Available online at www.sciencedirect.com
SCIENCE DIRECT®

TRANSPORTATION RESEARCH PART B

Transportation Research Part B 39 (2005) 659–678

www.elsevier.com/locate/trb

A first best toll pricing framework for variable demand traffic assignment problems

Mehmet Bayram Yildirim [a,*], Donald W. Hearn [b]

[a] Industrial and Manufacturing Engineering Department, Wichita State University, Wichita, KS 67260, United States
[b] Center for Applied Optimization, Industrial and Systems Engineering Department, University of Florida, United States

Received 21 June 2002; received in revised form 15 March 2004; accepted 10 August 2004

Abstract

In this paper, we present a toll pricing framework for a general variable demand traffic assignment problem with side constraints, where the demand between an origin destination pair is a function of the least total travel cost for making the trip. This general demand model unifies earlier toll pricing treatments of the variable demand models including elastic demand traffic assignment problems and combined distribution assignment problems. All of these models have the constant toll revenue property. Given that users experience the side constraints, we show that when they are charged by a toll vector in the first best toll set, the system optimal flows and demands are achieved. We then present a toll pricing framework by which a traffic planner might find the most appropriate toll vector given certain restrictions and objectives on the network. Finally, we derive the toll sets and illustrate the toll pricing framework for specific instances of the general variable demand models.

Keywords: Congestion toll pricing; Combined distribution assignment models; Elastic demand traffic assignment problem

* Corresponding author. Tel.: +1 316 978 3426.
E-mail addresses: bayram.yildirim@wichita.edu (M.B. Yildirim), hearn@ise.uf1.edu (D.W. Hearn).

0191-2615/$ - see front matter
doi:10.1016/j.trb.2004.08.001

1. Introduction

Congestion is becoming an inevitable part of everyday life in most metropolitan areas all over the world. Increasing population and wealth result in more automobiles than current transportation networks can handle. Due to limited expansion possibilities of the transportation network, congestion has increased drastically over the last decade. Arnott and Small (1994) estimate that the annual cost of driving in congested areas in 39 metropolitan areas of the US is around \$48 billion or \$640 per driver. Japan's international co-operation agency calculated that Bangkok loses one-third of its potential output due to congestion (The Economist, 1998). In 1995, 75% of San Francisco's, 66.5% of Los Angeles's, 63.8% of San Bernardino and Riverside's and 60% of San Jose's rush hour traffic was under congested conditions (Schiller, 1998).

Traffic planners often charge users in order to utilize the system resources more efficiently and restrain the number of travellers on the transportation network, based on the time, distance and congestion level. Economic theory argues that to achieve economic efficiency in a market, the price of a good or a service should be at its full cost to society, but this is not generally the case in transportation networks. Pigou (1920), Armstrong-Wright (1986), Beckmann et al. (1956), Elliot (1986), Johnson (1964), Luk and Chung (1997) and Arnott and Small (1994) recommend the marginal social cost pricing (MSCP) tolls that are equal to the negative externalities imposed on other users (such as cost of congestion, travel delays, air pollution, and accidents) in order to have an efficient utilization of the transportation system. The MSCP tolls are easy to compute by a formula which prices the time value of an additional user on the system. This makes MSCP tolls one of the most popular tools for road pricing applications. Economists define the MSCP tolls as the "first best" tolls since they achieve the optimal utilization of the transportation system by changing the user behavior to system optimal behavior. We extend the definition for the first best pricing by also including all toll vectors which achieve the most efficient utilization of the transportation system. We define the set of all such vectors as the *First Best Toll Set*. Bergendorff (1995), Bergendorff et al. (1997) and Hearn and Ramana (1998) show that there exist toll vectors other than the MSCP toll vector in the first best toll set for the fixed demand traffic assignment problems. A similar result for the elastic demand traffic assignment problems is given by Hearn and Yildirim (2002). Hearn and Ramana (1998) define the procedure for finding alternative toll vectors as the *Toll Pricing Framework*. For the fixed demand traffic assignment problems, Bergendorff (1995), Hearn and Ramana (1998) and Hearn et al. (2001) show that cheaper and more implementable toll vectors compared to MSCP tolls can be found among such solutions. Hearn and Yildirim (2002) and Larsson and Patriksson (1998) show that elastic demand traffic assignment problems have a constant toll revenue property.

This paper mainly focuses on traffic equilibrium models which have the constant toll revenue property. The system problem for these models usually has the network balance constraints and some side constraints. Side constraints can be used to describe the transportation authority's goals, control policies and physical constraints. Hearn (1980), Ferrari (1995), Larsson and Patriksson (1994, 1995a,b, 1998, 1999) and Yang and Bell (1997) analyze the side constrained traffic equilibrium models in detail to show that an unconstrained tolled user problem will have the same equilibrium flows as the side constrained one. Furthermore, Larsson and Patriksson (1998) present a toll pricing model based on Lagrange multipliers and show that the constant toll revenue property holds for elastic demand problems with side constraints.

In this paper, our primary goal is to generalize the first best elastic demand toll pricing framework to variable demand models. Variable demand traffic assignment problems model the situation where the demand between origin–destination (OD) pairs might change substantially based on the "service level," which is usually a function of the travel time on the transportation network. When service level varies, users might decide to take the trip, or not to take the trip at all. We propose a general variable demand (GVD) model which can encompass several traffic assignment problems such as the elastic demand (ED) traffic assignment problem, elastic demand with capacities on links (ED-C) and combined distribution assignment models (CDAM). We show that the toll set for each of these models is an instance of the toll set of the GVD model and all of these models have the constant toll revenue property.

2. General variable demand TA models

Let $\mathscr{G}$ denote the network model of a transportation system which consists of streets, $\mathscr{A}$, and intersections, $\mathscr{N}$. Mathematically, this is a network, $\mathscr{G} = (\mathscr{N}, \mathscr{A})$ where $\mathscr{N}$ is the node set and $\mathscr{A}$ is the link set. Let A be the node-arc incidence matrix of $\mathscr{G}$.We define a commodity by an origin, p, and a destination, q. Let $\mathscr{K}$ represent the set indexing all such origin–destination pairs $k = (p, q)$. The kth *commodity flow* vector is denoted by x^k and the sum of all the commodity flow vectors is the *aggregate flow vector* v. We assume that a continuously differentiable **cost map** $s : \mathscr{A} \to \mathscr{A}$ is given. ∇s denotes the Jacobian of s. When the aggregate flow on the network is v, the travel time for a user on arc a is given by $s_a(v)$. Let c_k be the generalized cost of travel for commodity k, and t_k denote a nonnegative invertible function of c_k. The demand for travel from some origin p to destination q is expressed as $t_k(c_k)$. It is generally assumed that $t_k(c_k)$ are monotonically decreasing and bounded from above. Let $w_k(t_k)$ indicate the inverse of t_k. The vector t has components t_k and the vector function $w(t)$ has $w_k(t_k)$ as components. Then the set of inequalities which define all possible feasible flows and demands can be stated as

$$
\begin{array}{rlll}
v & = \sum_{k\in\mathscr{K}} x^k & & : \mu \\
Ax^k & = E_k t_k & \forall k \in \mathscr{K} & : \rho^k \\
g_m(v) & \leqslant 0 & \forall m \in M & : \gamma_m \\
h_n(t) & = 0 & \forall n \in N & : \theta_n \\
-x^k & \leqslant 0 & \forall k \in \mathscr{K} & : \tau^k \\
-t_k & \leqslant 0 & \forall k \in \mathscr{K} & : \phi_k
\end{array}
$$

where $E_k = e_p - e_q$, a column incidence vector for commodity k, and e_p and e_q are unit vectors. The first constraint is the **aggregate flow** constraint and the second constraint is the **network balance** constraint. $g_m(v) \leqslant 0$, $m \in M$, are constraints on the aggregate flows, and $h_n(t) = 0$, $n \in N$, are constraints on the demand, and $(\mu, \rho, \gamma, \theta, \tau, \phi)$ are multipliers related to the constraints. Some of the constraints on flows and demands can be physical constraints on the network such as road capacity and environmental regulations. Others might be traffic management constraints usually imposed by regulators. We define the set of all feasible flows and demands, Ω, as

Table 1
$W(t)$, $g(v)$ and $h(t)$ for the special cases of the GVD model

Model	$W(t)$	$g(v)$	$h(t)$
ED	$-\sum_{k\in\kappa}\int_0^{t_k} w_k(z)\,\mathrm{d}z$		
ED-C	$-\sum_{k\in\kappa}\int_0^{t_k} w_k(z)\,\mathrm{d}z$	$v_a - C_a \leqslant 0$	
CDAM	$\frac{1}{\zeta}\sum_{k\in K} t_k \ln t_k$		$\sum_q t_k = O_p, \sum_p t_k = D_q, k=(p,q)$

$$\Omega = \{(v,t) : (v,t) \text{ satisfies the constraints above}\}.$$

The system and user objectives that we consider throughout this paper are of the form $S(v) - W(t)$. We use $S(v) = \sum_{a\in\mathscr{A}} s_a(v)v_a$ for the system problems and $S(v) = \sum_{a\in\mathscr{A}}\int_0^{v_a} s_a(z)\,\mathrm{d}z$ for the user problems. We assume that $S(v)$ and $-W(t)$ are convex. $W(t)$, $g(v)$ and $h(t)$ for the models that we consider are listed in Table 1. We assume that $g(v)$ and $h(t)$ are linear, and $s_a(v)$ and $w_k(t_k)$ are separable. The resulting system and user problems have a convex nonlinear objective and linear constraints. Thus a constraint qualification automatically holds and multipliers will exist at the local optima. [1] Note also that the results in this paper can be extended, under assumptions such as in Larsson and Patriksson (1998), when g and h are nonlinear, provided that some constraint qualification holds.

2.1. System problem

In the general variable demand model, when the system objective is to minimize the total system cost (or maximize the total social welfare), the goal is to utilize the network resources in the most efficient way. The system problem (SOPT-GVD) for the GVD model can be defined as:

$$\begin{aligned} \min \quad & s(v)^{\mathrm{T}}v - \sum_{k\in\mathscr{K}}\int_0^{t_k} w_k(z)\,\mathrm{d}z \\ \text{subject to} \quad & (v,t)\in\Omega. \end{aligned}$$

The system optimal flows and demands are characterized by the following lemma:

Lemma 1. *Assume that $(\bar{v},\bar{t})\in\Omega$ is an optimal solution to the SOPT-GVD problem. Then there exists $(\mu,\rho,\gamma,\theta,\tau,\phi)$ such that the following conditions hold:*

$$\begin{aligned} s_a(\bar{v}) + \tfrac{\partial s_a(\bar{v})}{\partial v_a}\bar{v}_a + \sum_{m\in M}\gamma_m\tfrac{\partial g_m(\bar{v})}{\partial v_a} &= \mu_a && \forall a\in\mathscr{A} \\ \mu_a + (\rho_i^k - \rho_j^k) &= \tau_a^k && \forall k\in\mathscr{K},\ \forall a=(i,j)\in\mathscr{A} \\ -w_k(\bar{t}_k) - \sum_{n\in N}\theta_n\tfrac{\partial h_n(\bar{t})}{\partial t_k} + (\rho_q^k - \rho_p^k) &= \phi_k && \forall k=(p,q)\in\mathscr{K} \\ \gamma_m g_m(\bar{v}) &= 0 && \forall m\in M \\ \sum_{k\in K}\left(w_k(\bar{t}_k) + \sum_{n\in N}\theta_n\tfrac{\partial h_n(\bar{t})}{\partial t_k}\right)\bar{t}_k &= \sum_{a\in\mathscr{A}}\mu_a\bar{v}_a \\ \tau,\phi,\gamma &\geqslant 0 \end{aligned}$$

[1] See Bazaraa et al. (1993, Chapter 5).

Proof. Since $(\bar{v}, \bar{t})$ is a KKT point for the SOPT-GVD problem, there exists $(\mu, \rho, \gamma, \theta, \tau, \phi)$ such that the following conditions hold:

$$\frac{\partial L}{\partial v_a} = s_a(\bar{v}) + \frac{\partial s_a(\bar{v})}{\partial v_a}\bar{v}_a - \mu_a + \sum_{m\in M}\gamma_m\frac{\partial g_m(\bar{v})}{\partial v_a} = 0 \qquad \forall a \in \mathscr{A}$$

$$\frac{\partial L}{\partial x_a^k} = \mu_a + (\rho_i^k - \rho_j^k) - \tau_a^k = 0 \qquad \forall k \in \mathscr{K}, \forall a \in \mathscr{A}$$

$$\frac{\partial L}{\partial t_k} = -w_k(\bar{t}_k) + (\rho_q^k - \rho_p^k) - \sum_{n\in N}\theta_n\frac{\partial h_n(\bar{t})}{\partial t_k} - \phi_k = 0 \quad \forall k = (p, q) \in \mathscr{K}$$

$$\gamma_m g_m(\bar{v}) = 0 \qquad \forall m \in M$$

$$\tau_a^k \bar{x}_a^k = 0 \qquad \forall k \in \mathscr{K}, \forall a = \in \mathscr{A}$$

$$\phi_k \bar{t}_k = 0 \qquad \forall k \in \mathscr{K}$$

$$\tau, \phi, \gamma \geqslant 0$$

where $L(x, v, t, \mu, \rho, \gamma, \theta, \tau, \phi)$ is the Lagrangian of the system problem. We can aggregate the complementarity conditions, $\tau_a^k \bar{x}_a^k = 0$ and $\phi_k \bar{t}_k = 0$ to

$$\sum_{k\in\mathscr{K}}\sum_{a\in\mathscr{A}}\tau_a^k\bar{x}_a^k = \sum_{k\in\mathscr{K}}\sum_{a\in\mathscr{A}}\mu_a\bar{x}_a^k - \sum_{k\in\mathscr{K}}(x^{-k})^{\mathrm{T}}A^{\mathrm{T}}\rho^k = 0$$

and

$$\begin{aligned}\sum_{k\in\mathscr{K}}\phi_k\bar{t}_k &= -\sum_{k\in\mathscr{K}}\left(w_k(\bar{t}_k) + \sum_{n\in N}\theta_n\frac{\partial h_n(\bar{t})}{\partial t_k}\right)\bar{t}_k + \sum_{k\in\mathscr{K}}\bar{t}_k E_k^{\mathrm{T}}\rho^k \\ &= -\sum_{k\in\mathscr{K}}\left(w_k(\bar{t}_k) + \sum_{n\in N}\theta_n\frac{\partial h_n(\bar{t})}{\partial t_k}\right)\bar{t}_k + \sum_{k\in\mathscr{K}}(A\bar{x}^k)^{\mathrm{T}}\rho^k = 0\end{aligned}$$

Summing up these equations we get,

$$\sum_{k\in\mathscr{K}}\sum_{a\in\mathscr{A}}\tau_a^k\bar{x}_a^k + \sum_{k\in\mathscr{K}}\phi_k\bar{t}_k = 0$$

$$\sum_{a\in\mathscr{A}}\mu_a\bar{v}_a - \sum_{k\in\mathscr{K}}(\bar{x}^k)^{\mathrm{T}}A^{\mathrm{T}}\rho_k - \sum_{k\in\mathscr{K}}\left(w_k(\bar{t}_k) + \sum_{n\in N}\theta_n\frac{\partial h_n(\bar{t})}{\partial t_k}\right)\bar{t}_k + \sum_{k\in\mathscr{K}}(A\bar{x}^k)^{\mathrm{T}}\rho_k = 0$$

$$\sum_{a\in\mathscr{A}}\mu_a\bar{v}_a - \sum_{k\in\mathscr{K}}\left(w_k(\bar{t}_k) + \sum_{n\in N}\theta_n\frac{\partial h_n(\bar{t})}{\partial t_k}\right)\bar{t}_k = 0. \qquad \square$$

Note that the multipliers $(\mu, \rho, \gamma, \theta, \tau, \phi)$ exist, since the system problem is a convex nonlinear programming problem with linear constraints (thus, a constraint qualification automatically holds).

2.2. User problem

As all deterministic traffic assignment models do, the user problem assumes that each traveller has complete and precise information about all routes available. The user problem is defined by Wardrop's first principle (1952). The underlying assumption is that all users taking a trip between an OD pair have the same travel time which is less than or equal to the travel time on any unutilized path.

In the system problem, there are side constraints which are used either to model the physical constraints or the management goals. However, it is usually assumed that the users perceive only conditions on the network such as the travel times, $s(v)$ (Hearn, 1980). In addition, they do not have any information on constraints imposed on the network, neither the physical ones nor the management goals. Although this might not be the case when there are physical link capacity constraints, as with bottlenecks, we will analyze the user problem without side constraints to model the user behavior.

Mathematically, the user problem can be formulated as a variational inequality. A given aggregate flow and demand vector $(\bar{v}, \bar{t})$ is a user equilibrium flow if and only if

$$s(\bar{v})^{\mathrm{T}}(v-\bar{v}) - w(\bar{t})^{\mathrm{T}}(t-\bar{t}) \geqslant 0 \quad \forall (v,t) \in V,$$

where V is the constraint set defined by the aggregate flow and network balance constraints.

Lemma 2 characterizes the user equilibrium (UOPT-GVD) flows and demands as follows:

Lemma 2. *A feasible point $(\bar{v}, \bar{t})$ is a user equilibrium flow if and only if there exists $(\mu, \rho, \gamma, \theta, \tau, \phi)$ such that the following holds*:

$$\begin{aligned}
s_a(\bar{v}) + (\rho_i^k - \rho_j^k) &= \tau_a^k && \forall k \in \mathscr{K},\ \forall a = (i,j) \in \mathscr{A} \\
-w_k(\bar{t}_k) + (\rho_q^k - \rho_p^k) &= \phi_k && \forall k \in \mathscr{K} \\
\sum_{k \in \mathscr{K}} w_k(\bar{t}_k)\bar{t}_k &= \sum_{a \in \mathscr{A}} \mu_a \bar{v}_a && \\
\tau, \phi &\geqslant 0 &&
\end{aligned}$$

Proof. Let $(\bar{v}, \bar{t})$ be the equilibrium flows and demand, then

$$s(\bar{v})^{\mathrm{T}} v - w(\bar{t})^{\mathrm{T}} t \geqslant s(\bar{v})^{\mathrm{T}} \bar{v} - w(\bar{t})^{\mathrm{T}} \bar{t} \quad \forall (v,t) \in V$$

holds. This implies

$$\min_{(v,t) \in V} s(\bar{v})^{\mathrm{T}} v - w(\bar{t})^{\mathrm{T}} t \geqslant s(\bar{v})^{\mathrm{T}} \bar{v} - w(\bar{t})^{\mathrm{T}} \bar{t}.$$

Consider now

$$\begin{aligned}
&\min && s(\bar{v})^{\mathrm{T}} v - w(\bar{t})^{\mathrm{T}} t \\
&\text{subject to} && (v,t) \in V
\end{aligned}$$

From the inequality above it is clear that $(\bar{v}, \bar{t})$ solves this linear program. Therefore, by linear programming duality, there exists $(\mu, \rho, \gamma, \theta, \tau, \phi)$ such that the following conditions hold:

$$\begin{aligned}
s_a(\bar{v}) - \mu_a &= 0 && \forall a \in \mathscr{A} \\
\mu_a + (\rho_i^k - \rho_j^k) - \tau_a^k &= 0 && \forall k \in \mathscr{K},\ \forall a \in \mathscr{A} \\
-w_k(\bar{t}_k) + (\rho_q^k - \rho_p^k) - \phi^k &= 0 && \forall k \in \mathscr{K} \\
\tau_a^k \bar{x}_a^k &= 0 && \forall k \in \mathscr{K},\ \forall a \in \mathscr{A} \\
\phi_k \bar{t}_k &= 0 && \forall k \in \mathscr{K} \\
\tau, \phi &\geqslant 0
\end{aligned}$$

where $\tau, \phi, \gamma \geqslant 0$. The last three equations are the complementary slackness conditions. Further, $\tau_a^k \bar{x}_a^k = 0$ and $\phi_k \bar{t}_k = 0$ can be aggregated in a manner similar to Lemma 1 to obtain

$$\sum_{k \in \mathscr{K}} w_k(\bar{t}_k)\bar{t}_k = \sum_{a \in \mathscr{A}} \mu_a \bar{v}_a$$

Then the lemma follows. □

3. First best toll pricing for the GVD model

In this section, we extend the toll set idea in Bergendorff (1995), Hearn and Ramana (1998) and Hearn and Yildirim (2002) to the GVD model. We will assume that the transportation authority aims to exactly replicate the system optimal flows and demands, and the effect of side constraints. We will first describe this framework for UOPT-GVD and then discuss the relation with other user problems.

3.1. Toll pricing theory for the GVD model

Suppose that $(\bar{\gamma}, \bar{\theta})$ are the multipliers for the side constraints in the SOPT-GVD problem. Assume that on each link users are charged by an amount equal to

$$\lambda_a = \beta_a + \sum_{m \in M} \bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}$$

where β is a toll vector and $\sum_{m \in \mathscr{M}} \bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}$ is the constraint cost. Furthermore, each user is rewarded (or penalized) for an amount equal to

$$\xi_k = \sum_{n \in N} \bar{\theta}_m \frac{\partial h_n(\bar{t})}{\partial t_k}$$

for making a trip between OD pair $k = (p, q)$.

Suppose $s_\lambda(v) = s(v) + \lambda$ is the perturbed cost map. Further, let $w_\xi(t) = w(t) + \xi$ be the perturbed inverse demand function. Then the perturbed user problem can be stated as a variational inequality:

$(\bar{v}, \bar{t})$ is a tolled user equilibrium flow if and only if

$$(s(\bar{v}) + \lambda)^{\mathrm{T}}(v - \bar{v}) - (w(\bar{t}) + \xi)^{\mathrm{T}}(t - \bar{t}) \geqslant 0 \quad \forall (v, t) \in V$$

Let U^*_β be the set of all perturbed equilibrium solutions, i.e., those $(\bar{v},\bar{t}) \in V$ that solve the above variational inequality, and let S^* be the optimal solution set for SOPT-GVD. We would like to identify all toll vectors such that the resulting perturbed user equilibrium problem has a system optimal solution, i.e.,

$$\emptyset \neq U^*_\beta \subseteq S^*$$

Any such β is defined to be a valid toll vector. Let the toll set denoted by $\mathscr{T}$ be the set of all such vectors, $\mathscr{T} := \{\beta \mid \emptyset \neq U^*_\beta \subseteq S^*\}$. Now, for a given vector $(\bar{v},\bar{t}) \in V$, define

$$W_{\mathrm{GVD}}(\bar{v},\bar{t},\bar{\gamma},\bar{\theta}) = \{\beta \mid (\bar{v},\bar{t}) \in U^*_\beta\}$$

as the set of all tolls which ensures that $(\bar{v},\bar{t})$ is a solution of $\mathrm{UOPT}_{(\lambda,\xi)}$-GVD. In fact, we can define $W_{\mathrm{GVD}}(\bar{v},\bar{t},\bar{\gamma},\bar{\theta})$ by the following result.

Lemma 3. *Given that $(\bar{v},\bar{t}) \in V$ and $(\bar{\gamma},\bar{\theta})$ are the system optimal multipliers, $W_{\mathrm{GVD}}(\bar{v},\bar{t},\bar{\gamma},\bar{\theta})$ is the polyhedron given by the β part of the linear system defined in (β,ρ):*

$$\begin{aligned}
\left(s_a(\bar{v}) + \beta_a + \sum_{m\in M} \bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}\right) + (\rho^k_i - \rho^k_j) \quad &\geqslant 0 \quad \forall k \in \mathscr{K},\ \forall a = (i,j) \in \mathscr{A}\\
\left(w_k(\bar{t}_k) + \sum_{n\in N} \bar{\theta}_n \frac{\partial h_n(\bar{t})}{\partial t_k}\right) - (\rho^k_q - \rho^k_p) \quad &\leqslant 0 \quad \forall k \in \mathscr{K}\\
\sum_{a\in\mathscr{A}} \left(s_a(\bar{v}) + \beta_a + \sum_{m\in M} \bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}\right)\bar{v}_a \quad &= \sum_{k\in\mathscr{K}} \left(w_k(\bar{t}_k) + \sum_{n\in N} \bar{\theta}_n \frac{\partial hn(\bar{t})}{\partial t_k}\right)\bar{t}_k
\end{aligned}$$

Proof. Using Lemma 2 and the perturbed cost vector $s_a(v_a) + \beta_a$ and the perturbed inverse demand function $w_\xi(t) = w(t) + \xi$, we observe that the solution to the perturbed user equilibrium satisfies the following:

$$\left(s_a(\bar{v}) + \beta_a + \sum_{m\in M} \bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}\right) = \mu_a = \rho^k_j - \rho^k_i + \tau^k_a \quad \forall k \in \mathscr{K},\ \forall a \in \mathscr{A}$$

and

$$\rho^k_q - \rho^k_p - \left(w_k(\bar{t}_k) + \sum_{n\in N} \bar{\theta}_n \frac{\partial h_n(\bar{t})}{\partial t_k}\right) = \phi^k \quad \forall k \in \mathscr{K}$$

hold. Combining the above with $\tau^k_a \geqslant 0$ and $\phi^k \geqslant 0$ yields the first two inequalities in the lemma. The lemma follows. □

When the optimality conditions of the system and tolled user problems are compared, it is clear that they differ in the definition of μ and the complementarity conditions, $\bar{\gamma}_m g_m(\bar{v}) = 0 \quad \forall m \in M$ and $\bar{\gamma} \geqslant 0$. The difference between μ for the system and tolled user problems is $\nabla_s(\bar{v})\bar{v}$, the MSCP toll vector. The complementarity conditions also hold for the tolled user problem since $(\bar{\gamma},\bar{\theta})$ are the multipliers from the system problem. We can conclude from these observations that if the individuals were charged with the MSCP tolls and constraint costs, the perturbed user equilibrium flows and demands would be the same as the system optimal flows and demands.

The toll set is defined by variables (β, ρ) and by parameters $(\bar{v}, \bar{t}, \bar{\gamma}, \bar{\theta})$. The interpretation of what these variables mean is as follows: β is the toll vector. ρ_i^k is the total travel cost from origin p to node i for the commodity $k = (p, q)$ on the transportation network. Furthermore, $\bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}$ is the amount that users will pay if there were no $g_m(v) \leqslant 0$ constraint and similarly, $\bar{\theta}_n \frac{\partial h_n(\bar{t})}{\partial t_k}$ is the amount that users will pay in the absence of experiencing the $h_n(t) = 0$ constraint.

Note that if the user and system problems (thus the tolled user problem) have unique solutions, then $\mathscr{T} = W_{\mathrm{GVD}}(\bar{v}, \bar{t}, \bar{\gamma}, \bar{\theta})$. A complete description of characterization of the toll sets is given in Hearn and Ramana (1998). In addition, the toll set where $(\bar{\gamma}, \bar{\theta})$ are fixed has similar characteristics as the elastic demand toll set (Hearn and Yildirim, 2002; Larsson and Patriksson, 1998). The following corollary shows that the total toll revenue for the GVD models is constant as it is in the elastic demand traffic assignment problems:

Corollary 1. *The toll revenue for the variable demand traffic assignment models is constant when the toll set is defined by* $W_{\mathrm{GVD}}(\bar{v}, \bar{t}, \bar{\gamma}, \bar{\theta})$.

Proof. This is obvious from the definition of $W_{\mathrm{GVD}}(\bar{v}, \bar{t})$. The total toll revenue is

$$\beta^{\mathrm{T}}\bar{v} - \sum_{k\in\mathscr{K}}\sum_{n\in N}\bar{\theta}_n \frac{\partial h_n(\bar{t})}{\partial t_k}\bar{t}_k + \sum_{a\in\mathscr{A}}\sum_{m\in M}\bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}\bar{v}_a = \sum_{k\in\mathscr{K}} w_k(\bar{t}_k)\bar{t}_k - \sum_{a\in\mathscr{A}} s_a(\bar{v})\bar{v}_a. \qquad \square$$

However, note that this is not the case in traffic assignment problems with fixed demand (Bergendorff et al., 1997; Hearn and Ramana, 1998; Larsson and Patriksson, 1998).

Wardrop's First Principle implies that at equilibrium the utilized paths for an origin–destination pair have the same travel time (travel costs) and the unused ones do not have lower travel times (travel costs). The next corollary verifies that the generalized cost version of this principle holds (a similar result for the fixed demand case is given by Larsson and Patriksson (1999). Let r_k be a path for commodity k, χ_{ar_k} be 1 if link a is on path r_k and zero otherwise,

$$s_a(\bar{v}_a) + \beta_a + \sum_{m\in M}\bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}$$

be the "generalized cost" of travelling on arc a and

$$w_k(\bar{t}_k) + \sum_{n\in N}\bar{\theta}_n \frac{\partial h_n(\bar{t})}{\partial t_k}$$

be the "generalized benefit" for commodity k.

Corollary 2. *At the* UOPT_β*-GVD solution* $(\bar{v}, \bar{t})$ *the total generalized cost on any path is greater than or equal to the generalized benefit for any commodity* k*, i.e.,*

$$\sum_{a\in\mathscr{A}}\chi_{ar_k}\left(s_a(\bar{v}_a) + \beta_a + \sum_{m\in M}\bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a}\right) \geqslant w_k(\bar{t}_k) + \sum_{n\in N}\bar{\theta}_m \frac{\partial h_n(\bar{t})}{\partial t_k}$$

For any path with positive flow, the inequality holds as an equality. Therefore the costs on the utilized paths are constant for any commodity.

Proof. For $k = (p,q)$ and $a = (i,j) \in r_k$,

$$s_a(\bar{v}_a) + \beta_a + \sum_{m\in M} \bar{\gamma}_m \frac{\partial g_m(\bar{v})}{\partial v_a} \geqslant \rho_j^k - \rho_i^k$$

holds as an equality when $\bar{x}_a^k > 0$ and similarly,

$$w_k(\bar{t}_k) + \sum_{n\in N} \bar{\theta}_n \frac{\partial h_n(\bar{t})}{\partial t_k} \leqslant \rho_j^k - \rho_i^k$$

holds as an equality when $\bar{t}_k > 0$. That is, the complementarity conditions force the inequality to hold as an equality when there is flow on that path. Thus, when they are summed over the arcs on the utilized paths, the inequality in the corollary holds as an equality. The conclusion of the corollary then follows. □

In other words, this corollary can be interpreted by saying that every commodity reaches an equilibrium exactly when the generalized path costs, including tolls, equals the generalized benefit at the final demand level.

Note that the Lagrangean multipliers related to side constraints are usually not unique (Larsson and Patriksson, 1998). Detailed discussions on traffic management through link tolls where the Lagrangean multipliers are allowed to vary can be found in Larsson and Patriksson (1998) and Yang and Bell (1997). In this case, β can absorb any contribution of the Lagrange multipliers for the side constraints. This leads to some simplification in notation. However, we have not used the conditions in Larsson and Patriksson (1998) to be compatible with Hearn (1980) and Yang and Huang (1998).

4. First best toll pricing framework for the GVD models

Assume that SOPT-GVD and UOPT-GVD models have unique solutions. Furthermore, assume that the traffic planner is interested in obtaining the system optimal flows and demands and the constraint costs $(\bar{\gamma}, \bar{\theta})$. Then, the first best toll pricing framework for GVD models can be summarized as follows:

Step 1: Solve the SOPT-GVD to obtain the system optimal solution $(\bar{v}, \bar{t})$ and $(\bar{\gamma}, \bar{\theta})$.
Step 2: Define a toll set which is the β part of $W_{\mathrm{GVD}}(\bar{v}, \bar{t}, \bar{\gamma}, \bar{\theta})$.
Step 3: Define and optimize an objective function over the toll set, possibly intersected with other constraints (see Table 2 for examples).

As in the fixed demand case, various objectives in Step 3 lead to linear or mixed integer programs. For the variable demand models, the natural choices can be D-MINSYS, D-MINREV, MINTB and MINMAX objectives (see Table 2). The first of these aims to minimize the total tolls collected while constraining the toll vector to be nonnegative. D-MINREV is similar to D-MINSYS, but tolls are free to be negative as well as positive. Note that the marginal social cost pricing tolls, $\beta_{\mathrm{MSCP-GVD}} = \nabla s(\bar{v})\bar{v}$, are optimal for the GVD model when the objective is minimization of the total toll revenue. This is because of the constant toll revenue property. MINTB minimizes the

Table 2
Alternative optimization formulations

Toll	Objective ($\overline{Z}$)	Extra constraints ($\widehat{T}$)
D-MINREV	$\beta^T \bar{v}$	
D-MINSYS	$\beta^T \bar{v}$	$\beta \geqslant 0$
MINMAX	z	$z \geqslant \beta_a + \bar{\gamma}_a \; \forall a \in \mathscr{A}, \beta \geqslant 0$
MINTB	$\sum_{a \in \mathscr{A}} y_a$	$\beta_a + \bar{\gamma}_a \leqslant M y_a \; \forall a \in \mathscr{A}, y_a \in \{0,1\}, \beta \geqslant 0$
MINTB/MINREV	$\sum_{a \in \mathscr{A}} y_a$	$\beta_a + \bar{\gamma} \leqslant M y_a \; \forall a \in \mathscr{A}, y_a \in \{0,1\}$

number of toll booths, and the MINMAX minimizes the maximum toll on the transportation network. Note that any valid toll vector including MSCP toll vector produces the same toll revenue. As a result, D-MINSYS and D-MINREV formulations are not very interesting for the GVD networks, however, having some alternative toll vectors might help the traffic planner to propose different toll pricing schemes for various scenarios. In all of these toll pricing problems, the transportation authority collects both β and $\bar{\gamma}$ as the total toll charge on a link. In this paper, we use GAMS optimization modeling package (1995) and linear and nonlinear solvers (CPLEX, 2001 and MINOS, 1983) to implement the GVD Toll Pricing Framework to obtain alternative toll vectors. The GAMS code is available upon request.

Actually, the toll pricing framework has interesting implications: The traffic managers do not need to put physical constraints (such as closing a lane, etc.) to restrict users on the transportation network. In fact, users can be charged by the "correct" toll amount for not only sustaining the management goals but also maintain the flows to be consistent with the constraints on the transportation system.

5. GVD models in the literature

In this section, we present special cases of the GVD model and define the toll set for each model. Furthermore, we illustrate the toll pricing framework for each model on the Nine Node network (Fig. 1) which has nine nodes and 18 links, and all of the links have cost functions with the same structure:

$$s_a(v) = s_a(v_a) = T_a(1 + 0.15(v_a/C_a)^4)$$

where T_a is a measure of travel time when there is zero flow and C_a is the practical capacity of link a. In fact $s(v)$ is strictly convex and separable. There are four OD pairs. The particular choices of $s(v)$ and $w(t)$ results in having unique solutions to the system and user problems. Thus, the toll set for each model is $\mathscr{T} = W_{\mathrm{GVD}}(\bar{v}, \bar{t}, \bar{\gamma}, \bar{\theta})$. In this paper, tolls are expressed in time units (Arnott and Small (1994) discuss how conversion to dollars can be made based on studies in the US).

5.1. Elastic demand TA models

The system model with elastic demand assumes that the goal of transportation planners is to maximize the net economic benefit (Hearn and Yildirim, 2002; Yang and Huang, 1998) which

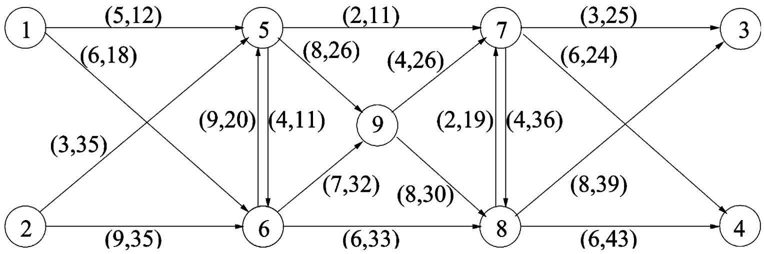

Fig. 1. Nine Node network: the tuple near link a is (T_a, C_a).

is the difference between the total network user benefit, $\sum_{k\in\mathscr{K}} \int_0^{t_k} w_k(z)\,\mathrm{d}z$, and the system cost, $s(v)^T v$. The system problem (SOPT-ED) can be stated as

$$\begin{aligned} &\max \quad \sum_{k\in\mathscr{K}} \int_0^{t_k} w_k(z)\,\mathrm{d}z - s(v)^{\mathrm{T}} v \\ &\text{subject to} \quad (v,t) \in V. \end{aligned}$$

where V denotes the set of all feasible flows and demands, which is defined by the aggregate flow, network balance and nonnegativity constraints, i.e., there are no g and h constraints.

The elastic demand user equilibrium problem (UOPT-ED) models Wardrop's first principle (1952). Mathematically, the user problem can be stated as a variational inequality:

Find $(\bar{v},\bar{t}) \in V$ such that

$$s(\bar{v})^{\mathrm{T}}(v-\bar{v}) - w(\bar{t})^{\mathrm{T}}(t-\bar{t}) \geqslant 0 \quad \forall (v,t) \in V.$$

Using Lemma 3 of the GVD model, we can extend the notion of toll pricing to the elastic demand traffic assignment models. [2] We will give a description of the toll set in the next section.

5.2. Elastic demand TA models with link capacities

Traditionally, traffic planners handle capacities using special social cost functions like

$$s_a(v_a) = T_a\left(1 + 0.15\left(\frac{v_a}{C_a}\right)^4\right).$$

However, using these functions will not guarantee that the attained flows do not exceed the capacity on each link. For example, in Fig. 1, the capacity of link $(5,7)$ is 11, but as it can be seen from Table 4, the uncapacitated user problem allows 26.44 users and the system problem has 17.98 users on link $(5,7)$, both of which are far above the capacity. Thus it might be important to use the capacity constraints explicitly in some cases.

[2] Refer to Hearn and Yildirim (2002) for a detailed description of first best toll pricing theory for elastic demand traffic assignment problems.

When GVD model has the objective function as maximizing the net economic benefit and constraints as aggregate flow, network balance and capacity constraints ($g_a(v_a) = v_a - C_a$, where C_a is the capacity of link a), the resulting model is the ED-C model.

Hearn (1980) show that for fixed demand traffic assignment problems, a tolled user problem will have the same flows as an untolled capacitated user problem. Ferrari (1995) extends this idea to elastic demand traffic assignment problems and Yang and Bell (1997) propose a bi-level toll pricing algorithm to find alternative toll vectors (Lagrangean multipliers) for the elastic demand problem.

Using the toll set $W_{\text{ED-C}}(\bar{v}, \bar{t}, \bar{\gamma})$ for the ED-C model, a traffic planner can achieve the most efficient utilization (i.e., the system optimal flows and demands) of the network while having users perceive exactly the same constraint costs, $\bar{\gamma}$, as those given by the system problem ($\bar{\gamma}$ is the KKT multiplier vector for the capacity constraints of the system problem). Based on the definitions and assumptions made in Section 3.1, $W_{\text{ED-C}}(\bar{v}, \bar{t}, \bar{\gamma})$ is:

$$\begin{aligned} s(\bar{v}) + \beta + \bar{\gamma} &\geqslant A^{\mathrm{T}} \rho^k \qquad \forall k \in \mathscr{K} \\ w_k(\bar{t}_k) &\leqslant E_k^{\mathrm{T}} \rho^k \qquad \forall k \in \mathscr{K} \\ (s(\bar{v}) + \beta + \bar{\gamma})^{\mathrm{T}} \bar{v} &= w(\bar{t})^{\mathrm{T}} \bar{t}. \end{aligned}$$

In this case, the total amount that users are charged is constant, $\beta^{\mathrm{T}}\bar{v} + \bar{\gamma}^{\mathrm{T}}\bar{v} = w(\bar{t})^{\mathrm{T}}\bar{t} - s(\bar{v})^{\mathrm{T}}\bar{v}$, which is unique for a given $(\bar{v}, \bar{t}) \in V$. Note that $W_{\text{ED-C}}(\bar{v}, \bar{t}, \bar{\gamma})$ reduces to $W_{\text{ED}}(\bar{v}, \bar{t})$ (i.e., the toll set for the uncapacitated elastic demand traffic assignment problems) when $\bar{\gamma} = 0$, since $\bar{v}_a < C_a$ for the ED problem.

In the literature, there have been several interpretations for the multiplier γ (Hearn, 1980; Larsson and Patriksson, 1994; Yang and Bell, 1997). For example, γ is interpreted as the link toll (queuing toll) that the travellers will pay for being allowed to use the links at capacity. It is also a measure of time gained by users of routes filled with capacity compared to the fastest route. It might be also interpreted as delays in the steady state link queue.

MSCP toll amount $\nabla s(\bar{v})\bar{v}$ and the system optimal multiplier $\bar{\gamma}$ are valid toll and constraint cost combination. This implies that the MSCP toll vector that Yang and Huang (1998) propose is a valid first best toll vector, i.e.,

$$\beta_{\text{MSCP-ED-YH}-C_a} = \beta_{\text{MSCP-ED-C}} + \bar{\gamma} = \begin{cases} \nabla s(\bar{v}_a)\bar{v}_a + \bar{\gamma} & \text{if } \bar{v}_a = C_a \\ \nabla s(\bar{v}_a)\bar{v}_a & \text{otherwise.} \end{cases}$$

In other words, $\beta_{\text{MSCP-ED-YH}-C_a}$ is a combination of the classical MSCP toll vector $\beta_{\text{MSCP-ED}-C}$ and the queuing toll/constraint cost, $\bar{\gamma}$ (which is obtained from the system optimal solution). But recall that Lemma 3 implies that the classical definition of the MSCP toll vector, $\nabla s(\bar{v})\bar{v}$, is a valid toll vector for all GVD models.

We use the capacitated Nine Node network to illustrate the toll pricing framework for the SOPT-ED-C problem, and exploit the effect of the queuing/capacities on the total user delay. The demand functions between OD pairs are, $t_{(1,3)}(c_{(1,3)}) = 10 - 0.5c_{(1,3)}$, $t_{(1,4)}(c_{(1,4)}) = 20 - 0.5c_{(1,4)}$, $t_{(2,3)}(c_{(2,3)}) = 30 - 0.5c_{(2,3)}$, and $t_{(2,4)}(c_{(2,4)}) = 40 - 0.5c_{(2,4)}$ where c_{pq} is the generalized cost between OD pair (p, q). Table 4 shows that adding capacity constraints of 20 on each link results in substantial changes in the network flows when compared to the SOPT-ED optimal flows and UOPT-ED equilibrium flows. The net user benefit for capacitated user and system problems differs only by 0.5%. In the absence of capacities, this difference is 10.24%. One can conclude that adding capacity constraints resulted in a system where the user and system optimal behaviors are quite similar in terms of the efficiency.

As can be seen in Table 3, the total number of users on the network, which is 48.46 for the UOPT-ED-C and 48.23 for the SOPT-ED-C, is much less than the uncapacitated case (60.75 for the UOPT-ED demand and 57.41 for the SOPT-ED demand). It is interesting to note that

Table 3
Demand and user benefit $(\bar{t}_k, w_k(\bar{t}_k))$ for the ED and ED-C Nine Node problems

OD pair	UOPT-ED solution		SOPT-ED solution	
	3	4	3	4
1	(0.15, 19.70)	(10.70, 18.61)	(0, 20.00)	(9.706, 20.61)
2	(20.67, 18.66)	(29.23, 21.54)	(19.48, 21.05)	(28.24,23.52)
	UOPT-ED-C solution		SOPT-ED-C solution	
	3	4	3	4
1	(0, 20.00)	(8.46, 23.08)	(0, 20.00)	(8.28, 23.45)
1	(15.75, 28.50)	(24.25, 31.50)	(15.75, 28.51)	(24.25,31.49)

Table 4
Nine Node problem-optimal solution to user and system problems for the ED and ED-C Nine Node problems

Link	ED		ED-C	
	$\bar{v}_a$	v_a^U	$\bar{v}_a$	v_a^U
(1, 5)			3.22	3.50
(1, 6)	9.70	10.85	5.06	4.96
(2, 5)	31.72	34.46	20.00	20.00
(2, 6)	16.00	15.45	20.00	20.00
(5, 6)				
(5, 7)	17.98	26.44	17.72	20.00
(5, 9)	13.74	8.02	5.50	3.50
(6, 5)				
(6, 8)	25.70	26.30	20.00	20.00
(6, 9)			5.06	4.96
(7, 3)	19.48	20.82	15.75	15.75
(7, 4)	12.24	13.79	12.53	12.71
(7, 8)				
(8, 3)				
(8, 4)	25.70	26.14	20.00	20.00
(8, 7)		0.15		0.15
(9, 7)	13.74	8.02	10.56	8.46
(9, 8)				
User benefit	2544.75	2613.50	2311.44	2315.67
Social cost	1005.47	1217.21	851.01	862.09
Net user benefit	1539.28	1396.29	1460.43	1453.57

the user and system problems do not allow any flow between OD pair (1, 3), since individuals do not have any incentive for making a trip between this OD pair. Another interesting fact is the distribution of flow between nodes 5 and 9 in the existence of capacities. The number of users utilizing the path (link) 5-7 in SOPT-ED-C is 11.40% (2.28 users) less than the number of users in the UOPT-ED-C problem. On the contrary, the number of users in SOPT-ED-C on path 5-9-7 is 2.00 units higher than the UOPT-ED-C flow. A traffic planner might charge on link 5-7 to divert some users 5-9-7 to have a better utilization of the network. In fact, users are being charged an amount of at least 8.00 units on link 5-7 when MINTB, MINMAX or D-MINSYS tolling schemes are employed.

As we have discussed before, the traffic planner might replicate exactly the system optimal conditions including the flows, demands and constraint costs using $W_{\text{ED-C}}(\bar{v}, \bar{t}, \bar{\gamma})$ as the toll set. Table 5 shows the results when the constraint costs are fixed to the system optimal $\bar{\gamma}$. Users have to pay 375.50 in terms of constraint costs for the waiting time in the queues and pay 180.23 in terms of the total toll revenue. As a result, the total toll cost, $(\beta + \bar{\gamma})^{\mathrm{T}}\bar{v}$, which users are paying both in monetary and constraint costs is 555.73. In Table 5, using the toll pricing framework, we obtain various first best toll vectors. The D-MINREV tolling scheme needs 14 toll booths while the MSCP tolling scheme needs 12 (including the toll booths where constraint costs are only collected). The MINMAX tolling scheme requires eight toll booths, while six toll booths are needed for the D-MINSYS, MINTB and MINTB/MINREV tolls. The maximum toll amount for the MSCP tolling scheme is 19.83. This amount is 14.60 for the D-MINREV scheme. The

Table 5
Nine Node problem-alternative tolls for the capacitated flow problem when $\bar{\gamma}$ is fixed

Link	$\bar{\gamma}_a$	β_{MSCP}	$\beta_{\text{D-MINSYS}}$	$\beta_{\text{D-MINREV}}$	β_{MINMAX}	β_{MINTB} ($\beta_{\text{MINTB/MRV}}$)
(1, 5)		0.02		−0.18		
(1, 6)		0.02		7.82		
(2, 5)	9.82	10.01	0.18		0.18	0.18
(2, 6)	4.35	4.93	0.55	8.38	0.55	0.55
(5, 6)						
(5, 7)		8.08	8.00	14.60	8.00	8.00
(5, 9)		0.01		10.62		
(6, 5)						
(6, 8)	4.60	0.96		−1.10		0.68
(6, 9)		0.00		2.62		
(7, 3)		0.28	0.37	−6.05	0.37	0.37
(7, 4)		0.27	0.36	−6.07	0.36	0.36
(7, 8)				−2.00		
(8, 3)						
(8, 4)		4.29	0.68	−6.04	0.68	
(8, 7)					0.56	
(9, 7)		0.07		−4.02		
(9, 8)				−8.00		
$\bar{\gamma}^{\mathrm{T}}\bar{v}$	375.50	375.50	375.50	375.50	375.50	375.50
$\beta^{\mathrm{T}}\bar{v}$		180.23	180.23	180.23	180.23	180.23
$\beta^{\mathrm{T}}\bar{v}$/NUB		12.00	12.00	12.00	12.00	12.00
Toll booths	4	12	6	14	8	6

D-MINSYS, MINTB, MINTB/MINREV and MINMAX require users be charged by a maximum amount of 10.00. As a result the traffic planner might decide that the D-MINSYS tolling scheme is the most appropriate tolling scheme to implement on the Nine Node network since D-MINSYS tolling scheme requires the least amount of toll booths and charges users a maximum amount same as the amount by the MINMAX tolling scheme.

Note that if all users have an identical value of time, the $\bar{\gamma}$ charge produces no losses for users on the transportation network, because it simply substitutes a charge for the wasted time in queues (Yang and Bell, 1997).

5.3. Combined distribution assignment model

In the traffic planning process, the trip distribution model (TDM) determines the number of trips per unit time between the OD pairs. Traffic assignment models take as input a complete description of the transportation system and an origin–destination matrix of trip demands or travel demand formulas, and output an estimate of traffic volumes, travel costs, and travel times on each link and demand between each origin–destination pair. Evans (1976) combined the trip distribution model (TDM) and traffic assignment models to determine simultaneously the distribution of trips between origins and destinations and assignment of trips to links. Balasubramaniam (1999) and Boyce et al. (2002) show that the MSCP toll pricing scheme is optimal for a more general combined distribution model where travel choices are also considered.

In this section, we present the combined distribution assignment model (CDAM) proposed by Evans (1976) and Lundgren and Patriksson (1998). We show that CDAM is a special case of the GVD model. We present the toll set and then give a numerical example.

The CDAM is a special instance of GVD model, where $h_1^p = \sum_q t_k - O_p$ (trip production constraint) and $h_2^q = \sum_p t_k - D_q$ (trip attraction constraints) are the side constraints and $w_k(t_k) = -\frac{1}{\zeta}(\ln t_k - 1)$ is the $W_k(t)$ of the objective function. Let Ω denote the set of all feasible flows and demands. Then, the system problem (SOPT-CDAM) is defined by the following mathematical program:

$$\begin{aligned} \min \quad & s(v)^{\mathrm{T}} v + \frac{1}{\zeta} \sum_{k \in K} t_k \ln t_k \\ \text{subject to} \quad & (v, t) \in \Omega \end{aligned}$$

Although the user problems for the GVD models assume that there are no side constraints, the CDAM has the trip production and attraction constraints. As a result, the user problem (UOPT-CDAM) can be stated as follows:

$$\begin{aligned} \min \quad & \sum_a \int_0^{v_a} s_a(z)\,\mathrm{d}z + \frac{1}{\zeta} \sum_{k \in K} t_k \ln t_k \\ \text{subject to} \quad & (v, t) \in \Omega \end{aligned}$$

Note that the only change between the system and user problems is in the first term in the objective function.

Let θ_p and θ_q be the KKT multipliers of the trip production arid the trip attraction constraints in the SOPT-CDAM. Using an approach similar to the one used in Lemma 1 and Lemma 3, the CDAM toll set is defined as follows.

Lemma 4. *$W_{\mathrm{CDAM}}(\bar{v},\bar{t},\bar{\theta})$ is the polyhedron given by the β part of the following linear system of equations where $(\bar{v},\bar{t}) \in \Omega$ and $\bar{\theta}$ is the system optimal multiplier for the trip production and the trip attraction constraints*:

$$\begin{aligned} A^{\mathrm{T}}\rho^k &\leqslant s(\bar{v})+\beta && \forall k \in \mathscr{K} \\ \bar{\theta}_p+\bar{\theta}_q-\tfrac{1}{\zeta}(\ln \bar{t}_k+1) &\leqslant E_k^{\mathrm{T}}\rho^k && \forall k=(p,q) \in \mathscr{K} \\ \sum_{k\in K}\left(\bar{\theta}_p+\bar{\theta}_q-\tfrac{1}{\zeta}(\ln \bar{t}_k+1)\right)\bar{t}_k &= (s(\bar{v})+\beta)^{\mathrm{T}}\bar{v} \end{aligned}$$

Table 6
SOPT-CDAM and UOPT-CDAM demands for the Nine Node problem

OD pair	SOPT-CDAM		UOPT-CDAM	
	3	4	3	4
1	12.00	18.00	5.35	24.65
2	28.00	42.00	34.65	35.35

Table 7
Nine Node problem SOPT-CDAM and UOPT-CDAM solutions

Link	SOPT-CDAM			UOPT-CDAM		
	$\bar{v}_a$	$s_a(\bar{v}_a)$	$\bar{v}_a s_a(\bar{v}_a)$	v_a^U	$s_a(v_a^U)$	$v_a^U s_a(v_a^U)$
(1,5)	9.41	5.28	49.72	5.35	5.03	26.91
(1,6)	20.59	7.54	155.26	24.65	9.17	225.94
(2,5)	38.33	3.65	139.83	49.90	4.86	242.53
(2,6)	31.67	9.90	313.63	20.10	9.15	183.81
(5,6)		9.00			9.00	
(5,7)	21.30	6.22	132.51	27.80	14.24	395.80
(5,9)	26.44	9.28	245.48	27.45	9.49	260.59
(6,5)		4.00			4.00	
(6,8)	39.47	7.84	309.57	44.75	9.04	404.62
(6,9)	12.78	7.03	89.81		7.00	
(7,3)	29.61	3.89	115.04	40.00	5.95	237.96
(7,4)	20.76	6.50	134.99	15.25	6.15	93.76
(7,8)		2.00			2.00	
(8,3)	10.39	8.01	83.20		8.00	
(8,4)	39.24	6.62	259.96	44.75	7.06	315.70
(8,7)		4.00			4.00	
(9,7)	29.06	4.94	143.46	27.45	4.75	130.29
(9,8)	10.16	8.02	81.45		8.00	
Total entropy			46.05			45.40
$s(v)^T v$			2253.92			2517.91
System objective (v)			2207.87			2472.37

Table 8
Nine Node CDAM network-alternative tolls

Link	β_{MSCP}	$\beta_{\mathrm{D\text{-}MINSYS}}$	$\beta_{\mathrm{D\text{-}MINREV}}$	β_{MINMAX}	β_{MINTB}	$\beta_{\mathrm{MINTB/MRV}}$
(1,5)	1.13	6.36	15.64	6.36		
(1,6)	6.16	6.36	13.38	5.56	2.54	
(2,5)	2.59	7.81	17.10	7.81	1.45	1.46
(2,6)	3.62	3.81	10.84	3.01		−2.54
(5,6)						
(5,7)	16.88	8.00	4.39	8.00	20.03	21.56
(5,9)	5.13		−9.28		2.55	
(6,5)						
(6,8)	7.37	7.20	0.17	8.00	11.01	16.03
(6,9)	0.11		−7.03	0.80		
(7,3)	3.54	7.20	1.53	7.20	1.53	
(7,4)	2.01	5.67		5.67		−1.52
(7,8)		1.08		1.08		
(8,3)	0.02					−2.47
(8,4)	2.50	2.47	2.47	2.47	2.47	
(8,7)						
(9,7)	3.75		5.67		9.49	13.56
(9,8)	0.06				3.81	8.83
$\beta^{\mathrm{T}}\bar{v}$	1493.53	1493.53	1493.53	1493.53	1493.53	1493.53
Toll booths	14	10	11	11	9	8
$\frac{\beta^{\mathrm{T}}\bar{v}}{s(\bar{v})^{\mathrm{T}}\bar{v}}$	67.67%	67.67%	67.67%	67.67%	67.67%	67.67%

The toll set inherits the characteristics of the GVD toll set. For example, the total toll revenue is constant. Furthermore, $\bar{\theta}_p + \bar{\theta}_q$ can be interpreted as the "benefit" gained (possibly negative) to a user who decides to take a trip from origin p to destination q.

To illustrate the first best toll pricing framework for the CDAM, the Nine Node problem (Bergendorff et al., 1997; Hearn and Ramana, 1998; Hearn and Yildirim, 2002) is modified by adding the trip attraction and trip production constraints. Let $O_1 = 30$, $O_2 = 70$, $D_3 = 40$, and $D_4 = 60$ be the total demand in the zones one through four. The dispersion parameter, ζ, is 10.

The demand vector for the system and user problems are presented in Table 6. The demand vector for both problems differ considerably, but the total entropy for each problem is approximately equal. The entropy term forces both problems to have trips between each OD pair.

In Table 7, the system and user optimal solutions are given. There is a 12% difference in the utilization of the network when the system objective at the user and system solutions are compared. Thus the traffic planner can utilize the toll pricing framework to make the user equilibrium flows be the same as the system optimal flows by imposing tolls on the transportation network. The system optimal $\bar{\theta}_1 = 37.68$, $\bar{\theta}_2 = 37.50$, $\bar{\theta}_3 = -1.09$ and $\bar{\theta}_4 = 0$. The $\zeta^{-1}(\ln \bar{t}_k + 1)$ is 0.36 for all OD pairs. For all utilized paths, the total path cost is equal to generalized user benefit $\bar{\theta} + \bar{\theta}_q - \zeta^{-1}(\ln \bar{t}_k + 1)$ for a commodity k.

Some alternative toll vectors using the toll set representation $W_{\mathrm{CDAM}}(\bar{v}, \bar{t}, \bar{\theta})$ are presented in Table 8. The total toll revenue is 1493.53 for each tolling scheme. Network users are charged 67.67% of their actual cost for the externalities they are imposing on others. This is achieved by 14 toll booths with MSCP tolls, 10 with D-MINSYS tolls, 11 with D-MINREV and MIN-MAX tolls, nine with MINTB and eight with MINTB/MINREV tolls.

6. Conclusion

This paper extends the first best toll pricing framework for the fixed demand traffic equilibrium model to the variable demand traffic assignment problems. Furthermore, this paper synthesizes earlier treatments of the elastic demand toll pricing theory, provides a unified toll pricing framework and analyzes toll pricing theory for CDAM, a special case of the GVD model. For the GVD model, all toll vectors in the first best toll set produce the same total toll revenue. It is important to emphasize that MSCP toll vector is not only an element of the toll set but also optimal when the objective is minimization of the total toll revenue. However, the major disadvantage of using MSCP tolls is that all of the congested links with traffic are being tolled and this is likely to be impractical due to high fixed costs and maintenance costs. By employing the toll pricing framework, a traffic engineer can provide several pricing alternatives to the transportation planning authorities. As a result, the authorities might achieve a better utilization of the roads by spreading the traffic over the system, and diverting trips to transit.

An interesting problem that should be considered is designing a toll pricing framework where (γ, θ) is not restricted to the system optimal multipliers. This pricing scheme enables the traffic planner not only control the constraint costs on the network but also design toll pricing schemes where the total toll revenue is not constant. This problem is very closely related to the one considered by Larsson and Patriksson (1994, 1995a,b, 1998) and Yang and Bell (1997). We will investigate this problem in a subsequent paper.

Acknowledgments

This research was partially supported by the National Science Foundation under grant no. DMII-9978642, and EPS-0236913 and matching support from the State of Kansas through Kansas Technology Enterprize Corporation.

References

Armstrong-Wright, A.T., 1986. Road pricing and user restraint: opportunities and constraints in developing countries. Transportation Research—A 20 (2), 123–127.

Arnott, R., Small, K., 1994. The economics of traffic congestion. American Scientist 82, 446–455.

Balasubramaniam, K., 1999. Marginal Cost Road Pricing in Combined Travel Choice Models. Master's Thesis, University of Illinois at Chicago.

Bazaraa, M.S., Sherali, H.D., Shetty, C.M., 1993. Nonlinear Programming: Theory and Algorithms, second ed. John Wiley & Sons, New York, NY.

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

Bergendorff, P., 1995. The Bounded Flow Approach to Congestion Pricing. University of Florida Center for Applied Optimization Report and Master's Thesis, Royal Institute of Technology, Stockholm.

Bergendorff, P., Hearn, D.W., Ramana, M.V., 1997. Congestion toll pricing of traffic networks. In: Pardalos, P.M., Hearn, D.W., Hager, W.W. (Eds.), Network Optimization. Springer-Verlag Series Lecture Notes in Economics and Mathematical Systems, pp. 51–71.

Boyce, D., Balasubramaniam, K., Tian, X., 2002. Implications of marginal social cost road pricing for urban travel choices and user benefits. In: Gendreau, M., Marcotte, P. (Eds.), Transportation and Network Analysis: Current Trends. Kluwer Academic Publishers, pp. 37–48.

CPLEX, 2001. ILOG Inc., Mountain View, CA.

Commuting. September 5th, 1998. The Economist, pp. 3–18.

Elliot, W., 1986. Fumbling toward the edge of history: California's quest for a road pricing experiment. Transportation Research—A 20 (2), 151–156.

Evans, S.P., 1976. Derivation and analysis of some models for combining trip distribution and assignment. Transportation Research 10, 37–57.

Ferrari, P., 1995. Road pricing and network equilibrium. Transportation Research—B 29 (5), 357–372.

GAMS, General Algebraic Modeling System, 1995. GAMS Development Corporation.

Hearn, D.W., 1980. Bounding flows in traffic assignment models. Research Report 80-4, Department of Industrial and Systems Engineering, University of Florida, Gainesville, FL.

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, pp. 109–124.

Hearn, D.W., Yildirim, M.B., 2002. A toll pricing framework for traffic assignment problems with elastic demand. In: Gendreau, M., Marcotte, P. (Eds.), Transportation and Network Analysis: Current Trends. Kluwer Academic Publishers, pp. 135–145.

Hearn, D.W., Yildirim, M.B., Ramana, M.V., Bai, L.H., 2001. Computational methods for congestion toll pricing models. The 4th International IEEE Conference on Intelligent Transportation Systems.

Johnson, B., 1964. On the economics of road congestion. Econometrica 32 (1–2), 137–150.

Larsson, T., Patriksson, M., 1994. Equilibrium characterizations of solutions to side constrained asymmetric traffic assignment models. Le Matematiche 49, 249–280.

Larsson, T., Patriksson, M., 1995a. An augmented lagrangean dual algorithm for link capacity side constrained traffic assignment problems. Transportation Research—B 29 (6), 443–455.

Larsson, T., Patriksson, M., 1995b. On side constrained models of traffic equilibria. In: Giannessi, F., Maugeri, A. (Eds.), Variational Inequalities and Network Equilibrium Problems, Proceedings of the International School of Mathematics G. Stampacchia 19th Course on Variational Inequalities and Network Equilibrium Problems, held June 19–25, 1994, in Erice, Italy. Plenum Press, New York, NY, pp. 169–178.

Larsson, T., Patriksson, M., 1998. Side constrained traffic equilibrium models-traffic management through link tolls. In: Marcotte, P., Nguyen, S. (Eds.), Equilibrium and Advanced Transportation Modeling. Kluwer Academic Publishers, pp. 125–151.

Larsson, T., Patriksson, M., 1999. Side constrained traffic equilibrium models—analysis, computation and applications. Transportation Research—B 33, 233–264.

Luk, J., Chung, E., 1997. Public acceptance and technologies for road pricing. Research Report, ARR 307, ARRB Transport Research.

Lundgren, J.T., Patriksson, M., 1998. An algorithm for the combined distribution and assignment problem. In: Bell, M.G.H. (Ed.), Transportation Networks: Recent Methadological Advances. Pergamon Press, Amsterdam, pp. 239–253.

Murtagh, A., Saunders, M.A., 1983. MINOS 5.0 Users Guide. Stanford University.

Pigou, A.C., 1920. The Economics of Welfare. Mac Millan, New York.

Schiller, E., 1998. The road ahead: the economic and environmental benefits of congestion pricing. Cache Copy.

Wardrop, J.G., 1952. Some theoretical aspects of road traffic research. Proceedings of the Institute of Civil Engineers. Part II. Engineering Divisions, Airport Maritime Railway Road 1, 325–378.

Yang, H., Bell, M.G.H., 1997. Traffic restraint, road pricing and network equilibrium. Transportation Research—B 31 (4), 303–314.

Yang, H., Huang, H., 1998. Principle of marginal-cost pricing: how does it work in a general road network? Transportation Research—A 32 (1), 45–54.