USER EQUILIBRIUM, SYSTEM OPTIMUM, AND EXTERNALITIES IN TIME-DEPENDENT ROAD NETWORKS

Andy H. F. Chow

Centre for Transport Studies
University College London
Gower Street
London WC1E 6BT
United Kingdom

Phone: +44 (020) 7679 1577
Fax: +44 (020) 7679 1567

Email: andy@transport.ucl.ac.uk

Submission Dates: July 31, 2006; November 14, 2006

Word count: 5,132 words + 7 figures = 6,882 total words

ABSTRACT

This paper develops a comprehensive framework for analysing and calculating user equilibrium, system optimum, and externalities in time-dependent road networks. Under dynamic user equilibrium, traffic is assigned such that for each origin-destination pair in the network, the individual travel costs experienced by each traveller, no matter which combination of travel route and departure time he/she chooses, are equal and minimal. The system optimal flow is determined by solving a state-dependent optimal control problem, which assigns traffic such that the total system cost of the network system is minimized. The externalities are derived by using a novel sensitivity analysis. The analyses developed in this paper can work with general travel cost functions. Numerical examples are provided for illustration and discussion. Finally, some concluding remarks are given.

1. INTRODUCTION

Time-dependent network models have several advantages over the conventional time-independent ones. On representing the characteristics of the transport system, the time-dependent models consider traffic flows and travel times to be time-varying. On the travel demand side, the time-dependent traffic models capture the temporal variation in travel demand over time. Such network models provide important insight into the dynamics of peak periods and sensitivity of travellers' behaviour in response to different transport policy measures. In the network model, the travellers' behaviour is represented by a dynamic traffic assignment. The dynamic traffic assignment follows two principles: dynamic user equilibrium and dynamic system optimum. Dynamic user equilibrium assignment has been the focus in the past two decades. As a result of previous research (see for example *1*, *2*, *3*), we have gained substantial knowledge on the formulations, properties, and solution methods of dynamic user equilibrium assignment. Dynamic system optimal assignment is an important yet relatively underdeveloped area. Dynamic system optimal assignment process suggests that there is a central "system manager" to distribute network traffic over time in a fixed study period. Consequently, the total, rather than individual, travel cost of all travellers through the network is minimised. Although system optimal assignment is not a realistic representation of network traffic, it provides a bound on how we can make the best use of the road system, and as such it is a useful benchmark for evaluating various transport policy measures.

It is also noted that each additional traveller entering the system at a certain time imposes an additional travel cost on the others who enter the system at that time and thereafter. We regard this additional cost as dynamic externality. Understanding the nature of this dynamic externality is important in managing time-dependent networks. Carey and Srinvasan (*4*) and many others performed comprehensive analyses on system optimizing flow and dynamic externality. Most of these previous studies adopted an outflow traffic model which was later found to violate causality and unable to capture the flow propagation behaviour properly (see *5*, *6)*. This paper revisits the dynamic externality in a more general and plausible way. We further develop a novel sensitivity analysis and apply it to derive and analyse the externality from the dynamic network optimization formulation.

This paper is organized as follows. In Section 2, we introduce the travel cost functions adopted in this study and the formulation of dynamic user equilibrium assignment. In Section 3, we present the formulation and optimality conditions of dynamic system optimal assignment. Dynamic system optimal assignment problem is formulated as a state-dependent optimal control problem. To understand and solve the dynamic system optimality conditions, Section 4 provides a detailed interpretation of various cost components appear at system optimality. We further develop a novel sensitivity analysis to derive and compute the dynamic externality. Section 5 presents the solution algorithms for solving the sensitivity analysis and the dynamic traffic assignments. The solution algorithms are developed using a dynamic programming approach. Following this, we show some numerical calculations and discuss the characteristics of the results in Section 6. Finally, some concluding remarks are given in Section 7.

2. TRAVEL COST AND DYNAMIC USER EQUILIBRIUM ASSIGNMENT

2.1 Travel cost functions

We consider the total travel cost $C_p(s)$ encountered by each traveller departs at time *s* and travels along route *p* between an origin-destination pair in the network has three distinct components. The first component is the time spent on travelling along the route, which is determined by the travel time model that is adopted. In addition to the travel time, we add a time-specific cost $f[\tau_p(s)]$ associated with arrival time $\tau_p(s)$ through route *p* at the destination. We consider this arrival time-specific cost to be a non-decreasing function of the associated arrival time. Finally, we add a time-specific cost $h(s)$ associated with departure from the origin at time *s*. This cost explicitly considers the value of time to travellers at the origin of a journey. We consider that travellers would gain continuing benefit from remaining at their origin but are drawn to their destination by a need to attend and hence to travel. Following this, $h(s)$ is considered to be a monotonic non-increasing function of departure time *s*. Consequently, the total travel cost $C_p(s)$ associated with entry time to route *p* at time s is determined as a linear combination of these costs as

$$C_p(s) = h(s) + [\tau_p(s) - s] + f[\tau_p(s)], \tag{1}$$

in which the term $[\tau_p(s) - s]$ represents the travel time along the route which can be calculated from the cost of using each link on the route at the time it will be reached by following the vehicle trajectory (see *7*, *8*). In addition, following (*5*) and (*6*), we consider the travel time $\tilde{c}_a(s)$ at the time of entry *s* to each link *a* on a route to be a linear non-decreasing function of link traffic volume $x_a(s)$ as

$$\tilde{c}_a(s) = \phi_a + \frac{x_a(s)}{Q_a}, \tag{2}$$

where ϕ_a and Q_a denote the free flow travel time and the capacity of the travel link respectively. The reason of adopting $\tilde{c}_a(s)$ as a linear function is that the first-in-first-out (FIFO) queue discipline, which is a crucial property for analytical dynamic traffic models (*9*, *10*), cannot be guaranteed for non-linear version of it (*10*, *11*).

2.2 Dynamic user equilibrium assignment

In this study, travellers' responses are represented by their choices of routes of travel and times of departure. It is assumed that all travellers make their travel decisions according to a common criterion that their individual costs associated with the travel are minimized. Under such mechanism, the system will reach a stable state which is called dynamic user equilibrium. This dynamic user equilibrium condition is used as a representation of existing network traffic, where each individual traveller is acting only in their own interests, but not the interest of the whole system. Following (*12*), for an assignment to be in dynamic user equilibrium of simultaneous choice of travel route and departure time, the total travel cost should be the same for all travellers between each origin-destination pair in the network, no matter what combinations of departure-time and route that the travellers have chosen. The dynamic user equilibrium assignment is stated as a complementary inequality for the inflow $e_p(s)$ to each route p at the entry time s as:

$$e_p(s) \quad \begin{Bmatrix} > 0 & \Rightarrow & C_p(s) = C_{od}^* \\ = 0 & \Rightarrow & C_p(s) \geq C_{od}^* \end{Bmatrix} \quad \forall p \in P_{od}, \forall s, \tag{3}$$

where P_{od} is the set of all routes between origin-destination pair *od*, C_{od}^* is the total travel cost with which travel will take place between origin-destination pair *od*. All travel between each origin-destination pair is achieved at the same cost C_{od}^* throughout the study period.

3. DYNAMIC SYSTEM OPTIMAL ASSIGNMENT

In contrast with dynamic user equilibrium, dynamic system optimal assignment assumes that travellers will cooperate in making their travel choices for the overall benefit of the whole system instead of their own individual benefits. The system optimal assignment with departure time choice for fixed travel demand can be formulated as the following optimal control problem, which seeks an optimal inflow profile $e_a(s)$ that minimizes the total system travel cost within the study period, *T*. The total travel demand with in the study period is fixed and given by J_{od}. The optimal control problem is formulated as:

$$\min_{e_a(s)} Z = \sum_{\forall a} \int_0^T C_a(s) e_a(s) ds \tag{4}$$

subject to:

$$\frac{dG_a[\tau_a(s)]}{ds} = e_a(s) \quad ,\forall s, \forall a \tag{5}$$

$$\frac{dx_a(s)}{ds} = e_a(s) - g_a(s) \quad ,\forall s, \forall a \tag{6}$$

$$\frac{dE_a(s)}{ds} = e_a(s) \quad ,\forall s, \forall a \tag{7}$$

$$\sum_{\forall a} E_a(T) = J_{od} \tag{8}$$

$$e_a(s) \geq 0 \quad ,\forall s, \forall a \tag{9}$$

Equation (5) ensures the proper flow propagation along each route, in which $G_a(s)$ denotes the cumulative outflow by the link time $\tau_a(s)$, where $\tau_a(s) = s + \tilde{c}_a(s)$ for each link *a*. Equation (6) is the state equations that govern the evolution of link traffic, $x_a(s)$. The variables $g_a(s)$ represents the link outflow rate at time *s*. Equation (7) defines the cumulative inflow

$E_a(s)$. Equation (8) specifies the amount of total throughput J_{od} generated in the system within the time horizon *T*. Condition (9) ensures the non-negativity of the control variable. Given a positive inflow $e_a(s)$, the corresponding outflow $g_a(s)$ and link traffic volume $x_a(s)$ is guaranteed to be positive (see for example, *9*). Hence, we do not add explicit constraints to ensure the non-negativity of $g_a(s)$ and $x_a(s)$. The traffic models considered in this paper satisfy FIFO structurally, hence we do not need to add any explicit constraint for this. In the present study, the formulation and analysis for the dynamic system optimal assignment are restricted to networks in which origin-destination pairs are connected with mutually distinct travel routes consisting of one single link.

One technical difficulty is that with the traffic models above, the time lag between changes to the control variable, $e_a(s)$, and the corresponding responses, $g_a(s)$, is state-dependent. This state-dependent control theoretic formulation is unorthodox. Its properties and application to dynamic user equilibrium were studied by (*2*). The necessary conditions for dynamic system optimal assignment are given by the following the proposition.

Proposition 1: The necessary conditions for the optimization problem (4) – (9) can be derived as

$$e_a(s)\begin{cases} > 0 \Rightarrow C_a(s) + \Psi_a(s) + \lambda_a(s) - \gamma_a(s) = \mu_a(s) = \nu_{od} \\ = 0 \Rightarrow C_a(s) + \Psi_a(s) + \lambda_a(s) - \gamma_a(s) \geq \mu_a(s) = \nu_{od} \end{cases} ,\forall a, \forall s \in [0,T], \tag{10}$$

Proof:
See Appendix A in (*13*). □

In condition (10), the notation $\gamma_a(s)$, $\lambda_a(s)$, $\mu_a(s)$, and ν_{od} are the *multipliers*, or called the *costate variables* in optimal control terminology, associated with constraints (5), (6), (7), and (8) respectively. From the other stationarity conditions at optimality, we determine that $\gamma_a(s) = \lambda_a[\tau_a(s)]$ and $\mu_a(s) = \nu_{od}$, where ν_{od} is constant with respect to time. Hence, condition (10) can be rewritten as

$$e_a(s)\begin{cases} > 0 \Rightarrow C_a(s) + \Psi_a(s) + \lambda_a(s) - \lambda_a[\tau_a(s)] = \nu_{od} \\ = 0 \Rightarrow C_a(s) + \Psi_a(s) + \lambda_a(s) - \lambda_a[\tau_a(s)] \geq \nu_{od} \end{cases} ,\forall a, \forall s \in [0,T], \tag{11}$$

The magnitude of ν_{od} is dependent on the total travel demand, J_{od}. The detail of this can be referred to (*13*).

Similar to the static counterpart (see *14*), proposition 1 shows that the system optimal assignment in dynamic setting can be reduced to an equivalent *new* dynamic user equilibrium assignment formulation in which additional components of the cost $[\Psi_a(s) + \lambda_a(s) - \lambda_a[\tau_a(s)]]$. This quantity is interpreted as the total cost or total *toll* that each traveller would have to pay in addition to the individual travel cost that they encounter in order to make the optimal use of the transport system at this *modified* equilibrium. The additional components are further interpreted in Section 4.

4. COSTATE VARIABLES, EXTERNALITY, AND SENSITIVITY ANALYSIS OF TRAVEL COST

4.1 Costate variables

The costate variables $\lambda_a(s)$ and $\gamma_a(s)$ in the optimal control formulation represents the sensitivity of the value of the objective function with respect to the changes in the state variables $x_a(s)$ and $g_a(s)$ in the corresponding constraints at the associated time (see *15*). In other words, the value of the costate variables in the system optimal control formulation equals to the total change in the value of the total system travel cost Z with respect to slight changes in the state variables (i.e. link traffic volume $x_a(s)$ and outflow profile $g_a(s)$) at time *s*. The costate variables $\lambda_a(s)$ and $\gamma_a(s) = \lambda_a[\tau_a(s)]$ can be calculated by the following costate equation which is derived from the optimality conditions (see *13*) as

$$\lambda_a(s) = \frac{1}{Q_a}\int_{t=s}^{T}(1 + f'[\tau_a(t)])e_a(t)dt\,. \tag{12}$$

The difference between costate variables $\lambda_a(s)$ and $\gamma_a(s)$ can be calculated as

$$\lambda_a(s) - \gamma_a(s) = \lambda_a(s) - \lambda_a[\tau_a(s)] = \frac{1}{Q_a}\int_{t=s}^{\tau_a(t)}(1 + f'[\tau_a(t)])e_a(t)dt \qquad ,\forall a, \forall s\,, \tag{13}$$

which can be interpreted as the change in the value of the total system cost Z related to the change in link traffic volume during the stay of a traveller who enters the system at time *s* and leaves the system at $\tau_a(s)$.

4.2 Dynamic externality

The notation $\Psi_a(s)$ in the cost components in condition (11), where $\Psi_a(s) = \int_0^T \left.\frac{\partial C_a}{\partial u_s}\right|_t e_a(t)dt$, refers to the additional travel cost imposed by an additional amount of traffic, u_s, at time *s* to existing travellers in the system. This additional cost is termed as "dynamic externality". In the notation, we define the parameters u_s to represent a perturbation in inflow profile for which

$$\frac{de_a(t)}{du_s} = \begin{cases} 1 & \text{if } t \in [s, s+ds) \\ 0 & \text{otherwise} \end{cases}, \tag{14}$$

in which *ds* represents the incremental time step. Mathematically, the value of $\Psi_a(s)$ equals to the total change in the value of the total system travel cost Z with respect to this change in the inflow profile during a particular time interval $[s, s+ds)$.

4.3 Determining dynamic externality

Differentiating both sides of Equation (1) with respect to u_s gives

$$\left.\frac{\partial C_a}{\partial u_s}\right|_t = \left(1 + f'[\tau_a(t)]\right)\left.\frac{\partial \tau_a}{\partial u_s}\right|_t . \tag{15}$$

As a result, calculating the externality $\Psi_a(s)$ requires the sensitivity $\left.\frac{\partial \tau_a}{\partial u_s}\right|_t$ of travel time with respect to perturbations in link traffic inflow. The derivation of this derivative is given in the following proposition.

Proposition 2: Suppose there is a change of u_s in the link inflow rate at a particular time *s*, the sensitivity of the time of exit τ_a at a time *t* with respect to this perturbation can be calculated as

$$\left.\frac{\partial \tau_a}{\partial u_s}\right|_t = \frac{1}{Q_a}\left\{ \int_{\kappa=\sigma_a(t)}^{t} \frac{de_a(\kappa)}{du_s} d\kappa + g_a(t)\left.\frac{\partial \tau_a}{\partial u_s}\right|_{\sigma_a(t)} \right\}, \tag{16}$$

in which $\sigma_a(t)$ is the time of entry to the link that leads to exit at time *t*. Indeed, $\sigma_a(\cdot)$ is defined as the inverse function of $\tau_a(\cdot)$.

Proof:
See Appendix B in (*13*). ☐

The derivative of exit time with respect to the perturbation u_s in inflow can then be expressed in terms of the dependence of the inflow profile $e_a(s)$ in which *s* lies between *s* and $\sigma_a(s)$, the current outflow $g_a(s)$, and the derivative of exit time at the time of entry, $\sigma_a(s)$.

Discussion: In the dynamic system optimal condition (11), the cost components $C_a(s)$ and $\Psi_a(s)$ are generated within the system, while last two cost components (i.e. the costate variables, $\lambda_a(s)$ and $\gamma_a(s)$) are external to the system. The quantity in (13) is interpreted as the external cost to be imposed on a traveller who enters the link at time *s* and leaves at time $\tau_a(s)$ in addition to the individual travel cost $C_a(s)$ that he/she encounters himself/herself and the externality $\Psi_a(s)$ that he/she generates to the system.

5. SOLUTION ALGORITHMS

In this section, we first present the algorithm to solve dynamic user equilibrium assignment in Section 5.1. Solving dynamic system optimal assignment requires calculating the dynamic externality as noted in Section 4. Hence, we present the algorithmic procedure in Section 5.2. Finally, the solution algorithm for dynamic system optimal assignment is presented in Section 5.3 by exploiting the externality.

5.1 Solving dynamic user equilibrium assignment

Step 0: Initialisation

0.1 Choose an initial equilibrium cost C_{od}^{*};

0.2 Set the overall iteration counter $n := 1$;

0.3 Set $e_a(k) := 0$ for all links *a* and all times *k*, $k \in [0, K]$. The notation $e_a(k)$ represents the assigned inflow to link *a* between times $k\Delta s$ and $(k+1)\Delta s$. The total number of simulated time steps is denoted as $K = T / \Delta s$ and the total number of parallel links is denoted by *A*; set time index $k := 0$;

0.4 Set the link index $a := 1$;

0.5 Set the time index $k := 0$;

0.6 Set the overall iteration counter $n^{i} := 1$.

Step 1: Network loading

Find $\tau_a(k+1)$ by loading the travel link using the inflow $e_a(k)$ at the current iteration. The network loading algorithm "Algorithm D2" described in (*10*) was adopted for this purpose.

Step 2: Updating the inflow

2.1 Calculate $C_a(k+1) = h(k+1) + [\tau_a(k+1) - (k+1)] + f[\tau_a(k+1)]$;

2.2 Calculate $\Omega = \dfrac{C_a(k+1) - C_a(k)}{\Delta s}$ and $\Omega' = \dfrac{\partial \Omega}{\partial e_a(k)} = (1 + f'[\tau_a(k+1)])\dfrac{1}{Q_a}$,

in which $f'[\tau_a(k)] \approx \dfrac{f[\tau_a(k+1)] - f[\tau_a(k)]}{\tau_a(k+1) - \tau_a(k)}$ using a finite difference approximation.

We note the equilibrium is achieved if and only if $\Omega = 0$ for all positive inflow $e_a(k)$;

2.3 Update the inflow as $e_a(k) := \max[(e_a(k) + \pi d), 0]$ using Newton's method. The second-order searching direction is denoted by $d = -\Omega / \Omega'$ and the step size π, which is interpolated linearly as

$$\pi = \frac{C_{od}^{*} - C_a^{0}(k+1)}{C_a^{1}(k+1) - C_a^{0}(k+1)},$$

where $C_a^{1}(k+1)$ and $C_a^{0}(k+1)$ represent the corresponding values of $C_a(k+1)$ when $e_a^{*}(k)$ is being updated with π is taken as 1 and 0 respectively. To determine π, two network loadings are required to calculate the values of $C_a^{1}(k+1)$ and $C_a^{0}(k+1)$ respectively.

Step 3: Stopping criteria

3.1. Check if $\left|C_a(k+1)-C_{od}^*\right|\le\varepsilon$ or n^i is greater than the predefined maximum number of inner iterations, then go to step 3.2; otherwise, set $n^i := n^i + 1$ and go to step 1;

3.2. If $k = K$, then go to step 3.3; otherwise $k := k + 1$ and go to step 1;

3.3. If $a = A$, then go to step 3.4; otherwise $a := a + 1$ and go to step 0.5;

3.4 Define $\xi = \dfrac{\sum_{k\in K}\sum_{a\in A} e_a(k)\left|C_a(k+1)-C_{od}^*\right|}{\sum_{k\in K}\sum_{a\in A} e_a(k)C_{od}^*}$ as a measure of disequilibrium, which is equal to zero at user equilibrium. If n is greater than the predefined maximum number of overall iterations or ξ is sufficiently small, i.e. $\xi\le\varepsilon$ where ε is a test value, then go to step 3.5; otherwise set $n := n+1$ and go to step 1.2;

3.5. check if the total throughput $E_{od} = \sum_{\forall a}\sum_{\forall k} e_a(k)$ from the system is equal to the predefined total demand J_{od} for the *o-d* pair. If yes, then *terminate* the algorithm; otherwise update $C_{od}^* := C_{od}^* + \left[\dfrac{J_{od}-E_{od}}{\frac{dE_{od}}{dC^*}}\right]$, and go back to step 0.3. The derivative $\dfrac{dE_{od}}{dC_{od}^*}$ is given by (*16*).

5.2 Calculating externality $\Psi_a(s)$

Step 1: Initialisation for calculating the derivatives of link exit time

1.1 Set the link index $a := 1$;

1.2 Set the time index $k := 0$, which represents the time when the inflow is perturbed;

1.3 Set the time index $\omega := 0$ to index the change in exit time due to the perturbation in inflow at time k;

1.4: Calculating the derivatives of link exit time:

If $\omega < k$, then $\left.\dfrac{d\tau_a}{du_k}\right|_{\omega} := 0$;

else if $k\le\omega\le\lceil\tau_a(k)\rceil$, then $\left.\dfrac{d\tau_a}{du_k}\right|_{\omega} := \dfrac{1}{Q_a}$;

else $\left.\dfrac{\partial\tau_a}{\partial u_k}\right|_{\omega} := \dfrac{g_a(\omega)}{Q_a}\left.\dfrac{\partial\tau_a}{\partial u_k}\right|_{\sigma_a(\omega)}$;

1.5 If $\omega = K$, then go to step 1.6; otherwise $\omega := \omega + 1$ and go to step 1.4;

1.6 If $k = K$, then go to step 1.7; otherwise $k := k + 1$ and go to step 1.3;

1.7 If $a = A$, then go to step 2; otherwise $a := a + 1$ and go to step 1.2.

Step 2: Calculating the derivatives of total travel cost function

2.1 Set the link index $a := 1$;

2.2 Set the time index $k := 0$;

2.3 Set the time index $\omega := 0$;

2.4 Calculate $\left.\dfrac{dC_a}{du_k}\right|_{\omega} = \left(1+f'[\tau_a(\omega)]\right)\left.\dfrac{d\tau_a}{du_k}\right|_{\omega}$;
2.5 If $\omega = K$, then go to step 2.6; otherwise $\omega := \omega + 1$ and go to step 2.4;
2.6 If $k = K$, then go to step 2.7; otherwise k:= k + 1 and go to step 2.3;
2.7 If $a = A$, then go to step 3; otherwise a:= a + 1 and go to step 2.2.

Step 3: Calculating the externality
3.1 Set the link index $a := 1$;
3.2 Set the time index $k := 0$;
3.3 Initialise $\Psi_a(k) := 0$;
3.4 Set the time index $\omega := 0$;
3.5 Calculate $\Psi_a(k) = \Psi_a(k) + e_a(\omega)\left.\dfrac{dC_a}{du_k}\right|_{\omega}$;
3.6 If $\omega = K$, then go to step 3.7; otherwise $\omega := \omega + 1$ and go to step 3.5;
3.7 If $k = K$, then go to step 3.8; otherwise k:= k + 1 and go to step 3.3;
3.8 If $a = A$, then stop; otherwise a:= a + 1 and go to step 3.2.

Note: In Algorithm 2, step 1.4, we note that the function $\sigma_a(\omega)$ does not necessarily give an integral value. To implement the sensitivity analysis into computer, a interpolation is needed to determine the value of $\left.\dfrac{d\tau_a}{du_k}\right|_{\sigma_a(\omega)}$. This study adopts a linear interpolation which approximates the value of $\left.\dfrac{d\tau_a}{du_k}\right|_{\sigma_a(\omega)}$ as

$$\left.\frac{d\tau_a}{du_k}\right|_{\sigma_a(\omega)} \approx \left.\frac{d\tau_a}{du_k}\right|_{\lfloor\sigma_a(\omega)\rfloor} + \left(\left.\frac{d\tau_a}{du_k}\right|_{\lceil\sigma_a(\omega)\rceil} - \left.\frac{d\tau_a}{du_k}\right|_{\lfloor\sigma_a(\omega)\rfloor}\right)\left(\sigma_a(\omega) - \lfloor\sigma_a(\omega)\rfloor\right), \tag{17}$$

where the notation $\lceil\sigma_a(\omega)\rceil$ represent the smallest integer not smaller than $\sigma_a(\omega)$, and $\lfloor\sigma_a(\omega)\rfloor$ is the greatest integer not larger than $\sigma_a(\omega)$.

5.3 Solving dynamic system optimal assignment

Step 0: Initialisation
0.1 Choose an initial equilibrium cost C^*_{od} ;
0.2 Set the overall iteration counter $n := 1$;
0.3 Set $e_a(k) := 0$ for all links a and all times k, $k \in [0, K]$. The notation $e_a(k)$ represents the assigned inflow to link a between times $k\Delta s$ and $(k+1)\Delta s$. The total number of simulated time steps is denoted as $K = T/\Delta s$ and the total number of parallel links is denoted by A; set time index $k := 0$;

0.4 Set costates $\lambda_a(k) := 0$ for all times $k \in [0, K]$;
0.5 Set the link index $a := 1$;
0.6 Set the time index $k := 0$;
0.7 Set the overall iteration counter $n^i := 1$.

Step 1: Network loading
Find $\tau_a(k+1)$ by loading the travel link using the inflow $e_a(k)$ at the current iteration. The network loading algorithm "Algorithm D2" described in (*10*) was adopted for this purpose.

Step 2: Calculating externality
Use **Algorithm 2** to calculate the externality $\Psi_a(k)$ associated with each $e_a(k)$.

Step 3: Determining the auxiliary inflow
3.1 Calculate

$$C_a(k+1) = h(k+1) + [\tau_a(k+1) - (k+1)] + f[\tau_a(k+1)] + \Psi_a(k+1) + \lambda_a(k) - \lambda_a[\tau_a(k)];$$

3.2 Calculate $\Omega(k) = \dfrac{C_a(k+1) - C_a(k)}{\Delta s}$ and $\Omega'(k) = \dfrac{\partial \Omega(k)}{\partial e_a(k)} = (1 + f'[\tau_a(k+1)])\dfrac{1}{Q_a}$;

3.3 Calculate the auxiliary inflow $d_a(k) = -\Omega_a(k) \big/ \Omega'_a(k)$;

3.4. If $a = A$, then go to step 3.5; otherwise a:= a + 1 and go to step 0.7;
3.5. If $k = K$, then go to step 4; otherwise k:= k + 1 and go to step 0.6.

Step 4: Determining step size
Search for an optimal step size θ by using golden section method and update the inflow as $e_a(k) := \max[e_a(k) + \theta d, 0]$ for all times s such that the total system cost is minimized.

Step 5: Calculating costate variables
5.1 Set the link index $a := 1$;
5.2 Set $\lambda_a(K) = 0$;
5.3 Set the time index $k := K - 1$;
5.4 Compute $\lambda_a(k) = \lambda_a(k+1) + (1 + f'[\tau_a(k)])\dfrac{e_a(k)}{Q_a}\Delta s$;
5.5 Calculate $\lambda_a[\tau_a(k)]$ from $\lambda_a(k)$ and $\tau_a(k)$ using linear interpolation as

$$\lambda_a[\tau_a(k)] \approx \lambda_a\big|_{\lfloor \tau_a(k) \rfloor} + \left(\lambda_a\big|_{\lceil \tau_a(k) \rceil} - \lambda_a\big|_{\lfloor \tau_a(k) \rfloor}\right)(\tau_a(k) - \lfloor \tau_a(k) \rfloor);$$

5.6. If $k = 0$, then go to step 6.7; otherwise k:= k - 1 and go to step 5.2;
5.7. If $a = A$, then go to step 7; otherwise a:= a + 1 and go to step 5.1.

Step 6: Stopping criteria

6.1 Define $\xi = \dfrac{\sum_{\forall a}\sum_{\forall k} e_a(k)\left|C_a(k+1) - C_{od}^*\right|}{\sum_{\forall a}\sum_{\forall k} e_a(k) C_{od}^*}$ as a measure of disequilibrium, which is equal to zero at system optimum. If n is greater than the predefined maximum number of overall iterations or ξ is sufficiently small, i.e. $\xi \le \varepsilon$ where ε is a test value, then go to Step 6.2; otherwise set n:=n+1 and go to step 0.5;

6.2. Check if the total throughput $E_{od} = \sum_{\forall a}\sum_{\forall k} e_a(k)$ from the system is equal to the predefined total demand J_{od} for the *o-d* pair. If yes, then *terminate* the algorithm; otherwise update $C_{od}^* := C_{od}^* + \left[\dfrac{J_{od} - E_{od}}{\dfrac{dE_{od}}{dC^*}}\right]$, and go back to step 0.2.

6. NUMERICAL EXAMPLES

6.1 Sensitivity analysis of link exit time

The critical step of determining the externality $\Psi_a(s)$ is calculating the derivative of link exit time $\left.\dfrac{\partial \tau_a}{\partial u_s}\right|_t$ with respect to perturbation in inflow. Hence, this section tests the accuracy of this derivative which is derived in proposition 2. We consider a single link, which has a free flow time 3 mins and a capacity 20 vehs/min, connecting a single origin-destination pair. The size of discretized time interval Δs is taken as 1 min. A parabolic inflow as specified in (18) is loaded into the travel link.

$$e_a(s) = \begin{cases} \dfrac{1}{8}(40 - s)s & \text{if } 0 \le s \le 40 \\ 0 & \text{otherwise} \end{cases}. \tag{18}$$

This profile has a peak inflow rate of 50 vehs/min, which equals to 2.5 times of the link capacity. To investigate the accuracy of the sensitivity analysis, we consider this parabolic inflow is increased by a unit of flow during time interval 1, and the associated derivative of link exit time are plotted in Figure 3. Each value of "derivatives" on the vertical axis over time represents the change in link exit time at that time due to the perturbation in inflow at time 1. The "analytical" derivatives are calculated according to Equation (16). The "numerical" derivatives are determined by using direct numerical finite difference method, and they are plotted in the same figure for comparison. To calculate the finite difference, one extra unit of inflow is added at time 1, while the inflow profile remains unchanged at other times. The "numerical" variations in travel times are then calculated by subtracting the link travel time loaded by the original inflow profile from that loaded by the perturbed inflow profile. The result shows that the analytical variations given by Equation (16) can represent the true numerical variations in travel time reasonably well. The value of the derivatives at each time represents the change in the link travel time at that time due to the perturbation in inflow at time 1. Both numerical and analytical variations drop to zero at time 83 when all traffic is cleared from the link.

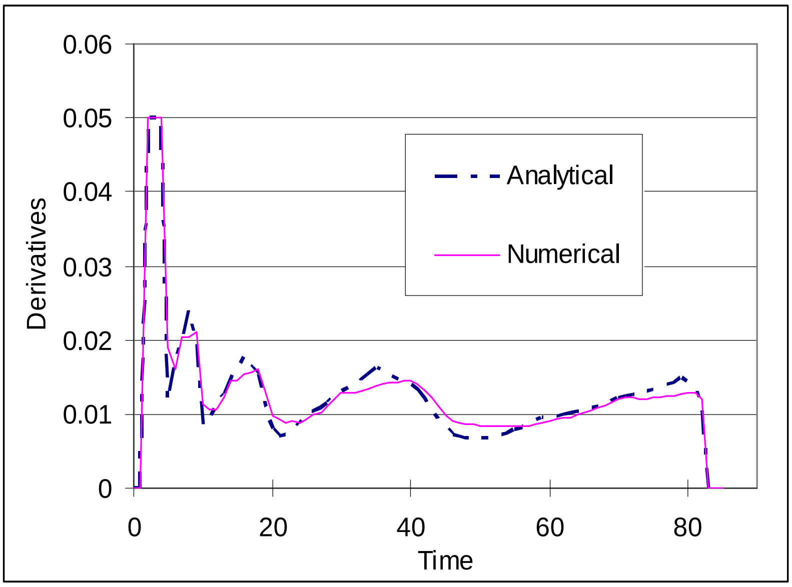

FIGURE 1 Sensitivity of travel time with respect to a perturbation in inflow

6.2 Dynamic traffic assignments

This section calculates the dynamic traffic assignments. Figure 2 shows a network with a single origin-destination pair connected with two parallel travel routes consisting of one single link. Link 1 has free flow time 3 mins and capacity 20 vehs/min, and link 2 has free flow time 4 mins and capacity 30 vehs/min. Furthermore, the origin-specific cost is specified to be a monotone linear function of time with a slope -0.4. The destination cost function is piecewise linear, with no penalty for arrivals before the preferred arrival time $t^* = 50$, and increases with a rate 2 afterwards. The length of the planning horizon $[0, T]$, where *T*=100, is set such that that all traffic can be cleared by time *T*. The total amount of traffic J_{od} is taken as 800 vehs.

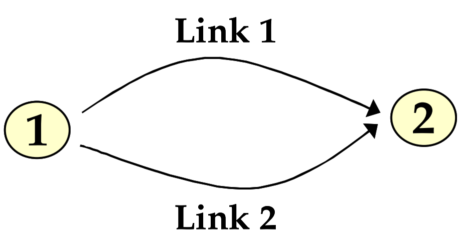

FIGURE 2 Example network

Figure 3 shows the corresponding profiles of link inflows and the total travel cost at dynamic user equilibrium. The assignment period to route 1 is from time 18 to time 49, and to route 2 is from time 21 to 49 which is shorter due to its higher capacity and hence traffic can be cleared more efficiently. The link flow volumes using route 1 and route 2 are 380.25 (vehs) and 419.75 (vehs) respectively. Figure 3 also shows good equilibration of travel cost in which the measure of disequilibrium ξ is less than 10^{-17}.

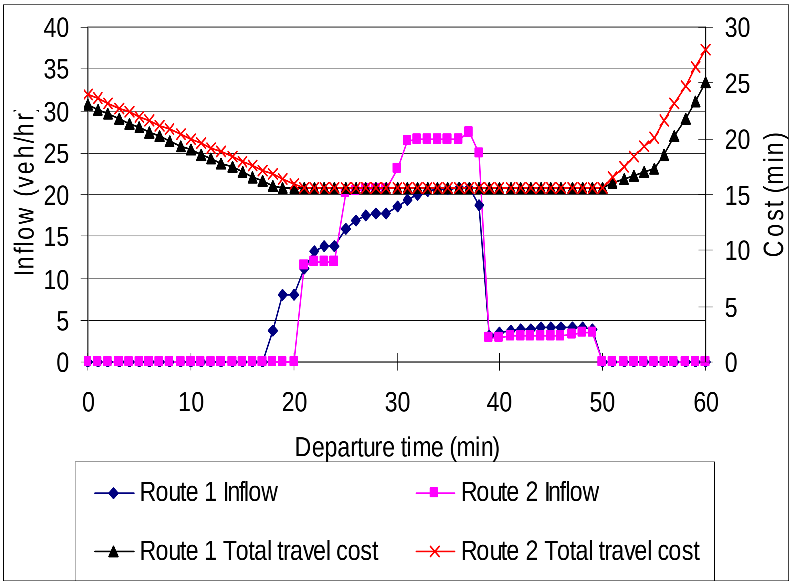

FIGURE 3 Dynamic user equilibrium assignment

As mentioned in Section 2.1, the total travel cost indeed consists of the travel time and the sum of time-specific costs. For further illustration, Figure 4 plots the cost components and the associated inflow and outflow profiles on route 1. The cost components on route 2 follows similar pattern and hence they are not included here for brevity. Figure 4 shows that the time-specific costs are decreasing over time until departure time 39. It is because travellers depart after time 39 will arrive at the destination at time 50.1 which is after the preferred arrival time 50. As a result, those travellers will be added a positive arrival specific cost. In addition, the figure shows that the link travel time increases with time when the inflow is higher than the outflow and vice versa after time 39. Overall, the sum of all these travel cost components is constant over time.

The total system travel cost Z is 12,465.2 veh-min at dynamic user equilibrium and it is understood that it is not the minimum yet. Figure 5 shows dynamic system optimal assignment. With the same total demand J_{od}, the period of assignment to link 1 expands from times [18, 49] to times [4, 56], while that to link 2 expands from times [21, 49] to times [6, 50]. In general, the inflow profiles are more spread at system optimum in order to reduce the intensity of congestion. In the figure, the legend "total travel cost" refers to values of $C_a(s)$ and the legend "total travel cost + toll" refers to the value of $C_a(s)+\Psi_a(s)+\lambda_a(s)-\lambda_a\left[\tau_a(s)\right]$ on each route. The associated total system travel cost Z at system optimum is decreased from 12,465.2 veh-min in user equilibrium to 11,447.3 veh-min. Contrast with dynamic user equilibrium, it is seen that the total travel cost $C_a(s)$ is not equal for all departure time s which implies some travellers can be better off while some of them have to be worse off for the good of the whole system.

Finally, as shown in the figure, the "total travel cost + toll" is not in good equilibration at system optimum in which the measure of disequilibrium can only reach 0.04. Indeed, solving the dynamic system optimal assignment is difficult, since the solution procedure involves solving two dynamic programmes simultaneously and consistently: solving the network loading forward in time for the state variables and solving the costate equations backward in time for the costate variables. Although the dynamic system optimal solution that we achieved shows a reduction of more than 8% over the dynamic user equilibrium assignment, we are still exploring a better algorithmic procedure for better quality solution.

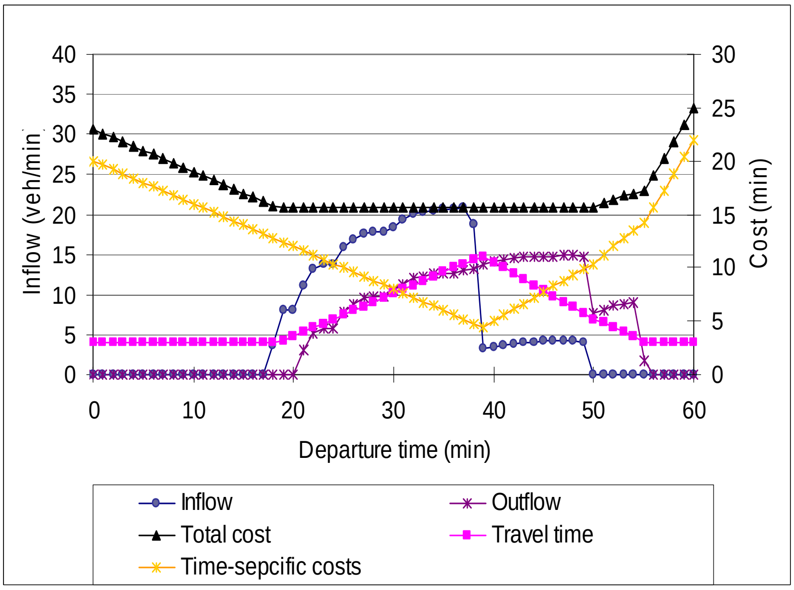

FIGURE 4 Inflow and travel cost components on route 1

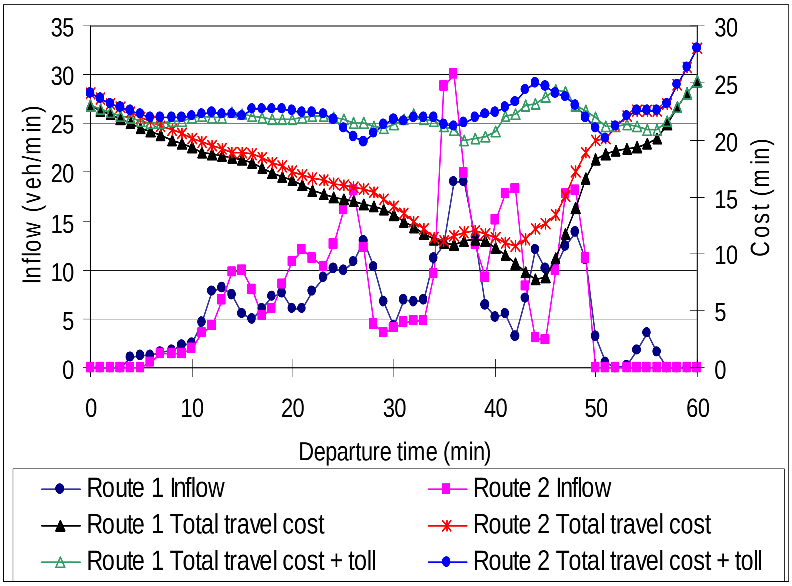

FIGURE 5 Dynamic system optimal assignment

To illustrate the cause of the decrease in total system travel cost, Figure 6 shows the link traffic volumes, which are directly related to congestion, at dynamic user equilibrium and dynamic system optimum. Interestingly, yet importantly, the results show that, with the link travel time function in Equation (2), the system optimal assignment has to allow congestion, which can only be managed and minimized, even at system optimum. This implies that the previous analyses on dynamic system optimum using bottleneck model (see for example *17*, *18*) in which congestion can be completely eliminated do not generally apply.

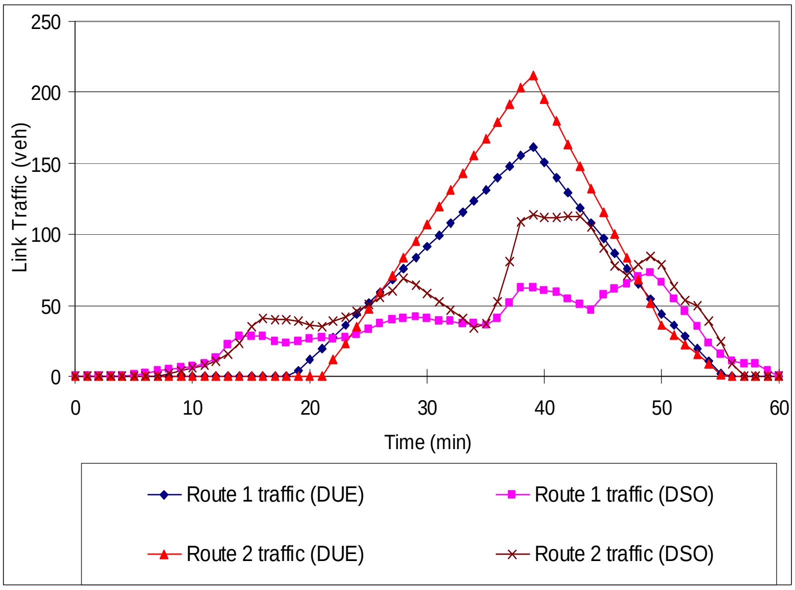

FIGURE 6 Link traffic volumes

Finally, to decentralize the system optimizing flow, a total amount of time-varying toll of $[\Psi_a(s)+\lambda_a(s)-\lambda_a[\tau_a(s)]]$ has to impose on each traveller in the system according to the departure time s of each traveller. The time-varying tolls on each route are calculated and plotted in Figure 7. In general, the tolls increase for travellers whose departure time would lead to an early arrival at the destination; decreases for travellers who would arrive late.

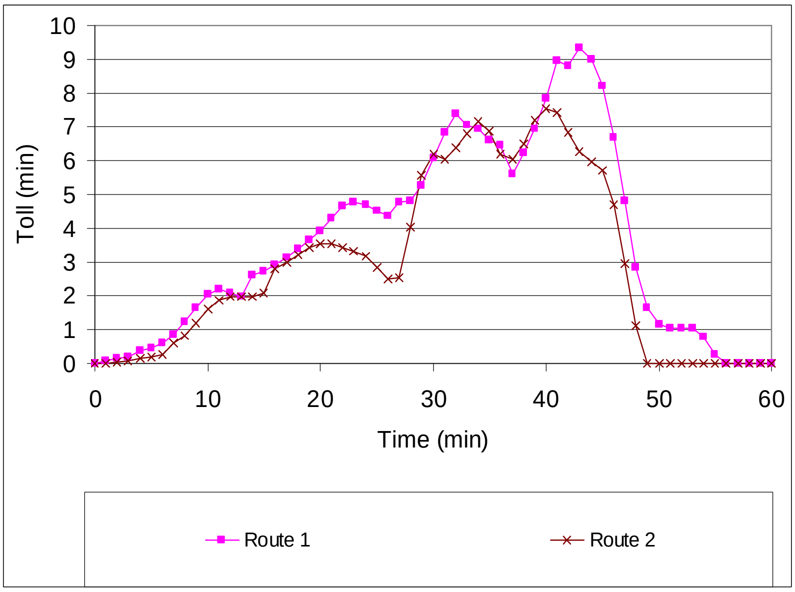

FIGURE 7 Tolls

7. CONCLUDING REMARKS

This paper developed a comprehensive framework for dynamic user equilibrium, system optimum, and dynamic externalities. In particular, we revisited the dynamic system optimal assignment and externality in a more general and plausible way. We also developed a novel sensitivity analysis of link exit time with respect to perturbations in inflow. The knowledge generated in this paper provides important insight into the management of peak traffic dynamics and travellers' behaviour. We also presented solution algorithms for implementing the sensitivity analysis and solving the dynamic traffic assignments. The solution algorithms were developed using a dynamic programming approach. We applied the algorithms to numerical calculations. The characteristics of the results were discussed. We showed that the system optimal assignment has to allow congestion which can only be managed but not eliminated even at system optimum. This implies that the previous analyses on dynamic system optimum using the bottleneck model do not apply generally. Nevertheless, further study is still required to improve the performance of the solution algorithm for calculating system optimal assignment.

In the present study, the formulation and analysis presented were restricted to networks with multiple origin-destination pairs connected with mutually distinct routes consisting of single links. In case of networks with multiple origin-destination pairs with overlapping routes, traffic entering the network during the journey time of a traveller from other origins downstream can influence the travel time of travellers from its upstream. As a result, some special computational technique, for example Guass-Seidel relaxation (see *14*, *19*), is required. The basic idea of such relaxation scheme is to decompose the assignment problem for networks with overlapping routes connecting multiple origin-destination pairs into several sub-problems. In each sub-problem, we calculate the assignments for one origin-destination pair, and temporarily neglect the influences from the flows between other origin-destination pairs. When equilibrium or system optimum is reached for the current origin-destination pair, we proceed with calculations for another pair. The procedure is repeated until equilibrium or system optimum is reached in the whole network. The relaxation scheme is not guaranteed to converge, but if it does, the solution will be the final assignment pattern (see *14*, p217). In case of travel route with multiple links, difficulties brought in when we have to calculate the derivatives of *route* exit time (see for example *20*). As shown earlier in proposition 2, changing the inflow to a link on the route during one time interval will induce perturbations in the link travel time, the link outflow, and hence the inflow to subsequent link(s) in several succeeding time intervals. Hence, the dimension of time intervals to be considered in calculating the derivatives will expand exponentially along the route. We are currently investigating the strategies to cope with this "curse of dimensionality". Efficient computing methods for system optimal assignments in general networks are still under investigation, however, the work reported in the present paper provide a solid and necessary foundation for future research on this.

ACKNOWLEDGEMENTS

I would like to thank Professor Benjamin Heydecker and Dr. JD Addison for their continuing encouragement and supervision of this study. I am also grateful for the helpful comments from the anonymous referees.

REFERENCES

1. Friesz, T.L., D. Bernstein, T.E. Smith, R.L. Tobin, and B.W. Wie. A variational inequality formulation of the dynamic network user equilibrium problem. *Operations Research*, Vol. 41(1), 1993, pp. 179-191.

2. Friesz, T.L., D. Bernstein, Z. Suo, and R.L. Tobin. Dynamic network user equilibrium with state-dependent time lags. *Network and Spatial Economics*, Vol. 1(3/4), 2001, pp. 319-347.

3. Heydecker, B.G. and J.D. Addison. Analysis of dynamic traffic equilibrium with departure time choice. *Transportation Science,* Vol. 39(1), 2005, pp. 39–57.

4. Carey M. and A. Srinivasan. Externalities, average and marginal costs, and tolls on congested networks with time-varying flows. *Operations Research*, Vol. 41, 1993, pp. 217-231.

5. Daganzo, C.F. Properties of link travel time functions under dynamic loads. *Transportation Research Part B*, Vol. 29B(2), 1995, pp. 95-98.

6. Mun, J.S. A divided linear travel time model for dynamic traffic assignment. *Proceedings of 9th World Conference on Transport Research*, Seoul, S.Korea, 2001.

7. Huang, H.J. and W. Lam (2002). Modeling and solving the dynamic user equilibrium route and departure time choice problem in network with queues. *Transportation Research Part B*, Vol. 36, 2002, pp. 253-273.

8. Heydecker, B.G. and J.D. Addison. Analysis of dynamic traffic assignment. *Proceedings of the 1st international symposium on dynamic traffic assignment*, Leeds, UK, 2006.

9. Carey, M. Non-convexity of the dynamic traffic assignment problem. *Transportation Research Part B*, Vol. 26B(2), 1992, pp. 127-33.

10. Nie, X. and H.M. Zhang. Delay-function-based link models: their properties and computational issues. *Transportation Research Part B,* Vol. 39B(8), 2005, pp. 729-752.

11. Astarita, A. A continuous time link model for dynamic network loading based on travel time functions. J.B. Lesort, ed. *Transportation and Traffic Theory*. Pergamon, Oxford, 1996, pp. 79-102.

12. Hendrickson, C. and G. Kocur. Schedule delay and departure time decisions in a deterministic model. *Transportation Science,* Vol. 15, 1981, pp. 62-77.

13. Chow, A.H.F. Analysis of dynamic system optimum and externalities with departure time choice. M.G.H Bell, B.G. Heydecker and R.E. Allsop, eds. *Transportation and Traffic Theory,* 2007, *In press.*

14. Sheffi, Y. *Urban transportation networks: equilibrium analysis with mathematical programming methods*. Englewood Cliffs, NJ: Prentice-Hall, 1985.

15. Dorfman, R. An Economic Interpretation of Optimal Control Theory, *The American Economics Review*, Vol. 59, 1969, pp. 817-831.

16. Heydecker, B.G. Dynamic equilibrium network design. MAP Taylor, ed. *Transportation and Traffic Theory*. Pergamon, Oxford, 2002, pp. 349-370.

17. Vickrey, W.S. Congestion theory and transport investment. *American Economics Review*, Vol. 72(3), 1969, pp. 467-479.

18. Arnott, R., A. de Palma, and R. Lindsey. Recent developments in the bottleneck model. K.J. Button and E.T. Verhoef, eds. *Road Pricing, Traffic Congestion and the Environment: Issues of Efficiency and Social Feasibility*. Edward Elgar, Aldershot, Hampshire, U.K., 1998, pp. 79-110.

19. Patriksson, M. The traffic assignment problem: models and methods. VSP, The Netherlands, 1994.

20. Balijepalli, N.C. and D. P. Watling. Doubly Dynamic Equilibrium Distribution Approximation Model for Dynamic Traffic Assignment. H. Mahmassani, ed. *Transportation and Traffic Theory: Flow, Dynamics and Human Interaction.* Elsevier, Oxford, U.K., 2005, pp 741-760.