvert K\rvert \times \lvert K\rvert$ diagonal matrix with elements defined as follows: $\Omega_{kk}^{(A)} = \begin{cases} 1, \text{ if OD movement } k \text{ has an origin in authority A}, \\ 0, \text{ otherwise} \end{cases}$
$\boldsymbol{\Omega}^{(\mathbf{B})}$	$\lvert K\rvert \times \lvert K\rvert$ diagonal matrix with elements defined as follows: $\Omega_{kk}^{(B)} = \begin{cases} 1, \text{ if OD movement } k \text{ has an origin in authority B}, \\ 0, \text{ otherwise} \end{cases}$

$\mathbf{h}(\boldsymbol{\tau})$	vector of welfare objective of each authority
$\mathbf{h}'(\boldsymbol{\tau})$	first order derivative of $\mathbf{h}(\boldsymbol{\tau})$
$\nabla\mathbf{h}'(\boldsymbol{\tau}^{\mathbf{0}})$	a 2 by 2 Jacobian matrix of $\mathbf{h}'(\boldsymbol{\tau})$ with respect to toll level evaluated at toll vector $\boldsymbol{\tau}^{\mathbf{0}}$. Diagonal elements are $\frac{\partial \mathbf{h}_i'(\tau_i^0)}{\partial \tau_i^0}, i \in \{\mathrm{A},\mathrm{B}\}$)off diagonal elements are given by $(\frac{\partial \mathbf{h}_i'(\tau_i^0)}{\partial \tau_j^0}), \mathbf{i} \in \{\mathrm{A},\mathrm{B}\} i \neq j$
$L(\boldsymbol{\tau}\vert\boldsymbol{\tau}^0)$	first order Taylor (linear) approximation of $\mathbf{h}'(\boldsymbol{\tau})$ at toll vector $\boldsymbol{\tau}^{\mathbf{0}}$

of social welfare (e.g. Verhoef et al., 1996; Yang and Lam, 1996; Hearn and Ramana, 1998). This literature is fundamentally based on the economic argument that drivers in congested conditions should be charged a toll so that they recognise the delay they cause to others, i.e. through "internalization" of the externality (Pigou, 1920; Yang and Huang, 2005). This literature assumes the regulator to be concerned with the social welfare of *all* users in the network, and that they are the single agent responsible for setting toll levels.

In reality, however, transportation systems may have several agents able to influence pricing in some way, each with its own jurisdiction and potentially conflicting objectives. Reflecting such considerations and the growing political inclination toward deregulation, a burgeoning literature (to which our work is closely related) is emerging in the toll pricing literature on the regulation and ownership of private toll roads, motivated by global interests in highway franchising (Engel et al., 1997; Tan et al., 2010; Tan and Yang, 2012). Apace with the development of theory, real-life systems of this nature are increasingly apparent. For example, there are several toll roads developed as part of Public Private Partnerships in competition with each other in the Australian cities of Brisbane, Melbourne and Sydney (Li and Hensher, 2010).

In terms of theoretical analysis of competition in highway networks, both Verhoef et al. (1996) and de Palma and Lindsey (2000) considered, amongst other scenarios, a private toll road operator in competition with a toll-free route. The authors demonstrated that a private toll road operator would be incentivized to internalize the congestion externality on the toll road. This conclusion was supported in further work by Engel et al. (2004) and Acemoglu and Ozdaglar (2007) who conjectured that when several operators competed on parallel routes to maximise revenue, toll competition would substitute for toll regulation and that in the limit, this would in fact be equal to the welfare maximising toll level. This body of work, however, was limited to the case of routes in parallel. On the other hand, it has been established in the literature (e.g. Mills, 1995; Small and Verhoef, 2007; Mun and Ahn, 2008; van den Berg, 2013) that if each link in a serial network was controlled by a different private operator, each operator would independently set tolls not only to internalize the externality of the link under their control but also the externalities of other links in the series. This would imply excessively high tolls and hence result in a welfare loss. This is a manifestation of "double marginalization", a phenomenon recognised in the industrial economics literature (Economides and Salop, 1992).

While the above works all focused on competition between several *private* operators, it is generally the case that highway networks span jurisdictional boundaries which are effectively artificial demarcations, meaning that even without private involvement there exists competition, but in this case between competing *public* organisations. To this end, De Borger et al. (2005) studied the case where several public organisations such as city authorities use tolls in competition with each other. They considered local and through traffic in a serial network setting, and found that governments would impose inefficient tolls on links which they control. They argue that this is a result of 'tax exporting' behaviour, whereby any authority charging tolls on non-residents of their jurisdiction is effectively transferring revenue across the jurisdictional boundaries. The analysis was extended in De Borger et al. (2007) to allow for simultaneous toll and capacity choices by the governments. These papers motivate the work presented in this paper, where (unlike the focus of much literature on competing private operators), the main actors in the game we shall formulate will be public city authorities that are in competition with each other.

Understanding the nature and outcome of the interactions between competing public bodies is important for several reasons. Many authors have argued that the state of institutional governance of transport is critical for the successful delivery of policy (Pemberton, 2000; Marsden and May, 2006). Pemberton illustrates the interaction of decision makers in local government during the design of integrated transport strategies for funding of transport improvements for the North East region of England. Officials in the areas of Sunderland, North Tyneside, South Tyneside and Gateshead perceived an "over dominance in terms of policy determination by Newcastle City Council" (Pemberton, 2000, p. 300). Marsden and May (2006) point to, on the one hand, the fear of competition between restrictive parking charges in city centres for traffic management purposes, and on the other hand, the loss of retail to outlying retail parks offering cheap and plentiful parking. We can thus see that there are several different kinds of competition possible between public bodies. In this sense, it is helpful in the first instance to distinguish between *horizontal* and *vertical* fiscal competition in the context of competition between authorities (De Borger and Proost, 2012). By vertical competition, we mean the potential for higher and lower level governments to tax the same tax base. In contrast, horizontal competition implies the desire of governments (at the same level such as city authorities in the UK or states in the US) to shift the tax burden to users from other jurisdictions.

We may then ask, how might these different kinds of public competition be evident in a transport policy context? Both Levinson (2001) and Rork (2009) addressed this question using an empirical approach to study the issue of 'tax exporting' behaviour across states in the US. They found econometric evidence to support the hypothesis that tolls were used as an instrument for horizontal competition, finding that the greater portion of non-resident commuters a state had, then apparently the more likely the state would be to levy tolls. Proost and Sen (2006), on the other hand, utilised a strategic model to explore the issue of tax-exporting behaviour and its potentially negative impacts, studying interactions between governments at two different levels, i.e. an example of vertical competition. In their model, a regional government was assumed to be in charge of setting tolls for a cordon, while a local government had control of parking charges. They concluded that there was the possibility for extensive welfare losses if the authorities did not cooperate in the application of these instruments. They showed that in such cases most of the explanation for the inefficiency was attributable to the desire for tax-exporting behaviour by the city government, which tended to extract very high parking charges from "commuters", that is, residents of the wider region who were not residents of the city and who travelled into the city area. While Proost and Sen's model had no representation of the road network, Ubbels and Verhoef (2008) considered horizontal competition in a serial network (with no route choice) where each government was assumed to have simultaneous toll and capacity setting decisions. They demonstrated, through numerical simulations, that horizontal competition between governments may have a negative impact on welfare.

Our focus in this paper is on horizontal competition, such as that which may occur between physically-adjacent jurisdictional regions, which we refer to as 'competing cities', with each 'city authority' able to determine the toll regime in its own city in order to tax private transportation. The 'competition' arises from two aspects. Firstly, the travellers who are resident in one city may use the transportation network of the non-resident city and therefore face the tolls set by either or both city authorities; therefore there is a dependence of the welfare of the residents of one city on the tolls set by both city authorities. Secondly, we allow for the potential for there to be some agreement between the authorities in terms of 'tax export', i.e. that a fraction of the toll revenue from non-residents of a city will be passed to their resident authority. Our interest is specifically in doing so in the context of congested networks where there are heterogeneous origin–destination movements, and the potential for travellers to potentially re-route or not travel by private transport, in order to avoid paying tolls. This provides an additional dimension to the two direct aspects of the competition between cities, in that the travellers themselves are also 'competing' (for scarce road space). This means that to effectively set tolls even from their own perspective (regardless of other city authorities), then a city must anticipate the impact of their policy on the behaviour of the travellers both inside and outside their jurisdiction. From the large body of academic work on tolling problems, we now know that these traveller responses are highly dependent on network structure and the location of activities relative to any tolled part of the network (May and Milne, 2000), since these critically affect the potential to re-route. Therefore, aside from our focus on horizontal as opposed to vertical competition, our work is unlike that of Proost and Sen (2006) in that we employ a detailed network representation with the potential for re-routing.

In our model formulation, we utilise tolls as the only strategic variable over which the authorities exercise control. This is in contrast to other authors (Xiao et al., 2007; Yang et al., 2009) who assume that the private operators play a single shot game in both determining tolls and capacities simultaneously based on a Nash game. Thus our work relates closely to that of Zhang et al. (2011), and extends our own recent work reported in Gühnemann et al. (2011, 2014) and Koh et al. (2012).

These studies also used detailed traffic network models to study the effects of competition between city authorities. A key element of all these studies was, as in the case of private tolling literature, that the traffic assignment was based on Wardrop's User Equilibrium (UE) assignment (Wardrop, 1952). However, there are several technical difficulties and issues that arise from such an assumption. Firstly, when we include the potential for partial tax-exporting agreements in the objective function for each city, then we need to know link flows disaggregated by originating city, but in general these are not uniquely defined by the UE model at given toll levels. Therefore such objective functions are typically not well-defined unless highly restrictive assumptions are made on the network structure and permitted tolls to enforce such uniqueness (see Koh et al., 2012). Secondly, even in the limited cases where uniqueness can be assured, the UE flows are non-smooth when viewed as an implicit mapping from the tolls to equilibrium flows. Thus, one cannot readily exploit the typically smooth nature of the social welfare functions with standard gradient-based optimisation methods. Thirdly, the overall nature of the resulting problem (as an Equilibrium Problem with Equilibrium Constraints, EPEC) is known to have an especially complex and special mathematical structure, which limits the opportunities for testing alternative solution strategies. We show that our approach, on the other hand, can be cast in a form to requires the solution only of a *single-level* variational inequality, thus opening up the problem to a range of possible solution algorithms Fourthly, it has been shown that such problems may have a range of 'local' and 'global' solutions that may emerge (Koh et al., 2012), yet it is unclear how sensitive these findings are to the UE assumption of identical, perfect information of the travellers; with even a small dispersion in behaviour and change in sensitivity of the travellers, would the same phenomena arise?

In the present paper, in order to address all four of the concerns above, we instead adopt a Stochastic User Equilibrium (SUE) model, focusing on the particular case of two cities competing over a cordon toll level. In Section 2, we develop the appropriate welfare functions for the competing cities, inclusive of partial tax-exporting agreements, which we show to be well-defined under the SUE model. These objective functions form the basis of the two-level game, with cities competing at the upper level and travellers competing for road space at the lower level. We then go on to develop an equivalent formulation, in which certain 'local' solutions of this two-level game may be cast as a *single-level* variational inequality problem. Having defined the solution algorithm (Section 3), Section 4 is devoted to numerical experiments, where we explore the sensitivity of the resulting solutions to various parameters of the model, and discuss the implications of the findings for policy. Section 5 concludes.

2. Notation: assumptions and mathematical formulation

2.1. Basic notation and assumptions

We represent the highway network as a graph consisting of links indexed by the set $\mathcal{L} = \{1, 2, \ldots, |\mathcal{L}|\}$. We suppose there to be two regulatory authorities (labelled A and B), each authority having their own pre-defined subset of network links over which they may charge a toll. These potentially-tolled links are identified by means of a link-authority incidence matrix Λ of dimension $|\mathcal{L}| \times 2$ with elements Λ_{li} equal to 1 only if link l may be tolled by authority i and equal to 0 otherwise ($l \in \{1, 2, \ldots, |\mathcal{L}|\}; i \in \{A, B\}$). We make the restriction in the present paper that each authority i has a *single*, non-negative toll level $\tau_i \in [0, \infty)$ that they may determine and levy on their tollable links. Since we assume that all users perceive the tolls in the same way, regardless of their socio-economic attributes, we shall assume the tolls to be in equivalent travel time units; that is to say, what we refer to as the 'toll' is effectively the monetary charge divided by the common value-of-time. Such an assumption is not necessary, but we believe considerably helps the appreciation of our later numerical experiments, where we can present tolls in such equivalent travel time units, without having to consider the important but quite separate issues of different units of currency and values-of-time (the assumption is justified by the fact that we consider only a single class of travellers, whereas a natural multi-class extension of our work would more easily be derived with tolls in monetary units and explicit values-of-time by class). Together, the two tolls to be determined can be collected in the vector $\boldsymbol{\tau} = (\tau_A, \tau_B)^{\#}$ with # denoting the transpose, with non-negativity constraints on the elements of the toll vector denoted $\boldsymbol{\tau} \geqslant \mathbf{0}$.

Aside from tolls, other factors will motivate travellers in terms of whether to travel and which route to take. We focus here only on travel time, since it is sufficient to illustrate our modelling approach.[1] We denote by $\mathbf{v}$ the vector of flows on the links of the network, with typical element v_l denoting the flow on link l (for $l = 1, 2, \ldots, |L|$). Then the function $t_l(\mathbf{v})$ denotes the travel time on link l when the link flows are $\mathbf{v}$. Then $\mathbf{t}(\mathbf{v}) = (t_1(\mathbf{v}), t_2(\mathbf{v}), \ldots, t_{|\mathcal{L}|}(\mathbf{v}))^{\#}$ denotes the vector mapping from link flows to link travel times. We suppose the vector function $\mathbf{t}(\cdot)$ to be differentiable and monotone in the sense of Cantarella (1997).

As well as links, the network consists of several other, mutually-dependent entities. Firstly, there are Origin–Destination (OD) movements which are indexed by the set $K = \{1, 2, \ldots, |K|\}$, with q_k ($k \in K$) denoting the travel demand for OD movement k. Secondly, the routes in the network are labelled such that $\mathcal{R} = \{1, 2, \ldots, |\mathcal{R}|\}$ denotes the index set of all acyclic routes, with $\mathcal{R}_k \subseteq \mathcal{R}$, the subset of routes serving OD pair $k \in \{1, 2, \ldots, |K|\}$. The relationship between routes and links is specified through the $|\mathcal{L}| \times |\mathcal{R}|$ link-route incidence matrix $\boldsymbol{\Delta}$, with elements Δ_{lr} equal to 1 only if link l is part of route r, and equal to 0 otherwise ($l = 1, 2, \ldots, |\mathcal{L}|; r = 1, 2, \ldots, |\mathcal{R}|$). The relationship between routes and OD movements is specified through the $|K| \times |\mathcal{R}|$ OD-route incidence matrix $\boldsymbol{\Gamma}$ with elements $\boldsymbol{\Gamma}_{kr}$ equal to 1 if $r \in \mathcal{R}_k$ and equal to 0 otherwise ($k = 1, 2, \ldots, |K|$). The convex set of feasible combinations of the OD travel demand vector $\mathbf{q}$ and the link flow vector $\mathbf{v}$ is denoted by:

$$D = \{(\mathbf{v}, \mathbf{q}) : \mathbf{v} = \Delta\mathbf{f} \text{ and } \mathbf{q} = \Gamma\mathbf{f}, \text{ where } \mathbf{f} \geqslant 0, \mathbf{f} \in \mathbb{R}^{|\mathcal{R}|}\}. \tag{1}$$

In terms of modelling route choice, we suppose that those individuals who decide to make a journey will choose between the available routes in proportions given by a random utility model. Our notational convention is to use a square bracket around an index to highlight cases where the index is referring to a partition of a vector rather than an individual element if it. Thus, we write q_k (without square brackets on the index) to refer to the *element* of the vector $\mathbf{q}$ that refers to OD movement k, whereas we write $\mathbf{c}_{[k]}$ with a square bracket around the index just to highlight to the reader that this is referring to a *partition* of the vector $\mathbf{c}$ (i.e. so $\mathbf{c}_{[k]}$ is itself a *vector*, not an individual element, as indicated also by the fact that $\mathbf{c}$ is in non-italic, bold-face), namely the *vector* of elements of $\mathbf{c}$ that relate to OD movement k.

For each OD movement $k \in \{1, 2, \ldots, |K|\}$ let ($\mathbf{p}_{[k]}(\mathbf{c}_{[k]}; \theta)$) denote the vector function with elements $p_r(\mathbf{c}_{[k]}; \theta)(r \in \mathcal{R}_k)$, these elements denoting the proportion of travellers who would select route r on OD movement k when the perceived distribution of route utilities is parameterized in a random utility model by $\mathbf{c}_{[k]}$ and θ. For the multinomial logit model in particular, these elements are given by:

$$p_r(\mathbf{c}_{[k]}; \theta) = \frac{\exp(-\theta c_r)}{\sum_{s \in \mathcal{R}_k} \exp(-\theta c_s)}, (r \in \mathcal{R}_k; \quad k = 1, 2, \ldots, |K|). \tag{2}$$

We then define $\mathbf{p}(\mathbf{c}; \theta)$ to be the vector function:

$$\mathbf{p}(\mathbf{c}; \theta) = (\mathbf{p}_{[1]}(\mathbf{c}_{[1]}; \theta), \mathbf{p}_{[2]}(\mathbf{c}_{[2]}; \theta), \ldots, \mathbf{p}_{[|K|]}(\mathbf{c}_{[|K|]}; \theta))^{\#}. \tag{3}$$

Associated with the choice probability function, we shall also define the functions $s_k(\mathbf{c}_{[k]}; \theta)$ for each OD movement k:

$$s_k(\mathbf{c}_{[k]}; \theta) = -\frac{1}{\theta} \ln\left(\sum_{r \in \mathcal{R}_k} \Gamma_{kr}(\exp(-\theta c_r))\right), \quad (k = 1, 2, \ldots, |K|). \tag{4}$$

[1] In the more general case, assuming the influences to be link-additive, the function $t_l(\cdot)$ for each link l contains the combined effect of all factors other than the tolls, scaled to be in equivalent travel time units, so can be termed the 'generalised travel time exclusive of tolls'. The stimulus to route choice and demand is still then, as it is here, the sum of the tolls (in equivalent time units) and these other factors, i.e. the generalised travel time.

There are several names in the literature that may be used to describe the quantity returned by the function $s_k(\cdot)$. Ortuzar (2001), following Williams (1977), referred to this as a *composite cost* (in the case when the c_k are measuring generalised cost), a term commonly used by those working with nested logit demand models. From a network assignment perspective, Sheffi (1985) developed such an expression from what he termed the *satisfaction function*. Sheffi considered two kinds of satisfaction; in the first, satisfaction (as one might expect by the word) is measuring an increasingly desirable characteristic. It is measured in the units of systematic utility, being the expected maximum utility, and for the logit model above is given by $\ln\left(\sum_{r=1}^{\mathcal{R}_k}\Gamma_{kr}(\exp(-\theta c_r))\right)$. In the second kind of "satisfaction", and the one we shall adopt, Sheffi effectively scaled the expected maximum utility to be in the same units as the c_k by dividing by $-\theta$, yielding expression (4). In this case "satisfaction" is a potentially misleading word, since (4) is no longer measuring a desirable attribute, However, these ambiguous names for the entity in (4) can be found throughout the networks literature. For example, Maher et al. (2005) refer to this expression as both a "satisfaction" and a composite travel cost; Ying and Yang (2005) refer to this entity as a *disutility*, which better captures its undesirability, but then is misleading in terms of units, since it is measuring something on the scale of c_k not of utility. Based on this discussion, we feel that the term 'composite cost' best captures both the nature and units of this measure, except that in our case c_k has been defined as a measure of generalised travel time rather than generalised travel cost; therefore, we shall use the term *composite generalised travel time* to describe the output of the functions in (4).

These composite generalised travel time functions can be neatly brought together in vector notation as:

$$\mathbf{s}(\mathbf{c};\theta) = -\theta^{-1}\ln(\boldsymbol{\Gamma}\ \exp(-\theta\mathbf{c})) \tag{5}$$

where $\ln(\mathbf{x})$ and $\exp(\mathbf{x})$ for vector argument $\mathbf{x}$ denote element-by-element application of the respective function.

Turning attention to the demand for travel, for each origin–destination movement k, a separable, bounded (uniformly above), differentiable and monotonically decreasing demand function $d_k(\cdot)$ is assumed, such that $d_k(s_k)$ is the demand level for OD movement k when that OD movement's composite generalised travel time is s_k, for $k \in K$. Under these assumptions, the demand function for each k has an inverse, and we refer to that inverse function as $g_k(\cdot)$, such that $g_k(q_k)$ is the OD composite generalised travel time that would give rise to a demand of q_k. Collectively we refer to the demand functions for all movements by the vector function $\mathbf{d}(\mathbf{s}) = (d_1(s_1), d_2(s_2), \ldots, d_{|K|}(s_{|K|}))^{\#}$.

It is useful to highlight some of the assumptions we have made, and particularly the potential for them to be generalised or extended:

- As we adopt an SUE approach we can move unambiguously between link-based and route-based formulations of the problem. In the analysis below we shall generally suppose that $\mathcal{R}_k$ denotes the *universal* set of acyclic routes available for OD movement k, but the analysis would hold equally in the case where a pre-defined subset of the universal set is defined (see Connors et al., 2007, for such an example).
- In our SUE approach below we shall restrict attention to the case of a multinomial logit choice model, which has the advantage that the satisfaction function is available in closed form. However, a similar approach could in principle be adopted for other kinds of SUE model, for example extending the work for a single city authority with probit SUE route choice reported in Connors et al. (2007).
- In the multinomial logit specification, the dispersion parameter θ is assumed to be common to all OD movements, but the approach below is trivially modified to permit a different θ for each OD movement. In terms of computational demands, such a modification would be analogous to moving to a multi-class SUE problem with different θ (Van Vuren and Watling, 1991; Lo and Szeto, 2002). While each SUE problem would, as a result, be longer to solve, the structure of the overall problems would be the same, and so we have no reason to expect the number of SUE evaluations to increase, and still we can use the same concepts, such as use of sensitivity analysis, and the single-level formulation of the resulting EPEC presented below.
- Our formulation assumes that each authority only has a single decision variable at their disposal, namely the common toll level to charge on all links pre-identified as tollable. We have in mind a situation where each authority has a cordon they have identified and will charge each time the cordon is crossed as this is a common way of implementing pricing in the few cities that have introduced it (namely Milan, Oslo, Stockholm, and Singapore). Limiting our consideration to the case of cordons with a common toll is obviously important for the way we present the results, but there is no methodological reason for this, we could just as easily have tolls differing by link.
- We restrict attention to a single user class specification. Under such an assumption, it is no further restriction to assume that the tolls are expressed in equivalent time units, since the value-of-time is common to all users.
- The major complication in our extending our two-authority model to three or more authorities would be the issue of how to handle any tax-exporting behaviour in the competing objective functions, given the various combinations that can arise of a resident of authority i paying a toll to travel on the network of authority j. Restricting attention to two authorities simplifies the conceptual side of understanding the competing behaviour, but the approach would be extensible in the future to cases of three or more authorities.

2.2. The global operator problem under SUE

Returning to the particular assumptions adopted in the present paper, then given any toll vector $\boldsymbol{\tau} \geqslant \mathbf{0}$, it is supposed that OD travel demand and routing patterns are determined by an elastic demand, Stochastic User Equilibrium (SUE) (e.g. Maher et al., 2005). Before moving on to the case of primary interest in this paper, that of competing authorities, we shall consider firstly the simpler case of a single authority, since a detailed understanding of that case is helpful for the extensions we make to multi-authority problems. The solutions to this problem also prove to be a useful reference point for our subsequent numerical analyses, reported in Section 4.

Therefore, let us imagine, contrary to the case described so far, that the toll vector $\boldsymbol{\tau}$ is to be decided by a *single* regulatory authority, aiming to maximise social welfare, which for an SUE-based model is given by the objective function:

$$\Psi_{\text{GLOBAL}}(\mathbf{v},\mathbf{q},\boldsymbol{\tau}) = \sum_{k=1}^{|K|}\int_0^{q_k} g_k(x)dx - (-\theta^{-1}\mathbf{q}^{\#}\ln\{\boldsymbol{\Gamma}\exp(-\theta\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v})+\boldsymbol{\Lambda}\boldsymbol{\tau}))\} - (\boldsymbol{\Lambda}\boldsymbol{\tau})^{\#}\mathbf{v}). \tag{6}$$

This objective function has been used in a variety of studies such as Huang et al. (2000) and Ying and Yang (2005). The first term on the right-hand side objective above is the Marshallian measure capturing the user benefit of the trips made. In order to obtain social welfare, we must then subtract the 'user cost' capturing the travel time impacts, namely the total 'price' stimulus (generalised travel time = travel time + tolls in equivalent time units) net of any tolls paid. Thus, the bracketed term is the total composite generalised travel time, based on (4) and (5), net of the total tolls paid (or total revenue). As a measure of welfare, it can be seen that (6) is equivalent to the sum of consumer surplus and revenue. A detailed discussion of the derivation of this measure can be found in the citations above, as is for the similar function derived in the case of UE (see, for example, Verhoef et al., 2010). Then we may define an SUE-based Global Regulatory Problem, to parallel the UE-based Global Regulatory Problem of Lawphongpanich and Hearn (2004). This will have the form of a Mathematical Program with Equilibrium Constraints (MPEC) given by:

$$\begin{aligned} \text{Maximise } & \Psi_{\text{GLOBAL}}(\mathbf{v},\mathbf{q},\boldsymbol{\tau}) \\ \text{subject to } & (\mathbf{v},\mathbf{q}) \in \boldsymbol{D} \\ & \boldsymbol{\tau} \geqslant 0 \\ \mathbf{v} = {} & \boldsymbol{\Delta}\mathbf{q}\ \mathbf{p}(\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v})+\boldsymbol{\Lambda}\boldsymbol{\tau});\theta) \\ \mathbf{q} = {} & \mathbf{d}(-\theta^{-1}\ln\{\boldsymbol{\Gamma}\ \exp(-\theta\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v})+\boldsymbol{\Lambda}\boldsymbol{\tau})\}). \end{aligned} \tag{7}$$

This problem can be solved as a standard non-linear programming problem by embedding the SUE conditions as constraints.

There are three key properties to remark on the constraints to this problem:

1. Under the stated assumptions above on the travel time functions, it follows that the vector of generalised link travel time functions, given by $\mathbf{t}(\mathbf{v}) + \boldsymbol{\Lambda}\boldsymbol{\tau}$, is continuous, separable and monotonically increasing in $\mathbf{v}$. Coupled with the assumptions above on the demand function $\mathbf{d}(\cdot)$, this implies that there is a unique elastic SUE solution $(\mathbf{v},\mathbf{q}) \in D$ for any given $\boldsymbol{\tau} \geqslant \mathbf{0}$ (Cantarella, 1997). Thus, the fixed point constraints to (7) define a *single* $(\mathbf{v},\mathbf{q})$ given $\boldsymbol{\tau}$, and so we may equivalently maximise only with respect to $\boldsymbol{\tau} \geqslant \mathbf{0}$ while leaving $(\mathbf{v},\mathbf{q})$ as free variables (with values to be determined by the constraints for any given $\boldsymbol{\tau}$).
2. Since by construction the functions $\mathbf{p}(\cdot)$ and $\mathbf{d}(\cdot)$ map to the feasible region D, then the only solution to the fixed point conditions over $(\mathbf{v},\mathbf{q}) \in \mathbb{R}^{|A|+|K|}$ (for given $\boldsymbol{\tau}$) is the unique SUE allocation in D, and therefore we can relax the condition $(\mathbf{v},\mathbf{q}) \in D$ as an explicit constraint.
3. The assumptions given ensure that the elastic SUE solution, viewed as an implicit function of $\boldsymbol{\tau}$, is smooth (differentiable); this result can be obtained, even for the case of non-separable travel time functions, by a trivial adaptation of the proof in Section 3 of Connors et al. (2007). This motivates the use of the differentiable implicit functions given by:

$$(\mathbf{v}^*(\boldsymbol{\tau}),\mathbf{q}^*(\boldsymbol{\tau})) = \text{SOL}\left\{\left\{(\mathbf{v},\mathbf{q}) \in \mathbb{R}^{|A|+|K|} \,\middle|\, \begin{aligned} \mathbf{v} &= \boldsymbol{\Delta}\mathbf{q}\ \mathbf{p}(\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v})+\boldsymbol{\Lambda}\boldsymbol{\tau});\theta) \\ \mathbf{q} &= \mathbf{d}(-\theta^{-1}\ln\{\boldsymbol{\Gamma}\exp(-\theta\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v})+\boldsymbol{\Lambda}\boldsymbol{\tau})\}) \end{aligned}\right\}\right\} \tag{8}$$

Since, by inspection of (7) the upper level objective Ψ_{GLOBAL} is a differentiable function of its arguments, then by function composition combined with the three properties above, it follows that by the Implicit Function Theorem we may obtain a smooth, single-level problem equivalent to (7), by using (8), as:

$$\begin{aligned} \text{Maximise } & h(\boldsymbol{\tau}) = \Psi_{\text{GLOBAL}}(\mathbf{v}^*(\boldsymbol{\tau}),\mathbf{q}^*(\boldsymbol{\tau}),\boldsymbol{\tau}) \\ & \text{subject to } \boldsymbol{\tau} \geqslant 0 \end{aligned} \tag{9}$$

Although the objective function h is non-convex, the possibility to evaluate gradients of h with respect to $\boldsymbol{\tau}$ (using sensitivity analysis of the fixed point problem (8)) provides the possibility to search for local stationary points using standard gradient-based solvers (Connors et al., 2007). This is something distinctive about adopting an SUE as opposed to UE constraint in the MPEC.

An SUE-based formulation of the equilibrium constraint, while at first appearing to add complexity, in fact leads to some considerable simplification of the resulting MPEC, yet the discussion above is restricted to the case of a single operator maximising the total welfare. It is therefore natural to investigate the properties of an SUE-based formulation in the context of interest to the present paper, namely that of multiple, competing authorities, which we begin to formulate in the next subsection.

2.3. Formulation of authority-specific objective functions

In the case of competing authorities, we assume that each authority has jurisdiction over setting tolls on its own set of links, but that its responsibility is only to trips that originate in its area. Thus, as a first step, we partition our OD demand and link flow variables, such that:

$$\mathbf{v} = \mathbf{v}^{(A)} + \mathbf{v}^{(B)}, \quad \mathbf{q} = \mathbf{q}^{(A)} + \mathbf{q}^{(B)} \tag{10}$$

where $\mathbf{v}^{(A)}$ and $\mathbf{v}^{(B)}$ are column vectors of dimension $|\mathcal{L}|$ respectively denoting the link flows from origins in authority A and B, and where $\mathbf{q}^{(A)}$ and $\mathbf{q}^{(B)}$ are column vectors of dimension $|K|$ respectively denoting the OD demands from origins in authority A and B.

Let us first consider Authority A. Authority A is assumed to be aiming to maximise the social welfare of its own residents by adjusting the toll level of links over which it has control, *anticipating* the impact of the toll on travellers' route and demand decisions, but *reacting* to the toll level levied by Authority B. That is to say, Authority A does not anticipate the effect that their own choice of toll will have on Authority B's response, but they simply react to the toll set by Authority B. Let us assume for the moment that Authority B has already decided its toll level $\tau_B \geqslant 0$, and that this is known to Authority A. Authority A then aims to maximise the social welfare of its own residents, given what Authority B has decided, through a counterpart objective to (6).

In fact, even given the details of this specification, there is still some ambiguity as to how we might define social welfare, depending on our assumptions on the existence of any revenue-sharing agreements between the city authorities, which correspond (using the terminology of the literature discussed in Section 1) to different assumptions on *tax exporting behaviour*. In the simplest case, there is no tax exporting behaviour and all toll revenues (across both authorities) are returned to the authority where the users who paid the tolls reside. That is to say, while the authorities decide their toll level individually, it is as if they put the revenues in a central fund, and share them according to the origin of the users who paid the toll. This concept of a central fund therefore mimics an element of the global regulator problem (6), and so might be expected to result in solutions which are closer to the global regulator problem (than other alternatives we explore) in terms of tolls and aggregate welfare. However, the case of individually-decided tolls with no tax export explicitly allows for full re-distribution of revenues to those who paid the toll; in contrast, the global regulator, even if obtaining the same tolls and aggregate welfare, would be free to distribute the revenues between the authorities via lump sum transactions in other proportions, so as to balance any welfare changes on a *per capita* basis.

In the case of *no tax exporting behaviour*, the social welfare of Authority A is, under the assumptions above, given by:

$$\Psi_A^0(\mathbf{v}^{(A)}, \mathbf{v}^{(B)}, \mathbf{q}^{(A)}, \mathbf{q}^{(B)}, \tau_A|\tau_B) = \sum_{k=1}^{|K|} \int_0^{q_k^{(A)}} g_k(x)dx - \left(-\theta^{-1}\left(\mathbf{q}^{(A)}\right)^{\#} \ln\left\{\boldsymbol{\Gamma}\exp\left(-\theta\boldsymbol{\Delta}^{\#}\left(\mathbf{t}(\mathbf{v}^{(A)} + \mathbf{v}^{(B)}) + \boldsymbol{\Lambda\tau}\right)\right)\right\} - (\boldsymbol{\Lambda\tau})^{\#}\mathbf{v}^{(A)}\right).$$

The first term in this objective function is the Marshallian measure of user benefit, but in this case only with respect to trips made from origins located within Authority A's jurisdiction; note that by definition $q_k^{(A)} = 0$ if OD pair k corresponds to an origin in Authority B, and so OD pairs for Authority B will make no contribution to the first term's summation over k. As for the global regulator problem earlier, we then subtract the 'user cost' net of tolls, the difference being again that here we restrict attention only to trips made by users with origins in Authority A. Thus, in the final term, we must subtract the element of the 'price' stimulus due to tolls imposed by either authority, but experienced only by Authority A travellers. Recall that the $\mathbf{q}^{(A)}$ vector in this second term has number of elements equal to number of OD pairs, but with elements equal to zero for any OD pair originating in authority B; therefore the second term overall does not contain any contribution of costs/prices experienced by Authority B residents. It is worth remarking a distinction with problem (6) is that we cannot now also give the final term an interpretation of revenue; in the notation given, the revenue to authority A from all flows would be $\left(\boldsymbol{\Lambda}\begin{pmatrix}\tau_A\\0\end{pmatrix}\right)^{\#}(\mathbf{v}^{(A)} + \mathbf{v}^{(B)})$ and the revenue to authority A only from its own residents would be $\left(\boldsymbol{\Lambda}\begin{pmatrix}\tau_A\\0\end{pmatrix}\right)^{\#}\mathbf{v}^{(A)}$.

At the other extreme to no tax export, we may assume *full tax exporting behaviour*, which means that revenues collected by Authority A (Authority B) remain within Authority A (Authority B), regardless of the originating authority of those who paid the tolls. As a result, Authority A travellers paying Authority B's toll will increase the social welfare of those in Authority B—and therefore *decrease* the social welfare of those in Authority A—by the toll revenue paid. Conversely, Authority B travellers paying Authority A's toll increase the social welfare of those in Authority A by the toll revenue they pay. This leads to two additional terms in the social welfare function of authority A, representing these two transfer payments:

$$\Psi_A^1(\mathbf{v}^{(A)},\mathbf{v}^{(B)},\mathbf{q}^{(A)},\mathbf{q}^{(B)},\tau_A|\tau_B)=\sum_{k=1}^{|K|}\int_0^{q_k^{(A)}}g_k(x)dx-(-\theta^{-1}(\mathbf{q}^{(A)})^{\#}\ln\left\{\boldsymbol{\Gamma}\exp\left(-\theta\boldsymbol{\Delta}^{\#}\left(\mathbf{t}(\mathbf{v}^{(A)}+\mathbf{v}^{(B)})+\boldsymbol{\Lambda\tau}\right)\right)\right\}-(\boldsymbol{\Lambda\tau})^{\#}\mathbf{v}^{(A)})$$
$$-\left(\boldsymbol{\Lambda}\begin{pmatrix}0\\\tau_B\end{pmatrix}\right)^{\#}\mathbf{v}^{(A)}+\left(\boldsymbol{\Lambda}\begin{pmatrix}\tau_A\\0\end{pmatrix}\right)^{\#}\mathbf{v}^{(B)}.$$

We may unify these two extreme cases of tax exporting behaviour by introducing a scalar tax exporting parameter α ($0\leqslant\alpha\leqslant1$), which we assume to be common value for both authorities (though this is not a necessary restriction), which leads after some slight simplification to our analogous objective to (6):

$$\begin{aligned}\Psi_A(\mathbf{v}^{(A)},\mathbf{v}^{(B)},\mathbf{q}^{(A)},\mathbf{q}^{(B)},\tau_A|\tau_B)&=\sum_{k=1}^{|K|}\int_0^{q_k^{(A)}}g_k(x)dx-(-\theta^{-1}(\mathbf{q}^{(A)})^{\#}\ln\{\boldsymbol{\Gamma}\exp(-\theta\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v}^{(A)}+\mathbf{v}^{(B)})+\boldsymbol{\Lambda\tau}))\}-(\boldsymbol{\Lambda\tau})^{\#}\mathbf{v}^{(A)})\\&\quad-\alpha\left(\boldsymbol{\Lambda}\begin{pmatrix}0\\\tau_B\end{pmatrix}\right)^{\#}\mathbf{v}^{(A)}+\alpha\left(\boldsymbol{\Lambda}\begin{pmatrix}\tau_A\\0\end{pmatrix}\right)^{\#}\mathbf{v}^{(B)}.\end{aligned}\tag{11}$$

With $\alpha = 1$ in (11) we have full tax exporting behaviour, whereas with $\alpha = 0$ there is no tax exporting behaviour. Cases with $0 < \alpha < 1$ allow some partial degree of tax exporting behaviour.

Authority A is then supposed to solve an MPEC (12) that is a variant of (7), given the value of $\tau_B\geqslant0$:

$$\begin{aligned}&\text{Maximise } \Psi_A(\mathbf{v}^{(A)},\mathbf{v}^{(B)},\mathbf{q}^{(A)},\mathbf{q}^{(B)},\tau_A|\tau_B)\\&\quad\text{subject to } (\mathbf{v},\mathbf{q})\in\boldsymbol{D}\\&\qquad\tau_A\geqslant0\\&\mathbf{v}=\mathbf{v}^{(A)}+\mathbf{v}^{(B)}\quad\mathbf{q}=\mathbf{q}^{(A)}+\mathbf{q}^{(B)}\\&\mathbf{v}=\boldsymbol{\Delta}\mathbf{q}\ \mathbf{p}(\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v})+\boldsymbol{\Lambda\tau});\theta)\\&\mathbf{q}=\mathbf{d}(-\theta^{-1}\ln\{\boldsymbol{\Gamma}\exp(-\theta\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v})+\boldsymbol{\Lambda\tau})\}).\end{aligned}\tag{12}$$

So far we have considered only authority A's perspective as a decision-maker, but authority B is also making decisions in the same way. Authority B thus aims to solve its own MPEC variant of (12) with a social welfare function of:

$$\begin{aligned}\Psi_B(\mathbf{v}^{(A)},\mathbf{v}^{(B)},\mathbf{q}^{(A)},\mathbf{q}^{(B)},\tau_B|\tau_A)&=\sum_{k=1}^{|K|}\int_0^{q_k^{(B)}}g_k(x)dx-(-\theta^{-1}(\mathbf{q}^{(B)})^{\#}\ln\left\{\boldsymbol{\Gamma}\exp(-\theta\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v}^{(A)}+\mathbf{v}^{(B)})+\boldsymbol{\Lambda\tau}))\right\}-(\boldsymbol{\Lambda\tau})^{\#}\mathbf{v}^{(B)})\\&\quad-\alpha\left(\boldsymbol{\Lambda}\begin{pmatrix}\tau_A\\0\end{pmatrix}\right)^{\#}\mathbf{v}^{(B)}+\alpha\left(\boldsymbol{\Lambda}\begin{pmatrix}0\\\tau_B\end{pmatrix}\right)^{\#}\mathbf{v}^{(A)}.\end{aligned}\tag{13}$$

2.4. Formulation of Nash Equilibria between competing authorities

Finally, in the present section, based on the building blocks of the preceding sections, we set out our overall model of competing authorities. As a first step, we seek a similar simplification to that obtained in the Global Regular case in the transition from formulation (7)–(9). It transpires that in our SUE-based formulation, this is possible for general networks without further assumptions required; this may be contrasted with the counterpart UE-based formulation, where additional special restrictions are needed to deal with partial tax-exporting behaviour, which requires knowledge of authority-specific flows (Koh et al., 2012). This is a particularly attractive feature of the alternative, SUE-based model we present here.

Let $\boldsymbol{\Omega}^{(A)}$ be the $|K|\times|K|$ diagonal matrix which has diagonal entries $\Omega_{kk}^{(A)}=1$ if OD movement k has an origin in authority A, and equal to zero otherwise. Clearly this matrix is definable *a priori* from the network structure. Suppose a counterpart matrix $\boldsymbol{\Omega}^{(B)}$ is also defined for authority B. Now suppose that $(\mathbf{v}^{*}(\boldsymbol{\tau}),\mathbf{q}^{*}(\boldsymbol{\tau}))$ is an elastic SUE solution as given by (8), for given $\boldsymbol{\tau}$. Then we may uniquely determine corresponding link flows and demand flows *disaggregated by authority* from:

$$\begin{aligned}&\mathbf{q}^{*(A)}(\boldsymbol{\tau})=\boldsymbol{\Omega}^{(A)}\mathbf{q}^{*}(\boldsymbol{\tau})\qquad\mathbf{v}^{*(A)}(\boldsymbol{\tau})=\boldsymbol{\Delta}\mathbf{q}^{*(A)}(\boldsymbol{\tau})\mathbf{p}(\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v}^{*}(\boldsymbol{\tau}))+\boldsymbol{\Lambda\tau});\theta)\\&\mathbf{q}^{*(B)}(\boldsymbol{\tau})=\boldsymbol{\Omega}^{(B)}\mathbf{q}^{*}(\boldsymbol{\tau})\qquad\mathbf{v}^{*(B)}(\boldsymbol{\tau})=\boldsymbol{\Delta}\mathbf{q}^{*(B)}(\boldsymbol{\tau})\mathbf{p}(\boldsymbol{\Delta}^{\#}(\mathbf{t}(\mathbf{v}^{*}(\boldsymbol{\tau}))+\boldsymbol{\Lambda\tau});\theta).\end{aligned}\tag{14}$$

Again, as in the Global Regulator case, the constraints to (12) define a unique allocation of link flows and demands (in this case disaggregated by authority). Clearly also since the disaggregated flows are differentiable functions of the total link flows and demands, then since $(\mathbf{v}^{*}(\boldsymbol{\tau}),\mathbf{q}^{*}(\boldsymbol{\tau}))$ given by (8) is differentiable in $\boldsymbol{\tau}$, then so is $\left(\mathbf{v}^{*(A)}(\boldsymbol{\tau}),\mathbf{v}^{*(B)}(\boldsymbol{\tau}),\mathbf{q}^{*(A)}(\boldsymbol{\tau}),\mathbf{q}^{*(B)}(\boldsymbol{\tau})\right)$ given by the combination of (8) and (14). Thus we may obtain a smooth, single-level problem equivalent to (12) based on a combination of the functions defined in (8), (11) and (14):

$$\begin{aligned} \text{Maximise } h_{\mathrm{A}}(\tau_{\mathrm{A}}|\tau_{\mathrm{B}}) = \Psi_{\mathrm{A}}(\mathbf{v}^{*(\mathrm{A})}(\boldsymbol{\tau}), \mathbf{v}^{*(\mathrm{B})}(\boldsymbol{\tau}), \mathbf{q}^{*(\mathrm{A})}(\boldsymbol{\tau}), \mathbf{q}^{*(\mathrm{B})}(\boldsymbol{\tau}), \tau_{\mathrm{A}}|\tau_{\mathrm{B}}) \\ \text{subject to } \tau_{\mathrm{A}} \geqslant 0 \end{aligned} \tag{15}$$

Based on a combination of the functions defined in (8), (13) and (14), the corresponding problem that authority B aims to solve may be expressed as:

$$\begin{aligned} \text{Maximise } h_{\mathrm{B}}(\tau_{\mathrm{B}}|\tau_{\mathrm{A}}) = \Psi_{\mathrm{B}}(\mathbf{v}^{*(\mathrm{A})}(\boldsymbol{\tau}), \mathbf{v}^{*(\mathrm{B})}(\boldsymbol{\tau}), \mathbf{q}^{*(\mathrm{A})}(\boldsymbol{\tau}), \mathbf{q}^{*(\mathrm{B})}(\boldsymbol{\tau}), \tau_{\mathrm{B}}|\tau_{\mathrm{A}}) \\ \text{subject to } \tau_{\mathrm{B}} \geqslant 0 \end{aligned} \tag{16}$$

The inter-play of the two authorities in each aiming to maximise its own welfare by setting a toll, conditional on the other authority's toll, while anticipating the impact on the travellers, leads us to an example of a so-called Equilibrium Problem with Equilibrium Constraints (EPEC) (Mordukhovich, 2005). Based on problems (15) and (16) we may write this as:

Find $\boldsymbol{\tau} = (\tau_{\mathrm{A}}, \tau_{\mathrm{B}})^{\#} \geqslant 0$ such that simultaneously:

$$\begin{aligned} h_{\mathrm{A}}(\tau_{\mathrm{A}}|\tau_{\mathrm{B}}) \geqslant h_{\mathrm{A}}(x|\tau_{\mathrm{B}}) \quad \forall x \geqslant 0 \\ h_{\mathrm{B}}(\tau_{\mathrm{B}}|\tau_{\mathrm{A}}) \geqslant h_{\mathrm{A}}(y|\tau_{\mathrm{A}}) \quad \forall y \geqslant 0 \end{aligned} \tag{17}$$

Problem (17) assumes that the authorities can only determine their own toll conditional on the other authority, but places no further restriction on the admissible tolls. That is to say, the conditions require that, as far as one authority is concerned, their toll gives (marginally, i.e. based only on optimising their own toll) a global optimum solution to their individual MPEC ((15) or (16)), conditional on the other authority's toll setting.

Taking the conditions for both authorities together, Eq. (17) defines a problem that we will henceforth simply refer to as a Nash Equilibrium (NE) (Nash, 1950, 1951), in reference to the 'game' implicitly played between the two authorities. However, we shall also be interested in Nash games that are variants of (17). Specifically, rather than each authority determining a *global* optimum toll conditional on the other authority's toll choice, we consider the possibility that each authority only determines a *local* optimum to their individual MPEC, requiring that the conditions (17) only hold within a local neighbourhood of the given toll vector. Following Son and Baldick (2004), we shall refer to such a solution as a Local Nash Equilibrium (LNE). Thus for an LNE, each authority only needs establish optimality within a neighbourhood of the given solution (see Son and Baldick, 2004 for such an example).

Since the LNE conditions are weaker, the solution set to the NE problem is contained within the solution set to the LNE. Such LNE solutions are local stationary points of problems (15) and (16); noting the differentiability of problems (15) and (16), such LNE points must simultaneously satisfy the first-order KKT conditions of the two problems:

$$\begin{aligned} \tau_{\mathrm{A}} h'_{\mathrm{A}}(\tau_{\mathrm{A}}|\tau_{\mathrm{B}}) = \tau_{\mathrm{B}} h'_{\mathrm{B}}(\tau_{\mathrm{B}}|\tau_{\mathrm{A}}) = 0 \\ \boldsymbol{\tau} = (\tau_{\mathrm{A}}, \tau_{\mathrm{B}})^{\#} \geqslant \mathbf{0} \\ h'_{\mathrm{A}}(\tau_{\mathrm{A}}|\tau_{\mathrm{B}}) \geqslant 0 \\ h'_{\mathrm{B}}(\tau_{\mathrm{B}}|\tau_{\mathrm{A}}) \geqslant 0 \end{aligned} \tag{18}$$

where $h'_{\mathrm{A}}(.|\tau_{\mathrm{B}})$ and $h'_{\mathrm{B}}(.|\tau_{\mathrm{A}})$ respectively denote the (partial) derivative functions of $h_{\mathrm{A}}(\cdot|\tau_{\mathrm{B}})$ and $h_{\mathrm{B}}(\cdot|\tau_{\mathrm{A}})$.

Expression (18) is in the form of a *complementarity problem* (Karamardian, 1972), which when stated in the generic form of such problems may be written as:

$$\text{Find } \boldsymbol{\tau} \in \mathbb{R}^n \text{ such that}: \quad \begin{aligned} \boldsymbol{\tau}^{\#}\mathbf{h}'(\boldsymbol{\tau}) = \mathbf{0} \\ \boldsymbol{\tau} \geqslant \mathbf{0} \\ \mathbf{h}'(\boldsymbol{\tau}) \geqslant \mathbf{0} \end{aligned} \tag{19}$$

where $\mathbf{h}' : \mathbb{R}^2 \to \mathbb{R}^2$ is given in (20):

$$\mathbf{h}'(\boldsymbol{\tau}) = \begin{pmatrix} h'_{\mathrm{A}}(\tau_{\mathrm{A}}|\tau_{\mathrm{B}}) \\ h'_{\mathrm{B}}(\tau_{\mathrm{B}}|\tau_{\mathrm{A}}) \end{pmatrix} \tag{20}$$

In the remainder of the paper, we will explore the determination of LNE solutions (i.e. the solution set to (19)/(20)), both as solutions of interest in their own right, and as a means for determining NE solutions (i.e. solutions to (17)), the latter being contained in the set of LNE solutions.

3. Solution algorithm for EPEC formulation

As discussed in Zhang (2010), there are two generic classes of methods for solving EPECs. The first is the class of Diagonalization Solution Methods (DSM). These include the Gauss–Jacobi and Gauss–Seidel methods that have been proposed in Harker (1984). The other category may be referred to as Simultaneous Solution Methods (SSM). While a DSM solves each MPEC individually until the entire system converges, an SSM method aims to solve the entire system of interrelated MPECs *simultaneously*. In this context, this paper proposes an algorithm based on solving a sequence of Linear Complementarity Problems (LCPs), which takes advantage of the problem formulation as given in (20). We thus tailor the approach originally proposed by Kolstad and Mathiesen (1991) to solve the EPEC in (17) *simultaneously*, i.e. our algorithm is an example of an SSM.

Kolstad and Mathiesen (1991) proposed to solve the complementarity problem presented in (18) by linearising the function $\mathbf{h}'(\boldsymbol{\tau})$ at some starting point and this results in a linear complementarity problem (LCP). This LCP can be solved by application of Lemke's algorithm and the function $\mathbf{h}'(\boldsymbol{\tau})$ is re-linearised at this solution and the process is repeated until convergence is achieved. Such an algorithm has been successfully applied to several kinds of problem in the transportation field (Friesz et al., 1983; Harker and Friesz, 1985; Ribeiroa and Simõesa, 2015).

Specifically, the linearisation of $\mathbf{h}'(\boldsymbol{\tau}^0)$ at some arbitrary starting vector of tolls $\boldsymbol{\tau}^{\mathbf{0}}$ using a first order Taylor expansion, results in expression (21), with $\nabla\mathbf{h}(\boldsymbol{\tau}^{\mathbf{0}})$ being the 2×2 Jacobian matrix of $\mathbf{h}'(\boldsymbol{\tau})$, evaluated at $\boldsymbol{\tau} = \boldsymbol{\tau}^0$:

$$L(\boldsymbol{\tau}|\boldsymbol{\tau}^{\mathbf{0}}) = \mathbf{h}'(\boldsymbol{\tau}^{\mathbf{0}}) + \nabla\mathbf{h}'(\boldsymbol{\tau}^{\mathbf{0}})(\boldsymbol{\tau} - \boldsymbol{\tau}^{\mathbf{0}}). \tag{21}$$

Expression (21) allows us to formulate a linearised counterpart of the complementarity problem in (19) obtaining thus an LCP (see Cottle et al., 2009) of the form defined in (22):

$$\text{Find } \boldsymbol{\tau} \in \mathbb{R}^2 \text{ such that}: \left.\begin{array}{r} L(\boldsymbol{\tau}|\boldsymbol{\tau}^{\mathbf{0}}) = \mathbf{z} + \mathbf{M}\boldsymbol{\tau} \geqslant \mathbf{0} \\ \boldsymbol{\tau} \geqslant \mathbf{0} \\ \boldsymbol{\tau}^{\#} L(\boldsymbol{\tau}/\boldsymbol{\tau}^0) = \mathbf{0} \end{array}\right\} \mathrm{LCP}(\mathbf{M}, \mathbf{z}) \tag{22}$$

where $\mathbf{z} = \mathbf{h}'(\boldsymbol{\tau}^0) - \nabla\mathbf{h}'(\boldsymbol{\tau}^{\mathbf{0}})\boldsymbol{\tau}^{\mathbf{0}}$ and $\mathbf{M} = \nabla\mathbf{h}'(\boldsymbol{\tau}^{\mathbf{0}})$

Our Sequential LCP (SLCP) algorithm, adapted from Kolstad and Mathiesen (1991) then operates as follows:

Sequential Linear Complementarity Problem (SLCP) Algorithm

Step 0:	Choose some starting vector of tolls $\boldsymbol{\tau}^{\mathbf{0}}$. Select a small positive convergence tolerance, ε ($\varepsilon > 0$). Set counter $j = 0$ and go to Step 1
Step 1:	Solve the logit traffic assignment problem (8) with τ^j
Step 2:	Obtain derivatives: $\mathbf{h}'(\boldsymbol{\tau}^j)$ and $\nabla\mathbf{h}'(\boldsymbol{\tau}^j)$ and thus compute $\mathbf{z}$ and $\mathbf{M}$
Step 3:	Solve LCP($\mathbf{M}$, $\mathbf{z}$) in (22)
Step 4:	Check convergence: If max $\mathbf{h}'(\boldsymbol{\tau}^j) < \varepsilon$, terminate else set $j = j + 1$ and go to Step 1

Thus at each iteration from an initial starting vector of tolls, the proposed algorithm iterates between solving an elastic demand SUE traffic assignment problem, in which the required derivative information is also computed, and the solution of a LCP, until convergence is attained. In terms of numerical implementation, we have found that sufficient accuracy is obtained regardless of whether the required derivatives are computed by a fine-level finite differencing or by analytical approaches (for the latter see, for example, Bell and Iida, 1997; Ying and Miyagi, 2001; Ying and Yang, 2005), though the latter is more computationally efficient. In our implementation, the LCP sub-problem in Step 3 was solved using PATH, a solver for complementarity problems (Ferris and Munson, 2000).

We remark that in addition to SLCP, alternative approaches based on minimising the Nikaido–Isoda function (Nikaidô and Isoda, 1955) have also been proposed for single level Nash games, e.g. Krawczyk and Uryasev (2000). Furthermore, since a Nash game can be posed (as we have shown) as a variational inequality problem, projection methods (see Tinti, 2005 for a review) are also applicable in this context. While testing such alternative methods is an interesting area for future research, in the present paper we restrict attention to the application of the SLCP algorithm. We used this method to generate the set of LNE solutions by specifying a grid of starting conditions (for the toll levels), and solving the SLCP for each.

4. Numerical examples

In this section we present two numerical examples. The first builds on an example previously reported in the literature in this context, but assuming a UE rather than SUE response of travellers. The second is an example adapted from Zhang et al. (2011) and provides results for a larger grid network.

In each example, the link travel time functions are assumed to follow the separable BPR form as shown in Eq. (23):

$$t_l(\mathbf{v}) = t_l^0\left(1 + 0.15\left(\frac{v_l}{C_l}\right)\right)^4 (l \in \mathcal{L}). \tag{23}$$

Regarding the demand function, while there are many choices of function that satisfy the conditions stated earlier, which ensure uniqueness of elastic demand SUE solutions (Cantarella, 1997), we wished to make comparisons with earlier work in which a DUE assumption was made with *unbounded* demand functions (Koh et al., 2012), of the power law form shown in Eq. (24):

$$d_k(s_k) = q_k^0\left(\frac{s_k}{s_k^0}\right)^{\eta} \quad (k = 1, 2, \ldots, |K|) \tag{24}$$

where the pair (q_k^0, s_k^0) are the base demands and composite generalised time in a reference case (in our case, with no tolls), $\eta < 0$ is the elasticity of demand (set in our tests as $\eta = -0.58$ for all OD pairs). In order to achieve such a goal, while maintaining our original stated assumption, the function in (24) was slightly modified, such that for any value of satisfaction less than some very small number $\varepsilon > 0$, the demand is linear in the satisfaction and differentiable at the break-point:

$$d_k(s_k) = \begin{cases} q_k^0\left(\frac{\varepsilon}{s_k^0}\right)^{\eta}\left(1 + \frac{\eta}{\varepsilon}(s_k - \varepsilon)\right) & \text{for } 0 \leqslant s_k \leqslant \varepsilon \\ q_k^0\left(\frac{s_k}{s_k^0}\right)^{\eta} & \text{for } s_k > \varepsilon \end{cases} \quad (k \in K).$$

In addition, we should note that it is possible—with very low values of the logit dispersion parameter θ—for the satisfaction to become negative, and in such cases the specification of the elastic demand function makes little sense. Such cases are excluded from our analysis, as we believe them to beyond an allowable sensible range for θ. Thus, across all tests for the range of θ values we consider, the minimum satisfaction for any OD movement is zero, and from the demand function specification above, it can be seen that $d_k(0) = q_k^0\left(\frac{\varepsilon}{s_k^0}\right)^{\eta}(1-\eta)$, i.e. there is a finite positive upper bound. In practice, by restricting the admissible θ values as above, we were able to choose such a small value of ε that no equilibrium fell on the linear section of the demand function, so that effectively our solutions coincide with using (24), except that we retain the theoretical condition of a bounded demand function.

In the following, subsections 4.1–4.3 focus on a small example previously studied in the literature in a similar context, but assuming a DUE rather than SUE model of traveller behaviour. Throughout subsections 4.1–4.5, it is assumed that there is full tax exporting (i.e. $\alpha = 1$). In this example, we explore the existence, multiplicity and computation of LNE and NE solutions, while varying the assumptions about traveller sensitivity/mis-perception. We investigate the robustness of the proposed SLCP algorithm for determining LNE under varying values of the dispersion parameter θ. In doing so we consider whether there exist multiple LNE and if so, whether the number of LNE varies with θ. We compare the LNE and NE solutions to the SUE-based global regulator solution.

Subsections 4.4–4.6 focus on a second, larger example, adapted from an example found in the literature. Beyond applying the SLCP algorithm to identify NE solutions, we investigate the influence of varying the tax exporting parameter α, and discuss its policy implications.

4.1. Example 1: Exploration of Nash Equilibria as the dispersion parameter (θ) is varied

The network used in Example 1 is taken from Koh et al. (2012), where full details of all parameters and functions can be found (see Fig. 1).

A cordon is a toll pricing scheme where "vehicles pay a toll to cross a cordon in the inbound direction, in the outbound direction, or possibly in both directions" (de Palma and Lindsey, 2011, p. 1381). This can be simulated in this model by assuming that Authority A sets a uniform common toll on Links 1 and 6, thus forming a cordon around its CBD zone 2, while Authority B sets a uniform common toll on Links 7 and 12 forming a cordon around its CBD zone 4. This sets the context of the game considered in where each city uses its cordon toll to maximise the welfare of its own residents according to (14) and (16).

Our previous research reported in Koh et al. (2012), demonstrated the existence of 4 LNE for this network in the UE case. In this section, we show that multiple LNE continue to exist under SUE for large θ (small dispersion), but as θ decreases and we move further from UE then only one LNE solution exists, which also satisfies the NE conditions (and so is the unique NE solution).

In order to explore the existence of LNE solutions for a given θ value, we evaluated the welfare for each authority for a given toll pair with tolls ranging between 0 and 1000 s. Given that we have only two uniform tolls in our example then it is possible to visualise the welfare surfaces and to estimate the gradient of an authority's welfare with respect to that authority's own toll at each point, given the other authority's toll. Now, we note that the LNE necessary conditions (18) for any $\boldsymbol{\tau} > \mathbf{0}$ are equivalent to requiring $h'_A(\tau_A|\tau_B) = h'_B(\tau_B|\tau_A) = 0$. Therefore to explore the location of the LNE solutions (i.e. solutions which both satisfy the LNE necessary first order conditions (18) and are a local maximum), a sensible strategy for such interior points is to produce 'best response function' plots for each authority and to identify toll combinations where these functions intersect. The best response function for authority A, for example, given authority B's toll τ_B, is a τ_A where (a) the gradient $h'_A(\tau_A|\tau_B) = 0$, and (b) the point is a local maximum with respect to τ_A. Note that any remaining cases, where one or other toll is equal to zero, can be separately verified and excluded from consideration.

Examples of such best response function plots for the cases $\theta = 0.005$ and $\theta = 0.5$ are shown in the left and right panes of Fig. 2. Where the best response functions are near horizontal or vertical, this is indicative of a low level of interaction between the two problems solved by the authorities. This is an artefact of the network and demand scenario selected in this example; this will not always be the case, as will be seen in Example 2. With $\theta = 0.005$ there is only one point of intersection,

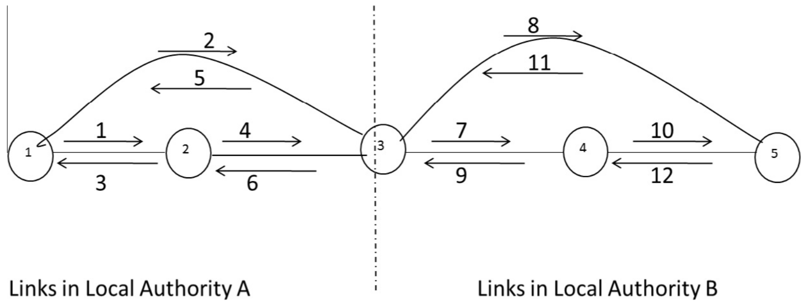

Fig. 1. Network used in Example 1, the numbers indicated are link numbers referred to in the text and direction of travel is indicated by the arrows. The dotted line that splits node 3 demarcates the limits of jurisdiction of each authority.

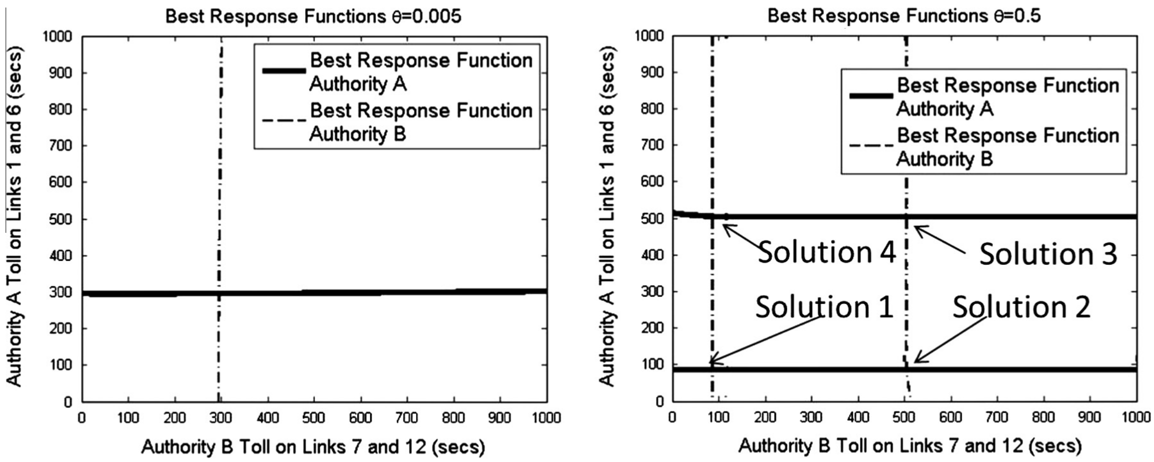

Fig. 2. Best response functions for $\theta = 0.005$ (left pane) and $\theta = 0.5$ (right pane).

which demonstrates there is only one LNE solution, and hence only one NE solution, at a toll of around 300 s set by each authority. With $\theta = 0.5$ there are four points of intersection, therefore four LNE solutions, and it can be verified (by simply checking the welfare levels at LNE solutions) that there is one NE solution, which is the one marked 'Solution 3' in Fig. 2.

4.2. Example 1: Solution by SLCP algorithm

The analysis in Section 4.2 required exhaustively computing a great many SUE solutions over a discrete grid of toll values. This would be computationally prohibitive for larger networks, and does not allow more refined estimates of the solutions to be obtained than the discretisation assumed. However, it gives a good reference point for testing the SLCP algorithm described in Section 3, which we propose as an efficient method for larger scale networks. Under the formulation given, this algorithm seeks solutions that satisfy necessary conditions for LNE, i.e. it aims to find *potential* LNE solutions, from which we can then test (by a small perturbation of tolls): does the solution satisfy the LNE conditions? By varying the starting conditions of the algorithm, we may then aim to find multiple potential LNE; of those that are LNE solutions, we may then test the satisfaction of the NE conditions (17) within this (typically small) set of LNE solutions. This raises several practical issues to test for the algorithm: Do the solutions it generates always satisfy the LNE conditions, or do we find other solutions emerging? By varying the initial conditions, do we get different LNE in cases where we know multiple LNE exist? How much effort (in terms of different initial conditions) might we have to expend to generate *all* the LNE?

As noted previously, the number of LNE solutions in this network depends on θ. Table 1 presents all LNE solutions identified in our experiments (from applying SLCP across a grid of starting conditions), for a range of θ values. As can be seen from the table, multiple LNE were seen to exist for $\theta \geqslant 0.02$, and in all other cases only a single LNE was found which is also the unique NE solution. Alongside the LNE solutions, we report in Table 1 the welfare *change* relative to an assumed base-case where neither authority tolls. While this base-case itself depends on θ, for any given θ the welfare of the base no-toll solution is clearly a constant, and so it is immaterial whether we assume the authorities are aiming to maximise absolute welfare or welfare change; the advantage in reporting the welfare change is that we feel it is a more sensible way to make comparisons over the θ values. For any of the LNE solutions it can be seen that the welfare change improves as θ increases; with lower sensitivity to travel costs (larger θ), more users choose the shorter town centre routes where they face a toll.

Table 1
Tolls (seconds) and Welfare Change by Authority (seconds) for LNE solutions, with NE solutions indicated by shaded cells.

θ	LNE Solution 1				LNE Solution 2				LNE Solution 3				LNE Solution 4			
	τ_A	τ_B	Welfare Change A	Welfare Change B	τ_A	τ_B	Welfare Change A	Welfare Change B	τ_A	τ_B	Welfare Change A	Welfare Change B	τ_A	τ_B	Welfare Change A	Welfare Change B
0.005	295.4	295.4	−41,845	−41,845	–	–	–	–	–	–	–	–	–	–	–	–
0.007	257.0	257.0	−33,459	−33,458	–	–	–	–	–	–	–	–	–	–	–	–
0.009	231.6	231.6	−28,190	−28,190	–	–	–	–	–	–	–	–	–	–	–	–
0.01	221.7	221.7	−26,157	−26,156	–	–	–	–	–	–	–	–	–	–	–	–
0.02	159.3	159.3	−12,691	−12,690	165.2	498.8	−103,358	−10,880	497.6	497.6	−98,661	−98,660	498.8	165.2	−10,880	−103,358
0.03	123.8	123.8	−3560	−3560	127.1	505.2	−105,498	5680	504.0	504.0	−93,868	−93,868	505.2	127.1	5680	−105,498
0.04	107.9	107.9	913	913	107.9	504.0	−105,168	14,265	504.0	504.0	−91,416	−91,416	504.0	107.9	14,265	−105,168
0.5	85.6	85.6	9646	9645	85.6	504.0	−102,092	23,822	504.0	504.0	−87,843	−87,843	504.0	85.6	23,822	−102,092
1	85.4	85.4	9372	9372	85.3	504.2	−102,475	23,960	504.0	504.0	−87,843	−87,843	504.2	85.3	23,960	−102,475

Table 2
Percentage of runs with alternative starting points out of 10,201 terminating at each LNE solution.

θ	LNE Solution 1 (%)	LNE Solution 2 (%)	LNE Solution 3 (%)	LNE Solution 4 (%)
0.02	27.36	24.70	23.08	24.86
0.05	0.54	6.62	86.31	6.54
0.5	0.08	2.33	94.61	2.98
1	0.02	0.69	98.30	0.99

For each of the LNE solutions identified in Table 1, we numerically tested whether the NE conditions were satisfied, by exploring the change in welfare for each authority when the other authority's toll was fixed at that LNE solution value (i.e. seeking a global optimum for that authority, conditional on the other's toll). In all cases a single NE was identified to exist, and this is indicated by the shading in Table 1. Clearly there is a significant change in the NE solution depending on θ. For $\theta \geqslant 0.02$ the NE solution leads to both authorities charging a high toll, approaching (as may be anticipated) the NE solution found under deterministic travel behaviour. For $\theta \leqslant 0.01$, on the other hand, the unique LNE/NE solution has both authorities charging a moderate toll. In all cases, the NE solutions give negative welfare changes; in spite of each authority aiming to maximise their own welfare, the competition between them leads in this case to overall disbenefits to both their own residents and to society as a whole.

In order to understand this apparent "paradox" of overall welfare disbenefits of welfare-maximising authorities, it is important to first remark that in no case was $\boldsymbol{\tau} = \mathbf{0}$ (neither authority tolls) found to be an LNE solution, although it is permitted. This implies that there must be some initial incentive for each authority to toll a non-zero amount. Exploring authority A's welfare as a function of τ_A, given a zero or very low authority B toll (say, $0 \leqslant \tau_B \leqslant 50$) (plot not shown here due to space limitations), positive welfare gains can certainly be obtained by authority A charging a very low toll (again, much lower than in any of the LNE solutions). This is borne out by the global regulator solutions, as discussed later (Section 4.4), which also have authorities charging some much lower toll than in any of the LNE solutions. Even at such lower tolls, there is an incentive (in the competing cities formulation) for cities to continue raising their toll levels. This may seem surprising, especially given the observation made earlier (Section 4.2) that in the network of Example 1 there is relative weak network interactions between the cities. In this case, the effect can be traced to be almost entirely attributable to the tax-exporting assumption; recall that with $\alpha = 1$ each authority's objective functions include two tax-export terms, which trade off the welfare loss from tolls paid by its citizens to the other authority with the welfare gain of tolls paid in its own network by the other authority's citizens. This means that any incentive for Authority B to increase its toll leads, *ceteris paribus*, to an indirect welfare loss to Authority A, for which Authority A then has the incentive to compensate through raising its own toll.

From Table 1, there are two kinds of LNE solutions, symmetric (equal toll) and asymmetric. In the asymmetric solutions, the highest welfare improvement (or, in some cases, least bad deterioration in welfare) is obtained for a city charging the higher toll, for all θ values. In all cases where multiple LNE exist, the *total* welfare change (given by the sum of Welfare Change in A and Welfare Change in B) for Solutions 2 and 4 is lower than that for Solution 1, while all three of these solutions are preferable to the NE (Solution 3) on total welfare grounds.

In terms of numerical implementation, a question of particular interest is: how does the SLCP algorithm perform, either in determining all LNE solutions or in locating the NE solution? In order to understand this, we generated a grid of initial conditions between 0 and 1000 s in steps of 10 for each authority thereby generating 10,201 distinct start points for the SLCP algorithm. Table 2 shows that when $\theta = 0.02$, the solutions found by SLCP were almost distributed evenly among each of the four LNE solutions. If we randomly chose the initial conditions we would in this case have almost equal probability of locating each of the LNE solutions. However for the case of $\theta = 1$, 98% of the solutions converged to Solution 3 (the NE solution), while Solution 1 was found in 0.02% of these cases. The performance of the algorithm in this respect is therefore rather sensitive to the value of θ.

To understand why this is the case, we examined welfare change for each authority as a function of that authority's toll, given that the other authority is tolling at an LNE solution. For illustration, and since of course it is of particular interest, we show the graphs in the neighbourhood of the NE solution; welfare changes for Authority A, assuming that Authority B sets the NE toll, are plotted in Figs. 3a–3d. These figures show how changes in the logit dispersion parameter θ change the shape of the welfare function. When $\theta = 0.5$ (Fig. 3a), the welfare function has multiple local optima,[2] but confirms that the NE solution for Authority A is indeed a global optimum, conditional on Authority B's toll. (The same graphs, but conditioned on Authority B tolling at one of the other LNE solutions, are not shown due to space limitations, but look very similar to these figures, and provide confirmation that the other mode of the function is indeed a local but not global optimum.) On the other hand, the global optimum of the welfare function corresponding to the NE solution is less obvious when $\theta = 0.02$ (Fig. 3b). It is clear

[2] In the toll range 0–100 s, non-captive users begin to move from the tolled routes to the bypass routes. For higher values of θ the route choice is more sensitive to costs and for tolls over 100 s practically no non-captive users take the tolled routes, and so the welfare curve for those particular OD pairs peaks for a toll less than 100 s. For the captive users, their welfare curves peak for tolls above 100 s as they have no alternate route choice. Finally we have the net change in revenue effect which tends to increase the optimal toll. The net effect of summing welfare over all OD pairs and accounting for revenue exportation results in multiple local optima. As θ is decreased, then the route choice is less sensitive and some non-captive users remain on tolled routes for higher tolls. This has the effect of flattening the welfare curve for these OD pairs and as θ is decreased further then the first local maximum is eventually smoothed out.

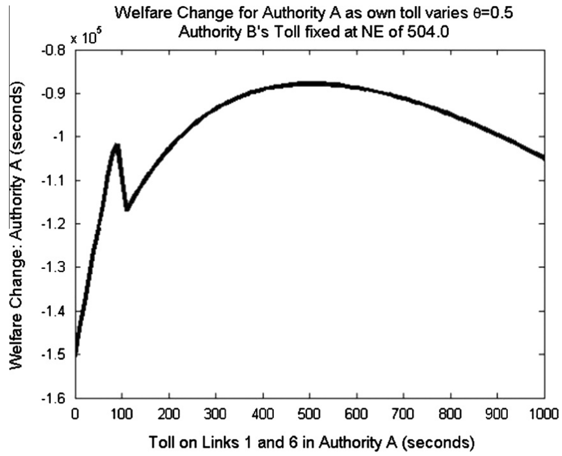

Fig. 3a. Welfare Change for Authority A as toll varies ($\theta = 0.5$) assuming Authority B sets the NE toll of 504.0.

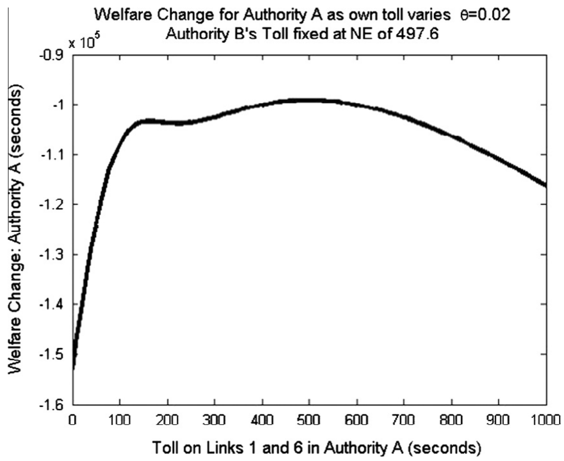

Fig. 3b. Welfare Change for Authority A as toll varies ($\theta = 0.02$) assuming Authority B sets the NE toll of 497.6.

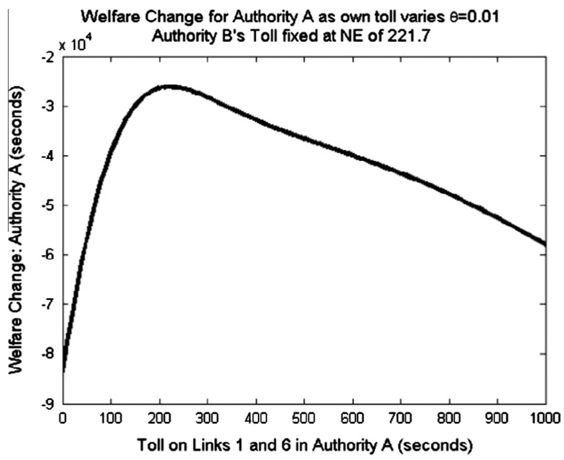

Fig. 3c. Welfare Change for Authority A as toll varies ($\theta = 0.01$) assuming Authority B sets the NE toll of 221.7.

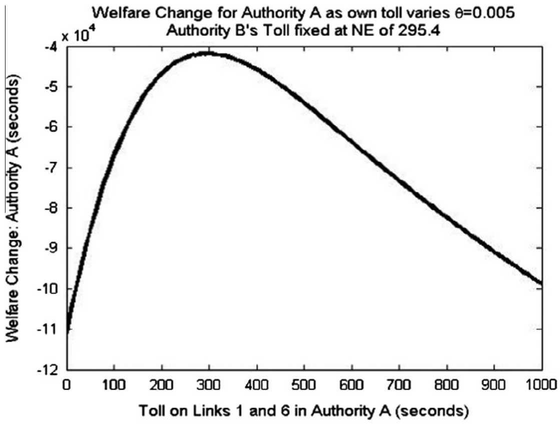

Fig. 3d. Welfare Change for Authority A as toll varies (θ = 0.005) assuming Authority B sets the NE toll of 295.4.

then that while SUE smooths the problem, it also makes the local optima of the welfare function less distinct and not easily distinguishable from the global optima. Thus the basin of attraction around the local optimum is larger with a lower θ, which implies that SLCP is not easily able to distinguish between these two solutions. Furthermore Figs. 3c and 3d show that as θ is further reduced, the welfare plots also begin to display a single local optimum which corresponds to the case where there is only one LNE and one NE.

4.3. Example 1: Comparison of LNE solutions with global regulator solution

As noted in Section 4.3, there exist LNE solutions across a range of θ values where both authorities set positive tolls that cause a reduction in welfare relative to an untolled base equilibrium. We refer to this as the Prisoner's Dilemma outcome because it reflects the paradox that, in equilibrium, both players are in fact worse off than if neither had tolled at all, the situation arising as a result of being incentivized to take an action by the other authority's toll decision. To avoid such a case, it may be necessary to regulate the tolls whereby a regulator aims to maximise global welfare. In this setting, a single regulator rather than the individual cities determines both cordon tolls with the aim of maximising total welfare. Due to the symmetry of our example, the tolls and welfare for each city will in the global regulator case be identical and each city obtains welfare equal to half the global welfare. To aid the discussion, we also include the maximum possible welfare obtainable from first-best pricing where all links may be tolled and include the ω index following Verhoef et al. (1996) which is the ratio of second-best welfare improvement over the first-best welfare improvement or relative efficiency measure.

Table 3 shows the solution of the Global Regulator Problem and First Best Problem. Comparing Table 1 with Table 3, it is evident that the total welfare (i.e. sum of Welfare A and Welfare B) at LNE Solution 1 moves further away from the global regulator solution as θ is decreased (users becoming less cost sensitive). As θ decreases, the tolls charged in the LNE Solution 1 become higher (and are always higher than the global regulator tolls) and total welfare decreases. In LNE Solutions 2 and 4, then for higher values of θ one authority tolls a low toll that is not so far from the global regulator toll, but this authority has overall disbenefits due to the high toll imposed by the other authority. In solution 3, which happens to be the only NE solution, both authorities charge a much higher toll than the global regulator solution, and both end up with significant disbenefits. For the cases where there is only one LNE/NE ($\theta \leqslant 0.01$), there is no solution for the Nash game which is "close" to the global regulator solution. This suggests that as dispersion increases (i.e. a smaller θ) then under competition there is no possibility of achieving a positive welfare outcome relative to the untolled base equilibrium for a given θ, and a regulator

Table 3
Tolls (seconds) and total welfare change (seconds) for global regulator problem and first best problem.

θ	Global regulator solution			First best solution	ω (=$\Delta W_1/\Delta W_2$)
	τ_A	τ_B	Welfare change (ΔW_1)	Welfare change (ΔW_2)	
0.005	61.2	61.2	9835	11,647	0.84
0.007	60.1	60.1	10,556	12,689	0.83
0.009	59.3	59.3	11,333	13,784	0.82
0.01	59.2	59.2	11,740	14,331	0.82
0.02	58.8	58.8	15,522	19,227	0.81
0.03	60.1	60.1	18,441	22,626	0.82
0.04	61.9	61.9	20,404	24,645	0.83
0.5	79.0	79.0	21,114	22,268	0.95
1	79.0	79.0	20,705	21,868	0.95

would be required to obtain this. Furthermore, with decreasing θ, as users are assumed to be less cost-sensitive in terms of their route choice and decision to travel, competition between cities would result in greater dis-benefits for both cities, with higher tolls imposed on the users.

For the higher θ value cases, LNE Solution 1 is seen to be closer to the regulator solution, and so there may be a greater chance that authorities in such a case would accept regulation and LNE Solution 1, if they were only adopting conservative/local optimising behaviour. As the gap between the global regulator solution and LNE Solution 1 increases with a decrease in θ, then there would be a greater need for regulation of the authorities. It is also of interest to note that as θ decreases even further, that the extent of welfare gain possible even under the globally regulated case decreases. This may suggest that if users are less cost-sensitive, then there may be less benefit to be gained from road pricing in general.

Regarding the first-best outcome (shown for comparison purposes in Table 3), as with the regulated case as θ decreases the absolute welfare improvement increases initially and then decreases with lower values of θ. This is due to the interaction between the terms in the welfare function, some of which increase with decreases in θ directly while others are dependent on the resulting demand, route flows and optimal toll which are dependent on θ indirectly. It should be noted that while the welfare gain from first best may increase initially, the gain is compared to a θ-dependent reference (no-toll equilibrium) which itself has a lower welfare outcome due to the users being less sensitive to cost. Our explanation is that as users become less cost-sensitive then route choices are more difficult to "correct" via tolling and the demand related terms in the welfare function become more dominant. Such findings may, however, be network-specific (backed up by the different pattern we shall discuss in our second example), so care should be taken in generalising these findings.

4.4. Example 2: Exploration of Local Nash Equilibria as dispersion parameter (θ) varied

Fig. 4 shows our second example, a grid network adapted from Zhang et al. (2011).

This network comprises 62 links and 20 nodes. The network parameters remain unchanged from that reported in Zhang et al. (2011) and each link adopts the BPR function of the form shown in Eq. (23).

Horizontal links have free flow travel time, t_l^0, of 2.5 min with capacity, C_l, of 700 vehicles per hour while vertical links have free flow travel time of 5 min with capacity of 1000 vehicles per hour.

On the demand side, a matrix is adopted that differs from that used in Zhang et al. (2011). This is because while the authors also studied toll revenue competition between two adjacent authorities, their setting assumed that Authority A was a residential suburb while Authority B was the employment zone. Furthermore, they did not consider cordon pricing as we do. In our example, the matrix comprising 56 OD pairs is shown in Table 4 and the demand function used was the power law form as shown in Eq. (24).

Note that only the diagonals of the OD matrix are zero (i.e. intrazonal trips were excluded from consideration), and that the base trip matrix is assumed to be entirely symmetric.

Authority A is interested in the welfare of residents in its jurisdiction (Zones 1, 7, 12 and 16) while Authority B is interested in the welfare of residents in its jurisdiction (Zones 5, 9, 14 and 20). The predefined cordon of Authority A is intended to toll traffic entering Zones 7 and 12 and a common toll is levied on the 6 links numbered: 9, 12, 17, 20, 36 and 55. On the other hand, Authority B's predefined cordon are links inbound into zones 9 and 14, namely links numbered 13, 16, 21, 24, 40 and 59. These cordon links are shown by the thicker lines in Fig. 4.

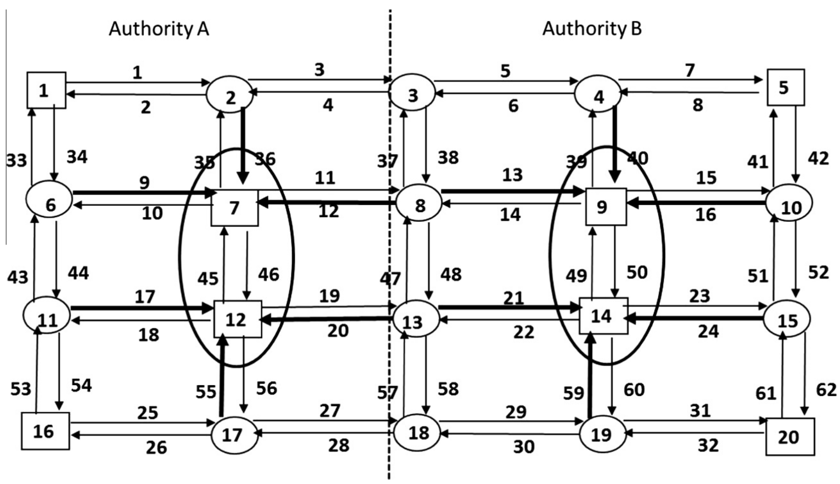

Fig. 4. Network with 20 nodes and 62 links from Zhang et al. (2011). Rectangular nodes represent origins and destinations while circular nodes only serve to connect links in the network. The dotted line demarcates the separation of jurisdiction between authorities. The thicker lines represent the closed cordon in each jurisdiction.

Table 4
Modified base trip matrix for example 2.

	1	**5**	**7**	**9**	**12**	**14**	**16**	**20**
Suburb A: Zone **1**	0	200	1000	500	1000	500	200	200
Suburb B: Zone **5**	200	0	500	1000	500	1000	200	200
CBD A: Zone **7**	200	200	0	200	100	200	200	200
CBD B: Zone **9**	200	200	200	0	200	100	200	200
CBD A: Zone **12**	200	200	100	200	0	200	200	200
CBD B: Zone **14**	200	200	200	100	200	0	200	200
Suburb A: Zone 1**6**	200	200	1000	500	1000	500	0	200
Suburb B: Zone **20**	200	200	500	1000	500	1000	200	0

Table 5
Tolls (min) and welfare change (min) for global regulator problem and first best problem.

θ	Global regulator solution					First best solution	ω (=$\Delta W_1/\Delta W_2$)
	τ_A	τ_B	Welfare change (ΔW_1)	Welfare Change A	Welfare Change B	Welfare change (ΔW_2)	
0.2	16.81	16.81	102,849	51,431	51,417	175,265	0.59
0.4	28.33	28.33	75,080	37,534	37,546	137,454	0.55
0.6	28.05	28.05	73,720	36,857	36,863	130,093	0.57
0.8	28.25	28.25	73,376	36,688	36,688	127,266	0.58
1	28.18	28.18	73,237	36,619	36,619	125,722	0.58
10	28.26	28.26	72,197	36,099	36,099	120,557	0.60

As the network and demands are entirely symmetric between the jurisdictions, the solution of the global regulator problem leads to equal toll levels for both jurisdictions; the results for this case are shown in Table 5.

First of all, it is noticeable that the optimal tolls are much higher in this example than in Example 1, despite using the same demand elasticity, similar free flow travel times and the same BPR function. To understand this, it is helpful to distinguish between captive trips and non-captive trips. Captive trips are those that have destinations within the tolled zones (i.e. CBDs of either authority) and are therefore captive to any cordon toll introduced. On the other hand, an example of a non-captive trip is the movement from Zone 16 to Zone 5; such trips are distinguished by the fact that they have a choice of avoiding the tolled links. Prior to the introduction of any toll, in the untolled base equilibrium for $\theta = 10$, around 120 (out of 200) trips from this particular OD pair utilise a route using one or more of the tollable links. With the introduction of the toll, these trips are suppressed as the toll is increased. This effect is also present in Example 1, however in the network of Example 2 the captive and non-captive trips share untolled entry links. This means that in Example 2, as the captive trips are suppressed, the non-captive trips benefit substantially from congestion relief. This effect does not occur in Example 1 where the introduction of a toll causes non-captive trips to re-route around the longer bypass, and as the toll increases more trips are routed around the bypass so that the cost of travel for non-captive trips is always increasing with tolls.

In the current example, non-captive trips experience a reduction in costs due to the structure of the network as the toll is increased; as we employ elastic demand, the number of these non-captive trips increases compared to the no toll case. These increases in non-captive trips, along with the decongestion benefits, then counter the suppression of the captive trips and so a greater welfare gain is possible with a much higher toll. Note that even with lower demands than the modified matrix that we have used and with different tollable links, Zhang et al. (2011) reported tolls of a similar order of magnitude (between 9 and 30 min) in their equivalent of our global regulator problem.

A second noticeable feature of Table 5 is that the toll level is much lower when $\theta = 0.2$ being 16 min versus a value around 28 min for higher θ values. In general, as θ decreases users are less sensitive to costs, and so for any given toll level, the suppression achieved is lower with a lower θ. At the same time, this means that the 'indirect generation' effect caused by the freeing up of untolled entry links will also be lower. In addition to the demand effect, there is a noticeable change in use of paths for certain OD pairs as users become less cost sensitive, and more dispersed across the available routes. This dispersion of traffic across paths combined with a lesser demand response impacts on different elements of the local welfare function, so that a relatively lower toll becomes optimal as θ is varied. On the contrary, with a higher θ, because users are then more sensitive to costs, more captive OD pairs are suppressed for a given toll level. This also implies that there is more freeing up of shared entry links. Essentially, the trade-off between benefits arising to different OD pairs within the welfare function varies as the demand and route choices vary with changes in cost sensitivity.

This different network structure also gives rise to a different pattern in first best welfare gains compared to Example 1. Here as sensitivity to cost is reduced, the potential benefit from first best tolling increases as θ decreases for all values tested. The benefit from the global regulator case also follows this pattern with the relative efficiency of the regulated case remaining between 0.55 and 0.6.

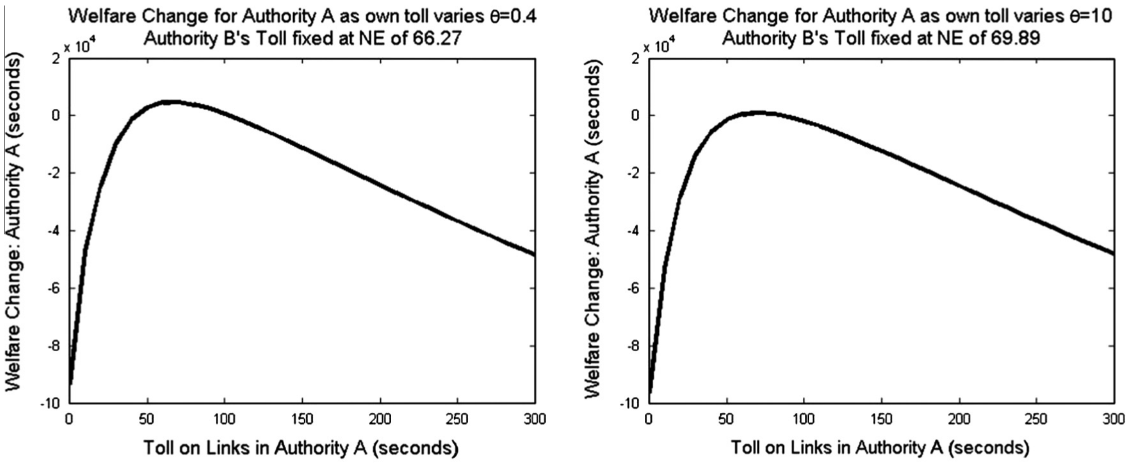

Fig. 5. (Left pane) Welfare Change for Authority A as toll varies ($\theta = 0.4$) assuming Authority B sets the NE toll of 66.27 (right pane) Welfare Change for Authority A as toll varies ($\theta = 10$) assuming Authority B sets the NE toll of 69.89.

Table 6
Total and jurisdictional welfare change with interjurisdictional competition.

θ	τ_A	τ_B	Welfare Change A	Welfare Change B	Total welfare change	ω
0.2	65.52	65.52	10,590	10,561	21,151	0.12
0.4	66.27	66.27	4678	4677	9355	0.07
0.6	67.31	67.31	3306	3306	6613	0.05
0.8	68.03	68.03	2641	2641	5281	0.04
1	68.47	68.47	2270	2270	4540	0.04
10	69.89	69.89	1125	1125	2250	0.02

4.5. Example 2: NE as dispersion parameter (θ) varies

In Example 2, the SLCP algorithm was again applied across a grid of starting conditions, in order to estimate potential LNE solutions, and this was repeated for a range of θ values. For each given θ tested, only a single LNE was found, which was indeed verified to satisfy the NE conditions. We believe that the existence of a single LNE can be explained by the common entry link for the captive and non-captive trips. As discussed above, when a toll is charged on the cordon, the non-captive trips actually benefit from decongestion and the welfare is increasing as the non-captive trips are tolled off the CBD routes. This is in contrast to Example 1, where non-captive trips see an increase in travel costs as they are tolled away from the CBD routes, which was seen to lead to an LNE at lower toll levels. Since with Example 2 we need not make the distinction in our discussion between LNE and NE solutions, we shall henceforth refer to the solutions obtained as NE solutions.

The left and right panes of Fig. 5 illustrate how the welfare for Authority A changes as its own toll varies for $\theta = 0.4$ and $\theta = 10$ respectively, if Authority B sets a toll at the NE solution, so that we see the incentive to charge alone around the NE solution. As the network is symmetric, a plot of B's welfare will display a similar pattern for a fixed value of Authority A's toll. We did not find any case of a 'double-peaked' welfare as we discussed with Example 1.

Table 6 gives the resulting NE toll levels and jurisdictional welfare change (measured relative to the untolled base equilibrium for a given θ) obtained by the SLCP algorithm with a convergence tolerance of $\boldsymbol{\tau}^{\#}\mathbf{h}'^{(\boldsymbol{\tau})} < 0.01$. Due to the symmetry of the example, we expect the tolls to be the same for the two authorities, and the final welfare functions; the differences seen are purely convergence error.

As an example to verify that the solutions are indeed NE solutions, Fig. 6 shows the best response functions for both authorities for the case of $\theta = 1$ (left panel) and $\theta = 10$ (right panel). In this figure, the best response function for Authority A to any toll level of Authority B is indicated by the continuous line. Similarly the best response function for Authority B to any toll level of Authority A is indicated by the broken lines. As the NE is the intersection of these best response functions, these figures numerically verify that the solution reported in Table 6 (indicated on these figures by a large asterisk) coincides with the intersection of these best response functions.

Comparing Tables 5 and 6, we see that the tolls under the NE solution are higher (under competition) than when a global regulator is in place. This is due to the desire of each jurisdiction to extract toll revenues from users from outside the jurisdiction, a result consistent with the previous network example. This shows that—at each dispersion parameter Θ value tested—the higher toll under inter-jurisdictional competition *vis-à-vis* the global regulator toll arises as a result of a combination of both the over-internalisation of externalities, as well as revenue maximising behaviour of each jurisdiction.

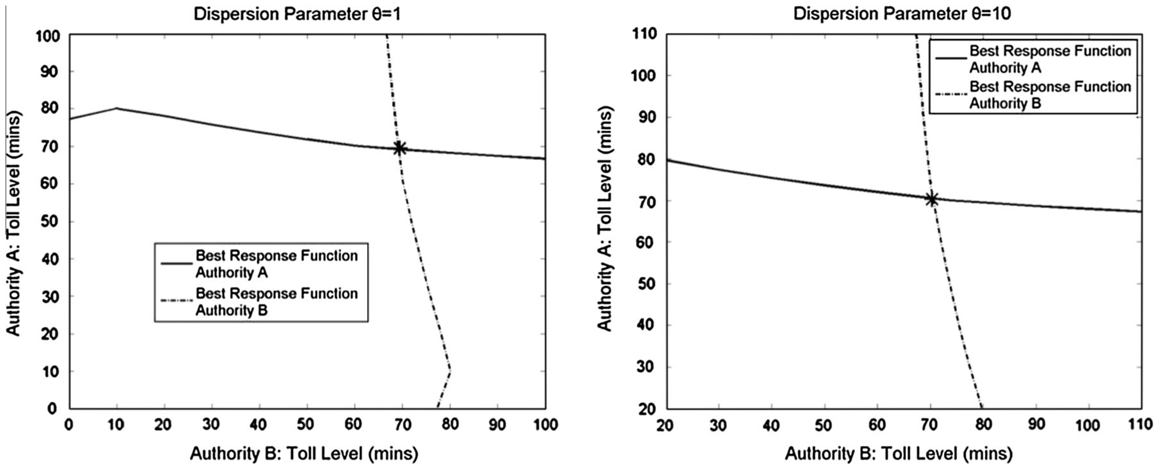

Fig. 6. (Left pane): Intersection of best response functions for $\theta = 1$ coincides with the cordon toll obtained by SLCP, of 68.47 min for both Authority A and B. (Right pane): Intersection of best response functions for $\theta = 10$ coincides with the cordon toll obtained by SLCP, of 69.89 min for both Authority A and B.

On the former (over-internalisation) issue, the double marginalisation problem arises because each authority does not take into account the toll set by the other authority when deciding its toll level in the game. This reflects research findings from the private operator literature (van den Berg, 2013). On the latter (revenue maximising) issue, this arises since extra-jurisdictional users who travel on a jurisdiction's road network contribute only to the toll revenue element of each player's objective function. Such a finding is consistent with those reported in Zhang et al. (2011). While the exact form of our (non-linear) demand function is likely to influence this result, we already know that such a phenomenon can occur even under linear demand functions (De Borger et al., 2007; Ubbels and Verhoef, 2008). In particular, Ubbels and Verhoef show that welfare under uncooperative tolling may be significantly lower than for first-best tolling in the case of linear demand functions. In their restricted setting, the first-best is the same as what we refer to as the global regulator problem (see Eq. (6)) and their model is a special case of ours where there is no route choice. Therefore, from this example we can say that it is possible that the welfare under non-cooperative tolling is significantly lower than under cooperative (global regulator) tolls, and that the tolls may also be relatively higher, even with a linear demand function.

We would also note that in our experiments, the welfare change under competition (relative to an untolled base equilibrium) is positive, and again we have a kind of Prisoner's Dilemma, i.e. players are worse off than under a co-operative solution. This is because each authority is still incentivised to begin the game (positive welfare compared with neither doing anything) but both end up worse off doing so, because the welfare gain is lower than that which would be possible under global regulation. In this case, competitive tolling is not worse than no tolling at all, but is worse than global regulation, only achieving between 0.02 and 0.12 of the relevant first best welfare gain. This does not compare favourably with the global regulator tolls which achieved relative efficiencies of between 0.55 and 0.6 of the welfare gain under first-best tolling.

4.6. Example 2: NE as tax exporting parameter (α) varies

In all cases presented in subsections 4.1–4.5 we restricted attention to results for the case $\alpha = 1$ (full tax exporting). We have tested both Examples for other values of α, and found qualitatively similar results in terms of algorithm convergence, existence and multiplicity of LNE/NE solutions, and impact of the dispersion parameter θ. On the other hand, some interesting policy impacts arise, and so in the present section, we analyse the effect of how changes in the tax exporting parameter α affects the NE toll levels, and offer some analysis of why this is the case.

The results we present relate to Example 2; again a single LNE/NE solution was found by the SLCP algorithm for all values of α and θ tested, regardless of the starting point of the algorithm, as so it is appropriate to refer to "the NE solution". Table 7 shows the NE tolls as the tax exporting parameter α is varied for three representative values of dispersion parameter θ. Fig. 7 shows the same NE tolls graphically.

To understand the results, let us first consider the two extreme points of $\alpha = 0$ (no tax exporting) and $\alpha = 1$ (full tax exporting). Firstly, the NE tolls when $\alpha = 0$ can be seen to be less than the tolls from the corresponding global regulator case; this can be seen by comparing each θ for $\alpha = 0$ with the corresponding case in Table 5. This is the case because in the NE solution, the welfare function (15) and (16) used by each authority does not take into account the full network-wide welfare effects, but only the congestion effect on its own residents. Secondly, in comparison with other values of α, under the full tax exporting regime, we may expect the highest tolls as the authority retains all revenues from the other authority's residents. Given these two points, it would then be expected that as α moves from 0 towards 1 that the NE tolls should increase. An increase in α increases the revenue retained by the authority, and in the extreme case of full tax exporting, the tolls are highest. The results in Table 6 confirm this expected phenomenon, and furthermore we can see that this continues to hold for each of the θ values tested.

Table 7
NE tolls and total welfare change (mins) as tax exporting parameter α changes.

α		Dispersion parameter		
		$\theta = 0.2$	$\theta = 0.4$	$\theta = 0.8$
0	τ_A	11.12	7.86	6.53
	τ_B	11.12	7.86	6.53
	Welfare change	93,512	58,276	45,621
0.2	τ_A	14.57	11.09	21.11
	τ_B	14.57	11.09	21.11
	Welfare change	101,574	65,525	69,109
0.4	τ_A	18.47	30.04	30.03
	τ_B	18.47	30.04	30.03
	Welfare change	102,315	74,704	73,075
0.6	τ_A	37.48	39.52	39.85
	τ_B	37.48	39.53	39.85
	Welfare change	79,649	66,325	64,691
0.8	τ_A	50.57	50.88	51.77
	τ_B	50.57	50.88	51.77
	Welfare change	56,524	46,033	43,668
1	τ_A	65.52	66.27	68.03
	τ_B	65.52	66.27	68.03
	Welfare change	21,151	9355	5281

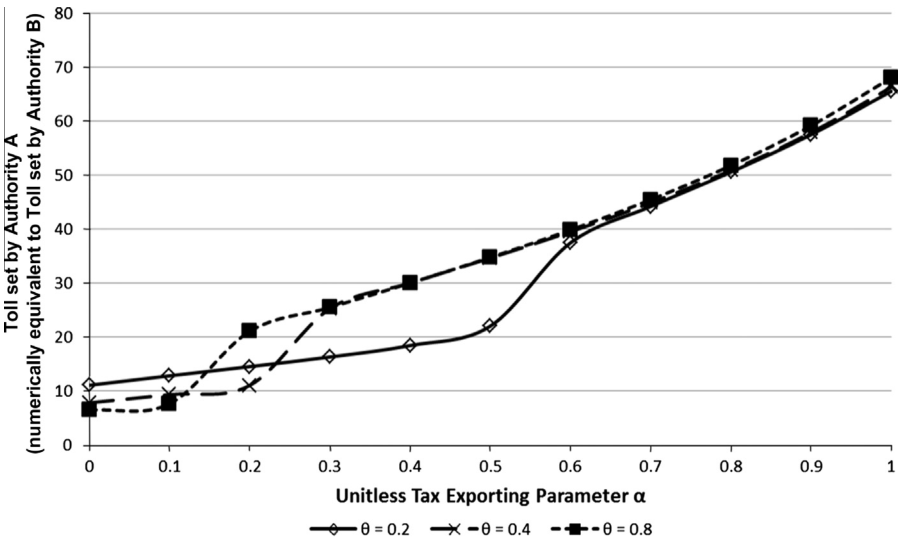

Fig. 7. NE tolls (min) as tax exporting parameter changes.

An additional effect is that as the toll levels rise with an increase in α, the number of paths effectively utilised by an OD pair decreases; that is to say, although all paths are used at least a small amount at any value of α, the number that are used by a significant amount of traffic decreases. From an analysis of the resulting path flows as the tolls increase, it is found that with an increase in α, non-captive OD pairs begin to avoid using the tolled links by rerouting away from these less attractive tolled links. This explains the apparent change in toll regime in Fig. 7 which notably appears at different levels of α for different dispersion parameters θ. That is to say, sensitivity to costs (as controlled by θ) has a complex interactive effect with the assumption on tax-exporting behaviour (as controlled by α), at the NE solution.

Where the phenomenon of a change in the toll regime precisely occurs depends on both θ and α, and this is difficult to predict *a priori*. But in general, as θ increases, because users become more sensitive to costs, this would be expected to take place at lower values of α. For example, for $\theta = 0.2$, there is a noticeable "break" at $\alpha = 0.6$ and an NE toll of 37.48 compared to the NE toll of 18.47 at $\alpha = 0.4$. However, when users become more cost sensitive this "break" occurs earlier. We can see a change between $\alpha = 0$ and $\alpha = 0.2$ for the case of $\theta = 0.8$.

If in the competitive (NE) case, some higher-level regulatory authority were able to influence the tax-exporting agreement through α (given knowledge of θ), so as to achieve the highest total welfare change, then we can see that the 'optimum' value occurs around $\alpha = 0.4$ for all cases of θ. The welfare gains achieved in such a way can be seen to be close to the benefits achieved for the relevant global regulator problem reported in Table 5. Another way to interpret this, from the viewpoint of the individual authorities, is that while there would be an incentive for the players to 'collude' in the non-cooperative game, full collusion (as represented by $\alpha = 1$) would result in lower welfare than partial collusion ($\alpha < 1$). As explained above, this is due to the local welfare function not taking full account of the total welfare effects. The fact that there is a close to global regulator solution with $\alpha = 0.4$ is most likely, we believe, to be network-specific and simply reflects that there is a trade-off being made by the local decision makers with respect to revenues from neighbours and congestion impacts on their own residents. Therefore, in any given network, any overall regulator of competition between cities would need to evaluate NE solutions for varying α, in order to determine an optimal level of cooperation for the region as a whole.

5. Summary and conclusions

In this paper, we have proposed a model and mathematical formulation of competition between two city authorities as a kind of two-level Nash game. On the upper level, each authority aims non-cooperatively to maximise social welfare of their residents by exerting tolls on their own jurisdiction, anticipating the reactions of travellers while reacting to the tolls set by the other authority. On the lower level, the travellers react to the tolls by non-cooperatively deciding whether to travel and if so by which route. A particular feature we exploit is that under an elastic demand SUE model for the lower level, necessary conditions for solutions to this problem can be written as a single-level variational inequality. We distinguish a form of local and global solution (referred to as LNE and NE respectively) that are contained within the set of solutions to the variational inequality problem. In our tests, we have shown how one general algorithm for solving such problems, namely SLCP, was successful at determining such solutions.

In our numerical experiments, we show that authorities, when allowed to retain the full stream of revenues from pricing ("full tax-exporting"), have an incentive to charge a toll higher than that obtained under the Global Regulatory Problem which is intended not only to internalise local congestion but also to extract revenues from users outside their jurisdiction. However, the resulting effect in physically adjacent regions (modelled as networks with serial dependencies) leads to so-called double marginalisation (Economides and Salop, 1992; De Borger et al., 2007). The consequence of double marginalisation is reduced benefits from pricing leading to a Prisoner's Dilemma outcome (Zhang et al., 2011). As shown in Example 1, toll competition could result in an even worse outcome than not tolling at all. As shown in Example 2, the benefits of toll competition could be lower than taking cooperative action, though higher than no tolls at all.

In our experiments on alternative tax-exporting assumptions, it is seen that even if local authorities share revenues from pricing, they might set tolls that are far from those that a global regulator would set. In these tests we saw that when authorities fully keep the revenues from pricing, they would set a toll higher than the (second best) socially optimal toll in order to retain the revenues which contribute to local welfare. The opposite occurs if they were to return the entire revenue stream back to those who paid the toll. This happens because the tolls decided at a local level do not, in contrast to a globally decided (second best) toll, imperfectly internalise externalities *wherever they occur in the network*. A locally decided toll does not therefore take into account the full spectrum of interaction effects that result in congestion across the entire network. Thus a potential policy implication of this research is that when pricing is decentralised to the local level, the authorities need to take into account the interactions of *all* users rather than solely that of local users. Furthermore, it is shown that even agreeing revenue recycling strategies (modelled with a common α parameter) will not necessarily offset the disbenefits associated with double marginalisation.

Acknowledgements

The work was funded by the Engineering and Physical Sciences Research Council of the UK under Grant EP/H021345/1. We are grateful to three anonymous referees for their constructive criticism of an earlier version of this paper.

References

Acemoglu, D., Ozdaglar, A., 2007. Competition and efficiency in congested markets. Mathematics of Operations Research 32 (1), 1–31.
Bell, M.G.H., Iida, Y., 1997. Transportation Network Analysis. Wiley, New York.
Cantarella, G.E., 1997. A general fixed-point approach to multimode multi-user equilibrium assignment with elastic demand. Transportation Science 31 (2), 107–128.
Connors, R.D., Sumalee, A., Watling, D.P., 2007. Sensitivity analysis of the variable demand probit stochastic user equilibrium with multiple user classes. Transportation Research Part B 41 (6), 593–615.
Cottle, R.W., Pang, J.-S., Stone, R.E., 2009. The Linear Complementarity Problem. Society for Industrial and Applied Mathematics, Philadelphia.
De Borger, B., Proost, S., 2012. Transport policy competition between governments: a selective survey of the literature. Economics of Transportation 1 (1–2), 35–48.
De Borger, B., Proost, S., van Dender, K., 2005. Congestion and tax competition on a parallel network. European Economic Review 49 (8), 2013–2040.
De Borger, B., Dunkerley, F., Proost, S., 2007. Strategic investment and pricing decisions in a congested transport corridor. Journal of Urban Economics 62 (2), 294–316.
de Palma, A., Lindsey, R., 2000. Private roads: competition under various ownership regimes. Annals of Regional Science 34 (1), 13–35.
de Palma, A., Lindsey, R., 2011. Traffic congestion pricing methodologies and technologies. Transportation Research Part C 19 (6), 1377–1399.
Economides, N., Salop, S., 1992. Competition and integration among complements, and network market structure. Journal of Industrial Economics 40 (1), 105–123.
Engel, E., Fischer, R., Galetovic, A., 1997. Highway franchising: pitfalls and opportunities. American Economic Review 87 (2), 68–72.
Engel, E., Fischer, R., Galetovic, A., 2004. Toll competition among congested roads. Topics in Economic Analysis 4 (1), 1–19.

Ferris, M., Munson, T., 2000. Complementarity problems in GAMS and the PATH solver. Journal of Economic Dynamics and Control 24 (2), 165–188.
Friesz, T.L., Tobin, R.L., Smith, T.E., Harker, P.T., 1983. A nonlinear complementarity formulation and solution procedure for the general derived demand network equilibrium problem. Journal of Regional Science 23 (3), 337–359.
Gühnemann, A., Koh, A., Shepherd, S., Lawler, M., 2011. Implications of interdependencies between charging strategies of local authorities for the protection of sensitive areas in the Trans-Pennine corridor. Utilities Policy 19 (1), 42–52.
Gühnemann, A., Koh, A., Shepherd, S., 2014. Optimal charging strategies under conflicting objectives for the protection of sensitive areas: a case study of the Trans-Pennine Corridor. Network and Spatial Economics. http://dx.doi.org/10.1007/s11067-013-9211-9.
Harker, P.T., 1984. A variational inequality approach for the determination of oligopolistic market equilibrium. Mathematical Programming 30 (1), 105–111.
Harker, P.T., Friesz, T.L., 1985. The use of equilibrium network models in logistics management with application to the U.S. coal industry. Transportation Research Part B 19 (5), 457–470.
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, Dordrecht, pp. 109–124.
Huang, H.J., Yang, H., Bell, M.G.H., 2000. The models and economics of carpools. Annals of Regional Science 34 (1), 55–68.
Karamardian, S., 1972. The complementarity problem. Mathematical Programming 2 (1), 107–129.
Koh, A., Shepherd, S.P., Watling, D.P., 2012. Competition between cities: an exploration of response surfaces and possibilities for collusion. Transportmetrica A 9 (10), 896–924.
Kolstad, C., Mathiesen, L., 1991. Computing Cournot-Nash equilibrium. Operations Research 39 (5), 739–748.
Krawczyk, J.B., Uryasev, S., 2000. Relaxation algorithms to find Nash equilibria with economic applications. Journal of Environmental Modeling and Assessment 5 (1), 63–73.
Lawphongpanich, S., Hearn, D., 2004. An MPEC approach to second-best toll pricing. Mathematical Programming 101 (1), 33–55.
Levinson, D.M., 2001. Why states toll: an empirical model of finance choice. Journal of Transport Economics and Policy 35 (2), 223–238.
Li, Z., Hensher, D.A., 2010. Toll roads in Australia: an overview of characteristics and accuracy of demand forecasts. Transport Reviews 30 (5), 541–569.
Lo, H.K., Szeto, W.Y., 2002. A methodology for sustainable traveler information services. Transportation Research Part B 36 (2), 113–130.
Maher, M., Stewart, K., Rosa, A., 2005. Stochastic social optimum traffic assignment. Transportation Research Part B 39 (8), 753–767.
Marsden, G., May, A.D., 2006. Do institutional arrangements make a difference to transport policy and implementation? Lessons for Britain. Environment and Planning Part C 24 (5), 771–789.
May, A.D., Milne, D.S., 2000. Effects of alternative road pricing systems on network performance. Transportation Research Part A 34 (6), 407–436.
Mills, G., 1995. Welfare and profit divergence for a tolled link in a road network. Journal of Transport Economics and Policy 29 (2), 137–146.
Mordukhovich, B.S., 2005. Optimization and equilibrium problems with equilibrium constraints. Omega – International Journal of Management Science 33 (5), 379–384.
Mun, S.-I., Ahn, K.-J., 2008. Road pricing in a serial network. Journal of Transport Economics and Policy 42 (3), 367–395.
Nash, J., 1950. Equilibrium points in N-person games. Proceedings of the National Academy of Science of the United States of America 36 (1), 48–49.
Nash, J., 1951. Non-cooperative games. Annals of Mathematics Second Series 54 (2), 286–295.
Nikaidô, H., Isoda, K., 1955. Note on non-cooperative convex games. Pacific Journal of Mathematics 5 (1), 807–815.
Ortuzar, J. de D., 2001. On the development of the nested logit model. Transportation Research Part B 35 (2), 213–216.
Pemberton, S., 2000. Institutional governance, scale and transport policy – lessons from Tyne and Wear. Journal of Transport Geography 8 (4), 295–308.
Pigou, A.C., 1920. Wealth and Welfare. Macmillan, London.
Proost, S., Sen, A., 2006. Urban transport pricing reform with two levels of government: a case study of Brussels. Transport Policy 13 (2), 127–139.
Ribeiroa, I.M., Simõesa, M., 2015. The fully actuated traffic control problem solved by global optimization and complementarity. Engineering Optimization. http://dx.doi.org/10.1080/0305215X.2014.995644.
Rork, J.C., 2009. Yardstick competition in toll revenues: evidence from US states. Journal of Transport Economics and Policy 43 (1), 123–139.
Sheffi, Y., 1985. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice Hall, Engelwood Cliffs, New Jersey.
Small, K.A., Verhoef, E.T., 2007. The Economics of Urban Transportation. Routledge, London.
Son, Y.S., Baldick, R., 2004. Hybrid coevolutionary programming for Nash equilibrium search in games with local optima. IEEE Transactions on Evolutionary Computation 8 (4), 305–315.
Tan, Z., Yang, H., 2012. Flexible build-operate-transfer contracts for road franchising under demand uncertainty. Transportation Research Part B 46 (10), 1419–1439.
Tan, Z., Yang, H., Guo, X., 2010. Properties of Pareto-efficient contracts and regulations for road franchising. Transportation Research Part B 44 (4), 415–433.
Tinti, F., 2005. Numerical solution for pseudomonotone variational inequality problems by extragradient methods. In: Giannessi, F., Maugeri, A. (Eds.), Variational Analysis and Applications. Springer, New York, pp. 1101–1129.
Ubbels, B., Verhoef, E.T., 2008. Governmental competition in road charging and capacity choice. Regional Science and Urban Economics 38 (2), 174–190.
van den Berg, V.A.C., 2013. Serial private infrastructures. Transportation Research Part B 56, 186–202.
Van Vuren, T., Watling, D.P., 1991. Multiple user class assignment model for route guidance. Transportation Research Record 1306, 22–32.
Verhoef, E.T., Nijkamp, P., Rietveld, P., 1996. Second-best congestion pricing: the case of an untolled alternative. Journal of Urban Economics 40 (3), 279–302.
Verhoef, E.T., Koh, A., Shepherd, S.P., 2010. Pricing, capacity and long-run cost functions for first-best and second-best network problems. Transportation Research Part B 44 (7), 870–885.
Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. Proceedings of the Institution of Civil Engineers, Part II 1 (3), 325–378.
Williams, H.C.W.L., 1977. On the formation of travel demand models and economic evaluation measures of user benefit. Environment and Planning A 9 (3), 285–344.
Xiao, F., Yang, H., Han, D., 2007. Competition and efficiency of private toll roads. Transportation Research Part B 41 (3), 292–308.
Yang, H., Huang, H.-J., 2005. Mathematical and Economic Theory of Road Pricing. Elsevier, Amsterdam.
Yang, H., Lam, W.H.K., 1996. Optimal road tolls under conditions of queueing and congestion. Transportation Research Part A 30 (5), 319–332.
Yang, H., Xiao, F., Huang, H.-J., 2009. Private road competition and equilibrium with traffic equilibrium constraints. Journal of Advanced Transportation 43 (1), 21–45.
Ying, J.Q., Miyagi, T., 2001. Sensitivity analysis for stochastic user equilibrium network flows—a dual approach. Transportation Science 35 (2), 124–133.
Ying, J.Q., Yang, H., 2005. Sensitivity analysis of stochastic user equilibrium flows in a bi-modal network with application to optimal pricing. Transportation Research Part B 39 (5), 769–795.
Zhang, X.-P., 2010. Overview of electricity market equilibrium problems and market power analysis. In: Zhang, X.-P. (Ed.), Restructured Electric Power Systems: Analysis of Electricity Markets with Equilibrium Models. IEEE Press, Piscataway New Jersey, pp. 99–137.
Zhang, X.N., Zhang, H.M., Huang, H.-J., Sun, L.-J., Tang, T.-Q., 2011. Competitive, cooperative and Stackelberg congestion pricing for multiple regions in transportation networks. Transportmetrica 7 (4), 297–320.