The Dynamic Efficient Toll Problem

Terry L. Friesz, *Penn State University, USA*, tfriesz@psu.edu
Changhyun Kwon[1], *Penn State University, USA*, chkwon@psu.edu
Andy H. F. Chow, *University College London, UK*, andy@transport.ucl.ac.uk
Benjamin Heydecker, *University College London, UK*, ben@transport.ucl.ac.uk

INTRODUCTION

The advent of new commitments by municipal, state and federal governments to construct and operate roadways whose tolls may be set dynamically has brought into sharp focus the need for a computable theory of dynamic tolls. Moreover, it is clear from the policy debates that surround the issue of dynamic tolls that pure economic efficiency is not the sole or even the most prominent objective of any dynamic toll mechanism that will be implemented. Rather, equity considerations as well as preferential treatment for certain categories of commuters must be addressed by such a mechanism. Accordingly, we introduce in this paper the dynamic efficient toll problem.

To study the dynamic efficient toll problem (DETP), it is necessary to employ some form of dynamic user equilibrium model. We elect the formulation due to Friesz et al (2001) and Friesz et al (2006) and its varieties analyzed by Ban et al (2006) and others. The dynamic DETP formulation will be constructed by direct analogy to the static efficient toll problem formulation of Hearn et al (2002). This approach to the formulation of the DETP leads directly to an the efficient toll pricing rule, provided appropriate necessary conditions that recognize time shifts are employed. The necessary conditions are those derived by Friesz et al (2004) for optimal control problems with state-dependent time shifts.

The main focus of this paper is the formulation and solution of the DETP. To this end, it will turn out that we need to solve a dynamic system optimum (DSO) problem and a dynamic user equilibrium (DUE) problem. Again using the DUE formulation reported in Friesz *et al.* (2001) and Friesz and Mookherjee (2006), we will provide the basis for the solution of the DETP. Also we show how to easily extend the formulation to include the day-to-day evolution of demand. Of course there are several ways such a model may be formulated. The dual-time scale formulation we shall emphasize is based on our prior work on differential

[1] Corresponding Author

variational inequalities and equilibrium network design and follows the qualitative theory conjectured (but not analyzed) by Friesz *et al.* (1996).

Central to the algorithmic study of the DETP in this paper is the descent in Hilbert spaces to the DSO problem. In the numerical approach, we employ an implicit fixed point scheme like that in Friesz and Mookherjee (2006) for dealing with time shifts in differential variational inequalities. In an example provided near the end of this paper, we numerically study a small network and determine its optimal dynamic tolls.

NOTATION AND MODEL FORMULATION

In this section we purposely repeat key portions of the time-lagged DUE formulation given in Friesz et al. (2001), because of its key role in this manuscript. The reader familiar with the notation and time-shifted DUE model presented in Friesz et al. (2001) may skip this section of the present paper.

Dynamic, Delay Operators and Constraints

The network of interest will form a directed graph $G(N, A)$, where N denotes the set of nodes and A denotes the set of arcs; the respective cardinalities of these sets are $|N|$ and $|A|$. An arbitrary path $p \in P$ of the network is

$$p \equiv \{a_1, a_2, ..., a_i, ..., a_{m(p)}\},$$

where P is the set of all paths and $m(p)$ is the number of arcs of p. We also let t_e denote the time at which flow exists an arc, while t_d is the time of departure from the origin of the same flow. The exit time function $\tau_{a_i}^p$ therefore obeys

$$t_e = \tau_{a_i}^p(t_d)$$

The relevant arc dynamics are

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \forall p \in P, \quad i \in \{1,2,...,m(p)\} \tag{1}$$

$$x_{a_i}^p(t) = x_{a_{i,0}}^p \qquad \forall p \in P, \quad i \in \{1,2,...,m(p)\} \tag{2}$$

where $x_{a_i}^p$ is the traffic volume of arc a_i contributed by path p, $g_{a_i}^p$ is flow exiting arc a_i and $g_{a_{i-1}}^p$ is flow entering arc a_i of path $p \in P$. Also, $g_{a_0}^p$ is the flow exiting the origin of path p; by convention we call this the flow of path p and use the symbolic name $h_p = g_{a_0}^p$. Furthermore

$$\delta_{a_i p} = \begin{cases} 1 & \text{if } a_i \in p \\ 0 & \text{if } a_i \notin p \end{cases}$$

so that $x_a(t) = \sum_{p \in P} \delta_{ap} \; x_a^p(t) \qquad \forall a \in A$ is the total arc volume.

Arc unit delay is $D_a(x_a)$ for each arc $a \in A$. That is, arc delay depends on the number of vehicles in front of a vehicle as it enters an arc. Of course total path traversal time is

$$D_p(t) = \sum_{i=1}^{m(p)} \left[\tau_{a_i}^p(t) - \tau_{a_{i-1}}^p(t) \right] = \tau_{a_{m(p)}}^p(t) - t \qquad \forall p \in P$$

It is expedient to introduce the following recursive relationships that must hold in light of the above development:

$$\tau_{a_1}^p(t) = t + D_{a_1}\left[x_{a_1}(t) \right] \quad \forall p \in P$$

$$\tau_{a_i}^p(t) = \tau_{a_{i-1}}^p(t) + D_{a_i}\left[x_{a_i}(\tau_{a_{i-1}}^p(t)) \right] \quad \forall p \in P, \quad i \in \{2,3,...,m(p)\}$$

from which we have the nested path delay operators first proposed by Friesz *et al.* (1993):

$$D_p(t,x) \equiv \sum_{i=1}^{m(p)} \delta_{a_i p} \Phi_{a_i}(t,x) \quad \forall p \in P,$$

where $x = (x_{a_i}^p : p \in P, i \in \{1,2,...,m(p)\}$

and

$$\begin{aligned} \Phi_{a_1}(t,x) &= D_{a_1}(x_{a_1}(t)) \\ \Phi_{a_2}(t,x) &= D_{a_2}(x_{a_2}(t + \Phi_{a_1})) \\ \Phi_{a_3}(t,x) &= D_{a_3}(x_{a_3}(t + \Phi_{a_1} + \Phi_{a_2})) \\ &\vdots \\ \Phi_{a_i}(t,x) &= D_{a_i}(x_{a_i}(t + \Phi_{a_1} + \cdots + \Phi_{a_{i-1}})) \\ &= D_{a_i}\left(x_{a_i}\left(t + \sum_{j=1}^{i-1} \Phi_{a_j} \right) \right). \end{aligned}$$

To ensure realistic behaviour, we employ asymmetric early/late arrival penalties

$$F\left[t + D_p(t,x) - t_A \right]$$

where t_A is the desired arrival time and

$$\begin{aligned} t + D_p(t,x) > t_A &\Rightarrow \quad F(t + D_p(t,x) - t_A) = \chi^L(x,t) > 0 \\ t + D_p(t,x) < t_A &\Rightarrow \quad F(t + D_p(t,x) - t_A) = \chi^E(x,t) > 0 \\ t + D_p(t,x) = t_A &\Rightarrow \quad F(t + D_p(t,x) - t_A) = 0 \end{aligned}$$

while $\qquad \chi^L(t,x) > \chi^E(t,x)$.

Let us further denote arc tolls by y_a for each arc $a \in A$. We assume that users pay any toll imposed on an arc at the entrance of the arc. Then the path tolls y_p for each path $p \in P$ are

$$y_p(t) = \sum_{i=1}^{m(p)} \delta_{a_i p} y_{a_i}\left(t + \sum_{j=1}^{i-1} \Phi_{a_j}(t,x)\right) \quad \forall p \in P,$$

where $\Phi_{a_0}(t,x) = 0$. If the tolls are paid when users exit arcs, then the path toll becomes

$$y_p(t) = \sum_{i=1}^{m(p)} \delta_{a_i p} y_{a_i}\left(t + \sum_{j=1}^{i} \Phi_{a_j}(t,x)\right) \quad \forall p \in P.$$

We now combine the actual path delays and arrival penalties to obtain the *effective delay operators*

$$\Psi_p(t,x) = D_p(t,x) + F\left(t + D_p(x,t) - T_A\right) \qquad \forall p \in P \tag{3}$$

According to the FIFO principle, the volume that has entered an arc up to time t will exit it by time $t + D_{a_i}(x_{a_i}(t))$ so that

$$\int_0^t g_{a_{i-1}}^p(t)\,dt = \int_{D_{a_i}(x_{a_i}(0))}^{t + D_{a_i}(x_{a_i}(t))} g_{a_i}^p(t)\,dt \quad \forall p \in P, i \in [1, m(p)], \tag{4}$$

where $g_{a_0}^p(t) = h_p(t)$. Differentiating both sides of (4) with respect to time t and using the chain rule, we have

$$h_p(t) = g_{a_1}^p(t + D_{a_1}(x_{a_1}(t)))(1 + D_{a_1}'(x_{a_1}(t))\dot{x}_{a_1}) \qquad \forall p \in P \tag{5}$$

$$g_{a_{i-1}}^p(t) = g_{a_i}^p(t + D_{a_i}(x_{a_i}(t)))(1 + D_{a_i}'(x_{a_i}(t))\dot{x}_{a_i}) \qquad \forall p \in P, \quad i \in [2, m(p)]. \tag{6}$$

These are *proper flow progression constraints* derived in a fashion that makes them completely *consistent with any traffic model that respects the FIFO discipline*. These constraints involve a state-dependent time lag $D_{a_i}(x_{a_i}(t))$ but make no explicit reference to the exit time functions. These flow propagation constraints describe the expansion and contraction of vehicle platoons; they were presented by Holden (1989); Friesz *et al.* (1995). Astarita (1995, 1996) independently proposed flow propagation constraints that may be readily placed in the above form.

The final constraints to consider are those of flow conservation and non-negativity:

$$\sum_{p \in P_{ij}} \int_{t_0}^{t_f} h_p(t)\,dt = Q_{ij} \quad \forall (i,j) \in W \tag{7}$$

$$h_p \geq 0 \quad \forall (i,j) \in P_{ij} \tag{8}$$

$$g_{a_i}^p \geq 0 \quad \forall p \in P, i \in [1, m(p)] \tag{9}$$

$$x_{a_i}^p \geq 0 \quad \forall p \in P, i \in [1, m(p)], \tag{10}$$

where W is the set of origin-destination pairs, P_{ij} is the set of paths connecting origin-destination pair (i,j), $t_f > t_0$, and $t_f - t_0$ defines the planning horizon. Furthermore, Q_{ij} is the travel demand (a volume) for the period $[t_0 . t_f]$. In what follows h will denote the vector of all path flows, g the vector of all arc exit flows. Finally, we denote the set of all feasible exit flow vectors (h,g) by Ω; that is

$$\Omega \equiv \{(h,g): (1),(2),(5),(6),(7),(8),(9),(10) \text{ are satisfied}\}. \tag{11}$$

Dynamic User Equilibrium

Given the effective unit travel delay Ψ_p for path p, the infinite dimensional variational inequality formulation for dynamic network user equilibrium itself is: find $(g^*, h^*) \in \Omega$ such that

$$\left\langle \Psi(t, x(h^*, g^*)), (h - h^*) \right\rangle = \sum_{p \in P} \int_{t_0}^{t_f} \Psi_p[t, x(h^*, g^*)] \cdot [h_p(t) - h_p^*(t)] dt \geq 0 \tag{12}$$

for all $(h, g) \in \Omega$, where Ψ denotes the vector of effective path delay operators. Friesz *et al*. (2001) show all solutions of (12) are dynamic user equilibria[2]. In particular the solutions of (12) obey

$$\Psi_p(t, x(g^*, h^*)) > \mu_{ij} \Rightarrow h_p^*(t) = 0 \tag{13}$$

$$h_p^*(t) > 0 \Rightarrow \Psi_p(t, x(g^*, h^*)) = \mu_{ij} \tag{14}$$

for $p \in P_{ij}$ where μ_{ij} is the lower bound on achievable costs for any ij -traveler, given by

$$\mu_p = ess \inf\{\Theta_p(t, x): t \in [t_0, t_f]\} \geq 0$$

and

$$\mu_{ij} = \min\{\mu_p : p \in P_{ij}\} \geq 0 .$$

We call a flow pattern satisfying (13) and (14) a *dynamic user equilibrium*. The behavior described by (13) and (14) is readily recognized to be a type of Cournot-Nash non-cooperative equilibrium. It is important to note that these conditions do not describe a stationary state, but rather a time varying flow pattern that is a Cournot-Nash equilibrium (or user equilibrium) at each instant of time.

THE DYNAMIC EFFICIENT TOLL PROBLEM (DETP)

Hearn and Yildrim (2002) studied the efficient toll in the static setting with the traveling cost which is linear in the traffic flow. The objective of the efficient toll is to make the user equilibrium traffic flow equivalent to the system optimum by appropriate congestion pricing. To study the dynamic efficient toll problem (DETP), we introduce the notion of a *tolled effective delay operator*:

$$\Theta_p(t, x, y_p) = D_p(t, x) + F\{t + D_p(x, t) - T_A\} + y_p(t) \qquad \forall p \in P ,$$

where y_p denotes the toll for path p . Of course we have the relationship

---

[2]Although we have purposely suppressed the functional analysis subtleties of the formulation, it should be noted that (12) involves an inner product in a Hilbert space, namely $\left(L^2[0,T]\right)^{|P|}$.

$$\Theta_p(t,x,y_p) = \Psi_p(t,x) + y_p(t). \tag{15}$$

To make the toll meaningful, we enforce the efficient toll non-negative:

$$y_p(t) \geq 0 \quad \forall t \in [t_0, t_f],\ p \in P.$$

Analysis of the System Optimum

The dynamic system optimum (DSO) is achieved by solving

$$\min J_1 = \int_{t_0}^{t_f} \sum_{p \in P} e^{-rt} \Psi_p(t,x)\, h_p(t)\, dt$$

subject to

$$\frac{dx_{a_i}^p(t)}{dt} = g_{a_{i-1}}^p(t) - g_{a_i}^p(t) \qquad \forall p \in P, \quad i \in [1, m(p)] \tag{16}$$

$$x_{a_i}^p(t_0) = x_{a_{i,0}}^p \qquad \forall p \in P, \quad i \in [1, m(p)]$$

$$g_{a_{i-1}}^p(t) = g_{a_i}^p(t + D_{a_i}(x_{a_i}(t)))(1 + D_{a_i}^{'}(x_{a_i}(t))\dot{x}_{a_i}) \qquad \forall p \in P, \quad i \in [1, m(p)] \tag{17}$$

$$\sum_{p \in P_{ij}} \int_{t_0}^{t_f} h_p(t)\, dt = Q_{ij} \quad \forall (i,j) \in W \tag{18}$$

$$x \geq 0 \qquad g \geq 0 \qquad h \geq 0, \tag{19}$$

where we have used the convention

$$g_{a_0}^p = h_p.$$

and r is the discount rate which must be defined according to the planning time horizon. It will be convenient to employ the following shorthand for shifted variables:

$$\bar{g}_{a_i}^p \equiv g_{a_i}^p(t + D_{a_i}(x_{a_i}(t))) \quad \forall p \in P, \quad i \in 0, m(p)].$$

Penaltizing (17) we obtain

$$J_1 = \int_{t_0}^{t_f} \left\{ \sum_{p \in P} e^{-rt} \Psi_p(t,x) h_p(t) + \sum_{p \in P} \sum_{i=1}^{m(p)} \frac{\mu_{a_i}^p}{2} \left[g_{a_{i-1}}^p(t) - \bar{g}_{a_i}^p(t)(1 + D_{a_i}^{'}(x_{a_i}(t))\dot{x}_{a_i}) \right]^2 \right\} dt, \tag{20}$$

where $\mu_{a_i}^p$ is the penalty coefficient. Let us then define the set of feasible controls

$$\Lambda \equiv \left\{ (h,g) : \sum_{p \in P_{ij}} \int_{t_0}^{t_f} h_p(t) dt = Q_{ij} \ \forall (i,j) \in W, h \geq 0, \quad g \geq 0 \right\}. \tag{21}$$

Optimal control problem (20) and (21) is an instance of the time-shifted optimal control problem analyzed in Friesz *et al.* (2001). We also employ the following notation for the state vector and control vector, respectively:

$$x = \left(x_{a_i}^{p} : p \in P, i \in [1, m(p)] \right)$$

$$g = \left(g_{a_i}^{p} : p \in P, i \in [1, m(p)] \right).$$

The DSO Hamiltonian is

$$H_1(t,x,h,g,\lambda;\mu) \equiv \sum_{p\in P} e^{-rt}\Psi_p(t,x)h_p(t) + \sum_{p\in P}\sum_{i=1}^{m(p)} \frac{\mu_{a_i}^{p}}{2}\left\{ g_{a_{i-1}}^{p}(t) - \overline{g}_{a_i}^{p}(t)(1+D_{a_i}^{'}(x_{a_i}(t))\dot{x}_{a_i}) \right\}^2$$
$$+ \sum_{p\in P}\sum_{i=1}^{m(p)} \lambda_{a_i}^{p}\left(g_{a_{i-1}}^{p}(t) - g_{a_i}^{p}(t) \right).$$

Let us introduce the vector

$$F(t,x,h,g,\lambda;\mu) = \left(F_{a_i}^{p}(t,x,h,g,\lambda;\mu) : p \in P, i \in [0, m(p)] \right),$$

where

$$F_{a_0}^{p}(t,x,h,g,\lambda;\mu) = \frac{\partial H_1(t,x,h,g,\lambda;\mu)}{\partial h_p} \quad \forall p \in P \tag{22}$$

$$F_{a_i}^{p}(t,x,h,g,\lambda;\mu) = \begin{cases} \dfrac{\partial H_1(t,x,h,g,\lambda;\mu)}{\partial g_{a_i}^{p}} & \text{if } t \in \left[t_0, D_{a_i}(x(t_0)) \right] \\[2ex] \dfrac{\partial H_1(t,x,h,g,\lambda;\mu)}{\partial g_{a_i}^{p}} + \left[\dfrac{\partial H_1(t,x,h,g,\lambda;\mu)}{\partial \overline{g}_{a_i}^{p}} \left(\dfrac{1}{1+D_{a_i}^{'}(x_{a_i}(t))\dot{x}_{a_i}} \right) \right]_{s_{a_i}(t)} & \text{if } t \in \left[D_{a_i}(x(t_0)), t_f + D_{a_i}(x(t_f)) \right] \end{cases} \tag{23}$$
$$\forall p \in P, \quad i \in [1, m(p)]$$

and each $s_{a_i}(t)$ is a solution of the fixed point problem $s_{a_i}(t) = \arg\left[s = t - D_{a_i}(x(s)) \right]$. We may write (22) and (23) in detail as

$$F_{a_0}^{p}(t,x,h,g,\lambda;\mu) = e^{-rt}\left[\Psi_p(t,x) + \frac{\partial \Psi_p(t,x)}{\partial h_p} h_p \right]$$
$$+ \mu_{a_1}^{p}\left[g_{a_0}^{p}(t) - \overline{g}_{a_1}^{p}(t)(1+D_{a_1}^{'}(x_{a_1}(t))\dot{x}_{a_1}) \right] + \lambda_{a_1}^{p} \quad \forall p \in P \tag{24}$$

$$F_{a_i}^{p}(t,x,h,g,\lambda;\mu)=\begin{cases}\mu_{a_{i+1}}^{p}\left\{g_{a_i}^{p}(t)-\bar{g}_{a_{i+1}}^{p}(t)(1+D_{a_{i+1}}^{'}(x_{a_{i+1}}(t))\dot{x}_{a_{i+1}})\right\}-\lambda_{a_i}^{p}+\lambda_{a_{i+1}}^{p} \\ \qquad \text{if } t\in\left[t_0,D_{a_i}(x(t_0))\right] \\ \mu_{a_{i+1}}^{p}\left\{g_{a_i}^{p}(t)-\bar{g}_{a_{i+1}}^{p}(t)(1+D_{a_{i+1}}^{'}(x_{a_{i+1}}(t))\dot{x}_{a_{i+1}})\right\}-\lambda_{a_i}^{p}+\lambda_{a_{i+1}}^{p} \\ \qquad -\left[\mu_{a_i}^{p}\left\{g_{a_{i-1}}^{p}(t)-\bar{g}_{a_i}^{p}(t)(1+D_{a_i}^{'}(x_{a_i}(t))\dot{x}_{a_i})\right\}\right]_{s_{a_i}(t)} \\ \qquad \text{if } t\in\left[D_{a_i}(x(t_0)),t_f+D_{a_i}(x(t_f))\right]\end{cases} \tag{25}$$

$$\forall p\in P,\quad i\in[1,m(p)-1]$$

$$F_{a_i}^{p}(t,x,h,g,\lambda;\mu)=\begin{cases}-\lambda_{a_i}^{p} \\ \qquad \text{if } t\in\left[t_0,D_{a_i}(x(t_0))\right] \\ -\lambda_{a_i}^{p}-\left[\mu_{a_i}^{p}\left\{g_{a_{i-1}}^{p}(t)-\bar{g}_{a_i}^{p}(t)(1+D_{a_i}^{'}(x_{a_i}(t))\dot{x}_{a_i})\right\}\right]_{s_{a_i}(t)} \\ \qquad \text{if } t\in\left[D_{a_i}(x(t_0)),t_f+D_{a_i}(x(t_f))\right]\end{cases} \tag{26}$$

$$\forall p\in P,\quad i=m(p).$$

Then a necessary condition for $(h^S,g^S)\in\Lambda$ to be the system optimum is

$$0\le\sum_{p\in P}\sum_{i=0}^{m(p)}F_{a_i}^{p}\left(t,x^S,h^S,g^S,\lambda^S;\mu\right)\left(g_{a_i}^{p}-g_{a_i}^{pS}\right)\quad\forall(h,g)\in\Lambda \tag{27}$$

for each time instant $t\in\left[t_{0,}\sup_{a_i\in A}\{t_f+D_{a_i}(x(t_f))\}\right]$, together with the state dynamics (16) and the following adjoint equations and boundary conditions

$$-\frac{d\lambda_{a_i}^{p,S}}{dt}=\frac{\partial H_1^S}{\partial x_{a_i}^{p}}=e^{-rt}\frac{\partial\Psi_p(t,x^S)}{\partial x_{a_i}^{p}}\qquad\forall p\in P,\quad i\in[1,m(p)]$$

$$\lambda_{a_i}^{p,S}(t_f)=0\qquad\forall p\in P,\quad i\in[1,m(p)],$$

where the superscript S denotes a trajectory corresponding to a system optimum.

Analysis of the User Equilibrium in the Presence of Tolls

However, a dynamic tolled user equilibrium must obey

$$\sum_{p\in P}\int_{t_0}^{t_f}e^{-rt}\left\{\Theta_p\left[t,x(h^U),y_p^U\right]\right\}\left[h_p(t)-h_p^U(t)\right]dt\ge0\quad\text{for all }(h,g)\in\Lambda, \tag{28}$$

where the state dynamics as well as all other state and control constraints are identical to those introduced above for DSO. In particular, the set of feasible controls Λ referred to in (28) remains unchanged. We formulate an optimal control problem[3] from the above dynamic user equilibrium variational inequality problem; its objective is

$$\min J_2 = \sum_{p\in P}\int_{t_0}^{t_f} e^{-rt}\Theta_p\left[t, x\left(h^U\right), y_p^U\right] h_p(t)dt$$

with the same constraints introduced previously. As previously done for the system optimum problem, we penalize the flow propagation constraints to obtain the modified criterion

$$J_2 = \sum_{p\in P}\int_{t_0}^{t_f}\left\{ e^{-rt}\Theta_p\left[t, x\left(h^U\right), y_p^U\right] h_p(t) + \sum_{p\in P}\sum_{i=1}^{m(p)}\frac{\mu_{a_i}^p}{2}\left[g_{a_{i-1}}^p(t) - \bar{g}_{a_i}^p(t)(1+D_{a_i}'(x_{a_i}(t))\dot{x}_{a_i})\right]^2\right\}dt \quad (29)$$

Then we have another standard form time-shifted optimal control problem, although it is subtly but importantly different than that for DSO. In particular, the Hamiltonian now becomes

$$\begin{aligned} H_2(t,x,h,g,\lambda;\mu) \equiv & \sum_{p\in P} e^{-rt}\Theta_p\left[t, x\left(h^U\right), y_p^U\right] h_p(t) + \sum_{p\in P}\sum_{i=1}^{m(p)}\frac{\mu_{a_i}^p}{2}\left\{ g_{a_{i-1}}^p(t) - \bar{g}_{a_i}^p(t)(1+D_{a_i}'(x_{a_i}(t))\dot{x}_{a_i})\right\}^2 \\ & + \sum_{p\in P}\sum_{i=1}^{m(p)}\lambda_{a_i}^p\left(g_{a_{i-1}}^p(t) - g_{a_i}^p(t)\right) \end{aligned}$$

An analysis of necessary conditions similar to that for DSO is now possible. The key difference is that the counterpart of (24) must in the user equilibrium case be written as follows:

$$\begin{aligned} G_{a_0}^p(t,x,h,g,\lambda;\mu) = & \, e^{-rt}\Theta_p\left[t, x\left(h^U\right), y_p^U\right] \\ & + \mu_{a_1}^p\left[g_{a_0}^p(t) - \bar{g}_{a_1}^p(t)(1+D_{a_1}'(x_{a_1}(t))\dot{x}_{a_1})\right] + \lambda_{a_1}^p \quad \forall p\in P \end{aligned} \quad (30)$$

$$G_{a_i}^p(t,x,h,g,\lambda;\mu) = F_{a_i}^p(t,x,h,g,\lambda;\mu) \qquad \forall p\in P, \quad i\in[1,m(p)]. \quad (31)$$

Then a necessary condition for $\left(h^S, g^S\right)\in\Lambda$ to be a dynamic user equilibrium (DUE) is

$$0 \le \sum_{p\in P}\sum_{i=0}^{m(p)} G_{a_i}^p\left(t, x^U, h^U, g^U, \lambda^U; \mu\right)\left(g_{a_i}^p - g_{a_i}^{pU}\right) \qquad g\in\Lambda \quad (32)$$

for each time instant $t\in\left[t_{0,}\ \sup_{a_i\in A}\left\{t_f + D_{a_i}\left(x\left(t_f\right)\right)\right\}\right]$, together with the state dynamics (16) and the following adjoint equations and boundary conditions:

$$-\frac{d\lambda_{a_i}^{p,U}}{dt} = \frac{\partial H_2^U}{\partial x_{a_i}^p} = e^{-rt}\frac{\partial\Theta_p\left[t, x\left(h^U\right), y_p^U\right]}{\partial x_{a_i}^p} \qquad \forall p\in P, \quad i\in[1,m(p)]$$

$$\lambda_{a_i}^{p,U}\left(t_f\right) = 0 \qquad \forall p\in P, \quad i\in[1,m(p)],$$

[3]This may not be used for numerical computation as its statement depends on knowledge of the dynamic user equilibrium being sought. However, it may be employed for qualitative analyses like those which follow.

where the superscript U denotes a trajectory corresponding to a dynamic user equilibrium in the presence of tolls.

Characterizing Efficient Tolls

It is the purpose of efficient tolls to make the criteria J_1 and J_2 identical along solution trajectories for which flow propagation and other constraints are satisfied, for then the system optimal total costs are identical to the tolled user optimal total costs. Furthermore, the vectors of path flows (departure rates) obey

$$h^U(t) = h^S(t). \tag{33}$$

There are as well identical arc exit flows and identical arc volumes. Therefore, along solution trajectories

$$\lambda_{a_1}^{p,S} = \frac{\partial J_1}{\partial x_{a_1}^{p,S}} = \frac{\partial J_2}{\partial x_{a_1}^{p,U}} = \lambda_{a_1}^{p,U} . \tag{34}$$

With (34) in mind and upon comparing (27) and (32), we find

$$\begin{aligned} e^{-rt}\left\{\Psi_p\left(t, x^S\right) + \frac{\partial \Psi_p\left(t, x^S\right)}{\partial h_p} h_p^S\right\} &= e^{-rt}\left\{\Theta_p\left[t, x\left(h^U\right), y_p^U\right]\right\} \\ &= e^{-rt}\left\{\Psi_p\left(t, x^U\right) + y_p^U(t)\right\}. \end{aligned}$$

The only toll constraint is non-negativity; hence applying the projection after the expression for $y_p{}^U(t)$ is derived with non-negativity relaxed will give an exact expression:

$$y_p^U(t) = \left[\frac{\partial \Psi_p\left(t, x^S\right)}{\partial h_p} h_p^S\right]^+ \qquad \forall t \in \left[t_0, t_f\right], \tag{35}$$

where $[\cdot]^+$ is the elementary projection operator defined by

$$[v]^+ = \begin{cases} v & \text{if } v \geq 0 \\ 0 & \text{if } v < 0 . \end{cases}$$

This result is completely analogous to that for an efficiently tolled static user equilibrium[4].

MULTIPLE TIME SCALES

We have investigated the within-day behavior of road network users so far. In this section we describe a day-to-day adjust process that sets daily travel demand. Our perspective is very simple: if today commuters experiences a level of congestion above a threshold representing

[4] We add the operator $[\cdot]^+$ to (35) to ensure non-negativity of the toll. In the corresponding static case, the derivative will never be negative. However, in the dynamic case, this depends on how the cost function is defined. In practice, the derivative is hard to calculate analytically, because **x** is an implicit function of **h**. Nevertheless, our numerical experiences has shown that the derivative does fall below zero under certain circumstances.

the budget or tolerance for congestion of the typical commuter, travel demand will be less tomorrow and more workers will elect to stay at home (telecommute). To operationalize this idea, we take the perspective of evolutionary game theory to describe the day-to-day demand learning process in terms of the *moving average* of congestion and difference equations.

Let $\tau \in \Upsilon \equiv \{1,2,...,L\}$ be one typical discrete day within the planning horizon, and take the length of each day to be Δ, while the continuous clock time t within each day is presented by $t \in [(\tau-1)\Delta, \tau\Delta]$ for all $\tau \in \{1,2,...,L\}$. The entire planning horizon spans L consecutive days. As noted above, we assume the travel demand for each day changes based on the moving average of congestion experienced over previous days. In fact we postulate that the travel demands Q_{ij}^{τ} for day τ between a given OD pair $(i,j) \in W$ are determined by the following system of difference equations:

$$Q_{ij}^{\tau+1} = \left[Q_{ij}^{\tau} - s_{ij}^{\tau} \left\{ \frac{\sum_{p \in P_{ij}} \sum_{j=0}^{\tau-1} \int_{j \cdot \Delta}^{(j+1)\cdot\Delta} \Psi_p\left[t, x\left(h^*, g^*\right)\right] dt}{\left|P_{ij}\right| \cdot \tau \cdot \Delta} - \chi_{ij} \right\} \right]^+ \quad \forall \tau \in [1, L-1] \tag{36}$$

with boundary condition

$$Q_{ij}^{1} = \tilde{Q}_{ij}, \tag{37}$$

where $\tilde{Q}_{ij} \in \Re_+$ is the fixed travel demand for the OD pair $(i,j) \in W$ for the first day and χ_{ij} is the representative threshold. The operator $[x]^+$ is shorthand from $\max[0,x]$. The parameter s_{ij}^{τ} is related to the rate of change of inter-day travel demand.

ALGORITHMS FOR SOLVING THE DETP

In this section, we provide the fixed point algorithm for solving the DETP.

The Implicit Fixed Point Perspective

In calculation of the efficient toll, state-dependent time shifts must and can be accommodated using an implicit fixed point perspective, as innovated for the dynamic user equilibrium by Friesz and Mookherjee (2006). More specifically, in such an approach, one employs control and state information from a previous iteration to approximate current time shifted functions. This perspective may be summarized as follows:

Step 1. Articulate the current approximate states (volumes) and controls (arc exit rates) by spline or other curve fitting techniques as continuous functions of time.

Step 2. Using the aforementioned continuous functions of time, express time shifted controls as pure functions of time, while leaving unshifted controls as decision functions to be updated within the current iteration.

Step 3. Update the states and controls, then repeat Step 2 and Step 3 until the control controls converge to a suitable approximate solution.

Step 4. Using the converged solutions, compute the dynamic efficient toll by the equation (35) derived in a previous section.

Descent in Hilbert Space

At each step 2 in the above numerical schemes, we need to solve the DSO problem, which is an instance of optimal control problems. Among algorithms to solve an optimal control problem, we will study descent method in Hilbert space. To articulate what is meant by descent in Hilbert space, it is much easier to study an abstract problem rather than the DSO because of the notational complexity of the problem. To that end, let us consider an abstract optimal control problem with mixed state-control constraints involving state-dependent time shifts from the point of view of infinite dimensional mathematical programming:

$$\min J = \int_{t_0}^{t_f} F(x,u,u_D,t)dt \tag{38}$$

subject to

$$x(u,u_D,t) \in \Lambda = \left\{ x : \frac{dx}{dt} = f(x,u,u_D,t), x(t_0) = 0, G(x,u,u_D,t) = 0, x \geq 0 \right\} \in (H^1[t_0,t_f])^n$$

where

$$\begin{aligned} u &\in U \subseteq (L^2[t_0,t_f])^m \\ u_D &\equiv u(t+D(x)) : (H^1[t_0,t_f])^n \times \Re^1_+ \to (L^2[t_0,t_f])^m \\ f &: (H^1[t_0,t_f])^n \times (L^2[t_0,t_f])^{2m} \times \Re^1_+ \to (L^2[t_0,t_f])^m \\ F &: (H^1[t_0,t_f])^n \times (L^2[t_0,t_f])^{2m} \times \Re^1_+ \to (L^2[t_0,t_f])^m \\ G &: (H^1[t_0,t_f])^n \times (L^2[t_0,t_f])^{2m} \times \Re^1_+ \to (L^2[t_0,t_f])^m . \end{aligned}$$

In the above, $(L^2[t_0,t_f])^m$ is the m-fold product of the space of square integrable functions $L^2[t_0,t_f]$ and $(H^1[t_0,t_f])^n$ is the n-fold product of the Sobolev space $H^1[t_0,t_f]$ for the real interval $[t_0,t_f] \subset \Re^1_+$. In applying descent in Hilbert space to this problem, it is convenient to use quadratic-loss penalty functions and a logarithmic barrier function to create the unconstrained program:

$$\min J_1 = \int_{t_0}^{t_f} F(x,u,u_D,t)dt + \frac{1}{2}\int_{t_0}^{t_f} \sum_i \eta_i (G_i(x,u,u_D,t))^2 dt + \frac{1}{2}\int_{t_0}^{t_f} \sum_i \rho_i \min(0,x_i)^2 dt \tag{39}$$

where it is understood that x denotes the operator

$$x(u,u_D,t)\in \Lambda_1 = \left\{ x : \frac{dx}{dt} = f(x,u,u_D,t), x(0) = x_0 \right\} \in (H^1[t_0,t_f])^n,$$

and η_i and ρ_i are penalty and barrier multipliers to be adjusted from iteration to iteration.

The resulting problem can be solved using a continuous time steepest descent method. For the penalized criterion (39), the algorithm can be stated as following:

Step 0. Initialization. Pick $u^0(t)\in U$ and set $k=1$.

Step 1. Finding state variables. Solve the state dynamics

$$\begin{aligned} \frac{dx}{dt} &= f(x,u^{k-1},u_D^{k-1},t) \\ x(0) &= x_0 \end{aligned}$$

and call the solution $x^k(t)$, using curve fitting to create an approximation to $x^k(t)$ when necessary.

Step 2. Finding adjoint variables. Solve the adjoint dynamics

$$\begin{aligned} -\frac{d\lambda}{dt} &= \left[\nabla_x H(x,u^{k-1},u_D^{k-1},\lambda,t)\right]_{x=x^k} \\ \lambda(t_f) &= 0 \end{aligned}$$

where the Hamiltonian is given by

$$H(x,u,u_D,\lambda,t) = F(x,u,u_D,t) + \frac{1}{2}\sum_i \rho_i \min(0,x_i)^2 + \frac{1}{2}\sum_i \eta_i (G_i(x,u,u_D,t))^2 + \lambda^T f(x,u,u_D,t)$$

Call the solution $\lambda^k(t)$, using curve fitting to create an approximation to $\lambda^k(t)$ when necessary.

Step 3. Finding the gradient. Determine

$$\nabla_u J^k \equiv \left[\nabla_u H(x^k,u,u_D^{k-1},\lambda^k,t)\right]_{u=u^k}$$

Step 4. Updating the current control. For a suitably small step size

$$\theta_k \in \Re^1_{++}$$

determine

$$u^{k+1}(t) = u^k(t) - \theta_k \nabla_u J^k$$

Step 5. Stopping Test. For $\varepsilon \in \Re^1_{++}$, a preset tolerance, stop if

$$\| u^{k+1} - u^k \| < \varepsilon$$

and declare

$$u^* \approx u^{k+1}$$

Otherwise set $k = k+1$ and go to Step1.

NUMERICAL EXAMPLE

In what follows, we consider a 3 arc, 3 node network shown in Figure 1. The arc labels and arc delay functions for this network are summarized in Table 1.

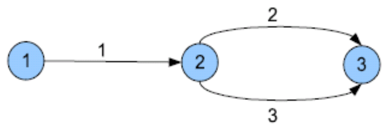

Figure 1. 3-arc 3-node traffic network.

Table 1. Arc labels and delay functions.

Arc name	From node	To node	Arc delay, $D_a(x_a(t))$
a_1	1	2	$2+\frac{1}{100}x_{a_1}$
a_2	2	3	$1+\frac{1}{150}x_{a_2}$
a_3	2	3	$3+\frac{1}{100}x_{a_3}$

There are 2 paths connecting the single OD pair formed by nodes 1 and 3, namely:

$$P_{13} = \{p_1, p_2\}, \qquad p_1 = \{a_1, a_2\}, \qquad p_2 = \{a_1, a_3\}.$$

The controls (path flows and arc exit flows) and states (path-specific arc traffic volumes) associated with the network are presented in Table 2.

Table 2. Control and state variables.

Path	Path flow	Arc exit flow	Traffic volume of arc
p_1	h_{p_1}	$g_{a_1}^{p_1}, g_{a_2}^{p_1}$	$x_{a_1}^{p_1}, x_{a_2}^{p_1}$
p_2	h_{p_2}	$g_{a_1}^{p_2}, g_{a_3}^{p_2}$	$x_{a_1}^{p_2}, x_{a_3}^{p_2}$

We consider a two-week toll planning in which each day is 24 hours, hence, $\Delta = 24$ and $L = 14$ (two weeks). We assume there is the initial travel demand $\tilde{Q} = 150$ units from node 1 (origin) to node 3 (destination). The threshold for travel cost is $\chi = 20000$ and the inter-day rate of change in travel demand is $s_{13} = 0.7$. We also assume the discount rate $r = 0$ as the planning horizon is relatively short in this example. The desired arrival time for each day is $t_A = 12$, and we employ the symmetric early/late arrival penalty

$$F\left[t + D_p(x,t) - t_A\right] = 5\left[t + D_p(x,t) - t_A\right]^2 .$$

Further, without any loss of generality, we take

$$x_{a_i}^{p}(0) = 0 \quad \forall i \in [1, m(p)], p \in P .$$

In what follows we forgo the detailed symbolic statement of this example, and, instead, provide numerical results in graphical form.

Computation of Tolls by the DETP

To compute the tolls, we suggest a computational scheme for DETP. Recall that the decision rule for the dynamic efficient toll is:

$$y_p^U(t) = \left[\frac{\partial \Psi_p\left(t, x^S\right)}{\partial h_p} h_p^S\right]^+ \qquad \forall t \in \left[t_0, t_f\right].$$

Note that the partial derivative of $\Psi_p\left(t, x^S\right)$ with respect to the path flow h_p is not zero, since the state variable x is an implicit function of the control h_p as the relationship is expressed in the state dynamics. Further we cannot calculate the derivative directly due to the nested delay operator appears in $\Psi_p(\cdot,\cdot)$. However, from the numerical study of the dynamic system optimum traffic assignment, it is known that the controls are zero or singular. When the departure rate is nonzero, it as well as the states obtained from it are smooth and the delay operator is differentiable, although the derivative $\frac{\partial \Psi_p\left(t, x^S\right)}{\partial h_p}$ does not exist at the time moments where there are kinks in the controls. The derivative is numerically approximated as:

$$\frac{\partial \Psi_p\left[t, x\left(h^*, g^*\right)\right]}{\partial h_p} \cong \frac{\Psi_p\left[t, x\left(h + \delta, g\right)\right] - \Psi_p\left[t, x\left(h, g\right)\right]}{\delta} .$$

In our calculation, we have used $\delta = 10$. The resulting tolls at paths p_1 and p_2 are presented in Figures 2 and 3 for the first day. When, for path p_1, we compare the effective path delays (including tolls) with path flows (origin departure rates) by plotting both for three days, Figure 4 is obtained. This figure shows that departure rate peaks when the associated effective path delay achieves a local minimum, thereby demonstrating that a dynamic user equilibrium has been found. Similar comparisons are made for paths p_2 in Figure 5. The daily changes of travel demand from the origin to destination according to the difference equation (36) are given in Figure 6. The step size $\theta_k = 1/k$ to assist the convergence.

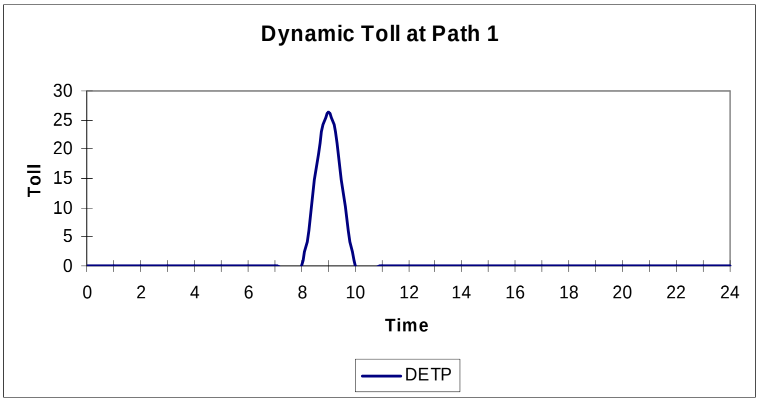

Figure 2. Dynamic Toll by DEPT for path p_1.

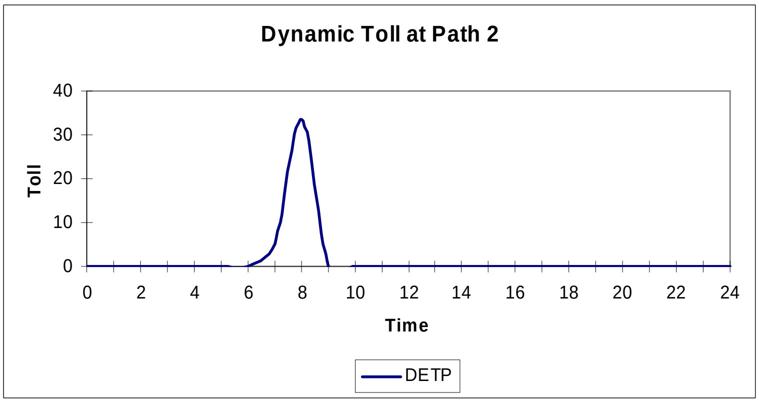

Figure 3. Dynamic Toll by DEPT for path p_2.

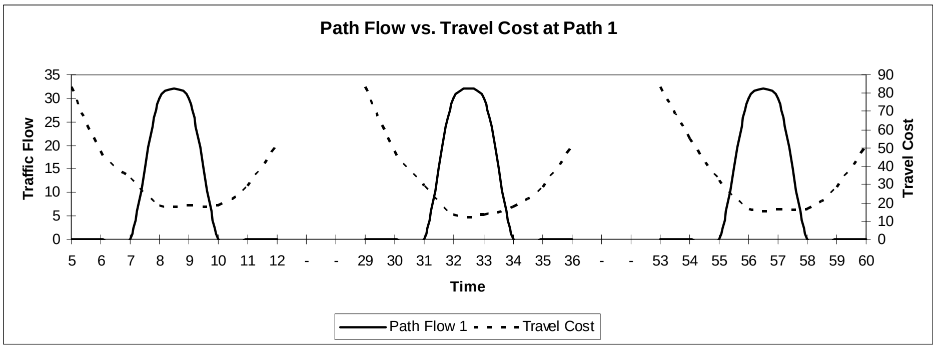

Figure 4. Comparison of path flow and associated unit travel costs for path p_1.

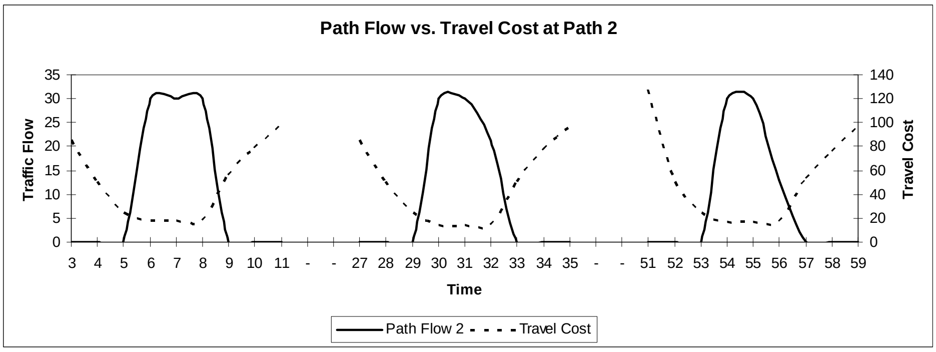

Figure 5. Comparison of path flow and associated unit travel costs for path p_2.

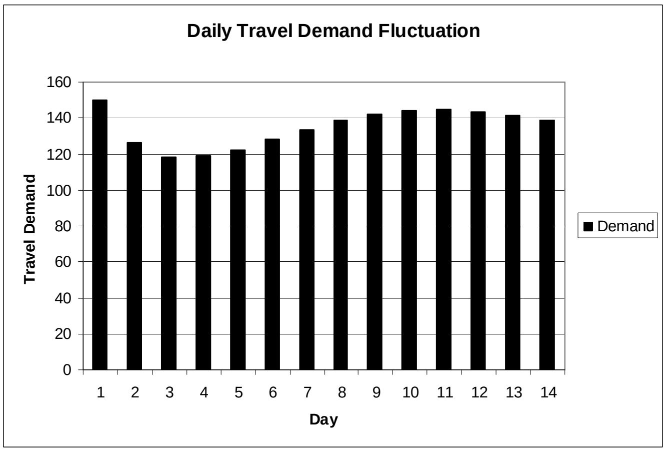

Figure 6. Daily changes of travel demand from the origin (node 1) to the destination (node 3)

CONCLUDING REMARKS

We have presented a mathematical formulation of the DETP and have shown how it may be directly solved using the notion of the implicit fixed point algorithm for a small illustrative problem. Clearly, in-depth testing and comparison of these solution methods is required before we measure the performance of the algorithm. We close by commenting that analytical DUE models — in our opinion — are far and away the best starting point for studies of the theoretical aspects of dynamic efficient tolls and dynamic congestion pricing. In particular, we have shown in this paper that an intuitive generalization to a dynamic setting of the efficient static toll rule is correct — something that could not be established in such a definitive way with a simulation model.

REFERENCES

Astarita, V (1995). Flow propagation description in dynamic network loading models. In: *Proceedings of the IV International Conference on Application of Advanced Technologies in Transportation Engineering* (Y J Stephanedes and F Filippi, eds.) 599-603.

Astarita, V (1996). A Continuous Time Link Based Model for Dynamic Network Loading Based on Travel Time Function. *Proceedings of the 13th International Symposium on Transportation and Traffic Theory*, (J-B Lesort, ed.), 79-102.

Ban, J, H Liu, M Ferris and B Ran (2005). A link based quasi-vi formulation and solution algorithm for dynamic user equilibria. In: *INFORMS 2005*, San Francisco, CA USA.

Beckmann, M, C B McGuire and C B Winsten (1956). *Studies in the Economics of Transportation*. Yale University Press.

Friesz, T L, D Bernstein and N Kydes (2002). Congestion pricing in disequilibrium. *Networks and Spatial Economics,* **4**, 181-202.

Friesz, T L, D Bernstein, Z Suo and R Tobin (2001). Dynamic network user equilibrium with state-dependent time lags. *Networks and Spatial Economics*, **1**, 319-347.

Friesz, T L, R Tobin, D Bernstein and Z Suo (1995). Proper propagation constraints which obviate exit functions in dynamic traffic assignment. *INFORMS Spring National Meeting*, Los Angeles, April 23-26.

Friesz, T L, D Bernstein, T Smith, R Tobin and B Wie (1993). A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research*, **41**, 179-191.

Friesz, T L, D Bernstein and R Stough (1996). Dynamic systems, variational inequalities and control theoretic models for predicting urban networks. *Transportation Science*, **30**(1), 14-31.

Friesz, T L and R Mookherjee (2006). Solving the dynamic network user equilibrium with state-dependent time shifts. *Transportation Research*, **40B**, 207-229.

Hearn, D W and M B Yildirim (2002). A Toll Pricing Framework for Traffic Assignment Problems with Elastic Demand, In: *Current Trends in Transportation and Network Analysis—Papers in Honor of Michael Florian* (M Gendreau and P Marcotte, eds.), 135-145. Kluwer Academic Publishers.

Holden, DJ (1989) Wardrop’s third principle: urban traffic congestion and traffic policy. *Journal of Transport Economics and Policy*, **23**, 239-262.

Liu, L N (2004). Multi-period congestion pricing models and efficient tolls in urban road. *Review of Network Economics*, **3**, 381-391.