Price Competition in Communication Networks

Daron Acemoglu
Department of Economics
Massachusetts Institute of Technology
Cambridge, MA 02142
Email: daron@mit.edu

Asuman Ozdaglar
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02142
Email: asuman@mit.edu

***Abstract*— We study the efficiency properties of oligopoly equilibria in congested networks. Our measure of efficiency is the difference between users' willingness to pay and delay costs. Previous work has demonstrated that in networks consisting of parallel links, efficiency losses from competition are bounded. In contrast, in this paper we show that in the presence of serial links, the efficiency loss relative to the social optimum can be arbitrarily large because of the double marginalization problem, whereby each serial provider charges high prices not taking into account the effect of this strategy on the profits of other providers along the same path. Nevertheless, when there are no delay costs without transmission (i.e., latencies at zero are equal to zero), irrespective of the number of serial and parallel providers, the efficiency of strong oligopoly equilibria can be bounded by 1/2, where strong oligopoly equilibria are equilibria in which each provider plays a strict best response and all of the traffic is transmitted. However, even with strong oligopoly equilibria, inefficiency can be arbitrarily large when the assumption of no delay costs without transmission is relaxed.**

I. INTRODUCTION

There has been growing interest in pricing as a method of allocating scarce network resources (see, e.g., [17], [19], [25]). Although prices may be set to satisfy some network objectives, in practice many prices are controlled by for-profit service providers that charge prices, at least in part, to increase their revenues and profits.

Research to date suggests that profit-maximizing pricing may improve the allocation of resources in communication networks. Let the metric of efficiency be the difference between users' willingness to pay and delay costs in the equilibrium relative to that in the social optimum (which would be chosen by a a network planner with full information and full control over users). Acemoglu and Ozdaglar [2] show that with inelastic and homogeneous users, pricing by a monopolist controlling all links in a parallel-link network always achieves efficiency (i.e., the efficiency metric is equal to one). Huang, Ozdaglar and Acemoglu [15] extend this result to a general network topology. Acemoglu and Ozdaglar [1] show that in a parallel-link network with inelastic and homogeneous users, the efficiency metric with an arbitrary number of competing network providers is always greater than or equal to 5/6. More recently, Hayrapetyan, Tardos, and Wexler [13] studied pricing in a parallel-link network with elastic and homogeneous users, and provided bounds on the efficiency metric.

This paper shows that the efficiency loss with competing service providers is considerably higher when we consider more general network topologies, suggesting that unregulated consumption in general networks may have significant costs in terms of resource allocation.

To illustrate the potential inefficiencies of price competition in congested communication networks, we consider a parallel-serial topology where an origin-destination pair is linked by multiple parallel paths, each potentially consisting of an arbitrary number of serial links. Congestion costs are captured by link-specific non-decreasing convex latency functions, denoted by $l_i(\cdot)$ for link i. Each link is owned by a different service provider. All users are inelastic and homogeneous.

This environment induces the following two-stage game: each service provider simultaneously sets the price for transmission of bandwidth on its link, denoted by p_i. Observing all the prices, in the second stage users route their information through the path with the lowest effective cost, where effective cost consists of the sum of prices and latencies of the links along a path [i.e., sum of $p_i + l_i(\cdot)$'s over the links comprising a path]. Our objective is to study the efficieny properties of the subgame perfect equilibria of this game.

The main novel aspect of this model compared to the parallel-link case is the pricing decisions of different (serial) service providers along a single path. When a particular provider charges a higher price, it creates a negative externality on other providers along the same path, because this higher price reduces the transmission they receive. This is the equivalent of the *double marginalization* problem in economic models with multiple monopolies, and leads to significant degradation of the equilibrium performance of the network.

In its most extreme form, the double marginalization problem leads to a type of "coordination failure", whereby all providers, expecting others to charge high prices, also charge prohibitively high prices, effectively killing all data transmission on a given path, and leading to arbitrarily low efficiency. This type of pathological behavior can happen in subgame perfect equilibria (what we refer to as oligopoly equilibria, OE), but we show that it cannot happen in strict subgame perfect equilibria, *strict OE*, which follows the notion of strict equilibrium introduced in Harsanyi [12]. In strict OE, each service provider must play a strict best response to the pricing strategies of other service providers, which is sufficient to rule out the pathological coordination failures mentioned above.

Nevertheless, we show that strict OE can also have arbitrarily large efficiency losses because the double marginalization problem can again prevent any transmission on a particular path, even when such transmission is socially optimal.

Instead, we define an even stronger notion of equilibrium, *strong OE*, as a strict OE in which all traffic is transmitted.[1] We show that when latency without any traffic is equal to zero [i.e., $l_i(0) = 0$], there is a tight bound of 1/2 on the efficiency of strong OE irrespective of the number of paths and service providers in the network. This bound is reached by simple examples. In strong OE, the double marginalization problem is still present, and this is the reason why the bound of 1/2 is lower than the 5/6 bound in our previous work, [1].

Furthermore, we show that even with strong OE, when the assumption that $l_i(0) = 0$ is relaxed, the efficiency loss optimum can be arbitrarily large.

These results shed doubt on the conjecture that unregulated competition among service providers might achieve satisfactory network performance in general. Nevertheless, it has to be borne in mind that the examples that have very poor performance relative to the social optimum are somewhat pathological, and this begs the question of whether much better performance results could be obtained in more realistic topologies, which is an area left for future work.

[1]Models of selfish routing without prices, e.g., [22] or [7], assume that all traffic is always transmitted. Our model incorporates a reservation utility for users, so that this is not necessarily the case. All traffic will be transmitted in equilibrium when this reservation utility is sufficiently large.

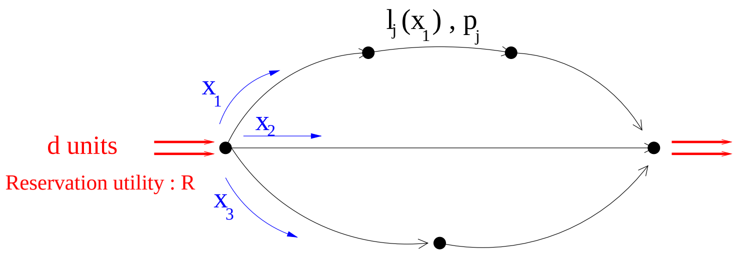

Fig. 1. A network with serial and parallel links.

Work related to our paper includes studies quantifying efficiency losses of selfish routing without prices (e.g., Koutsoupias and Papadimitriou [18], Roughgarden and Tardos [22], Correa, Schulz, and Stier-Moses [7], Perakis [21], and Friedman [10]); of resource allocation by different market mechanisms (e.g., Johari and Tsitsiklis [16], Sanghavi and Hajek [23]); and of network design (e.g., Anshelevich et. al. [4]). More closely related are the works of Basar and Srikant [5], who analyze monopoly pricing in a network context under specific assumptions on the utility and latency functions; He and Walrand, [14], who study competition cooperation among Internet service providers under specific demand models; as well as Acemoglu, Ozdaglar, and Srikant [3], who study resource allocation in a wireless network under fixed pricing. None of these papers, except our previous work, Acemoglu and Ozdaglar [1] and the recent work, Hayrapetyan, Tardos, and Wexler [13], consider the performance of a network with competing providers. No other paper, to the best of our knowledge, has investigated price competition in the presence of serial providers or more general topologies.

II. Model

We consider a network with I parallel paths that connect two nodes. Each path i consists of n_i links. Let $\mathcal{I} = \{1, \ldots, I\}$ denote the set of paths and $\mathcal{N}_i$ denote the set of links on path i. Let x_i denote the flow on path i, and $x = [x_1, \ldots, x_I]$ denote the vector of path flows. Each link in the network has a flow-dependent latency function $l_i(x_i)$, which measures the delay as a function of the total flow on link i (see Figure 1). We denote the price per unit flow (bandwidth) of link j by p_j. Let $p = [p_j]_{j \in \mathcal{N}_i, i \in \mathcal{I}}$ denote the vector of prices.

We are interested in the problem of routing d units of flow across the I paths. We assume that this is the aggregate flow of many "small" users and thus adopt the Wardrop's principle (see [24]) in characterizing the flow distribution in the network; i.e., the flows are routed along paths with minimum effective cost, defined as the sum of the latencies and prices of the links along that path (see the definition below). Wardrop's principle is used extensively in modelling traffic behavior in transportation networks ([6], [8], [20]) and communication networks ([22], [7]). We also assume that users have a *reservation utility* R and decide not to send their flow if the effective cost exceeds the reservation utility. This implies that user preferences can be represented by the piecewise linear aggregate utility function $u(\cdot)$ depicted in Figure 2.[2]

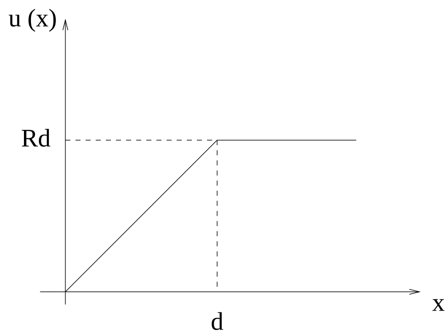

Fig. 2. Aggregate utility function.

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

$$\begin{aligned} \sum_{j\in\mathcal{N}_i} \Big(l_j(x_i^{WE}) + p_j\Big) &= \min_{k\in\mathcal{I}} \Big\{ \sum_{j\in\mathcal{N}_k} \Big(l_j(x_k^{WE}) + p_j\Big)\Big\} \\ &\qquad \forall\ i \text{ with } x_i^{WE} > 0, \quad (1) \\ \sum_{j\in\mathcal{N}_i} (l_j(x_i^{WE}) + p_j\Big) &\leq R, \quad \forall\ i \text{ with } x_i^{WE} > 0, \\ \sum_{i\in\mathcal{I}} x_i^{WE} &\leq d, \end{aligned}$$

with $\sum_{i\in\mathcal{I}} x_i^{WE} = d$ if $\min_{k\in\mathcal{I}}\{\sum_{j\in\mathcal{N}_k} l_j(x_k^{WE}) + p_j\} < R$. We denote the set of WE at a given p by $W(p)$.

We adopt the following assumption on the latency functions throughout the paper except in Section IV-F.

Assumption 1: For each $i \in \mathcal{I}$, the latency function $l_i : [0,\infty) \mapsto [0,\infty)$ is convex, continuously differentiable, nondecreasing, and satisfies $l_i(0) = 0$.

Proposition 1: **(Existence and Continuity)** Let Assumption 1 hold. For any price vector $p \geq 0$, the set of WE, $W(p)$, is nonempty. Moreover, the correspondence $W : \mathbb{R}_+^I \rightrightarrows \mathbb{R}_+^I$ is upper semicontinuous.

Proof sketch: Given any $p \geq 0$, the proof is based on using Assumption 1 (in particular the nondecreasing assumption on the latency functions) to show that the set of WE is given by the set of optimal solutions of the following optimization problem

$$\begin{aligned} \max_{x\geq 0} \quad & \sum_{i\in\mathcal{I}}((R - \sum_{j\in\mathcal{N}_i} p_i)x_i - \int_0^{x_i} \sum_{j\in\mathcal{N}_i} l_j(z)dz) \quad (2) \\ \text{s.t.} \quad & \sum_{i\in\mathcal{I}} x_i \leq d. \end{aligned}$$

Q.E.D.

For a given price vector p, the WE need not be unique in general. Under further restrictions on the l_i, we obtain:

Proposition 2: **(Uniqueness)** Let Assumption 1 hold. Assume further that for all $i \in \mathcal{I}$, there exists some $j \in \mathcal{N}_i$, such that l_j is strictly increasing. For any price vector $p \geq 0$, the set of WE, $W(p)$, is a singleton. Moreover, the function $W : \mathbb{R}_+^I \mapsto \mathbb{R}_+^I$ is continuous.

Proof: Under the given assumptions, for any $p \geq 0$, the objective function of problem (2) is strictly convex, and therefore has a unique optimal solution. This shows the uniqueness of the WE at a given p. Since the correspondence W is upper semicontinuous from Proposition 1 and single-valued, it is continuous. **Q.E.D.**

We next define the social problem and the social optimum, which is the routing (flow allocation) that would be chosen by a central network planner that has full control and information about the network.

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

$$\begin{aligned} \text{maximize}_{x\geq 0} \quad & \sum_{i\in\mathcal{I}} \Big(R - \sum_{j\in\mathcal{N}_i} l_j(x_i)\Big)x_i \quad (3) \\ \text{subject to} \quad & \sum_{i\in\mathcal{I}} x_i \leq d. \end{aligned}$$

[2]This simplifying assumption implies that all users are "homogeneous" in the sense that they have the same reservation utility, R. We discuss potential issues in extending this work to users with elastic and heterogeneous requirements in the concluding section.

By Assumption 1, the social problem has a continuous objective function and a compact constraint set, guaranteeing the existence of a social optimum, x^S. Moreover, using the optimality conditions for a convex program, we see that a vector $x^S \in \mathbb{R}^I_+$ is a social optimum if and only if $\sum_{i\in\mathcal{I}} x_i^S \le d$ and there exists a $\lambda^S \ge 0$ such that $\lambda^S(\sum_{i=1}^I x_i^S - d) = 0$ and for each $i \in \mathcal{I}$,

$$R - \sum_{j\in\mathcal{N}_i} l_j(x_i^S) - x_i^S \sum_{j\in\mathcal{N}_i} l_j'(x_i^S) \quad \begin{aligned} &\le \lambda^S \quad \text{if } x_i^S = 0, \\ &= \lambda^S \quad \text{if } x_i^S > 0. \end{aligned} \tag{4}$$

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

$$\mathbb{S}(x) = \sum_{i\in\mathcal{I}} \Big(R - \sum_{j\in\mathcal{N}_i} l_j(x_i)\Big) x_i, \tag{5}$$

as the *social surplus*, i.e., the difference between the users' willingness to pay and the total latency.

III. Oligopoly Pricing and Equilibrium

We assume that there are multiple service providers, each of which owns one of the links on the paths in the network. Service provider j charges a price p_j per unit bandwidth on link $j \in \mathcal{N}_i$. Given the vector of prices of links owned by other service providers, $p_{-j} = [p_k]_{k\neq j}$, the profit of service provider j with $j \in \mathcal{N}_i$ is

$$\Pi_j(p_j, p_{-j}, x) = p_j x_i,$$

where $x \in W(p_j, p_{-j})$.

The objective of each service provider is to maximize profits. Because their 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 they do according to the notion of subgame perfect Nash equilibrium. We refer to the game among service providers as the *price competition game*.

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

$$\Pi_j(p_j^{OE}, p_{-j}^{OE}, x^{OE}) \ge \Pi_j(p_j, p_{-j}^{OE}, x), \qquad \forall\, p_j \ge 0,\ \forall\, x \in W(p_j, p_{-j}^{OE}). \tag{6}$$

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

The next proposition shows that for linear latency functions, there exists a pure strategy OE.

Proposition 3: Let Assumption 1 hold and assume that the latency functions are linear. Then the price competition game has a pure strategy OE.

Proof: Let $l_j(x) = a_j x$ for some $a_j \ge 0$. Define the set

$$\mathcal{I}_0 = \{i \in \mathcal{I} \mid \sum_{j\in\mathcal{N}_i} a_j = 0\},$$

(or equivalently, $\mathcal{I}_0$ is the set of $i \in \mathcal{I}$ such that $a_j = 0$ for all $j \in \mathcal{N}_i$). Let I_0 denote the cardinality of set $\mathcal{I}_0$. There are two cases to consider:

I. $I_0 \ge 2$: Then it can be seen that a vector (p^{OE}, x^{OE}) with $p_j^{OE} = 0$ for all $i \in \mathcal{I}_0$, $j \in \mathcal{N}_i$ and $x^{OE} \in W(p^{OE})$ is an OE.

II. $I_0 \le 1$: For some $j \in \mathcal{N}_i$, let $B_j(p_{-j}^{OE})$ be the set of p_j^{OE} such that

$$(p_j^{OE}, x^{OE}) \in \arg \max_{\substack{p_j \ge 0 \\ x\in W(p_j, p_{-j}^{OE})}} p_j x_i. \tag{7}$$

Let $B(p^{OE}) = [B_j(p_{-j}^{OE})]$. In view of the linearity of the latency functions, it follows that $B(p^{OE})$ is an upper semicontinuous and convex-valued correspondence. Hence, we can use Kakutani's fixed point theorem to assert the existence of a p^{OE} such that $B(p^{OE}) = p^{OE}$. To complete the proof, it remains to show that there exists $x^{OE} \in W(p^{OE})$ such that (6) holds.

If $\mathcal{I}_0 = \emptyset$, we have by Proposition 2 that $W(p^{OE})$ is a singleton, and therefore (6) holds and $(p^{OE}, W(p^{OE}))$ is an OE.

Assume finally that $I_0 = 1$, and that without loss of generality $1 \in \mathcal{I}_0$. We show that for all $\bar{x}$, $\tilde{x} \in W(p^{OE})$, we have $\bar{x}_i = \tilde{x}_i$, for all $i \neq 1$. Let

$$EC(x, p^{OE}) = \min_{k\in\mathcal{I}} \{ \sum_{j\in\mathcal{N}_k} l_j(x_k) + p_j^{OE} \}.$$

If at least one of

$$EC(\tilde{x}, p^{OE}) < R, \qquad \text{or} \qquad EC(\bar{x}, p^{OE}) < R$$

holds, then one can show that $\sum_{i=1}^I \tilde{x}_i = \sum_{i=1}^I \bar{x}_i = d$. Substituting $x_1 = d - \sum_{i\in\mathcal{I},\ i\neq 1} x_i$ in problem (2), we see that the objective function of problem (2) is strictly convex in $x_{-1} = [x_i]_{i\neq 1}$, thus showing that $\tilde{x} = \bar{x}$. If both $EC(\tilde{x}, p^{OE}) = R$ and $EC(\bar{x}, p^{OE}) = R$, then

$$\sum_{j\in\mathcal{N}_i} l_j(\bar{x}_i) = \sum_{j\in\mathcal{N}_i} l_j(\tilde{x}_i),$$

which, by the assumption that l_j is strictly increasing for some $j \in \mathcal{N}_i$, implies that $\bar{x}_i = \tilde{x}_i$ for all $i \neq 1$, establishing our claim.

For some $x \in W(p^{OE})$, consider the vector $x^{OE} = (d - \sum_{i \neq 1} x_i, x_{-1})$. Since x_{-1} is uniquely defined and x_1 is chosen such that the providers on link 1 have no incentive to deviate, it follows that (p^{OE}, x^{OE}) is an OE. **Q.E.D.**

The existence result cannot be generalized to general convex latency functions as shown in the following example.

Example 1: Consider a two link network. Let the total flow be $d = 1$. Assume that the latency functions are given by

$$l_1(x) = 0, \qquad l_2(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq \delta \\ \frac{x-\delta}{\epsilon} & x \geq \delta, \end{cases}$$

for some $\epsilon > 0$ and $\delta > 1/2$, with the convention that when $\epsilon = 0$, $l_2(x) = \infty$ for $x > \delta$. It can be easily verified that there exists no pure strategy OE for small ϵ (see [1] for details).

Nevertheless, a mixed strategy OE always exists. We define a mixed strategy OE as a mixed strategy subgame perfect equilibrium of the price competition game (see Dasgupta and Maskin, [9]). Let $\mathcal{B}$ be the space of all (Borel) probability measures on $[0, R]$. Let T denote the total number of links in the network, i.e., $T = \sum_{i=1}^{I} n_i$. Let $\mu_j \in \mathcal{B}$ be a probability measure, and denote the vector of these probability measures by μ and the vector of these probability measures excluding j by μ_{-j}.

Definition 4: $(\mu^*, x^*(p))$ is a mixed strategy *Oligopoly Equilibrium* (OE) if the function $x^*(p) \in W(p)$ for every $p \in [0, R]^T$ and

$$\int_{[0,R]^T} \Pi_j(p_j, p_{-j}, x^*(p_j, p_{-j})) d\left(\mu_j^*(p_j) \times \mu_{-j}^*(p_{-j})\right)$$

$$\geq \int_{[0,R]^T} \Pi_j(p_j, p_{-j}, x^*(p_j, p_{-j})) d\left(\mu_j(p_j) \times \mu_{-j}^*(p_{-j})\right)$$

for all j and $\mu_j \in \mathcal{B}$, where $\Pi_j(p, x^*(p))$ is the profit of service provider j at price vector p and $x^*(p) \in W(p)$.

A mixed strategy OE thus requires that there is no profitable deviation to a different probability measure for each oligopolist. The existence of a mixed strategy OE can be established along the lines of the analysis in [1], which leads to the following result (proof omitted):

Proposition 4: Let Assumption 1 hold. Then the price competition game has a mixed strategy OE, $(\mu^{OE}, x^{OE}(p))$.

IV. Efficiency Analysis

A. Inefficiency of OE

In this section, we study the efficiency properties of OE, and strict and strong OE (defined below). We take as our measure of efficiency the ratio of the social surplus at the equilibrium flow allocation to the social surplus at the social optimum, $\mathbb{S}(x^{OE})/\mathbb{S}(x^S)$ [cf. (5)]. We consider price competition games that have OE or strict OE (this set includes, but is larger than, games with linear latency functions, see Section III). Given a parallel-link network with I paths, n_i links on path i, and latency functions $\{l_j\}_{(j \in \mathcal{N}_i, i \in \mathcal{I})}$, let $\overrightarrow{OE}(\{l_j\})$ denote the set of flow allocations $x^{OE} = [x_i^{OE}]_{i \in \mathcal{I}}$ at an OE (or strict OE depending on the context). We define the efficiency metric at some $x^{OE} \in \overrightarrow{OE}(\{l_j\})$ as

$$r_I(\{l_j\}, x^{OE}) = \frac{R \sum_{i \in \mathcal{I}} x_i^{OE} - \sum_{i \in \mathcal{I}} \left(\sum_{j \in \mathcal{N}_i} l_j(x_i^{OE}) \right) x_i^{OE}}{R \sum_{i \in \mathcal{I}} x_i^S - \sum_{i \in \mathcal{I}} \left(\sum_{j \in \mathcal{N}_i} l_j(x_i^S) \right) x_i^S}, \quad (8)$$

where x^S is a social optimum given the latency functions $\{l_j\}$ and R is the reservation utility. Following the literature on the "price of anarchy," (see [18]), we are interested in the worst performance of an oligopoly equilibrium, so we look for a lower bound on

$$\inf_{\{l_j\}} \inf_{x^{OE} \in \overrightarrow{OE}(\{l_j\})} r_I(\{l_j\}, x^{OE}).$$

We first show that the performance of an OE can be arbitrarily bad.

Example 2: Consider a two path network, which has 3 links on path 1 with identically 0 latency functions and one link on path 2 with latency function $l(x_2) = kx_2$, where $k \geq 0$. Let the total flow be $d = 1$ and the reservation utility be $R = 1$.

The unique social optimum for this example is $x^S = (1, 0)$. Now consider the following strategy combination. Each of the three service providers on path 1, denoted by $i = 1, 2$, and 3, charge price $p_1^i = 1$, while the service provider on path 2 charges $p_2 = 1/2$. It can be verified that there is no deviation that is profitable for any of the service providers. First, consider the serial providers on path 1; given the prices of two of the serial service providers, there will always be zero traffic on path 1, so the remaining service provider is playing a best response (since any price for this provider would lead to zero profits). Moreover, it can be verified that these strategies are not weakly dominated, since if $i = 1, 2$ were to play

$p_1^i = 0$ and the provider on path 2 were to set a high enough p_2, $i = 3$ would choose to play $p_1^3 = 1$. (This also establishes that the OE will be trembling hand perfect, see [11], pp. 351-356 for a definition). Finally, let us consider the provider on path 2. Given the strategies of the serial providers on path 1 and a fixed $k > 0$, it can be verified that the optimal strategy of this provider is to set $p_2 = 1/2$. The resulting equilibrium flow allocation is

$$x^{OE} = \left[0, \frac{1}{2k}\right],$$

which involves routing all the admitted traffic on path 2 (though not all of the traffic is necessarily admitted). Therefore, the efficiency metric for this example is

$$r_2(\{l_j\}, x^{OE}) = \frac{\sum_{i=1}^{2} x_i^{OE} - l_2(x_2^{OE})x_2^{OE}}{1} = \frac{1/4k}{1},$$

which goes to 0 as $k \to \infty$.

Example 2 establishes that pure strategy OE with the parallel-serial link topology can be arbitrarily inefficient. This result is at some level pathological, however. The reason why there is so much inefficiency is because service providers on path 1 charge unreasonably high prices. It is a best response (even weakly undominated) strategy for them to do so, because other providers also charge unreasonably high prices, so there is no transmission on this path and they suffer no adverse consequences from charging unreasonable prices.

We may expect this pathological situation not to arise for a number of reasons. First, firms may not coordinate on such an equilibrium (especially when other equilibria exist). In this case, for example, we may expect them to realize that if they all reduced their prices, they would all make higher profits and would still be playing equilibrium actions. Second, we may even expect providers on a path to form a "coalition" and coordinate their pricing decisions. It can be verified that a special form of coalition-proof subgame perfect Nash equilibrium, where only providers along a given path can form coalitions, would lead to the results of our previous work, [1], in particular to the tight bound on efficiency of $5/6$. This is because, once serial providers along a path form a subcoalition, their optimal strategy would be to behave as a single provider along that path, thus removing the double marginalization problem. In this case, the analysis in our previous work [1] applies and leads to a bound of $5/6$.

B. Strict OE and Price Characterization

In this paper, rather than allowing coalitions to form, we study a stronger concept of equilibrium, Harsanyi's strict equilibrium (see [12], or [11], pp. 11-12), which requires each player's best response to be unique. Recall that the standard Nash equilibrium and our OE concept only require each player, in particular each service provider, to play a weak best response. We now strengthen this condition.

Definition 5: A vector $(p^{OE}, x^{OE}) \geq 0$ is a *strict OE* (Oligopoly Equilibrium) if $x^{OE} \in W\left(p_j^{OE}, p_{-j}^{OE}\right)$ and for all $i \in \mathcal{I}$, $j \in \mathcal{N}_i$,

$$\Pi_j(p_j^{OE}, p_{-j}^{OE}, x^{OE}) > \Pi_j(p_j, p_{-j}^{OE}, x),$$
$$\forall\, p_j \geq 0,\ p_j \neq p_j^{OE},\ \forall\, x \in W(p_j, p_{-j}^{OE}). \quad (9)$$

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

In the remainder of this paper, we focus on strict OE and we use the notation $\overrightarrow{OE}(\{l_j\})$ to denote the set of flow allocations $x^{OE} = [x_i^{OE}]_{i\in\mathcal{I}}$ at a strict OE for a network with latency functions $\{l_j\}_{(j\in\mathcal{N}_i,\ i\in\mathcal{I})}$.

The difference between Definitions 3 and 5 is obvious. The latter requires service providers to play a strict best response, while the former does not. Notice that in both equilibria, we have not changed the behavior of the users given by the WE (as in Definition 1). Notice also that we have removed the qualifier "pure strategy," since as is well known, strict equilibria always have to be pure strategy (because mixed strategy equilibria, by definition, involve players being indifferent among the strategies over which they are mixing). Therefore, there are situations in which a mixed strategy OE exists, but strict OE does not. Moreover, it can be verified that there are also situations in which a pure strategy OE exists, but a strict OE does not.

We do not view this as a serious shortcoming, since, as Example 1 above showed, even pure strategy OE do not always exist. Moreover, we have the following result for linear latency functions. The proof is similar to the proof of Proposition 3, and therefore is omitted.

Proposition 5: Let Assumption 1 hold. Assume further that the latency functions are linear and for all $i \in \mathcal{I}$, there exists some $j \in \mathcal{N}_i$ such that l_j is strictly increasing. Then the price competition game has a strict OE.

We next provide an explicit characterization of the strict OE prices, which will be essential for the subsequent efficiency analysis. The following lemma establishes that all path flows are positive at a strict OE.

Lemma 1: Let (p^{OE}, x^{OE}) be a strict OE. Let Assumption 1 hold. Then $p_j^{OE} x_i^{OE} > 0$ for all $i \in \mathcal{I}$ and $j \in \mathcal{N}_i$.

Proof: Assume to arrive at a contradiction that $p_j^{OE} x_i^{OE} = 0$ for some $i \in \mathcal{I}$ and $j \in \mathcal{N}_i$. Then, at any price $\bar{p}_j$ with $\bar{p}_j > p_j^{OE}$, we have

$$\Pi_j(\bar{p}_j, p_{-j}^{OE}, x) = \Pi_j(p_j^{OE}, p_{-j}^{OE}, x^{OE}),$$

contradicting the definition of the strict OE (cf. Definition 9). **Q.E.D.**

As shown in Example 2, the result of the preceding lemma does not extend to non-strict OE prices, i.e., there may be OE in which some of the providers make zero profit while others are making positive profits. We have shown in [1] that for parallel-link topology, if at any OE one of the providers makes positive profit, all of the providers make positive profits (see [1], Lemma 4). Example 2 shows that this result no longer holds for non-strict OE for the parallel-serial topology. Lemma 1, on the other hand, ensures that it holds for strict OE and allows us to write the optimization problems for each provider in terms of equality and inequality constraints. We can then use the first order optimality conditions to obtain an explicit characterization of the strict OE prices.

Proposition 6: Let Assumption 1 hold. Then, for all $i \in \mathcal{I}$, $j \in \mathcal{N}_i$, we have
(a)

$$p_j^{OE} \geq x_i^{OE} \sum_{k \in \mathcal{N}_i} l_k'(x_i^{OE}).$$

(b)

$$p_j^{OE} = \begin{cases} x_i^{OE} \sum_{k\in\mathcal{N}_i} l_k'(x_i^{OE}), & \text{if } l_k'(x_s^{OE}) = 0, \\ & \text{for } k \neq j,\ s \neq i, \\ \min\left\{ \frac{1}{n_i}\left[R - \sum_{k\in\mathcal{N}_i} l_k(x_i^{OE})\right], \right. & \\ \quad \left. x_i^{OE}\left[\sum_{k\in\mathcal{N}_i} l_k'(x_i^{OE}) + \frac{1}{\sum_{s\neq i} \frac{1}{\sum_{k\in\mathcal{N}_s} l_k'(x_s^{OE})}}\right]\right\}, & \\ & \text{otherwise.} \end{cases} \tag{10}$$

In particular, for two links, when the minimum effective cost is less than R, for $i = 1, 2,\ j \in \mathcal{N}_i$, the strict OE prices are given by

$$p_j^{OE} = x_i^{OE}\left[\sum_{k\in\mathcal{N}_1} l_k'(x_1^{OE}) + \sum_{k\in\mathcal{N}_2} l_k'(x_2^{OE})\right].$$

The price characterization in Proposition 6 is a generalization of the price characterization in [1], and as in that paper, it will be useful in providing bounds on the inefficiency of price competition. However, the next example shows that even with strict OE, efficiency losses can be arbitrarily large.

C. Inefficiency of Strict OE

Example 3: Consider a one path network, which has n links with identical latency functions $l(x) = x/n$. Let the total flow be $d = 1$ and the reservation utility be $R = 1$.

For any n, the unique social optimum for this example is $x^S = 1/2$, with a corresponding social surplus $\mathbb{S}(x^S) = 1/4$. Using the price characterization given in Proposition 6 and the definition of a WE, it follows that there exists a unique strict OE, in which all providers charge the price $p^{OE} = 1/(n+1)$, and the equilibrium flow is $x^{OE} = 1/(n+1)$. The efficiency metric for this example is therefore

$$r_1(\{l_j\}, x^{OE}) = \frac{\left(1 - \frac{1}{n+1}\right)\frac{1}{n+1}}{\left(1 - \frac{1}{2}\right)\frac{1}{2}} = \frac{4n}{(n+1)^2},$$

which goes to 0 as $n \to \infty$.

This example establishes that even with strict OE, which rules out the pathological coordination failures discussed above, efficiency losses can be arbitrarily large. The reason for this is again the double marginalization problem, which increases the cost of transmission so much that there is no transmission in equilibrium along certain paths (e.g., along the single path in the example as $n \to \infty$). This type of behavior is also pathological at some level, especially when we think of networks where the reservation utility, R, of users is high enough. This leads us to define an even stronger notion of equilibrium, *strong OE*.

Definition 6: A vector $(p^{OE}, x^{OE}) \geq 0$ is a *strong OE* (Oligopoly Equilibrium) if it is a strict OE, and $\sum_{i\in\mathcal{I}} x_i^{OE} = d$. In this case, we refer to p^{OE} as the *strong OE price* and denote the set of strong OE flow allocations in a network with latency functions by $\{l_j\}$ by $\overrightarrow{OE}^d(\{l_j\})$.

The only difference between Definition 5 and Definition 6 is that in the latter we require all of the potential flow, d, to be transmitted. This will be the case when the reservation utility, R, of users is large enough.

D. Efficiency of Strong OE with Two Paths

We now characterize the efficiency properties of strong OE. We start with a two path network, with n_i links on path $i = 1, 2$, where each link is owned by a different provider. First, consider the following example, which illustrates that even with strong OE the efficiency loss can be worse than that in parallel link networks (which was shown to be bounded below by 5/6 in [1]).

Example 4: Consider a two path network, which has n links on path 1 with identically 0 latency functions and one link on path 2 with latency function $l(x_2) = x_2/2$. Let the total flow be $d = 1$ and the reservation utility be $R = 1$.

The unique social optimum for this example is $x^S = (1, 0)$. Using Proposition 6 and the definition of a WE, OE flows x^{OE} must satisfy

$$\sum_{j\in\mathcal{N}_1} l_j(x_1^{OE}) + x_1^{OE}\Big[\sum_{j\in\mathcal{N}_1} l'_j(x_1^{OE}) + \sum_{j\in\mathcal{N}_2} l'_j(x_2^{OE})\Big]$$

$$= \sum_{j\in\mathcal{N}_2} l_j(x_2^{OE}) + x_2^{OE}\Big[\sum_{j\in\mathcal{N}_1} l'_j(x_1^{OE}) + \sum_{j\in\mathcal{N}_2} l'_j(x_2^{OE})\Big].$$

Substituting for the latency functions and solving the above together with $x_1^{OE} + x_2^{OE} = 1$ shows that unique strong OE involves

$$x^{OE} = \Big(\frac{2}{n+2}, \frac{n}{n+2}\Big),$$

which goes to $(0, 1)$ as $n \to \infty$. The social surplus at the social optimum is 1, while the social surplus at the strong OE goes to $1/2$ as $n \to \infty$.

We next present two lemmas, which will be useful in providing a bound on the efficiency metric for strong OE. Note that these lemmas are valid for all OE as well. The first lemma allows us to assume without loss of generality that $R\sum_{i=1}^I x_i^S - \sum_{i=1}^I l_i(x_i^S)x_i^S > 0$ in the subsequent analysis.

Lemma 2: Given a set of latency functions $\{l_j\}_{j\in\mathcal{N}_i,\ i\in\mathcal{I}}$, assume that

$$\sum_{i\in\mathcal{I}}\Big(\sum_{j\in\mathcal{N}_i} l_j(x_i^S)\Big)x_i^S = R\sum_{i\in\mathcal{I}} x_i^S,$$

for some social optimum x_s. Then every $x^{OE} \in \overrightarrow{OE}(\{l_j\})$ is a social optimum, implying that $r_I(\{l_j\}, x^{OE}) = 1$.

Proof: Assume that $\sum_{i\in\mathcal{I}}\Big(\sum_{j\in\mathcal{N}_i} l_j(x_i^S)\Big)x_i^S = R\sum_{i\in\mathcal{I}} x_i^S$. Since x^S is a social optimum and every $x^{OE} \in \overrightarrow{OE}(\{l_j\})$ is a feasible solution to the social problem [problem (3)], we have

$$\begin{aligned} 0 &= \sum_{i\in\mathcal{I}}\Big(R - \sum_{j\in\mathcal{N}_i} l_j(x_i^S)\Big)x_i^S \\ &\geq \sum_{i\in\mathcal{I}}\Big(R - \sum_{j\in\mathcal{N}_i} l_j(x_i^{OE})\Big)x_i^{OE}, \quad \forall\ x^{OE} \in \overrightarrow{OE}(\{l_j\}). \end{aligned}$$

By the definition of a WE, we have $x_i^{OE} \geq 0$ and $R - \sum_{j\in\mathcal{N}_i} l_j(x_i^{OE}) \geq \sum_{j\in\mathcal{N}_i} p_j^{OE} \geq 0$ (where p_j^{OE} is the price of link $j \in \mathcal{N}_i$ at the OE) for all i. This combined with the preceding relation shows that x^{OE} is a social optimum. **Q.E.D.**

The following lemma provides a relation between the total flow admitted at an OE and at a social optimum.

Lemma 3: For a set of latency functions $\{l_j\}_{(j\in\mathcal{N}_i,\ i\in\mathcal{I})}$, let Assumption 1 hold. Let (p^{OE}, x^{OE}) be an OE 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 $x_i^{OE} > x_i^S$ for some i. Hence,

$$l_j(x_i^{OE}) \geq l_j(x_i^S), \qquad \forall\ j \in \mathcal{N}_i.$$

We also have $l_j(x_i^{OE}) > l_j(x_i^S)$ for some $j \in \mathcal{N}_i$. [Otherwise, we would have $l_j(x_i^S) = l'_j(x_i^S) = 0$ for all $j \in \mathcal{N}_i$, which yields a contradiction by the optimality conditions (4) and the fact that $\sum_{i\in\mathcal{I}} x_i^S < d$.] Using the definition of the WE and the optimality conditions (4), we obtain

$$R - \sum_{j\in\mathcal{N}_i}\Big(l_j(x_i^{OE}) - p_j^{OE}\Big) \geq R - \sum_{j\in\mathcal{N}_i}\Big(l_j(x_i^S) - x_i^S l'_j(x_i^S)\Big).$$

Combining the preceding with $l_j(x_i^{OE}) \geq l_j(x_i^S)$ for all $j \in \mathcal{N}_i$, with strict inequality for some j, and

$$p_j^{OE} \geq x_i^{OE} l'_j(x_i^{OE}) \geq x_i^S l'_j(x_i^S),$$

[using Proposition 6(a) and the fact that $xl'(x)$ is nondecreasing, cf. Assumption 1], we obtain a contradiction. **Q.E.D.**

The next theorem provides a tight lower bound on $r_2(\{l_j\}, x^{OE})$ [cf. (8)] for a strong OE. In the following, we assume without loss of generality that $d = 1$.

Theorem 1: Consider a two path network, with n_i links on path $i = 1, 2$, where each link is owned by a different provider, and link $j \in \mathcal{N}_i$ has a latency function l_j. Suppose that Assumption 1 holds and the price competition game has a strong OE. Then

$$r_2(\{l_j\}, x^{OE}) \geq \frac{1}{2}, \qquad \forall\ x^{OE} \in \overrightarrow{OE}^d(\{l_j\}). \quad (11)$$

Moreover, the bound is tight, i.e., there exists $\{l_j\}$ and $x^{OE} \in \overrightarrow{OE}^d(\{l_j\})$ that attains the lower bound in (11).

Proof: The proof follows a number of steps:

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

$$\inf_{\{l_j\}} \inf_{x^{OE} \in \overrightarrow{OE}^d(\{l_j\})} r_2(\{l_j\}, x^{OE}). \quad (12)$$

Given $\{l_j\}$, let $x^{OE} \in \overrightarrow{OE}(\{l_j\})$ and let x^S be a social optimum. By Lemma 3 and the fact that $x^{OE} \in \overrightarrow{OE}^d(\{l_j\})$ (i.e., it is a strong OE), we have

$$\sum_{i=1}^{2} x_i^{OE} = \sum_{i=1}^{2} x_i^S = 1.$$

This implies that there exists some i such that $x_i^{OE} < x_i^S$. Since the problem is symmetric, we can restrict ourselves to $\{l_j\}$ for which $x_1^{OE} < x_1^S$. We claim that

$$\inf_{\{l_i\} \in \mathcal{L}_2} \inf_{x^{OE} \in \overrightarrow{OE}^d(\{l_i\})} r_2(\{l_i\}, x^{OE}) \geq r_2^{OE}, \quad (13)$$

where

$$r_2^{OE} = \text{minimize}_{\substack{l_{i,j}^S, (l_{i,j}^S)' \geq 0 \\ l_{i,j}, l'_{i,j} \geq 0 \\ y_i^S, y_i^{OE} \geq 0}} \frac{R - y_1^{OE}\Big(\sum_{j \in \mathcal{N}_1} l_{1,j}\Big) - y_2^{OE}\Big(\sum_{j \in \mathcal{N}_2} l_{2,j}\Big)}{R - y_1^S\Big(\sum_{j \in \mathcal{N}_1} l_{1,j}^S\Big) - y_2^S\Big(\sum_{j \in \mathcal{N}_2} l_{2,j}^S\Big)} \quad (E)$$

subject to

$$l_{i,j}^S \leq y_i^S (l_{i,j}^S)', \qquad i = 1, 2,\ j \in \mathcal{N}_i, \quad (14)$$

$$\Big(\sum_{j \in \mathcal{N}_2} l_{2,j}^S\Big) + y_2^S \Big(\sum_{j \in \mathcal{N}_2} (l_{2,j}^S)'\Big) = \Big(\sum_{j \in \mathcal{N}_1} l_{1,j}^S\Big) + y_1^S \Big(\sum_{j \in \mathcal{N}_1} (l_{1,j}^S)'\Big), \quad (15)$$

$$\Big(\sum_{j \in \mathcal{N}_1} l_{1,j}^S\Big) + y_1^S \Big(\sum_{j \in \mathcal{N}_1} (l_{1,j}^S)'\Big) \leq R, \quad (16)$$

$$\sum_{i=1}^{2} y_i^S = 1, \quad (17)$$

$$l_{1,j} + l'_{1,j}(y_1^S - y_1^{OE}) \leq l_{1,j}^S, \qquad \forall\ j \in \mathcal{N}_1, \quad (18)$$

$$l_{i,j} \leq y_i^{OE} l'_{i,j}, \qquad i = 1, 2,\ j \in \mathcal{N}_i, \quad (19)$$

$$\sum_{i=1}^{2} y_i^{OE} = 1,$$

+ Strict OE Constraints.

Problem (E) can be viewed as a finite dimensional problem that captures the equilibrium and social optimum characteristics of the infinite dimensional problem given in (12). This implies that instead of optimizing over the entire function l_j for some $j \in \mathcal{N}_i$, $i \in \mathcal{I}$, we optimize over the possible values of $l_j(\cdot)$ and $l'_j(\cdot)$ at the equilibrium and the social optimum, which we denote by $l_{i,j}, l'_{i,j}, l_{i,j}^S, (l_{i,j}^S)'$. The constraints of the problem guarantee that these values satisfy the necessary optimality conditions for a social optimum and a strict OE (which are the same as the conditions for a strong OE). In particular, conditions (14) and (19) capture the convexity assumption on $l_j(\cdot)$ by relating the values $l_{i,j}, l'_{i,j}$ and $l_{i,j}^S, (l_{i,j}^S)'$ [note that the assumption $l_j(0) = 0$ is essential here]. Condition (15) is the optimality condition for the social optimum. Condition (18) uses the nondecreasing and the convexity assumption on the latency functions; since we are focusing on $\{l_j(\cdot)\}$ such that $x_1^{OE} \leq x_1^S$, we must have

$$l_{1,j} + l'_{1,j}(y_1^S - y_1^{OE}) \leq l_{1,j}^S,$$

for all $j \in \mathcal{N}_1$. Finally, the last set of constraints are the necessary conditions for a pure strategy OE. In particular, for a two path network, using Proposition 6, the Strict OE Constraints are given by

$$n_1 y_1^{OE}\Big[\sum_{j \in \mathcal{N}_1} l'_{1,j} + \sum_{j \in \mathcal{N}_2} l'_{2,j}\Big] + \sum_{j \in \mathcal{N}_1} l_{1,j} = n_2 y_2^{OE}\Big[\sum_{j \in \mathcal{N}_1} l'_{1,j} + \sum_{j \in \mathcal{N}_2} l'_{2,j}\Big] + \sum_{j \in \mathcal{N}_2} l_{2,j},$$

[and therefore n_1 and n_2 are also decision variables in problem (E)]. Note that given any feasible solution of problem (12), we have a feasible solution for problem (E) with the same objective function value. Therefore, the optimum value of problem (E) is indeed a lower bound on the optimum value of problem (12).

Step 2: Consider the following change of variables for problem (E)

$$l_1^S = \sum_{j\in\mathcal{N}_1} l_{1,j}^S, \qquad l_2^S = \sum_{j\in\mathcal{N}_2} l_{2,j}^S$$
$$l_1 = \sum_{j\in\mathcal{N}_1} l_{1,j}, \qquad l_2 = \sum_{j\in\mathcal{N}_2} l_{2,j},$$
$$(l_1^S)' = \sum_{j\in\mathcal{N}_1} (l_{1,j}^S)', \qquad (l_2^S)' = \sum_{j\in\mathcal{N}_2} (l_{2,j}^S)',$$
$$l_1' = \sum_{j\in\mathcal{N}_1} l_{1,j}', \qquad l_2' = \sum_{j\in\mathcal{N}_2} l_{2,j}',$$

and rewrite problem (E) as

$$r_2^{OE} = \text{minimize}_{\substack{l_i^S,(l_i^S)'\geq 0\\ l_i,l_i'\geq 0\\ y_i^S,y_i^{OE}\geq 0}} \frac{R - l_1 y_1^{OE} - l_2 y_2^{OE}}{R - l_1^S y_1^S - l_2^S y_2^S} \qquad (E')$$

subject to

$$\begin{aligned}
&l_i^S \leq y_i^S (l_i^S)', \qquad i = 1,2,\\
&l_2^S + y_2^S (l_2^S)' = l_1^S + y_1^S (l_1^S)',\\
&l_1^S + y_1^S (l_1^S)' \leq R,\\
&\sum_{i=1}^{2} y_i^S = 1,\\
&l_1 + l_1'(y_1^S - y_1^{OE}) \leq l_1^S,\\
&l_i \leq y_i^{OE} l_i', \qquad i = 1,2,\\
&\sum_{i=1}^{2} y_i^{OE} = 1,\\
&+ \text{ Strict OE Constraints.}
\end{aligned}$$

Note that this problem has a very similar structure to the finite-dimensional problem considered in the proof of Theorem 1 of [1] for parallel-link networks. Let $(\bar{l}_i^S, (\bar{l}_i^S)', \bar{l}_i, \bar{l}_i', \bar{y}_i^S, \bar{y}_i^{OE})$ denote the optimal solution of problem (E'). We have shown in [1] that $\bar{l}_i^S = 0$ for $i = 1, 2$.

Step 3: Using $\bar{l}_i^S = 0$ for $i = 1, 2$, and $\bar{l}_1 = 0$, $\bar{l}_1' = 0$, we see that

$$r_2^{OE} = \min_{\substack{l_2,l_2'\\ y_1^{OE},\ y_2^{OE}\geq 0\\ n_1,\ n_2\geq 1}} \quad 1 - \frac{l_2 y_2^{OE}}{R} \qquad (20)$$

$$\begin{aligned}
\text{subject to} \quad & l_2 \leq y_2^{OE} l_2',\\
& l_2 + n_2 y_2^{OE} l_2' = n_1 y_1^{OE} l_2',\\
& n_1 y_1^{OE} l_2' \leq R.\\
& \sum_{i=1}^{2} y_i^{OE} \leq 1.
\end{aligned}$$

Next, using the transformation $m_1 = n_1 y_1^{OE}$ and $m_2 = n_2 y_2^{OE}$ to write:

$$r_2^{OE} = \min_{\substack{l_2,l_2'\\ y_1^{OE},\ y_2^{OE}\geq 0\\ m_1,\ m_2\geq 0}} \quad 1 - \frac{l_2 y_2^{OE}}{R} \qquad (21)$$

$$\begin{aligned}
\text{subject to} \quad & l_2 \leq y_2^{OE} l_2',\\
& l_2 + m_2 l_2' = m_1 l_2',\\
& m_1 l_2' \leq R.\\
& \sum_{i=1}^{2} y_i^{OE} = 1,
\end{aligned}$$

though we also have to ensure that the solution to this program ensures that n_1 and n_2 are integers.

Now it can verified that $(\bar{l}_2, \bar{l}_2', \bar{y}_1^{OE}, \bar{y}_2^{OE}, \bar{m}_1, \bar{m}_2) = (\frac{R}{2}, \frac{R}{2}, 0, 1, 2, 1)$ is an optimal solution to the program (21), and moreover, it satisfies $n_1, n_2 \geq 1$, thus it is also a solution to (20). The corresponding optimum value is $r_2^{OE} = 1/2$. By (13), this implies that

$$\inf_{\{l_j\}} \inf_{x^{OE}\in\overrightarrow{OE}(\{l_j\})} r_2(\{l_j\}, x^{OE}) \geq \frac{1}{2}.$$

Finally, Example 4 shows that this bound is tight, i.e.,

$$\min_{\{l_j\}} \min_{x^{OE}\in\overrightarrow{OE}(\{l_j\})} r_2(\{l_j\}, x^{OE}) = \frac{1}{2}.$$

Q.E.D.

Therefore, when we focus on strong OE, there exists a tight bound of 1/2. In contrast to the case in Example 3, strong OE ensures that all of the traffic is transmitted in equilibrium, which is the key to the existence of a bound on the inefficiency of equilibrium.

The bound with strong OE is nonetheless worse than the efficiency bound in the parallel-link topology considered in [1]. This is again because of the double marginalization problem: each provider along path 1 has a greater incentive to increase its price (relative to the benchmark where all these links are owned by the same provider), because it does not internalize the reduction in the profits of the other link owners along the same path. Consequently, in Example 4, there are higher prices along path 1, and this induces greater fraction of users to choose path 2, increasing inefficiency. To see the role of serial links more clearly, consider a modified version of Example 4, where all n links along path 1 are owned by the same service provider. This would make the example equivalent to a parallel-link topology. In this case the unique strict OE flows are given by $x_1^{OE} = 2/3$ and $x_2^{OE} = 1/3$, and this example reaches the 5/6 bound of [1] rather than $1/2$ bound of Example 4.

E. Efficiency of Strong OE with Multiple Paths

We next consider an I path network, with n_i links on path i, where each link is owned by a different provider. The following example illustrates the efficiency properties of a strong OE in an I path network.

Example 5: Consider an I path network, which has n links on path 1 with identically 0 latency functions and one link on each of the paths $2, \dots, I$ with the same latency function $l(x) = x(I-1)/2$. Let the total flow be $d = 1$ and the reservation utility be $R = 1$.

Clearly, the unique social optimum for this example is $x^S = [1, 0, \dots, 0]$. Using Proposition 6 and the definition of a WE, it can be seen that the flow allocation at the unique strict (strong) OE is

$$x^{OE} = \Big[\frac{2/n}{1+2/n}, \frac{1}{(I-1)(1+2/n)}, \dots, \frac{1}{(I-1)(1+2/n)}\Big].$$

Hence the efficiency metric for this example is

$$r_I(\{l_j\}, x^{OE}) = 1 - \frac{1}{2}\left(\frac{1}{1+2/n}\right)^2,$$

which goes to 1/2 as $n \to \infty$.

The next theorem generalizes Theorem 1. The proof is similar to that of Theorem 1 and is omitted.

Theorem 2: Consider a general I path network, with n_i links on path $i \in \mathcal{I}$, where each link is owned by a different provider, and link j, $j \in \mathcal{N}_i$, has a latency function l_j. Suppose that Assumption 1 holds and the price competition game has a strong OE. Then

$$r_I(\{l_j\}, x^{OE}) \geq \frac{1}{2}, \qquad \forall\ x^{OE} \in \overrightarrow{OE}^d(\{l_j\}). \tag{22}$$

Moreover, the bound s tight, i.e., there exists $\{l_j\}$ and $x^{OE} \in \overrightarrow{OE}^d(\{l_j\})$ that attains the lower bound in (22).

F. Positive Latency at 0 Congestion

Unfortunately, the bound on the efficiency loss of strong OE does not generalize once we relax the assumption that $l_i(0) = 0$.

Example 6: Consider a two path network, which has n links on path 1 with identically 0 latency functions and one link on path 2 with latency function $l(x_2) = \epsilon x_2 + b$ for some scalars $\epsilon > 0$ and $b > 0$. Again the unique social optimum is $\bar{x}^S = (1, 0)$. The flows at the unique strict (strong) OE are given by

$$\bar{x}^{OE} = \left(\frac{2\epsilon + b}{\epsilon(n+2)}, \frac{\epsilon n - b}{\epsilon(n+2)}\right).$$

Let $\epsilon = b/\sqrt{n}$. Then, as $b \to 1$ and $n \to \infty$, we have that $\bar{x}^{OE} \to (0, 1)$, and the efficiency metric $r_2(\{l_j\}, x^{OE}) \to 0$.

This example shows that the efficiency loss could be arbitrarily high even at a strong OE for a network that involves parallel and serial links if the assumption $l_i(0) = 0$ is relaxed. This establishes:

Proposition 7: In the presence of positive latency at zero congestion, strong OE with the parallel-serial topology can be arbitrarily inefficient.

It is useful to note that in the same example with the parallel-link topology (i.e., all n links along path 1 owned by the same provider), we would have

$$x^{OE} = \begin{cases} \left(\frac{b+2\epsilon}{3\epsilon}, \frac{\epsilon - b}{3\epsilon}\right), & \text{if } \epsilon \geq b, \\ (1, 0), & \text{otherwise.} \end{cases}$$

Consequently, $b \to 1$ and $\epsilon \to 0$, we have that $x^{OE} \to (1, 0)$, and $r_2(\{l_j\}, x^{OE}) \to 1$. Therefore, the highly inefficient equilibrium is a result of the parallel-serial topology, not of the assumption that there is positive latency at 0 congestion. In fact, [1] shows that with parallel topology, but positive latency at 0 congestion, there is again a tight bound of $2\sqrt{2} - 2$ on efficiency, which is quite close to, but slightly lower than 5/6.

V. CONCLUSIONS

In this paper, we presented an analysis of price competition in communication networks with congestion. The focus has been the efficiency implications of price competition in networks with the serial-parallel topology.

Our major result is that contrary to the case of pure parallel-link topology studied in [1], the parallel-serial topology leads to significant efficiency losses relative to the social optimal. In particular, OE can now be arbitrarily inefficient. This is partly due to an extreme (pathological) form of double marginalization, whereby all serial providers on a particular path charge prohibitively high prices expecting others on that path to do so as well.

We showed that the concept of strict OE, which requires all service providers to play strict best responses, removes this pathological behavior, but the efficiency loss of strict OE is also unbounded because of the related double marginalization problem. In particular, the total cost of transmission on a path consisting of many serial providers can be sufficiently high that most of the users do not transmit in equilibrium.

Yet, when users value transmission sufficiently, we may expect them to transmit even with high costs. Motivated by this, we defined a stronger notion of equilibrium, strong OE, which is a strict OE with all of the traffic transmitted in equilibrium. For strong OE, we showed that as long as there is zero latency at zero congestion, there is a tight bound of 1/2 on the inefficiency resulting from price competition.

Once the zero latency at zero congestion assumption is removed, however, there is no such tight bound even with strong OE, and the equilibrium can once again be arbitrarily inefficient.

In all the examples of extreme inefficiency, there is a flavor of pathological results, however. Therefore, we suspect that these worst-case results are not informative as to whether for realistic network structures such high levels of inefficiency can emerge. This is an area for future research and methods similar to those in Friedman's analysis of genericity of inefficiency of selfish routing may be useful in this context as well (see [10]).

Acknowledgments: We thank Attila Ambrus, Muhamet Yildiz and various seminar participants for very useful comments and suggestions.

REFERENCES

[1] Acemoglu, D., and Ozdaglar, A.,“Competition and efficiency in congested markets,” submitted for publication, January 2005.

[2] Acemoglu, D., and Ozdaglar, A.,“Flow control, routing, and performance from service provider viewpoint,” *LIDS report*, WP1696, May 2004.

[3] Acemoglu, D., Ozdaglar, A., and Srikant, R., “The marginal user principle for resource allocation in wireless networks,” *Proceedings of the 43rd IEEE Conference on Decision and Control*, 2004.

[4] Anshelevich E., Dasgupta A., Kleinberg J., Tardos E., Wexler T., and Roughgarden T., “The price of stability for network design with selfish agents,” *IEEE Symposium on Foundations of Computer Science,* pp. 295-304, 2004.

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

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

[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] Dasgupta, P. and Maskin E., “The existence of equilibrium in discontinuous economic games. 2: Theory,” *Review of Economic Studies*, vol. 53, pp. 1-26, 1986.

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

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

[12] Harsanyi, J., “Games with Randomly Distributed Payoffs: a New Rationale for Mixed Strategy Equilibrium Points” *International Journal of Game Theory*, vol. 1, pp. 1-23.

[13] Hayrapetyan, A., Tardos, E., and Wexler T., “A network pricing game for selfish traffic,” *Journal of Distributed Computing*, 2005.

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

[15] Huang, X., Ozdaglar, A., and Acemoglu, D., “Efficiency and Braess' paradox under pricing in general networks,” forthcoming *JSAC Special issue: Price-Based Access Control and Economics for Communication Networks*.

[16] Johari, R. and Tsitsiklis, J., “Network resource allocation and a congestion game,” *Mathematics of Operations Research*, 2004.

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

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

[19] 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.

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

[21] Perakis, G., “The Price of anarchy when costs are non-separable and asymmetric”, manuscript.

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

[23] Sanghavi S. and Hajek B., “Optimal allocation of a divisible good to strategic buyers,” 43rd IEEE Conference on Decision and Control - CDC 2004.

[24] Wardrop, J. G., “Some theoretical aspects of road traffic research,” *Proceedings of the Institute of Civil Engineers*, Pt. II, vol. 1, pp. 325-378, 1952.

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