# Price Competition with Elastic Traffic *

Asuman Ozdaglar
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology

August 7, 2006

**Abstract**

In this paper, we present a combined study of price competition and traffic control in a congested network. We study a model in which service providers own the routes in a network and set prices to maximize their profits, while users choose the amount of flow to send and the routing of the flow according to Wardrop's principle. When utility functions of users are concave and have concave first derivatives, we characterize a tight bound of 2/3 on efficiency in pure strategy equilibria of the price competition game. We obtain the same bound under the assumption that there is no fixed latency cost, i.e., the latency of a link at zero flow is equal to zero. These bounds are tight even when the numbers of routes and service providers are arbitrarily large.

*I thank Daron Acemoglu for his comments and support in writing this paper.

# 1 Introduction

Many of today's large-scale networks, including the Internet, are not controlled by a single network planner with a single objective, but have emerged from the interconnection of multiple independent and autonomous (for-profit) administrative domains. The decentralized structure of these networks precludes regulation of traffic through centrally designed optimal control strategies. Two main changes from the traditional network control environment are particularly noteworthy. First, there is an emerging trend to consider a new routing paradigm ("source routing" or "selfish routing"), which allow end hosts to choose the amount and the routing of their own traffic for their "selfish" objective (i.e., to minimize their own costs).[1] Second, for-profit administrative domains will charge prices for transmission of information through their subnetworks (or routes).[2]

From the viewpoint of modeling and design of flow control and routing schemes, this has two fundamental implications. First, the system needs to be analyzed not as an optimization problem as in traditional network analyses, but as an equilibrium problem among multiple agents with different objectives. Second, the well-known potential inefficiencies related to game-theoretic interactions imply that the performance of the resulting equilibrium may be far worse than the optimum of a "social" objective (such as, minimization of total delay, maximization of total user satisfaction, or fairness). This has motivated a large literature investigating the so-called *price of anarchy*, which is a measure of performance degradation of the equilibrium relative to the social optimum (see [14]).

In this paper, we analyze price competition among multiple service providers in a parallel-link network with elastic traffic, i.e., in an environment featuring both flow control and routing by users. Our objective is both to provide a characterization of the equilibrium and to provide a bound on the extent of the inefficiency resulting from the interaction between selfish users and multiple profit-maximizing service providers in this context (i.e., to determine the price of anarchy).

More explicitly, we consider the following environment. Multiple users decide how much flow to send and how to route these flows across $I$ alternative (parallel) routes. More flow on a particular route causes delays, exerting a negative congestion externality on existing flows. Congestion costs are captured by a route-specific non-decreasing convex latency function, $l_i(\cdot)$, while the utility that users receive from transmitting flow is captured by a concave utility function $u(\cdot)$. Each route is owned by a different service provider, who sets a price (toll) of $p_i$, so that the effective unit cost of using route $i$ is $l_i(x_i) + p_i$, when the total flow on route $i$ is $x_i$. Given these costs, users choose the amount of flow and the routing pattern optimally. Service providers choose prices to maximize profits (anticipating users' behavior).

There is a large literature on models of congestion both in transportation and com-

[1] Selfish routing is first studied in the well-known paper by Roughgarden and Tardos [20]. Other papers in this literature include [7], [18], [9], and [21].

[2] A number of papers, in particular, [13], [15], [23] use prices in communication networks as control parameters. The case of for-profit pricing with multiple service providers and inelastic traffic was analyzed in [2]. Recent work by [11] analyzes the interaction between selfish users and pricing with multiple service providers in a model with elastic traffic.

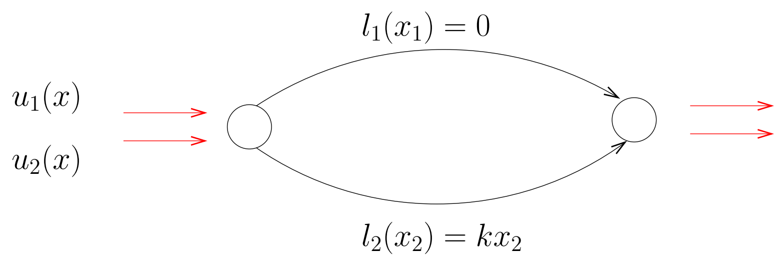


Figure 1: A two link network with congestion-dependant latency functions.

munication networks (e.g. [5], [17], [19], [16], [20]). However, very few studies have investigated the interaction between selfish-users and pricing by service providers. In particular, in [4], Basar and Srikant analyze pricing by a single service provider under specific assumptions on the utility and latency functions. He and Walrand [12] study competition and cooperation among internet service providers under specific demand models. Issues of efficient allocation of traffic across routes do not arise in these papers. The impact of the selfish user-service provider interaction on efficiency was first studied in [3] in a model with both flow control and routing, but a single service provider. Previous work in [2] studies competition among multiple service providers with inelastic traffic (i.e., a fixed user demand) and provides a tight bound of 5/6 on efficiency when latency without congestion is zero and a bound of $2 - 2\sqrt{2}$ on efficiency (which is quite close to 5/6) with possibly positive latency without congestion [i.e., $l_i(0) \geq 0$]. Finally, using a different mathematical approach, Hayrapetyan, Tardos and Wexler [11] provide the first analysis of an environment with multiple service providers and elastic traffic with a single user class. Their analysis provides non-tight bounds on the efficiency loss using the assumption that the utility function is concave and has a concave first derivative.[3] In this paper, we analyze the same environment, but using mathematical tools similar to those used in [2]. This approach enables us to provide a full characterization of pure strategy oligopoly equilibria and a tight bound on efficiency in such equilibria.

The next example illustrates the efficiency implications of price competition in networks with inelastic and elastic traffic.

**Example 1** Consider the two-route network illustrated in Figure 1. The latency functions are given by

$$l_1(x) = 0, \qquad l_2(x) = kx,$$

where $k$ is a positive scalar. For any amount of incoming flow, the efficient allocation [i.e., the routing that minimizes the total delay cost $\sum_i l_i(x_i)x_i$] routes all flow on route 1.

We first consider inelastic traffic and study the routing pattern that emerges at the equilibrium when the two routes are owned by different profit-maximizing service providers. We use the inelastic traffic model given in [2], i.e., we assume that two units of traffic will travel from origin to destination using either route 1 or route 2, and users

[3] For example, they provide the non-tight bound of 1/5.064 in general, and the bound of 1/3.125 for the case when latency without congestion is zero.

have a reservation utility $R = 3/4$ and decide not to send their traffic if the effective cost [i.e., $p_i + l_i(x_i)$] exceeds the reservation utility. This implies that the user preferences can be represented by the piecewise linear aggregate utility function depicted in Figure 1(a). It can be shown that the equilibrium prices take the form $p_i = x_i(l_1' + l_2')$, when the reservation utility is sufficiently high (or equivalently when $k$, the slope of $l_2$, is sufficiently small), and take the form $p_i = R - l_i(x_i)$, otherwise, i.e.,

$$p_i = \min \{R - l_i(x_i),\ x_i(l_1' + l_2')\}, \tag{1}$$

(see Proposition 9 in [2]). The resulting (Wardrop) equilibrium allocation that equates the effective costs on the two routes is given by $x_1 = 4/3$, $x_2 = 2/3$, when $k$ is sufficiently small, and $x_1 = 2 - (3/8k)$, $x_2 = (3/8k)$, otherwise. We take as our efficiency metric $r$ to be the ratio of the social surplus [defined as the difference between the user utility and total delay cost; see Eq. (7)] in the equilibrium and the maximum social surplus. It can be seen that for inelastic traffic, the efficiency metric $r$ satisfies $r = 5/6 \approx .833$ for $k = 9/16$ (i.e., the social surplus in the equilibrium is 5/6 of the maximum social surplus), and $r = .953$ for $k = 2$.

We next consider elastic traffic and assume that user preferences can be represented by the utility function

$$u(x) = \begin{cases} x & \text{if } 0 \le x \le 1, \\ -\frac{x^2}{2} + 2x - \frac{1}{2} & \text{if } 1 \le x \le 2, \\ \frac{3}{2} & \text{if } 2 \le x, \end{cases}$$

[see Figure 1(b)]. In this case, it can be shown that the equilibrium prices are given by

$$p_i = x_i \left[ l_i'(x_i) + \frac{1}{\frac{1}{l_j'(x_j)} - \frac{1}{u''(x_1+x_2)}} \right], \tag{2}$$

where $j \neq i$ and $u''$ denotes the second derivative of the utility function (see Proposition 2). The resulting equilibrium allocation is given by $x_1 = \frac{4k+4}{4k+3}$, and $x_2 = \frac{2}{4k+3}$. Hence, for elastic traffic, the efficiency metric $r$ takes the values $r \approx .88$ for $k = 9/16$, and $r \approx .78$ for $k = 2$.

The intuition for the inefficiency of price competition with inelastic traffic is due to a new source of differential monopoly power: when provider 1 charges a higher price for route 1, some traffic will be pushed from route 1 to route 2, raising the congestive level there and making it less attractive. As a result, the optimal price that each service provider charges includes an additional markup over the Pigovian markup, which are $x_1 l_2'$ for route 1 and $x_2 l_1'$ for route 2, distorting the traffic allocation away from the efficient allocation [cf. Eq. (1)]. Nevertheless, it was shown in [2] that, in an environment with inelastic traffic, the efficiency loss of any pure strategy equilibrium is bounded below by 5/6 when the latency at zero flow (traffic) is equal to 0. In the case of elastic traffic, there is an additional source of distortion in view of the change in the total demand in response to effective costs, which is similar to the distortion observed in standard

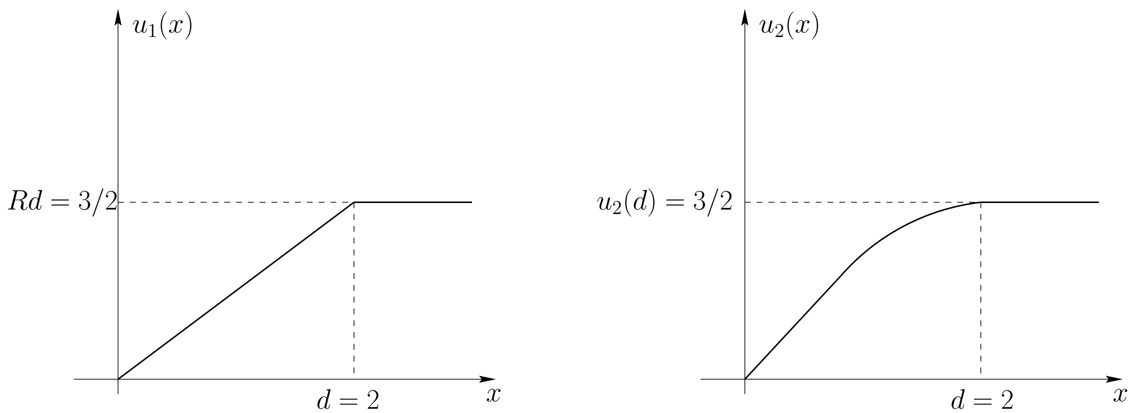


Figure 2: Aggregate utility functions representing user preferences for Example 1: part (a) illustrates the utility function for inelastic traffic, and part (b) illustrates the utility function for elastic traffic.

monopoly models. In this paper, we show that despite the additional distortion, it is possible to bound the efficiency loss under some assumptions on the utility function.

In particular, we show that a pure strategy equilibrium exists when all the latency functions are affine, and provide a characterization of equilibrium prices in any pure strategy equilibrium. Our main result is that when the utility function $u(\cdot)$ is not only concave but also has concave first derivative, we can provide a tight bound of $2/3$ on efficiency resulting from the interaction between selfish users and profit-maximizing service providers. Compared to the case of inelastic traffic, this is a worse bound. But what is remarkable is that the deterioration relative to the fixed demand case is not too large.[4]

The rest of the paper is organized as follows. Section 2 introduces the basic model. Section 3 introduces the notion of equilibrium and characterizes the equilibrium prices. Section 4 analyzes the efficiency of equilibrium and provides a tight bound of $2/3$ on efficiency under the assumption that latency of a link at zero flow is equal to zero. Section 5 relaxes this assumption and establishes that the efficiency bound of $2/3$ is also valid in this case. Section 6 contains our concluding remarks.

Regarding notation, all vectors are viewed as column vectors, and inequalities are to be interpreted componentwise. We denote by $\mathbb{R}^I_+$ the set of nonnegative $I$-dimensional vectors. Let $C$ be a closed subset of $[0, \infty)$ and let $f : C \mapsto \mathbb{R}$ be a convex function. We use the notation $f'(x)$ to denote a subgradient of the convex function $f$ at a point

[4]Nevertheless, it has to be emphasized that these bounds are for networks consisting of parallel structures. In [1], it was shown that the price of anarchy can become much worse in more general network topologies.

$x$. Note that for any $x$ and any subgradient $f'(x)$, we have

$$f^-(x) \leq f'(x) \leq f^+(x),$$

where $f^-$ and $f^+$ denote the left and right derivatives of the function $f$ at point $x$, respectively.

# 2 Model

We consider a network with $I$ parallel links. Let $\mathcal{I} = \{1, \ldots, I\}$ denote the set of links. Let $x_i$ denote the total flow on link $i$, and $x = [x_1, \ldots, x_I]$ denote the vector of link flows. Each link in the network has a flow-dependent latency function $l_i(x_i)$, which measures the travel time (or delay) as a function of the total flow on link $i$. We assume that each link is owned by a different service provider. We denote the price per unit flow (bandwidth) of link $i$ by $p_i$. Let $p = [p_1, \ldots, p_I]$ denote the vector of prices.

We assume that the user preferences can be represented by an aggregate utility function $u\Big(\sum_{i \in \mathcal{I}} x_i\Big)$ which represents the amount of utility gained from sending a total amount of flow $\sum_{i \in \mathcal{I}} x_i$ through the network.[5]

We adopt the following assumption on the latency functions and the utility function throughout the paper:

**Assumption 1**

(a) For each $i \in \mathcal{I}$, the latency function $l_i : [0, \infty) \mapsto [0, \infty)$ is convex, continuously differentiable [6], and strictly increasing.

(b) The utility function $u : [0, \infty) \mapsto [0, \infty)$ is concave, continuously differentiable, nondecreasing, satisfies $u(0) = 0$, $u'(0) < \infty$, and there exists some $d \geq 0$ such that $u(x) = u(d)$ for all $x \geq d$.

(c) The derivative of the utility function $u' : [0, d] \mapsto [0, \infty)$ is concave.

The convexity and the monotonicity assumptions on the latency functions and the utility function are standard (see Shenker [21]). The analysis below also extends to the case when the latency functions are nondecreasing, but we adopt the strictly increasing assumption here to simplify some of our arguments. We assume that there is a threshold on the total flow, denoted by $d$, over which the utility gained by flow transmission remains constant. The assumption that the function $u'$ is concave is non-standard. We will show in Section 4 that this assumption cannot be dispensed with in obtaining a nonzero bound for efficiency of our price competition model.

We also adopt the following assumption on the latency functions for some of our results:

[5] This assumption implies that all users are "homogeneous". We discuss possible extensions of this work to multiple user classes in the concluding section.

[6] We assume throughout this paper that the derivatives at 0 are right derivatives.

**Assumption 2** For each $i \in \mathcal{I}$, the latency function $l_i : [0, \infty) \mapsto [0, \infty)$ satisfies $l_i(0) = 0$.

The assumption of zero latency at zero flow implies that all latency is due to flow of traffic, and there are no fixed latency costs. This assumption would be a good approximation to communication networks where queueing delays are more substantial than propagation delays. It is adopted to simplify the discussion, especially the characterization of equilibrium prices in Proposition 2 below. We will discuss the implications of relaxing this assumption in Section 5, where we show that the same tight bound on the efficiency of oligopoly equilibria hold even when we relax this assumption.

We assume that at a price vector, the amount of flow and the distribution of flow across the links is given by the Wardrop Equilibrium (see [22]), which we define next.

**Definition 1** For a given price vector $p \geq 0$, a vector $x^{WE} \in \mathbb{R}^I_+$ is a *Wardrop equilibrium* (WE) if

$$x^{WE} \in \arg\max_{x \geq 0} \left\{ u\Big(\sum_{i \in \mathcal{I}} x_i\Big) - \sum_{i \in \mathcal{I}} (l_i(x_i^{WE}) + p_i)x_i \right\}. \tag{3}$$

We denote the set of WE at a given $p$ by $W(p)$.

Wardrop equilibrium is used extensively in modeling traffic behavior in transportation networks [5], [8], [17], and more recently in communication networks [20], [7]. In view of the concavity of the utility function (cf. Assumption 1), we have the following useful characterization of the WE.

**Lemma 1** For a given price vector $p \geq 0$, a vector $x^{WE} \in \mathbb{R}^I_+$ is a Wardrop equilibrium if and only if

$$\begin{aligned} l_i(x_i^{WE}) + p_i &= u'\Big(\sum_{j \in \mathcal{I}} x_j^{WE}\Big), \qquad \forall\, i \text{ with } x_i^{WE} > 0, \\ l_i(x_i^{WE}) + p_i &\geq u'\Big(\sum_{j \in \mathcal{I}} x_j^{WE}\Big), \qquad \forall\, i \in \mathcal{I}. \end{aligned} \tag{4}$$

This lemma states that for any price vector $p$, in a Wardrop equilibrium $x^{WE} \in W(p)$, *the effective costs*, defined as $l_i(x_i^{WE}) + p_i$, are equalized on all links with positive flows. The following properties of the WE are well-known (see for example, [5], [8]): Assume that the latency functions $l_i$ are convex, continuously differentiable, and nondecreasing. For any $p \geq 0$, the set $W(p)$ is nonempty. Moreover, the correspondence $W : \mathbb{R}^I_+ \rightrightarrows \mathbb{R}^I_+$ is upper semicontinuous. If we further assume that the $l_i$ are strictly increasing, then for any $p \geq 0$, the set $W(p)$ is a singleton and the function $W(p)$ is a continuous function.

We next define the social problem and the social optimum, which is the flow distribution that would be chosen by a planner that has full information and full control over the network.

**Definition 2** A flow vector $x^S$ is a *social optimum* if it is an optimal solution of the *social problem*

$$\max_{x\geq 0} \quad u\Big(\sum_{i\in\mathcal{I}} x_i\Big) - \sum_{i\in\mathcal{I}} l_i(x_i)x_i. \tag{5}$$

In view of Assumption 1, we have the following lemma.

**Lemma 2** A vector $x^S \in \mathbb{R}^I_+$ is a social optimum if and only if

$$\begin{aligned} l_i(x_i^S) + x_i^S l_i'(x_i^S) &= u'\Big(\sum_{j\in\mathcal{I}} x_j^S\Big), \qquad \forall\ i \text{ with } x_i^S > 0, \\ l_i(x_i^S) + x_i^S l_i'(x_i^S) &\geq u'\Big(\sum_{j\in\mathcal{I}} x_j^S\Big), \qquad \forall\ i \in \mathcal{I}. \end{aligned} \tag{6}$$

Note the presence of the term $x_i^S l_i'(x_i^S)$ on the left hand side of equation (6), which can be thought of as *the social marginal cost of transmission*. An additional unit of flow through link $i$ increases the cost of using link $i$ by an amount $l_i'(x_i^S)$ per unit flow, or a total of $x_i^S l_i'(x_i^S)$. An optimal allocation requires this social marginal cost to be internalized by the users in choosing their routing and total amount of transmission, which is what equation (6) ensures. The inefficiency of oligopoly equilibria will be due to the fact that the equilibrium prices will differ from the social marginal cost of transmission, $x_i^S l_i'(x_i^S)$.

For a given vector $x \in \mathbb{R}^I_+$, we define the value of the objective function in the social problem,

$$\mathbb{S}(x) = u\Big(\sum_{i\in\mathcal{I}} x_i\Big) - \sum_{i\in\mathcal{I}} l_i(x_i)x_i, \tag{7}$$

as the *social surplus*, i.e., the difference between the utility and the total latency.

# 3 Price Competition Equilibrium

We assume that each of the links is owned by a different service provider. Service provider $i$ charges a price $p_i$ per unit bandwidth on link $i$. Given the vector of prices of links owned by other service providers, $p_{-i} = [p_j]_{j\neq i}$, the profit of service provider $i$ is

$$\Pi_i(p_i, p_{-i}) = p_i x_i,$$

where $x$ is the Wardrop equilibrium at the price vector $(p_i, p_{-i})$, which by the assumption that the latency functions $l_i$ are strictly increasing (cf. Assumption 1), is uniquely defined, i.e., $x = W(p_i, p_{-i})$.

The objective of each service provider is to maximize profits. Since the demand and therefore profits depend on the prices set by other service providers, each service provider forms conjectures about the actions of other service providers, as well as the behavior of users, which, we assume, they do according to the Nash Equilibrium notion.

**Definition 3** A vector $(p^{OE}, x^{OE}) \geq 0$ is a (pure strategy) *Oligopoly Equilibrium* (OE) if $x^{OE} = W\left(p_i^{OE}, p_{-i}^{OE}\right)$ and for all $i \in \mathcal{I}$,

$$\Pi_i(p_i^{OE}, p_{-i}^{OE}) \geq \Pi_i(p_i, p_{-i}^{OE}), \qquad \forall\, p_i \geq 0. \tag{8}$$

We refer to $p^{OE}$ as the *OE price*.

We refer to the game among service providers as the *price competition game*.[7] In the next proposition, we establish the existence of a pure strategy OE for affine latency functions. Similar existence results were proven in [2] and [11] under different assumptions. The proof uses optimal price characterizations that will be provided below, and therefore is provided in the Appendix.

**Proposition 1** Let Assumption 1 hold, and assume further that the latency functions are affine functions, i.e., for each $i \in \mathcal{I}$, $l_i(x) = a_i x + b_i$ for some scalars $a_i > 0$ and $b_i \geq 0$. Then the price competition game has a pure strategy OE.

We next provide an explicit characterization of the OE prices, which will be essential in the efficiency analysis of the next section. We need the following lemma, which allows us to write the optimization problem for each service provider in terms of a set of equality constraints.

**Lemma 3** Let Assumption 1 and Assumption 2 hold. Let $(p^{OE}, x^{OE})$ be a pure strategy OE. If $p_{\bar{i}}^{OE} x_{\bar{i}}^{OE} > 0$ for some $\bar{i} \in \mathcal{I}$, then $p_i^{OE} x_i^{OE} > 0$ for all $i \in \mathcal{I}$.

**Proof.** Define $K = p_{\bar{i}}^{OE} + l_{\bar{i}}(x_{\bar{i}}^{OE})$, which is positive by assumption. Assume $p_i^{OE} x_i^{OE} = 0$ for some $i$. Consider the price $\tilde{p}_i = K - \epsilon > 0$ for some small $\epsilon > 0$. By Assumption 2, it follows that for the price vector $(\tilde{p}_i, p_{-i}^{OE})$ and any $\tilde{x} \in W(\tilde{p}_i, p_{-i}^{OE})$, we have $\tilde{x}_i > 0$. Hence, the service provider that owns link $i$ has an incentive to deviate to $\tilde{p}_i$ at which he will make positive profit, contradicting the fact that $(p^{OE}, x^{OE})$ is a pure strategy OE. **Q.E.D.**

**Proposition 2** Let Assumption 1 and Assumption 2 hold. Let $(p^{OE}, x^{OE})$ be an OE such that $p_{\bar{i}}^{OE} x_{\bar{i}}^{OE} > 0$ for some $\bar{i} \in \mathcal{I}$. Then, for all $i \in \mathcal{I}$, there exists a subgradient of the function $u'$ at $\sum_{j \in \mathcal{I}} x_j^{OE}$, denoted by $u''\Big(\sum_{j \in \mathcal{I}} x_j^{OE}\Big)$, such that

$$p_i^{OE} = \begin{cases} x_i^{OE} l_i'(x_i^{OE}), & \text{if } u''\Big(\sum_{j \in \mathcal{I}} x_j^{OE}\Big) = 0, \\ x_i^{OE} l_i'(x_i^{OE}) + \dfrac{x_i^{OE}}{\left(\sum_{j \neq i} \frac{1}{l_j'(x_j^{OE})}\right) - \left(\frac{1}{u''(\sum_j x_j^{OE})}\right)}, & \text{otherwise.} \end{cases} \tag{9}$$

[7]It is possible to extend the definition of the Oligopoly Equilibrium to price competition games with nondecreasing latency functions by requiring that the profit of provider $i$ is maximized over all Wardrop equilibria in the set $W(p_i, p_{-i}^{OE})$; see [2], where it was also shown that the oligopoly equilibria coincide with the subgame perfect equilibria in the price competition game.

**Proof.** By the assumption that $p_{\bar{i}}^{OE} x_{\bar{i}}^{OE} > 0$ for some $\bar{i} \in \mathcal{I}$, Lemma 3 implies that $p_i^{OE} x_i^{OE} > 0$ for all $i \in \mathcal{I}$. Then by the definition of a Wardrop Equilibrium (cf. Lemma 1), it follows that, for any $i \in \mathcal{I}$, the vector $(p_i^{OE}, x^{OE})$ is an optimal solution of the following problem:

$$
\begin{aligned}
\text{maximize}_{p_i, x \geq 0} \quad & p_i x_i \\
\text{subject to} \quad & p_i + l_i(x_i) = p_j^{OE} + l_j(x_j), \qquad \forall\, j \neq i, && (10) \\
& p_i + l_i(x_i) = u'\Big(\sum_{j \in \mathcal{I}} x_j\Big). && (11)
\end{aligned}
$$

We assign Lagrange multipliers $\lambda_j$ and $\mu$ to constraints in (10) and (11) respectively. By Assumption 1, the $l_i$ are continuously differentiable and $u'$ is concave. Therefore, using the first order optimality conditions, we obtain

$$
\begin{aligned}
x_i^{OE} + \sum_{j \neq i} \lambda_j + \mu = 0, \\
p_i^{OE} + \Big(\sum_{j \neq i} \lambda_j + \mu\Big) l_i'(x_i^{OE}) - \mu u''\Big(\sum_{j \in \mathcal{I}} x_j^{OE}\Big) = 0, \qquad (12) \\
-\lambda_j l_j'(x_j^{OE}) - \mu u''\Big(\sum_{j \in \mathcal{I}} x_j^{OE}\Big) = 0, \qquad \forall\, j \neq i,
\end{aligned}
$$

where $u''(\sum_{j \in \mathcal{I}} x_j^{OE})$ denotes a subgradient of the concave function $u'$ at the point $\sum_{j \in \mathcal{I}} x_j^{OE}$. If $u''(\sum_{j \in \mathcal{I}} x_j^{OE}) = 0$, it follows from the preceding relations that

$$p_i^{OE} = x_i^{OE} l_i'(x_i^{OE}).$$

Assume that $u''(\sum_{j \in \mathcal{I}} x_j^{OE}) < 0$. Since, by Assumption 1, the latency functions are strictly increasing, solving for $\mu$ in the above relations yields

$$\mu = \frac{\frac{x_i^{OE}}{u''(\sum_j x_j^{OE})}}{\Big(\sum_{j \neq i} \frac{1}{l_j'(x_j^{OE})}\Big) - \Big(\frac{1}{u''(\sum_j x_j^{OE})}\Big)},$$

Substituting for $\mu$ in Eq. (12), we obtain the desired result. **Q.E.D.**

# 4 Efficiency Analysis

In this section, we study the efficiency properties of pure strategy oligopoly equilibria. We consider price competition games that have pure strategy equilibria (this set includes, but is larger than, games with affine latency functions, see Section 3). We say that $(u, \{l_i\}_{i \in \mathcal{I}}) \in \mathcal{P}_I$ when the utility and latency functions satisfy Assumption 1 and Assumption 2, and the associated price competition game has a pure strategy OE. Let

$\overrightarrow{OE}(u, \{l_i\})$ denote the set of flow allocations at an OE. We define the efficiency metric at some $x^{OE} \in \overrightarrow{OE}(u, \{l_i\})$ as

$$r_I(u, \{l_i\}, x^{OE}) = \frac{u\left(\sum_{i \in \mathcal{I}} x_i^{OE}\right) - \sum_{i=1}^{I} l_i(x_i^{OE}) x_i^{OE}}{u\left(\sum_{i \in \mathcal{I}} x_i^{S}\right) - \sum_{i=1}^{I} l_i(x_i^{S}) x_i^{S}}, \tag{13}$$

where $x^S$ is a social optimum given the utility function $u$ and the latency functions $\{l_i\}_{i \in \mathcal{I}}$. In other words, our efficiency metric is the ratio of the social surplus in the oligopoly equilibrium relative to the surplus in the social optimum. Following the literature on the "price of anarchy", in particular [14], we are interested in the worst performance in an oligopoly equilibrium, so we look for a lower bound on $r_I(u, \{l_i\}, x^{OE})$, i.e.,

$$\inf_{(u,\{l_i\}) \in \mathcal{P}_I} \inf_{x^{OE} \in \overrightarrow{OE}(u,\{l_i\})} r_I(u, \{l_i\}, x^{OE}).$$

The next example illustrates that when Assumption 1, in particular, the assumption that $u'$ is concave, fails to hold, the efficiency metric can be arbitrarily close to 0.

**Example 2** Consider a network with a single link with latency function $l(x) = \epsilon x$, where $\epsilon$ is a small positive scalar. Assume that the aggregate utility function $u : [0, \infty) \mapsto [0, \infty)$ is given by

$$u(x) = \begin{cases} \frac{x^{1-\frac{1}{\alpha}}}{1-\frac{1}{\alpha}} - \frac{1}{\theta} x & \text{if } 0 \le x \le \theta^\alpha, \\ \theta^{\alpha-1} & \text{otherwise,} \end{cases}$$

where $\alpha$ and $\theta$ are scalars with $\alpha > 1$ and $\theta > 1$. This utility function is concave. Note that the threshold on the total flow, denoted by $d$ in Assumption 1, is given by $d = \theta^\alpha$. The derivative of the utility function $u'(x) = x^{-1/\alpha} - 1/\theta$ is not concave. As $\epsilon \to 0$, it can be shown that the oligopoly equilibrium is given by

$$x^{OE} = \left(\frac{\theta(\alpha - 1)}{\alpha}\right)^\alpha,$$

and the social optimum is given by

$$x^S = \theta^\alpha.$$

As $\alpha \to 1$, the efficiency metric for this example satisfies

$$\lim_{\alpha \to 1} r_1(u, l, x^{OE}) = \lim_{\alpha \to 1} \frac{u(x^{OE})}{u(x^S)} = \lim_{\alpha \to 1} (\alpha - 1)^{\alpha - 1} \frac{2\alpha - 1}{\alpha^\alpha} = 0.$$

In the subsequent analysis, we show that when Assumption 1 holds, there exists a tight bound on the efficiency metric $r_I(u, \{l_i\}, x^{OE})$. The next lemma allows us to use the oligopoly price characterization given in Proposition 2 in bounding the efficiency metric.

**Lemma 4** For utility function $u$ and latency functions $\{l_i\}_{i\in\mathcal{I}}$, let Assumption 1 hold. Let $(p^{OE}, x^{OE})$ be an OE such that $p_i^{OE} x_i^{OE} = 0$ for all $i \in \mathcal{I}$. Then $x^{OE}$ is a social optimum.

**Proof.** We first show that $x_i^{OE} = 0$ for all $i \in \mathcal{I}$. Assume that $x_j^{OE} > 0$ for some $j$, which implies that $p_j^{OE} = 0$. Since by Assumption 1, the latency functions are strictly increasing, it follows that for all $p \geq 0$, the set of Wardrop equilibria $W(p)$ is a singleton and a continuous function of $p$. Therefore, there exists some $\epsilon > 0$ such that the Wardrop equilibrium $x^\epsilon \in W(\epsilon, p_{-j}^{OE})$ satisfies $x_j^\epsilon > 0$, yielding a positive profit of $\epsilon x_j^\epsilon$ for provider $j$, contradicting the assumption that $(p^{OE}, x^{OE})$ is an OE.

If $l_i(0) \geq u'(0)$ for all $i \in \mathcal{I}$, then by Lemma 2, it follows that the vector $x^{OE} = 0$ is a social optimum. Suppose that $l_j(0) < u'(0)$ for some $j$. Let $\epsilon > 0$ be some scalar such that $l_j(0) + \epsilon < u'(0)$. It can be seen that at the price vector $(\epsilon, p_{-j}^{OE})$, the Wardrop equilibrium $x^\epsilon \in W(\epsilon, p_{-j}^{OE})$ satisfies $x_j^\epsilon > 0$, yielding a positive profit of $\epsilon x_j^\epsilon$ for provider $j$, and contradicting the assumption that $(p^{OE}, x^{OE})$ is an OE. **Q.E.D.**

Note that Assumption 2 is not used in establishing the preceding result, i.e., the result holds with differential levels of $l_i(0)$. This lemma will be used in Section 5 in providing a bound on the efficiency metric without Assumption 2.

The next lemma shows that the total flow at an oligopoly equilibrium is always less than or equal to the total flow at a social optimum. This is an intuitive result; in an oligopoly equilibrium, profit-maximizing service providers charge markups over and above the social marginal cost of transmission, $x_i^S l_i'(x_i^S)$, reducing the total amount of equilibrium transmission below the socially-optimal level. This relationship will be useful in providing a bound on the efficiency of OE in the next section.

**Lemma 5** For utility function $u$ and latency functions $\{l_i\}_{i\in\mathcal{I}}$, let Assumption 1 hold. Let $(p^{OE}, x^{OE})$ be an OE such that $p_j^{OE} x_j^{OE} > 0$ for some $j \in \mathcal{I}$, and $x^S$ be a social optimum. Then

$$\sum_{i\in\mathcal{I}} x_i^{OE} \leq \sum_{i\in\mathcal{I}} x_i^S.$$

**Proof.** Assume to arrive at a contradiction that $\sum_{i\in\mathcal{I}} x_i^{OE} > \sum_{i\in\mathcal{I}} x_i^S$. This implies that there exists some $j \in \mathcal{I}$ such that

$$x_j^{OE} > x_j^S. \tag{14}$$

By the concavity of the $u$, we have

$$u'\Big(\sum_{i\in\mathcal{I}} x_i^{OE}\Big) \leq u'\Big(\sum_{i\in\mathcal{I}} x_i^S\Big).$$

Using Lemmas 1 and 2 together with $x_j^{OE} > 0$, we obtain

$$l_j(x_j^{OE}) + p_j^{OE} \leq l_j(x_j^S) + x_j^S l_j'(x_j^S).$$

Using the oligopoly price characterization [cf. Eq. (9)] and the fact that $u''(x) \leq 0$ for all $x$, it follows that $p_j^{OE} \geq x_j^{OE} l_j'(x_j^{OE})$. Substituting this in the preceding relation, we obtain

$$l_j(x_j^{OE}) + x_j^{OE} l_j'(x_j^{OE}) \leq l_j(x_j^S) + x_j^S l_j'(x_j^S).$$

Since, for all $i \in \mathcal{I}$, the latency function $l_i(x_i)$ is convex and nondecreasing, it follows that the function $x_i l_i(x_i)$ is convex, and therefore the function $l_i(x_i) + x_i l_i'(x_i)$ is nondecreasing. Hence, the preceding relation implies that $x_j^{OE} \leq x_j^S$, a contradiction. **Q.E.D.**

## 4.1 Efficiency Bound for Two Links:

We first consider a parallel link network with two links owned by two service providers and provide a bound on the efficiency metric. In addition to being an interesting result on its own, the result for parallel link networks with multiple links will be proven by reducing the proof to the two link case.

Our main result is given in the following theorem, which provides a tight bound of 2/3 on the efficiency metric. The structure of the proof is similar to the proof of Theorem 1 in [2]. The problem of finding a lower bound on $r_2(u, \{l_i\}, x^{OE})$ is an infinite-dimensional problem, since the minimization is over utility and latency functions. The proof first lower-bounds the infinite-dimensional problem by the optimal value of a finite-dimensional optimization problem using the relations between the flows at social optimum and equilibrium, and convexity of the latency functions. It then shows that the solution will involve one of the links having zero latency. Finally, using this fact, the price characterization from Proposition 2, and concavity of the utility and derivative of the utility function, it reduces the problem of characterizing the bound on inefficiency to a simple minimization problem, with optimal value 2/3. Note that the concavity of the derivative of the utility function is a key assumption in our proof.

Recall that $\mathcal{P}_2$ is the set of utility and latency functions that satisfy Assumption 1 and Assumption 2, and for which the associated price competition game has a pure strategy OE.

**Theorem 1** Consider a two link network where each link is owned by a different provider. Then

$$\inf_{(u,\{l_i\}) \in \mathcal{P}_2} \inf_{x^{OE} \in \overrightarrow{OE}(u,\{l_i\})} r_2(u, \{l_i\}, x^{OE}) = \frac{2}{3}. \tag{15}$$

**Proof.** The proof follows a number of steps:

*Step 1:* We are interested in finding a lower bound for the problem

$$\inf_{(u,\{l_i\}) \in \mathcal{P}_2} \inf_{x^{OE} \in \overrightarrow{OE}(u,\{l_i\})} r_2(u, \{l_i\}, x^{OE}). \tag{16}$$

Given $(u, \{l_i\}) \in \mathcal{P}_2$, let $x^{OE} \in \overrightarrow{OE}(u, \{l_i\})$ and let $x^S$ be a social optimum. By Lemma 5, we have $\sum_{i=1}^2 x_i^{OE} \leq \sum_{i=1}^2 x_i^S$. This implies that there exists some $j$ such that $x_j^{OE} < x_j^S$

(otherwise, we would have $x_j^{OE} = x_j^S$ for all $j$). Without loss of generality, we restrict ourselves to $(u, \{l_i\}) \in \mathcal{P}_2$ such that $x_1^{OE} < x_1^S$, i.e., we consider $(u, \{l_i\}) \in \mathcal{P}_2$ such that $x_1^{OE} \leq x_1^S - \epsilon$ for some $\epsilon > 0$. We claim that optimal value of problem (16) can be lower bounded by

$$\inf_{(u,\{l_i\})\in\mathcal{P}_2} \inf_{x^{OE}\in\overrightarrow{OE}(u,\{l_i\})} r_2(u, \{l_i\}, x^{OE}) \geq \inf_{\epsilon>0} r_2^{OE}(\epsilon), \tag{17}$$

where $r_2^{OE}(\epsilon)$ denotes the optimal value of problem $\mathrm{E}_2^\epsilon$ given by

$$r_2^{OE}(\epsilon) = \quad \underset{\substack{0\leq u\leq u^S\leq R,\ u'\geq 0,\ u''\leq 0 \\ l_i^S,\ (l_i^S)'\geq 0 \\ l_i,\ l_i',\ l_i^0\geq 0 \\ y_i^S,\ y_i\geq 0 \\ R,\ d\geq 0}}{\text{minimize}} \quad \frac{u - l_1 y_1 - l_2 y_2}{u^S - l_1^S y_1^S - l_2^S y_2^S} \tag{$\mathrm{E}_2^\epsilon$}$$

$$\begin{aligned}
\text{subject to} \quad & l_i^S \leq y_i^S (l_i^S)', \qquad i = 1, 2, && (18)\\
& l_i \leq y_i l_i', \qquad i = 1, 2, && (19)\\
& l_1^S + y_1^S (l_1^S)' \leq l_2^S + y_2^S (l_2^S)', && (20)\\
& \sum_{i=1}^{2} y_i^S \leq d, && (21)\\
& l_1 + l_1'(y_1^S - y_1) \leq l_1^S, && (22)\\
& y_1 \leq y_1^S - \epsilon && (23)\\
& \sum_{i=1}^{2} y_i \leq d && (24)\\
& + \text{ Oligopoly Equilibrium Constraints.}
\end{aligned}$$

Problem ($\mathrm{E}^\epsilon$) is a finite-dimensional problem in which instead of optimizing over the entire function $l_i$, we optimize over the possible values of $l_i(\cdot)$ and $l_i'(\cdot)$ at the equilibrium and the social optimum, which we denote by $l_i, l_i'$ and $l_i^S, (l_i^S)'$ respectively, and also over the possible values of $u(\cdot)$ at the social optimum, which we denote by $u^S$, and the possible values of $u(\cdot)$, $u'(\cdot)$, $u''(\cdot)$ at the equilibrium, which we denote by $u, u', u''$. The variables $y_i$ and $y_i^S$ denote the values of flows on link $i$ at the equilibrium and the social optimum. The decision variable $d$ denotes the total flow threshold in Assumption 1, i.e., the value of total flow over which the utility function remains at a constant value, denoted by $R = u(d)$ in this formulation (thus we have the constraint $0 \leq u \leq u^S \leq R$).

The constraints of problem ($\mathrm{E}^\epsilon$) capture the relations that must be satisfied between these values. In particular:

- Since the latency functions $l_i$ are convex and $l_i(0) = 0$, we have

$$l_i(x) \leq x l_i'(x), \qquad \forall\ x \geq 0,$$

  (see Figure 1). The constraints in (18) and (19) express the preceding relation at the social optimum and the equilibrium.

- The constraint in (20) follows by the necessary optimality conditions for a social optimum and the assumption that $x_1^{OE} < x_1^S$, which implies that $x_1^S > 0$.

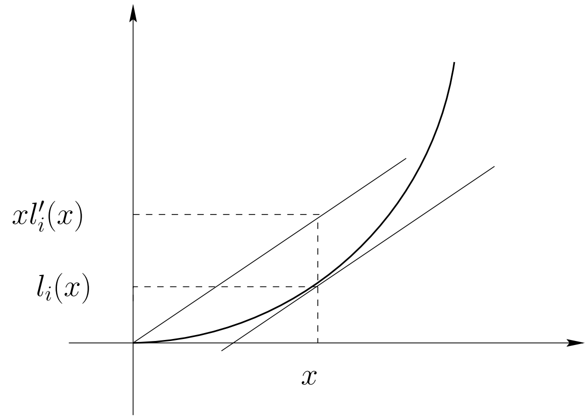


Figure 3: Overestimating a convex latency function by a linear approximation.

- By the convexity of the $l_i$, we also have $l_1(x_1^{OE}) + l_1'(x_1^{OE})(x_1^S - x_1^{OE}) \leq l_1(x_1^S)$, which is expressed in (22).
- Finally, the oligopoly equilibrium constraints represent the relations between $l_i(\cdot)$, $l_i'(\cdot)$, and $u(\cdot)$, $u'(\cdot)$, $u''(\cdot)$ at the oligopoly equilibrium as expressed in the price characterization given in Proposition 2.

Given any feasible solution of problem (15), there exists some $\epsilon > 0$ and a feasible solution for problem ($\mathrm{E}^\epsilon$) with the same objective function value. Therefore the optimum value for problem $\inf_{\epsilon>0} r_2^{OE}(\epsilon)$ is indeed a lower bound on the optimum value of problem (16).

*Step 2:* Let $(\bar{u}^S, \bar{u}, \bar{u}', \bar{u}'', \bar{l}_i^S, (\bar{l}_i^S)', \bar{l}_i, \bar{l}_i', \bar{y}_i^S, \bar{y}_i^{OE})$ denote an optimal solution of problem ($\mathrm{E}^\epsilon$). Following a similar argument to that used in the proof of Theorem 1 in [2], we can show that $\bar{l}_i^S = 0$ for $i = 1, 2$. Moreover, it is straightforward to see that $\bar{u}^S = R$.

*Step 3:* Since $\bar{l}_i^S = 0$ for $i = 1, 2$, we must have $\bar{l}_1 = 0$ [cf. Eq. (22)] and $\bar{l}_1' = 0$ [cf. Eqs. (22) and (23)]. Substituting these in problem ($\mathrm{E}^\epsilon$) and expressing the oligopoly equilibrium constraints using the price characterization given in Eq. (2) as $l_1' \to 0$, we obtain for all $\epsilon > 0$

$$r_2^{OE}(\epsilon) \geq \min_{\substack{0\leq u\leq R,\ u'\geq 0,\ u''\leq 0 \\ l_2,\ l_2'\geq 0 \\ y_1,\ y_2\geq 0 \\ R,\ d\geq 0}} \frac{u - l_2 y_2}{R} \tag{25}$$

$$\text{subject to} \quad l_2 \leq y_2 l_2', \tag{26}$$

$$l_2 + y_2 l_2' = \frac{y_1}{\frac{1}{l_2'} - \frac{1}{u''}}, \tag{27}$$

$$l_2 + y_2 l_2' = u', \tag{28}$$

$$u \geq u'(y_1 + y_2), \tag{29}$$

$$u' \geq -u''(d - (y_1 + y_2)), \tag{30}$$

$$R \leq u + u'(d - (y_1 + y_2)) + \frac{u''}{2}(d - (y_1 + y_2))^2, \quad (31)$$

$$\sum_{i=1}^{2} y_i \leq d.$$

In addition to expressing the constraints given by the price characterization in Eq. (2), the preceding formulation incorporates relations that must be satisfied due to the concavity assumptions on the utility function (cf. Assumption 1). In particular:

- Since the utility function is concave and continuously differentiable, we must have

$$u(x) \geq xu'(x), \qquad \forall\ x \geq 0,$$

  [see Figure 2(a)]. The constraint in (29) expresses the preceding relation at the oligopoly equilibrium, $x = x_1^{OE} + x_2^{OE} = y_1 + y_2$.

- Since the derivative of the utility function $u'$ is concave, we must have

$$u'(x) \geq -(d - x)u''(x), \qquad \forall\ x \geq 0,$$

  where $u''(x)$ denotes a subgradient of the concave function $u'$ at the point $x$ [see Figure 2(b)]. The constraint in (30) expresses the preceding relation at the oligopoly equilibrium, $x = x_1^{OE} + x_2^{OE} = y_1 + y_2$.

- Finally, we express the relation between $R = u(d)$ and the values $u(x)$, $u'(x)$, and $u''(x)$. Using the concavity of the function $u'$, for all $0 \leq x \leq d$, we obtain

$$\begin{aligned} R &= u(x) + \int_x^d u'(z)dz \\ &\leq u(x) + u'(x)(d - x) + \frac{u''(x)}{2}(d - x)^2, \end{aligned}$$

  (cf. Figure 3). The constraint in (31) expresses the preceding relation at the oligopoly equilibrium, $x = x_1^{OE} + x_2^{OE} = y_1 + y_2$.

We show that the optimal value of Problem (25) is 2/3. It is immediate to see that at an optimal solution, the constraints (29) and (31) are binding. We also assume that the constraints (26) and (30) are binding at an optimal solution [8] By Eqs. (26) and (28), we have $u' = 2l_2$. Combined with Eq. (30), this yields $y_1 + y_2 = d + \frac{2l_2}{u''}$. Together with Eq. (27), this implies that

$$y_2 = \frac{1}{3}\left[d + \frac{4l_2}{u''}\right].$$

Since $u'' < 0$, this imposes an upper bound on $l_2$, i.e.,

$$l_2 \leq \frac{-u''d}{4}. \qquad (32)$$

[8] The following argument can be repeated using slack variables for these constraints. Optimizing over the slack variables yields the same optimal value.

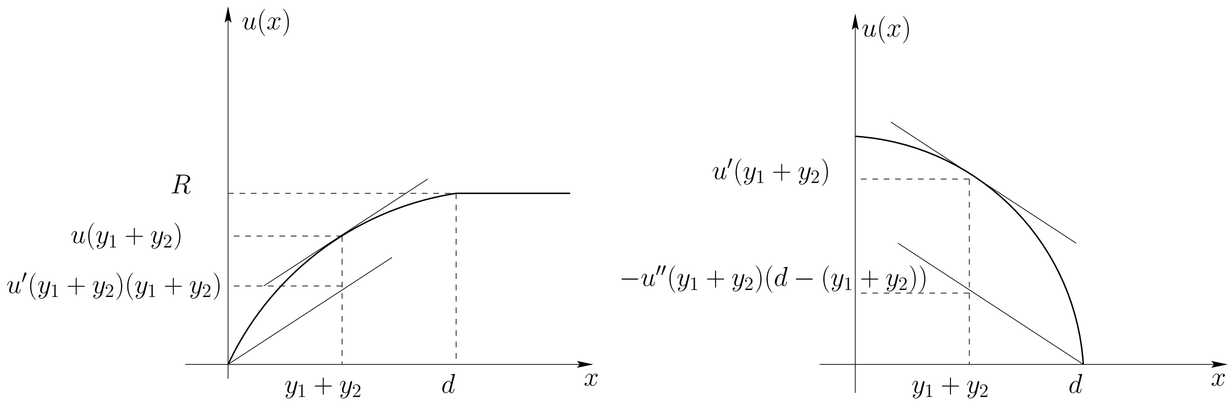


Figure 4: Underestimating a concave utility function and its concave derivative by linear approximations.

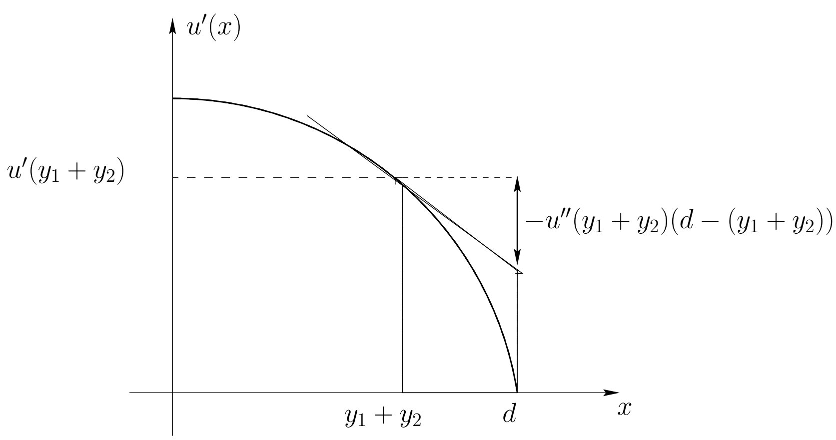


Figure 5: Overestimating the area under the function $u'(x)$.

Substituting these relations in (29) and (31), we obtain

$$u = 2l_2 d + \frac{4l_2^2}{u''},$$

$$R = 2l_2 d + \frac{2l_2^2}{u''}.$$

Hence, the value of the objective function is given by

$$\frac{u - l_2 y_2}{R} = \frac{2l_2 d + \frac{4l_2^2}{u''} - \frac{l_2}{3}\left[d + \frac{4l_2}{u''}\right]}{2l_2 d + \frac{2l_2^2}{u''}}.$$

It can be seen that the above expression is strictly decreasing in $l_2$, therefore the minimum value is attained at $l_2 = \frac{-u''d}{4}$ [cf. Eq. (32); this also implies that at the optimum $\bar{y}_2 = 0$ and $\bar{y}_1 = d/2$]. Substituting in the preceding relation, we obtain

$$\min_{0 \le l_2 \le \frac{-u''d}{4}} \frac{u - l_2 y_2}{R} = \frac{d/2}{3d/4} = \frac{2}{3}.$$

It follows that

$$\inf_{\{l_i\} \in \mathcal{L}_2} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_i\})} r_2(u, \{l_i\}, x^{OE}) \ge \frac{2}{3}.$$

We next show that this bound is tight. Consider the latency functions

$$l_1(x) = \epsilon x, \qquad l_2(x) = kx,$$

where $\epsilon$ is a small positive scalar and $k$ is an arbitrarily large positive scalar, and the utility function

$$u(x) = \begin{cases} x & \text{if } 0 \le x \le 1, \\ -\frac{x^2}{2} + 2x - \frac{1}{2} & \text{if } 1 \le x \le 2, \\ \frac{3}{2} & \text{if } 2 \le x. \end{cases}$$

Using the price characterization in (9), it can be shown that as $k \to \infty$ and $\epsilon \to 0$, the OE flow $x^{OE} \to (1, 0)$ and the social optimum $x^S \to (2, 0)$. Hence the efficiency metric for this problem is $r_2(u, \{l_i\}, x^{OE}) \to 2/3$ as $k \to \infty$, thus showing that

$$\inf_{(u,\{l_i\}) \in \mathcal{P}_2} \inf_{x^{OE} \in \overrightarrow{OE}(u,\{l_i\})} r_2(u, \{l_i\}, x^{OE}) = \frac{2}{3}.$$

**Q.E.D.**

The preceding result shows that the efficiency loss is maximized when one of the links, say link 1, has a latency function that is identically equal to 0 (or a latency function that is arbitrarily close to 0 in view of the assumption that the latency functions are strictly increasing), and consequently, at the social optimum, all $d$ units of traffic is admitted and routed on link 1. This is analogous to the worst case for the inelastic traffic (see [2]), in which the loss is maximized when link 1 has 0 latency and as much of the total traffic as possible goes through link 2 in the equilibrium, leading to a loss of 1/6. Hence, with inelastic traffic, the efficiency loss is mainly due to the difference between the marginal latency costs of the two links leading to a misallocation of traffic. More interesting here is the fact that with elastic traffic, the worst case efficiency loss is realized with a very high latency function for link 2, essentially shutting down all traffic on link 2, and reducing the amount of total admitted traffic to $d/2$, which leads to a higher efficiency loss of 1/3.

## 4.2 Efficiency Bound for Multiple Links

We next consider the general case where we have a parallel link network with $I$ links, where each link is owned by a different provider. The next theorem generalizes Theorem 1 to a parallel link network with $I \geq 2$ links. The proof is based on reducing the problem into the case with two links.

**Theorem 2** Consider a parallel link network with $I$ links, where each link is owned by a different provider. Then

$$\inf_{(u,\{l_i\})\in\mathcal{P}_I} \inf_{x^{OE}\in\overrightarrow{OE}(u,\{l_i\})} r_I(u,\{l_i\},x^{OE}) = \frac{2}{3}. \tag{33}$$

**Proof.** As in the proof of Theorem 1, we have

$$\inf_{(u,\{l_i\})\in\mathcal{P}_I} \inf_{x^{OE}\in\overrightarrow{OE}(u,\{l_i\})} r_I(u,\{l_i\},x^{OE}) \geq \inf_{\epsilon>0} r_I^{OE}(\epsilon), \tag{34}$$

where $r_I^{OE}(\epsilon)$ denotes the optimal value of problem $\mathrm{E}_I^\epsilon$ given as

$$r_I^{OE}(\epsilon) = \underset{\substack{0\leq u\leq u^S\leq R,\ u'\geq 0,\ u''\leq 0 \\ l_i^S,\ (l_i^S)'\geq 0 \\ l_i,\ l_i',\ l_i^0\geq 0 \\ y_i^S,\ y_i\geq 0 \\ R,\ d\geq 0}}{\text{minimize}} \quad \frac{u-\sum_{i\in\mathcal{I}} l_i y_i}{u^S-\sum_{i\in\mathcal{I}} l_i^S y_i^S} \tag{$\mathrm{E}_I^\epsilon$}$$

$$\begin{aligned}
\text{subject to} \quad & l_i^S \leq y_i^S (l_i^S)', & i=1,\ldots,I, \quad (35)\\
& l_i \leq y_i l_i', & i=1,\ldots,I, \quad (36)\\
& l_1^S + y_1^S (l_1^S)' \leq l_i^S + y_i^S (l_i^S)', & i=2,\ldots,I,\\
& \sum_{i\in\mathcal{I}} y_i^S \leq d, &\\
& l_1 + l_1'(y_1^S - y_1) \leq l_1^S, &\\
& y_1 \leq y_1^S - \epsilon, &\\
& \sum_{i\in\mathcal{I}} y_i \leq d &\\
& + \text{Oligopoly Equilibrium Constraints.} &
\end{aligned}$$

Let $(\bar{u}^S, \bar{u}, \bar{u}', \bar{u}'', \bar{l}_i^S, (\bar{l}_i^S)', \bar{l}_i, \bar{l}_i', \bar{y}_i^S, \bar{y}_i^{OE})$ denote an optimal solution of problem $(\bar{E})$. Similar to the proof of Theorem 1, it can be shown that $\bar{l}_i^S = 0$ for all $i = 1, \ldots, I$, $\bar{l}_1 = 0$, $\bar{l}_1' = 0$, and $\bar{u}^S = R$. Substituting these in problem $\mathrm{E}_I^\epsilon$, we obtain for all $\epsilon > 0$

$$
\begin{aligned}
r_I^{OE}(\epsilon) \geq \quad & \text{minimize}_{\substack{u,\ u' \geq 0,\ u'' \leq 0 \\ l_2,\ l_2' \geq 0 \\ y_1,\ y_2 \geq 0 \\ R,\ d \geq 0}} && \frac{u - \sum_{i=2}^I l_i y_i}{R} && && (37) \\
& \text{subject to} && l_i \leq y_i l_i', && i = 2, \ldots, I, && (38) \\
& && l_i + y_i l_i' = \frac{y_1}{\left(\sum_{i=2}^I \frac{1}{l_i'} - \frac{1}{u''}\right)}, && i = 2, \ldots, I, && (39) \\
& && l_2 + y_2 l_2' = u' && && (40) \\
& && u \geq u' \sum_{i \in \mathcal{I}} y_i \\
& && u' \geq -u'' \left(d - \sum_{i \in \mathcal{I}} y_i\right) \\
& && R \leq u + u' \left(d - \sum_{i \in \mathcal{I}} y_i\right) + \frac{u''}{2} \left(d - \sum_{i \in \mathcal{I}} y_i\right)^2 \\
& && \sum_{i \in \mathcal{I}} y_i \leq d.
\end{aligned}
$$

Using the first order optimality conditions, it can be shown that for all $i = 2, \ldots, I$, constraint (38) is binding (see the analogous argument in the proof of Theorem 2 in [2]). Hence, $\bar{l}_i = \bar{y}_i \bar{l}_i'$ for all $i = 2, \ldots, I$, which by Eqs. (28) and (40) implies that $\bar{l}_i = \bar{y}_i \bar{l}_i' = \frac{\bar{u}'}{2}$. We use the transformation of variables

$$y = \sum_{i=2}^I y_i, \quad \text{and} \quad \frac{1}{l'} = \sum_{i=2}^I \frac{1}{l_i'}.$$

Using $\bar{l}_i' = \frac{\bar{u}'}{2} \frac{1}{\bar{y}_i}$, we obtain

$$\frac{1}{\bar{l}'} = \frac{2}{\bar{u}'} \sum_{i=2}^I \bar{y}_i = \frac{2}{\bar{u}'} \bar{y},$$

and therefore,

$$\bar{y} \bar{l}' = \frac{\bar{u}'}{2} = \bar{y}_i \bar{l}_i', \qquad \forall\ i = 2, \ldots, I.$$

Substituting $l_i = l$ for all $i = 2, \ldots, I$ and using the preceding relations, it follows that the optimal value of problem (37) is the same as the optimal value of the following problem:

$$\text{minimize}_{\substack{0 \leq u \leq R,\ u' \geq 0,\ u'' \leq 0 \\ l,\ l' \geq 0 \\ y_1,\ y \geq 0 \\ R,\ d \geq 0}} \qquad \frac{u - ly}{R}$$

$$
\begin{aligned}
\text{subject to} \quad & l \le y l', \\
& l + y l' = \frac{y_1}{\frac{1}{l'} - \frac{1}{u''}}, \\
& l + y l' = u', \\
& u \ge u'(y_1 + y), \\
& u' \ge -u''(d - (y_1 + y)), \\
& R \le u + u'(d - (y_1 + y)) + \frac{u''}{2}(d - (y_1 + y))^2, \\
& y_1 + y \le d.
\end{aligned}
$$

which is identical to problem (25) in the two-link case, showing that

$$
\inf_{(u,\{l_i\}) \in \mathcal{P}_I} \inf_{x^{OE} \in \overrightarrow{OE}(u,\{l_i\})} r_I(u, \{l_i\}, x^{OE}) \ge \frac{2}{3}.
$$

We next show that the bound is tight. Consider an $I$-link network, where each link is owned by a different provider. Assume that the utility function is given by

$$
u(x) = \begin{cases} x & \text{if } 0 \le x \le 1, \\ -\frac{x^2}{2} + 2x - \frac{1}{2} & \text{if } 1 \le x \le 2, \\ \frac{3}{2} & \text{if } 2 \le x, \end{cases}
$$

and the latency functions are given by

$$
l_1(x) = \epsilon x, \qquad l_i(x) = kx, \qquad i = 2, \ldots, I,
$$

where $\epsilon$ is a small positive scalar and $k$ is an arbitrarily large positive scalar. Then it can be verified that as $\epsilon \to 0$ and $k \to \infty$, we have $r_I(u, \{l_i\}, x^{OE}) \to 2/3$. **Q.E.D.**

A notable feature of this result is that the tight bound on the efficiency metric is independent of the number of links $I$.

# 5 Bound for Positive Latency at Zero Flow

In this section, we relax Assumption 2, i.e., $l_i(0) = 0$ for all $i \in \mathcal{I}$. Using Lemma 4, we focus on oligopoly equilibria $(p^{OE}, x^{OE})$ such that $p_i^{OE} x_i^{OE} > 0$ for some $i \in \mathcal{I}$ (otherwise the efficiency metric is equal to 1). We first provide an equilibrium price characterization, which generalizes Proposition 2.

**Proposition 3** Let $(p^{OE}, x^{OE})$ be an OE such that $p_i^{OE} x_i^{OE} > 0$ for some $i \in \mathcal{I}$. Define the index set

$$
\mathcal{N} = \{ j \in \mathcal{I} \mid p_i^{OE} + l_i(x_i^{OE}) < p_j^{OE} + l_j(0) \}. \tag{41}
$$

Let Assumption 1 hold. Then, for all $i \in \mathcal{I}$, there exists a subgradient of the function $u'$ at $\sum_{j\in\mathcal{I}} x_j^{OE}$, denoted by $u''\Big(\sum_{j\in\mathcal{I}} x_j^{OE}\Big)$, such that

$$p_i^{OE} = \begin{cases} x_i^{OE} l_i'(x_i^{OE}), & \text{if } u''\Big(\sum_{j\in\mathcal{I}} x_j^{OE}\Big) = 0, \\ & \text{or } \mathcal{I} - \mathcal{N} - \{i\} = \emptyset, \\ x_i^{OE} l_i'(x_i^{OE}) + \dfrac{x_i^{OE}}{\Big(\sum_{\substack{j\neq i \\ j\notin\mathcal{N}}} \frac{1}{l_j'(x_j^{OE})}\Big) - \Big(\frac{1}{u''(\sum_j x_j^{OE})}\Big)}, & \text{otherwise.} \end{cases} \tag{42}$$

The proof of this proposition follows immediately from the proof of Proposition 2. In particular, $\mathcal{N}$ is the set of all latencies where $x_j^{OE} = 0$, so that any $j \in \mathcal{N}$ can be discarded when considering the individual optimization problem of each service provider. In what follows, let $\mathcal{P}_I^*$ denote the set of utility and latency functions that satisfy Assumption 1 and for which the associated price competition game has a pure strategy OE.

**Theorem 3** Consider a parallel link network with $I$ links, where each link is owned by a different provider. Then

$$\inf_{(u,\{l_i\})\in\mathcal{P}_I^*} \inf_{x^{OE}\in\overrightarrow{OE}(u,\{l_i\})} r_I(u,\{l_i\},x^{OE}) = \frac{2}{3}.$$

*Proof Sketch:* The proof follows those of Theorems 1 and 2 closely. Once again, the problem (16) is lower-bounded by a modified version of the finite dimensional problem ($\mathrm{E}_I^\epsilon$) (see the proof of Theorem 2), in which we introduce additional variables $l_i^0 \geq 0$, which represent the value of the latency function, $l_i(\cdot)$ at 0. Using the convexity of the latency functions, we replace constraint (35) by

$$l_i^S \leq y_i^S (l_i^S)' + l_i^0.$$

Following the same line of argument, it can be seen that problem (16) can further be bounded below by a problem identical to (25) except that constraint (36) is replaced by

$$l_i \leq y_i l_i' + l_i^0.$$

Using a similar transformation, this problem can be seen to be equivalent to

$$\begin{aligned} \operatorname*{minimize}_{\substack{0\leq u\leq R,\ u'\geq 0,\ u''\leq 0 \\ l,\ l',\ l^0\geq 0 \\ y_1,\ y\geq 0 \\ R,\ d\geq 0}} \quad & \frac{u - ly}{R} \\ \text{subject to} \quad & l \leq yl' + l^0, \\ & l + yl' = \frac{y_1}{\frac{1}{l'} - \frac{1}{u''}}, \\ & l + yl' = u', \\ & u \geq u'(y_1 + y), \\ & u' \geq -u''(d - (y_1 + y)), \\ & R \leq u + u'(d - (y_1 + y)) + \frac{u''}{2}(d - (y_1 + y))^2, \\ & y_1 + y \leq d. \end{aligned}$$

Using an argument similar to that in the proof of Theorem 1, the optimal value of this problem can be shown to be equal to 2/3. The example provided in the proof of Theorem 2 shows that the bound is tight. **Q.E.D.**

The preceding result shows that the tight bound of 2/3 on the efficiency metric holds irrespective of the assumption that $l_i(0) = 0$ for all $i$. This is not the case for inelastic traffic, where relaxing the assumption $l_i(0) = 0$ leads to a slightly lower bound on the efficiency metric (see Theorem 3 in [2]). This is due to the fact that the main source of inefficiency with inelastic traffic is because of the misallocation of traffic, and allowing for differential levels of $l_i(0)$ increases this effect. With elastic traffic the main source of inefficiency is due to the reduction in the amount of admitted traffic, and is realized when one of the links has an arbitrarily high latency, effectively stopping all transmission on this link. Therefore, allowing for positive $l_i(0)$ does not lead to further deterioration in performance.

# 6 Concluding Remarks

In this paper, we presented an analysis of competition in congested networks. Despite the potential inefficiencies of selfish flow control, routing, and service provider pricing, we are able to provide a tight bound of 2/3 on the efficiency of pure strategy oligopoly equilibria. This bound is still tight for the special case when the latency without congestion is equal to zero (this contrasts with the inelastic traffic case, where the tight bound is different depending on whether or not latency without congestion is equal to zero, see [2]).These bounds apply even for arbitrarily large parallel link networks.

A number of concluding comments are useful:

- While the analysis above demonstrated that there are no bounds on the inefficiency of oligopoly equilibria for general concave utility functions, whether there are other classes of utility functions besides those with concave first derivatives that also provide bounds on inefficiency is an open question.

- Our analysis has been simplified by our focus on parallel link networks. The paper [1] provides an analysis of efficiency of oligopoly equilibria with inelastic traffic for topologies with parallel-serial structure. The analysis for more general topologies with inelastic traffic and the study of the parallel-serial structure with elastic traffic are areas for future research.

- It is possible to extend our model to multiple user classes represented by different utility functions, and obtain similar bounds on efficiency under the assumption that the derivative of each of the utility functions is a concave function at the expense of introducing extra notation (see [3] for a model with multiple user classes).

- The analysis was carried out under differentiability assumptions on the utility and the latency functions. These assumptions were adopted to simplify the expositions and the results remain valid even when the differentiability assumptions are relaxed (see [2] for an analysis without differentiability assumptions on latency functions).

- The focus throughout has been on the efficiency of pure strategy equilibria. The efficiency of mixed strategy equilibria is an open research question.

# 7 Appendix A: Proof of Proposition 1

**Proof.** For all $i \in \mathcal{I}$, the latency functions are given by $l_i(x) = a_i x + b_i$ for some scalars $a_i > 0$ and $b_i \geq 0$. Let $B_i(p^*_{-i})$ be the set of $p^*_i$ such that

$$(p^*_i, \bar{x}) \in \arg \max_{\substack{p \geq 0 \\ x \in W(p_i, p^*_{-i})}} p_i x_i. \tag{43}$$

Let $B(p^*) = [B_i(p^*_{-i})]_{i \in \mathcal{I}}$. By the Theorem of the Maximum [6], it follows that $B(p^*)$ is an upper semicontinuous correspondence. We next show that it is convex-valued. For some $i \in \mathcal{I}$ and $p^*_{-i} \geq 0$, consider two price vectors $p_i \in B_s(p^*_{-i})$ and $\bar{p}_i \in B_s(p^*_{-i})$ and corresponding Wardrop equilibria $x \in W(p_i, p^*_{-i})$ and $\bar{x} \in W(\bar{p}_i, p^*_{-i})$ such that the vectors $(p_i, x)$ and $(\bar{p}_i, \bar{x})$ are optimal solutions of problem (43).

Suppose first that $p_i x_i = \bar{p}_i \bar{x}_i = 0$. Since $(p_i, x)$ is an optimal solution of problem (43), for any $\gamma \in [0, 1]$, the scalar $p^\gamma_i = \gamma p_i + (1 - \gamma)\bar{p}_i$ must satisfy $p^\gamma_i x^\gamma_i = 0$ for all $x^\gamma \in W(p^\gamma_i, p^*_{-i})$, implying that $p^\gamma_i \in B_i(p^*_{-i})$.

Suppose next that $p_i x_i = \bar{p}_i \bar{x}_i > 0$. We show that in this case $p_i = \bar{p}_i$. Assume to arrive at a contradiction that $p_i > \bar{p}_i$, which implies that $x_i < \bar{x}_i$. Using the first order conditions of problem (43), we obtain

$$p_i = a_i x_i + \frac{x_i}{\Big(\sum_{\substack{j \neq i \\ j \notin \mathcal{N}}} \frac{1}{a_j}\Big) - \Big(\frac{1}{u''(\sum_{j \in \mathcal{I}} x_j)}\Big)}, \tag{44}$$

$$\bar{p}_i = a_i \bar{x}_i + \frac{\bar{x}_i}{\Big(\sum_{\substack{j \neq i \\ j \notin \bar{\mathcal{N}}}} \frac{1}{a_j}\Big) - \Big(\frac{1}{u''(\sum_{j \in \mathcal{I}} \bar{x}_j)}\Big)}, \tag{45}$$

where

$$\mathcal{N} = \{j \in \mathcal{I} \mid p_i + a_i x_i < p^*_j + l_j(0)\}, \tag{46}$$

$$\bar{\mathcal{N}} = \{j \in \mathcal{I} \mid \bar{p}_i + a_i \bar{x}_i < p^*_j + l_j(0)\}, \tag{47}$$

(see the proof of Proposition 3). There are two cases to consider:

- $p_i + a_i x_i \geq \bar{p}_i + a_i \bar{x}_i$: Since both $x_i > 0$ and $\bar{x}_i > 0$, by Lemma 1, this implies that

  $$u'\Big(\sum_{j \in \mathcal{I}} x_j\Big) \geq u'\Big(\sum_{j \in \mathcal{I}} \bar{x}_j\Big).$$

  In view of the assumption that $u$ is concave, it follows that $\sum_{j \in \mathcal{I}} x_j \leq \sum_{j \in \mathcal{I}} \bar{x}_j$. Since $u'$ is also concave, this implies that

  $$-u''\Big(\sum_{j \in \mathcal{I}} x_j\Big) \leq -u''\Big(\sum_{j \in \mathcal{I}} \bar{x}_j\Big),$$

  (where as usual $u''(y)$ refers to a subgradient of $u'$ at $y$.) Moreover, since $p_i + a_i x_i \geq \bar{p}_i + a_i \bar{x}_i$, we have $\mathcal{N} \subset \bar{\mathcal{N}}$. Using these relations in Eqs. (44) and (45) together with the assumption $x_i < \bar{x}_i$, we obtain $p_i < \bar{p}_i$, hence a contradiction.

- $p_i + a_i x_i < \bar{p}_i + a_i \bar{x}_i$: Since both $x_i > 0$ and $\bar{x}_i > 0$, using Lemma 1, we obtain

$$\sum_{j \in \mathcal{I}} x_j \geq \sum_{j \in \mathcal{I}} \bar{x}_j. \tag{48}$$

  For all $j \in \mathcal{I}$ such that $j \neq i$ and $x_j > 0$, the assumption $p_i + a_i x_i < \bar{p}_i + a_i \bar{x}_i$ and Lemma 1 imply that

$$\bar{p}_j^* + a_j x_j = u'(\sum_{k \in \mathcal{I}} x_k) < u'\Big(\sum_{k \in \mathcal{I}} \bar{x}_k\Big) \leq p_j^* + a_j \bar{x}_j,$$

  implying that $x_j \leq \bar{x}_j$ for all $j \neq i$. Together with the assumption $x_i < \bar{x}_i$, this contradicts the relation in (48). This shows that $B(p)$ is convex-valued.

Since $B(p)$ is an upper semicontinuous and convex-valued correspondence, we can use Kakutani's fixed point theorem to assert the existence of a $\bar{p}$ such that $B(\bar{p}) = \bar{p}$ (see [6]). Since by assumption the latency functions are strictly increasing, we have that $W(\bar{p})$ is a singleton, thus completing the proof. **Q.E.D.**

# References


[1] Acemoglu, D. and Ozdaglar, A., "Price competition in communication networks," *Proceedings of INFOCOM*, 2006.

[2] Acemoglu, D. and Ozdaglar, A., "Competition and efficiency in congested markets," to appear in *Mathematics of Operations Research*, 2006.

[3] Acemoglu, D. and Ozdaglar, A.,"Flow control, routing, and performance from service provider viewpoint," *LIDS report*, WP-1696, September 2004.

[4] Basar, T. and Srikant R., "Revenue-maximizing pricing and capacity expansion in a many-users regime," *Proceedings of INFOCOM*, 2002.

[5] Beckmann, M., Mcguire, C. B., and Winsten, C. B., Studies in the Economics of Transportation. Yale University Press, 1956.

[6] Berge, C., *Topological Spaces,* Dover reprint, 1963.

[7] Correa, J. R., Schulz, A. S., and Stier Moses, N., "Selfish Routing in Capacitated Networks," *Mathematics of Operations Research*, 29:4, pp. 961-976, Nov. 2004.

[8] Dafermos, S. and Sparrow F. T., "The traffic assignment problem for a general network," *Journal of Research of the National Bureau of Standards-B. Mathematical Sciences,* vol. 73(2), pp. 91-118, 1969.

[9] Friedman, E., "A Generic Analysis of Selfish Routing," *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004.

[10] Fudenberg, D. and Tirole J., Game Theory. The MIT Press, 1991.

[11] Hayrapetyan, A., Tardos, E., and Wexler T., "A network pricing game for selfish traffic," *PODC*, 2005, to appear in a special issue of the journal *Distributed Computing.*

[12] He, L. and Walrand, J., "Pricing internet services with multiple providers," *Proceedings of Allerton Conference*, 2003.

[13] Kelly, F. P., Maulloo A. K., and Tan D. K.,"Rate control for communication networks: shadow prices, proportional fairness, and stability," *Journal of the Operational Research Society*, vol. 49, pp. 237-252, 1998.

[14] Koutsoupias, E. and Papadimitriou, C.,"Worst-case Equilibria," *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science,* pp. 404-413, 1999.

[15] Low, S. and Lapsley, D. E., "Optimization flow control, I: Basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7(6), pp. 861-874, 1999.

[16] Milchtaich, I., "Congestion games with player-specific payoff functions," *Games and Economic Behavior,* vol. 13, pp. 11-124, 1996.

[17] Patriksson, M., The Traffic Assignment Problem: Models and Methods, VSP BV, Netherlands, 1994.

[18] Perakis, G., "The Price of anarchy when costs are non-separable and asymmetric", manuscript.

[19] Rosenthal, R., W., "A class of games possessing pure-strategy Nash equilibria," *International Journal of Game Theory,* vol. 2, pp. 65-67, 1973.

[20] Roughgarden, T. and Tardos, E., "How bad is selfish routing?" *Journal of the ACM*, vol. 49(2), pp. 236-259, 2002.

[21] Shenker, S., "Fundamental design issues for the future Internet," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1176-1188, 1995.

[22] Wardrop, J. G., "Some theoretical aspects of road traffic research," *Proceedings of the Institute of Civil Engineers*, Pt. II, vol. 1, pp. 325-378.

[23] Yaïche, H., Mazumdar, R. R., and Rosenberg, C., "A Game Theoretic Framework for Bandwidth Allocation and Pricing in Broadband Networks", *IEEE/ACM Transactions on Networking*, vol. 9, no. 5, 2000.