NET Institute*

www.NETinst.org

Working Paper #07-21

Revised January 2008

Competition and Contracting in Service Industries

Ramesh Johari
Stanford University

Gabriel Weintraub
Columbia Business School

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Competition and Contracting in Service Industries

Ramesh Johari • Gabriel Y. Weintraub

Management Science and Engineering, Stanford University

Columbia Business School

ramesh.johari@stanford.edu • gweintraub@columbia.edu

First Version: September 2007; This Version: January 2008

Two very different contractual structures are commonly observed in service industries with congestion effects: *service level guarantees* (SLGs) and *best effort* (BE) service. We analyze the impact of these contractual agreements on market outcomes in oligopolistic industries. First, we consider a model where firms compete by setting prices and SLGs simultaneously. The SLG is a contractual obligation on the part of the service provider: regardless of how many customers subscribe, the firm is responsible for investing so that the congestion experienced by all subscribers is equal to the SLG. We then consider the BE contractual model where firms compete by setting prices and investment levels simultaneously. With BE contractual agreements, firms provide the best possible service given their infrastructure, but without an explicit guarantee. Using the Nash equilibria (NE) of the games played by firms, we compare these competitive models in terms of the resulting prices, service levels, firms' profits, and consumers' surplus.

We first show that the SLG game can be reduced to a standard pricing game, greatly simplifying the analysis of this otherwise complex competitive scenario. We then compare the SLG game with the BE game; equilibria for the BE is characterized in a previous paper. Using these results we show that in the case of constant returns to investment, while the NE price for the SLG game is perfectly competitive, firms obtain positive markups in the unique NE for the BE game. We also study the firms' choice of the strategy space, i.e., whether to offer SLG or BE contracts to the consumer, and find that competition is intensified if even one firm chooses to offer SLG contracts. Our results contribute to the basic understanding of competition and contracting in service industries and yield insight into business and policy considerations.

1. Introduction

Two very different contractual structures are commonly observed in service industries with congestion effects: *service level guarantees* (SLGs), and *best effort* (BE) service. The SLG is a contractual obligation on the part of the service provider: regardless of how many customers subscribe, the firm is responsible for investing in infrastructure, capacity, or service quality so that the congestion experienced by all subscribers is equal to the SLG. Examples of such guarantees abound; for example, UPS and FedEx offer overnight delivery guarantees for their premium services. Web service providers will often deliver service level guarantees on web server availability. By contrast, in other industries firms offer a *best effort* contractual agreement, where firms provide the best possible service given their infrastructure, but without an explicit guarantee. For example, in telecommunications, for the average consumer of Internet service, network service providers typically do not offer guarantees to purchasers of network service.

This paper studies the following range of questions: In a competitive setting, which of these two contractual agreements give rise to more intense competition, and thus lower prices and higher service levels? Which contractual agreement is most beneficial for firms? Which one is most beneficial for consumers? When do SLGs emerge endogenously as a result of competition? We address these questions by comparing the two competitive models through the Nash equilibria (NE) of the respective games played by firms. Our results contribute to the basic understanding of competition and contracting in service industries and yield insight into business and policy considerations as we explain below. For example, while the Internet has traditionally employed a BE service model, network service providers face increasing pressure to deliver end-to-end SLGs to users, as these are required to meet the increasing needs of real-time traffic such as video and voice applications (The Economist 2006). Since the emergence of SLGs would be a shift in the underlying competitive paradigm of the industry, we believe that any industry analysis over these strategic decisions should consider models such as those in our paper.

Our oligopoly model consists of a finite collection of providers that compete to deliver a service with *congestion effects*: the benefits consumers experience are offset by a negative externality that is increasing in the total volume of customers served. We consider a model where a consumer's demand depends on the *full price*, which is a sum of two components: the price of the service, and a congestion cost that increases with the total number of consumers subscribing to the same firm. Firms can invest in their service (e.g., through infrastructure or capacity); investment lowers the congestion cost experienced by consumers. This model is relevant to a wide range of services, including transportation, cellular telephony, data networking services, and call centers.

We first define and study a *SLG game* where firms simultaneously choose two decision variables: price and a SLG. Demands are allocated based on the chosen prices and SLGs, and then firms make the necessary investment expenditures to meet their SLGs. We show that under very general conditions, the SLG game can be reduced to a standard pricing game, but with a modified cost function that can be explicitly characterized in terms of the model parameters. Additionally, we show that for a broad class of relevant congestion cost functions, the resulting equivalent pricing game is in fact pricing with convex costs. Moreover, we find that when the original game exhibits *constant returns to investment*—a property satisfied by a wide range of congestion models, including loss systems—the resulting equivalent pricing game yields constant marginal costs. If, additionally, services are perfect substitutes, we obtain standard Bertrand competition among firms. These analytical insights dramatically reduce the complexity of analyzing competition in SLGs in service industries, as standard results from pricing games can be used to characterize the NE for the SLG game. Our analytical results here are a novel contribution to the study of such games in their own right.

Having characterized the SLG game, we undertake a comparison of the equilibrium outcomes of the SLG game with those of a *BE game*, formulated and studied by Weintraub et al. (2007) and in a special case by Xiao et al. (2007). In the BE game firms set prices and investment levels simultaneously, and then consumers choose their service providers. There are no explicit guarantees on service levels. If firms are ex-ante symmetric, and services are perfect substitutes and exhibit constant returns to investment, we find a sharp prediction: while the NE price for the SLG game is perfectly competitive and all firms obtain zero profits, firms obtain positive markups in the unique NE for the BE game. In the SLG game, the total number of consumers served is socially efficient. In the BE game, because the full price is higher, the number of consumers is below the socially efficient level. Price competition is more intense in the SLG game than in the BE game; in the former firms cannot exercise market power. We also compare both models for alternative cost structures. Our comparisons are similar in spirit to the well known difference between price competition and quantity competition in standard oligopoly games (Vives 2001). We emphasize, however, that while the SLG game can be reduced to an equivalent pricing game, the BE game *cannot* be reduced to an equivalent quantity competition game.

We conclude by considering the firms' choice of contractual agreement. In many cases, one of the most important strategic decisions made by firms is *which* contractual agreement they will offer to customers: a SLG or BE service. Motivated by this fact, we also consider a multistage game where firms can first choose their strategic variable to be either SLG or investment, in addition to price. In the second stage, each firm chooses its actions: either a price and investment level, or a price and SLG. For constant returns to investment, we find that if even one firm chooses to set price and a SLG, then in any resulting second stage NE, all firms obtain zero profits. Thus the only subgame perfect equilibrium where all firms make positive profits is the one where all firms choose prices and investments at the second stage. Our result implies that if all other firms have chosen to compete by offering prices and investment levels, a deviation to offer SLGs is not profitable.

Our model and results provide a broad framework within which a range of service industries can be studied. Our attention to both the cost structure and the form of contracts with consumers offers guidance for policy analysis and business strategy. In particular, we have obtained strong insights for industries where constant returns to investment can be surmised to hold. As previously mentioned, such models include examples where loss and blocking probabilities are measures of user satisfaction. This is the case in many important telecommunication markets, including those for wireless services (Campo-Rembado and Sundararajan 2004). In such industries, our insights suggest that competition in prices and SLGs is stronger than competition in prices and investment levels. When firms compete in prices and SLGs, prices are perfectly competitive, firms obtain zero profits, and social welfare is maximized. When firms compete in prices and investment levels, they can sustain mark-ups and make positive profits, full prices are higher, and the number of consumers served is below the socially efficient level. Thus competition in prices and SLGs is harsh for firms, and beneficial for consumers, relative to competition in prices and investment levels. Additionally, our results suggest that in this case, providers would not have incentives to deviate from a BE contracting model to offer SLGs, if customers are homogeneous in how they trade off money and congestion (as we have assumed in our model). Thus we expect that SLGs can only be profitable if customers have heterogeneous preferences, in which case we expect that SLGs can help to segment the market.

The remainder of the paper is organized as follows. In Section 2, we discuss related results in the literature. We introduce our model in Section 3. In Section 4, we analyze equilibria for the SLG game. In Section 5 we compare the SLG game with the BE game. In Section 6 we analyze the firms' choice of contractual agreement. Finally, Section 7 presents conclusions and directions for future research.

2. Related Work

Our results contribute to the recent literature in operations management that studies competition in service industries with SLGs (see Allon and Federgruen 2005a,b, 2006, Cachon and Harker 2002, So 2000, for a survey). These papers, with the exception of Cachon and Harker (2002), do not assume that price and congestion cost can be aggregated in a full price. Thus our equivalency with a pricing game result is not valid in their settings. Additionally, none of these papers *compare* different contractual models (SLGs versus BE) as we do in this paper. In their series of papers, Allon and Federgruen study competition between oligopolistic firms that compete by setting prices and service levels (defined as the difference between an industry benchmark and the firm's actual SLG). They characterize equilibria for a rich set of models that consider: (1) different timing of decisions; (2) different congestion cost functions derived from queuing waiting time expressions; and (3) multiple customer segments.

While the objectives are very different, our model and analysis has similarities with Cachon and Harker (2002). Their work analyzes competition and outsourcing in industries that exhibit increasing returns to investment. In our work, we compare two different contractual structures in industries that exhibit nonincreasing returns to investment. Cachon and Harker (2002) assume that consumers' demand depends on the full price, as we do in our paper. They consider a duopoly pricing model with scale economies; firms incur a cost that is a concave monomial in the served demand. The authors show that two different specific game structures can be fit into this model: a queueing game where consumers' congestion cost is given by the M/M/1 delay function; and an EOQ (economic order quantity) game. The former game is a special case of the general class of games studied in this paper.[1] Indeed, in our paper, we derive a systematic and general way of reducing the SLG game to an equivalent pricing game with a modified cost function under weak assumptions. In particular, we derive an explicit expression for the modified cost function, and discuss its convexity/concavity properties for a broad range of congestion models.

Our paper is also related to the growing recent literature on welfare analysis in congestion games in transportation and communication networks; see Acemoglu and Ozdaglar (2006), Ozdaglar (2006) (who analyze a pricing game), and Roughgarden (2005) for an overview. Most closely related to our paper is the work of Weintraub et al. (2007) and Xiao et al. (2007), who, as discussed in the introduction, analyze the BE game. We use these results to compare the BE game with the SLG game.

Miller and Pazgal (2006) study a two-stage game where in the first stage firms can choose between specifying an "output" strategy (a goal that they will reach at any cost) or an "input" strategy (the cost that they are willing to incur in pursuit of their goal). While our model and purpose is very different, our analysis of the two-stage game in Section 6 is similar in spirit to Miller and Pazgal (2006), where SLG is like an output strategy and investment is like an input strategy. Our analysis is also related to the price vs. quantity competition studies of Singh and Vives (1984) and Cheng (1985). They show that for a broad range of reasonable demand structures, it is a dominant strategy for the players to choose to set quantities (as opposed to prices) when goods

[1]In the EOQ game studied by Cachon and Harker (2002), there is an industry standard such that firms offer a zero SLG. Hence, firms compete via explicit prices only.

are substitutes.

3. Model

In this section we introduce our model of service provision. In Sections 3.1 and 3.2 we use this model to define the SLG game and the BE game, respectively. Suppose there are N profit-maximizing incumbent firms that compete for consumers. The benefits consumers experience are offset by a negative externality that is increasing in the total volume of consumers served. However, firms can invest to mitigate this congestion effect. We assume there are no externalities among firms; therefore, consumers served by firm j are unaffected by the investments of firms other than j.

For firm j, we let p_j, I_j, and q_j denote, respectively, the price per consumer charged, the investment level chosen, and the number of consumers served. We let $b_j(q_j)$ be the nonnegative production cost for firm j. Each investment level I_j is measured in currency units, and the resulting physical capacity can be a nonlinear function of this investment expenditure. Hence, the profit of firm j is given by

$$\pi_j(p_j, I_j, q_j) = p_j q_j - b_j(q_j) - I_j. \tag{1}$$

Thus profits of firms are determined by the price, investment expenditure, and number of consumers served.

The demand side of the market formalizes a congestion externality among consumers subscribed to the same service provider. We assume that when firm j invests at a level I_j and serves q_j consumers, all customers of that firm experience a *congestion cost* $\ell_j(q_j, I_j)$. The congestion function $\ell_j(q_j, I_j)$ represents the disutility perceived by consumers due to congestion. For each j, the function $\ell_j(q_j, I_j)$ is nonnegative; it is increasing in q_j, incorporating the congestion effect into the model; and it is decreasing in I_j, incorporating the effect of investment.

We assume that the disutility from congestion is measured in monetary units; thus a consumer will be sensitive to the *sum* of price and congestion, which we call the *full price*. The *full price* experienced by a customer of firm j is defined as $f_j = p_j + \ell_j(q_j, I_j)$. The full price is the actual price paid plus the disutility perceived due to congestion. The assumption implies that consumers are homogenous in how they trade off money and congestion. This is a common simplifying assumption in competition models with congestion effects (e.g., Xiao et al. 2007, Weintraub et al. 2007, Acemoglu and Ozdaglar 2006, Cachon and Harker 2002).

For later reference, we define v_j as follows:

$$v_j(q_j) \equiv \min_{I_j \geq 0} \quad q_j \ell_j(q_j, I_j) + I_j \tag{2}$$

The function $\ell_j(q_j, I_j)$ represents the congestion cost experienced per unit mass of consumers. Hence, $q_j \ell_j(q_j, I_j)$ represents the sum of congestion costs experienced by all subscribers to firm j. Therefore, $v_j(q_j)$ is the minimum sum of total congestion cost and investment cost for a given mass of consumers q_j. We introduce the following assumption.

Assumption 1. *For all j and for all $q_j \geq 0$, the minimization problem (2) admits a unique optimal solution.*

Assumption 1 is kept throughout the paper. For all j and $q_j \geq 0$, we let $I_j(q_j)$ be the solution of problem (2). Because $I_j(q_j)$ minimizes the sum of total congestion cost and investment cost for a given mass of consumers q_j, $I_j(q_j)$ is called the *efficient investment level* given a consumer mass q_j. We define the *total cost function* for firm j as $t_j \equiv b_j + v_j$.

3.1 Service Level Guarantee Game

In the service level guarantee (SLG) game, firms compete by setting prices p_j and SLGs $\overline{\ell}_j$ *simultaneously*. Given prices p_j and SLGs $\overline{\ell}_j$, demand quantities for each firm are realized. Given our model for consumers' utility, demands will be a function of the full prices $f_j = p_j + \overline{\ell}_j$. Therefore, given full prices, the demand quantity for firm j can be denoted $d_j(f_j, \boldsymbol{f}_{-j})$, where $\boldsymbol{f}_{-j}$ denotes the vector of full prices set by the competitors of firm j. (Throughout the paper we use boldface type to denote vectors.) Finally, firms make the necessary investments to meet their SLGs given their demand quantities. Hence, if q_j consumers require service from firm j, the firm will invest I_j to ensure that $\ell_j(q_j, I_j) = \overline{\ell}_j$. We analyze this game through its pure strategy Nash equilibria, abbreviated NE.

Definition 1. *A* pure strategy Nash equilibrium *(NE) for the SLG game is a tuple consisting of prices $\boldsymbol{p}^{SLG}$, SLGs $\overline{\boldsymbol{\ell}}^{SLG}$, full prices $\boldsymbol{f}^{SLG}$, investment levels $\boldsymbol{I}^{SLG}$, and demand quantities $\boldsymbol{q}^{SLG}$,* *such that each firm maximizes profit given prices and SLGs of other firms; i.e., for all $j = 1, \ldots, N$,*

$$(3) \qquad (p_j^{SLG}, \bar{\ell}_j^{SLG}, f_j^{SLG}, I_j^{SLG}, q_j^{SLG}) \in \underset{p_j, \bar{\ell}_j, f_j, I_j, q_j}{\text{argmax}} \ \pi_j(p_j, I_j, q_j)$$

$$\textit{s.t.}$$

$$q_j = d_j(f_j, \boldsymbol{f}_{-j}^{SLG})$$

$$f_j = p_j + \bar{\ell}_j$$

$$\bar{\ell}_j = \ell_j(q_j, I_j)$$

$$p_j, \bar{\ell}_j, f_j, I_j, q_j \geq 0 \, .$$

Firm j finds its optimal price and service level guarantee, given the full prices $\boldsymbol{f}_{-j}$ of its competitors. Firm j maximizes profits subject to several constraints. The first constraint states that the number of consumers served by firm j is given by its demand function, which in turn depends on the full prices of all the firms. The second constraint defines the full price. The third constraint ensures firm j invests the necessary amount to meet its SLG. The last constraint ensures that all variables are non-negative.

3.2 Best Effort Game

In the best effort (BE) game, firms compete by setting prices and investment levels simultaneously. Given prices p_j and investment levels I_j, demand quantities q_j are realized; let the demand quantities be denoted by $\boldsymbol{q} = W(\boldsymbol{p}, \boldsymbol{I})$. The main setting we consider, introduced later in Section 5, considers a demand model where services being offered are perfect substitutes. In this case, demands in the BE game are given by a static equilibrium in which all firms that attract a positive mass of customers offer the same full price. Such an equilibrium is commonly known as a *Wardrop equilibrium*, particularly in the transportation literature (Wardrop 1952). In this case $W(\boldsymbol{p}, \boldsymbol{I})$ would denote the set of Wardrop equilibria given prices $\boldsymbol{p}$ and investment levels $\boldsymbol{I}$.

We also analyze this game through its NE.

Definition 2. *A* pure strategy Nash equilibrium *(NE) for the BE game is a triple consisting of prices $\mathbf{p}^{BE}$, investment levels $\boldsymbol{I}^{BE}$, and demand quantities $\mathbf{q}^{BE}$, such that, each firm maximizes* *profit given prices and investment levels of other firms; i.e., for all $j = 1, \ldots, N$,*

$$(p_j^{BE}, I_j^{BE}, \boldsymbol{q}^{BE}) \in \underset{p_j, I_j, \boldsymbol{q}}{\text{argmax}} \quad \pi_j(p_j, I_j, q_j) \qquad (4)$$

$$\textit{s.t.}$$

$$\boldsymbol{q} = W(p_j, \boldsymbol{p}_{-j}^{BE}, I_j, \boldsymbol{I}_{-j}^{BE})$$

$$p_j, I_j \geq 0 \; .$$

Thus in a NE, firm j finds its optimal price and investment level, given the prices $\boldsymbol{p}_{-j}$ and investment levels $\boldsymbol{I}_{-j}$ of its competitors.

4. Equilibrium Analysis of the SLG Game

In order to compare the NE of the SLG game and the BE game, we must first analyze equilibria in each of the two games. As previously discussed, equilibrium behavior in the BE game was studied by Weintraub et al. (2007) (henceforth, WJV). In this section, we provide an equilibrium analysis of the SLG game. Our results here constitute a novel contribution to the analysis of such games in their own right.

The centerpiece of our approach is to first show in Section 4.1 that the SLG game is in fact equivalent to a standard pricing game, but where costs for each firm j are given by the total cost functions t_j. We use this characterization to study the impact of cost structure in Section 4.2. As we discuss, WJV have shown that for many congestion models of interest the function t_j is smooth and convex; in this case the SLG game is equivalent to a pricing game with convex costs. We then show that for a class of models that exhibit *constant returns to investment*, the resulting pricing game is in fact one with *linear* cost functions t_j. These results significantly simplify the analysis of this otherwise complex competitive scenario, as standard results for pricing games can be applied.

4.1 Equivalence with Pricing Game

In this section, we derive a key insight for our analysis: the simultaneous pricing and SLG game played by firms is equivalent to a standard pricing game where costs for each firm j are given by the total cost functions t_j.

We make the following assumption.

Assumption 2. *For all j and for all $\boldsymbol{f}_{-j}$, firm j can choose a full price $f_j \geq 0$, a price $p_j \geq 0$, and an investment level $I_j \geq 0$, so that $\pi_j\left(p_j, I_j, d_j(f_j, \boldsymbol{f}_{-j})\right) \geq 0$.*

Assumption 2 is satisfied if, for example, for all j and for all $\boldsymbol{f}_{-j}$, $d_j(f_j, \boldsymbol{f}_{-j}) = 0$, for large enough f_j, and if $b_j(0) = 0$, as is the case in many models of interest. In this situation, firm j can choose not to serve any consumers and to make zero profits. *Assumption 2 is kept throughout the paper.*

The problem that firm j solves to find its optimal price and service level guarantee, given the full prices $\boldsymbol{f}_{-j}$ of its competitors (problem (3)), is equivalent to:

$$
\begin{aligned}
(5) \qquad & \max_{f_j, I_j, q_j} && (f_j - \ell_j(q_j, I_j))\, q_j - b_j(q_j) - I_j \\
& \text{s.t.} && \\
& q_j &=& \; d_j(f_j, \boldsymbol{f}_{-j}) \\
& f_j, I_j, q_j &\geq& \; 0\,.
\end{aligned}
$$

By Assumption 2 the omitted constraint $p_j \geq 0$ must be satisfied in an optimal solution. If not, the firm would make negative profits. Additionally, the omitted constraint $\overline{\ell}_j \geq 0$ must be satisfied, because ℓ_j is non-negative.

The optimization over I_j can be decoupled from the other variables, hence the program (5) is equivalent to:

$$
(6) \qquad \max_{f_j \geq 0} \quad \pi_j(f_j, \boldsymbol{f}_{-j}),
$$

where:

$$
(7) \qquad \pi_j(f_j, \boldsymbol{f}_{-j}) = f_j d_j(f_j, \boldsymbol{f}_{-j}) - b_j\left(d_j(f_j, \boldsymbol{f}_{-j})\right) - v_j\left(d_j(f_j, \boldsymbol{f}_{-j})\right).
$$

Here (through an abuse of notation) we let $\pi_j(f_j, \boldsymbol{f}_{-j})$ denote the profits firm j makes if it invests optimally, when it sets full price f_j, and all other firms set full prices $\boldsymbol{f}_{-j}$. If f_j^* solves the above problem, we can recover the optimal solution of the original problem by setting $q_j^* = d_j(f_j^*, \boldsymbol{f}_{-j})$, $I_j^* = I_j(q_j^*)$, $\overline{\ell}_j^* = \ell_j(q_j^*, I_j^*)$, and $p_j^* = f_j^* - \overline{\ell}_j^*$.

It is interesting to note that in any best response calculation, firms optimally choose an investment level that minimizes the sum of total congestion cost and investment cost. In other words, firms invest efficiently given the mass of consumers they serve. Intuitively, firms invest at the efficient investment level because they can extract any additional social surplus generated by investment through appropriate pricing. WJV find a similar result when firms compete by setting prices and investment levels simultaneously; Xiao et al. (2007) also employ such a result in their analysis of a special case of the pricing and investment game. Scotchmer (1985) finds a similar result in pricing and investment games for club goods.

Note that problem (6) is equivalent to the optimal pricing problem of a firm that faces a demand function $d_j(f_j, \boldsymbol{f}_{-j})$, has a cost function t_j, and where its competitors prices are given by $\boldsymbol{f}_{-j}$. Therefore, we have proved the following key result.

Proposition 1. *The SLG game is equivalent to a pricing game, with demand functions d_j and cost functions t_j for all firms j: i.e., firms compete by simultaneously choosing full prices f_j, and payoffs are given by π_j as defined in* (7)*.*

Pricing games are one of the most studied models in oligopoly theory and industrial organization. Proposition 1 is useful because it permits use of the wide set of existing results in pricing games to characterize Nash equilibria for the game where firms compete in prices and SLGs simultaneously in an industry with congestion effects. It is worth noting that the equivalence is valid under weak conditions (Assumptions 1 and 2). Neither ex-ante symmetric firms, nor more specific assumptions on the demand or cost functions are required.

4.2 Cost Structure

In this section we consider the impact of cost structure on NE in the SLG game. For many congestion models of interest, the resulting total cost function t_j is smooth and convex. Moreover, for an important subset of this class of models, t_j is linear. These insights have significant consequences for the SLG game due to Proposition 1, because an important subset of results in pricing games assume that cost functions are smooth and convex.

We introduce the following assumptions over the functions b_j and ℓ_j.

Assumption 3. *For all j, the production cost function b_j is strictly increasing, twice differentiable, and convex, with $b_j(0) = 0$.*[2]

Assumption 4.
For all j,

1. *The congestion cost function $\ell_j(q, I)$ is finite for all $q \geq 0$ and $I > 0$, and is twice differentiable in this region. Further, for all $q > 0$ and $I > 0$, $\partial \ell_j(q, I)/\partial q > 0$, $\partial \ell_j(q, I)/\partial I < 0$ and $\ell_j(0, I) = 0$. In addition, $\ell_j(0, 0) = 0$, and $\lim_{I \downarrow 0} \ell_j(q, I) = \ell_j(q, 0) = \infty$ for all $q > 0$.*

[2]We assume throughout the paper that derivatives at zero are right directional derivatives.

2. *The total congestion cost $q\ell_j(q,I)$ is convex in (q,I) and strictly convex in I.*[3]

The first part of Assumption 4 implies that congestion increases with the mass of subscribers, incorporating the congestion effect, and decreases with investment expenditures. It also incorporates natural boundary conditions. The second part of the assumption imposes a form of convexity over the congestion cost functions. As we discuss later in this section, many important congestion cost models satisfy Assumption 4. We note that Assumption 4 guarantees that Assumption 1 holds.

The following straightforward result was first observed by WJV. The proof can be found in the Appendix.

Proposition 2. *Suppose that Assumptions 3 and 4 hold. For all j, the function t_j is strictly increasing, convex, and continuous, with $t_j(0) = 0$. Additionally, $t_j(q)$ is twice differentiable for $q > 0$.*

The result implies that if Assumptions 3 and 4 hold, the SLG game can be reduced to a pricing game with convex costs (by Propositions 1 and 2). Hence, standard results from pricing games yield conditions over the model primitives that guarantee existence and uniqueness of NE for the SLG game. For example if, for all j, $d_j(f_j, \boldsymbol{f}_{-j})$ is log-concave in f_j (whenever $d_j(f_j, \boldsymbol{f}_{-j}) > 0$), and additional minor technical conditions hold, then a NE is guaranteed to exist for the SLG game (Vives 2001). Similarly, uniqueness of NE can be guaranteed using a standard diagonal strict dominance condition (Rosen 1965).

We now demonstrate that for a wide variety of convex congestion cost models, the resulting total cost function is *linear*. We consider cost models of the following form.

Assumption 5. *For all j, $b_j(q_j) = \beta_j q_j$, with $\beta_j \geq 0$, and Assumption 4 holds. Further, for all j, and all $K > 0$, there holds:*

$$\ell_j(q_j, I_j) = \ell_j(Kq_j, KI_j).$$

Under Assumption 5, for all j, there exists a function h_j such that $\ell_j(q_j, I_j) = h_j\left(q_j/I_j\right)$. Thus we assume that the congestion cost depends only on the *ratio* of quantity to investment expenditure, q/I. Congestion cost functions satisfying Assumption 5 exhibit *constant returns to investment* as defined by WJV: if both the quantity q and investment I are scaled by a constant K, then the total

[3] $q\ell(q,I)$ being (strictly) convex refers to $q\ell(q,I)$ being (strictly) convex on the set $[0,\infty) \times (0,\infty)$.

congestion cost $q\ell(q, I)$ is scaled by the same constant K. We have the following important result that we prove in the Appendix.[4]

Proposition 3. *Suppose Assumption 5 holds. Then, for all j, there exists $c_j > 0$ such that $t_j(q) = c_j q$.*

As a result of Propositions 1 and 3, we conclude that if the marginal cost of production is constant, and the congestion cost exhibits constant returns to investment, then the SLG game is in fact equivalent to a pricing game with constant marginal costs. This is a significant simplification from the complexity initially presented by such models. For example, as is standard in such games, if ex-ante symmetric firms provide perfect substitute goods then perfectly competitive pricing emerges as the unique symmetric equilibrium.

WJV discuss several important congestion cost models that satisfy Assumptions 4 and 5, including service provision or make-to-order facilities that can be modeled as loss systems (i.e., consumers leave if they find all servers busy) and where firms invest to increase the service rate. As one example, Hall and Porteus (2000) use loss system models to analyze competition in capacitated systems. Xiao et al. (2007) use a similar model to analyze competition among private toll roads. Loss systems are also a plausible model for wireless service provision, where we expect that consumers are most sensitive to the fraction of times they are unable to connect to a base station after paying a subscription fee to a given provider.[5] Indeed, constant returns to investment are exhibited, for example, when the marginal productivity of investment expenditure in building capacity is constant and $\ell(q, I)$ represents Erlang's formula for a loss system with mean arrival rate q, service rate I, and a fixed number of servers. Constant returns to investment are also exhibited for alternative loss models like the loss probability of an $M/M/1/s$ system or the exceedance probability of an $M/M/1$ queue. Assumption 4 is satisfied if $\ell(q, I)$ is derived from one of the previous loss models and the marginal productivity of investment expenditure in building capacity is nonincreasing.

We conclude by considering the analytical consequences of a congestion cost model that is not captured by Assumptions 4 and 5; for example, suppose $\ell(q, I)$ represents the M/G/1 mean steady-

[4]Equation (12) in the proof was also derived for a similar model studied by Xiao et al. (2007); they consider a pricing and investment game where Assumption 5 is also satisfied. However, that paper does not establish that such cost models lead to linear total cost functions.

[5]Campo-Rembado and Sundararajan (2004) use such a model to study competition among wireless service providers.

state delay function. WJV show that in this case $v(q_j)$ is concave in q_j.[6] As a consequence, if congestion is measured as the delay in an M/G/1 queueing system, and b_j is linear, the resulting total cost function t_j is *concave*. As also pointed out by Cachon and Harker (2002), even existence of pure strategy equilibrium may fail for such pricing games. This points to an important dichotomy in analytical simplicity between service provision models where consumers are *loss* sensitive (e.g., where congestion cost involves the Erlang formula), and those where consumers are *delay* sensitive (e.g., where congestion cost involves a queueing delay function); a similar observation is made by WJV.

The rest of the paper analyzes industries for which Assumption 3 and Assumptions 4 or 5 hold. In this case, the SLG game can be reduced to a pricing game with convex or linear costs.

5. Comparison of SLG and BE Games

In this section we compare competition in prices and SLGs with competition in prices and investment levels through the NE of the respective games played by firms. To characterize the SLG game we use the results of the preceding section. The BE model was studied by WJV, and in a special case by Xiao et al. (2007).

We begin by specializing our model: we assume that firms are ex-ante symmetric, and that there are no production costs.

Assumption 6. *For all i, j, $b_j = 0$ (hence $t_j = v_j$), and $\ell_i = \ell_j = \ell$.*

For technical simplicity, we also specialize our demand model to one where services are homogeneous and perfect substitutes, and demand is downward sloping. We describe this model in Section 5.1, and provide a technical specification of demand for both the SLG and BE games. In Section 5.2, we recapitulate the main results of WJV for the BE game under this demand model.

In Section 5.3, we compare the NE for the SLG and BE games in the case of constant returns to investment. In Section 5.4 we compare both models in industries where the total cost function is convex, but not necessarily linear.

[6] They assume that increasing investment expenditure reduces mean service time while holding the coefficient of variation of the service time fixed. Cachon and Harker (2002) show a similar result for the M/M/1 delay function.

5.1 Demand Model

We assume that services are homogeneous and perfect substitutes, a common simplifying assumption made in competitive congestion models similar to the ones described in Section 5.1.2 (see, e.g., WJV, Xiao et al. 2007, Acemoglu and Ozdaglar 2006). We further assume that consumers generate a downward sloping demand function $Q(f)$, where f is the full price in the market. Our goal in this section is to provide a formal specification of demand for both the SLG and BE games with the property that services are homogeneous and perfect substitutes. In Section 5.1.1, we specify our demand model in the SLG game. In Section 5.1.2, we specify our demand model in the BE game.

5.1.1 SLG Game

Because services are perfect substitutes, all demand will go to the firm charging the lowest full price. In case of ties, we assume that firms that set the lowest price split demand equally.

Assumption 7. *In the SLG game, for all j, the demand function is given by:*

$$d_j(f_j, \boldsymbol{f}_{-j}) = \begin{cases} Q(f_j) & \text{if for all } i \neq j, f_j < f_i \\ Q(f_j)/n & \text{if for all } i, f_j \leq f_i, \text{ and } n = |\{i : f_j = f_i\}| \\ 0 & \text{if for some } i, f_j > f_i \,, \end{cases} \tag{8}$$

where $Q'(f) < 0$ and $\lim_{f \to \infty} Q(f) = 0$.

Recall that by Proposition 1, to analyze the SLG game it is enough to consider the pricing game with profit functions $\pi_j(f_j, \boldsymbol{f}_{-j})$.

5.1.2 BE Game

To make a fair comparison between the SLG and BE games, for the BE game we consider a static equilibrium in which firms that attract customers offer the same full price. This assumption corresponds to the model of homogeneous and perfect substitute services. Such an equilibrium is commonly known as a *Wardrop equilibrium*, particularly in the transportation literature; we adopt the same terminology here, with the abbreviation WE (Wardrop 1952).

Let $P(q)$ denote the inverse demand function: i.e., $P(Q(f)) = f$ for all $f > 0$. We interpret $P(q)$ as the marginal utility obtained by an additional infinitesimal consumer when the total number of consumers being served is q.

Assumption 8. *In the BE game, for given price and investment vectors* $\boldsymbol{p}$ *and* $\boldsymbol{I}$*, the vector of demand quantities* $\boldsymbol{q} \geq 0$ *is given by a* Wardrop equilibrium*, that is,*

$$\begin{aligned} p_j + \ell(q_j, I_j) &= P(Q), \qquad \text{for all } j \text{ with } q_j > 0; \\ p_j + \ell(q_j, I_j) &\geq P(Q), \qquad \text{for all } j, \end{aligned}$$

where $Q = \sum_{i=1}^{N} q_i$*,* $P'(q) < 0$*, and* $\lim_{q \to \infty} P(q) = 0$*.*

Under Assumption 4, given price and investment vectors $\boldsymbol{p}, \boldsymbol{I}$, if $I_j > 0$ for at least one j, then a WE exists and is unique. The NE for the BE game is given by Definition 2, with $W(\boldsymbol{p}, \boldsymbol{I})$ determined by a WE.

5.2 Equilibrium Analysis of the BE Game: Recap

In this section we recapitulate the main results regarding the BE game, with the demand model given in Section 5.1.2. For the BE game, WJV show that for an important class of models satisfying Assumptions 4, 6, and 8, if a NE exists, then it is unique and symmetric.[7] Notably, any model that also satisfies Assumption 5, and thus exhibits constant returns to investment, falls in this class.

Moreover, for all firms j, the demand quantities and investment levels are given by $q_j^{BE} = Q^{BE}/N$, and $I_j^{BE} = I\left(Q^{BE}/N\right)$, while prices are given by

$$p_j^{BE} = q_j^{BE} \left(\partial \ell \left(q_j^{BE}, I_j^{BE} \right) / \partial q + \frac{1}{(N-1)/\left(\partial \ell \left(q_j^{BE}, I_j^{BE} \right) / \partial q \right) - 1/P'(Q^{BE})} \right). \tag{9}$$

WJV also show that a NE for the BE game can fail to exist, and provide conditions for which a NE is guaranteed to exist for models that satisfy Assumptions 4, 6, and 8. Motivated by this fact, we introduce the following assumption.

Assumption 9. *A NE* $(\boldsymbol{p}^{BE}, \boldsymbol{I}^{BE}, \boldsymbol{q}^{BE})$ *for the BE game exists. Moroever, it is unique and symmetric. For all* j*,* $q_j^{BE} = Q^{BE}/N$ *and* $I_j^{BE} = I\left(Q^{BE}/N\right)$*. Prices* p_j *are given by equation (9).*

[7] The class of models considered by WJV is characterized by Assumptions 4, 6, and 8, as well as an additional technical condition related to the *marginal rate of substitution* (MRS) of the congestion cost function. The MRS of I for x is defined as follows:

$$\mathrm{MRS}(I; x) = -\frac{\partial \ell(x, I)/\partial x}{\partial \ell(x, I)/\partial I}.$$

The following technical condition is required:

$$\frac{\partial}{\partial I} \mathrm{MRS}(x; I) \geq \frac{1}{x}.$$

The examples derived from loss systems discussed in Section 4.2 are included in this class; indeed, it is satisfied with equality for any model that exhibits constant returns to investment according to Assumption 5.

5.3 Comparison: Constant Returns to Investment

In this section we prove one of our main results: we give a sharp comparison between the NE for the SLG and BE game in the case of constant returns to investment. While the NE price for the SLG game is perfectly competitive and all firms obtain zero profits, firms obtain positive markups in the unique NE for the BE game.

Suppose Assumption 5 holds, that is, the industry exhibits constant returns to investment. In this case, the SLG game is equivalent to a standard symmetric Bertrand game with constant marginal costs (by Propositions 1 and 3). As a result, the unique symmetric NE price coincides with the perfectly competitive price (the marginal cost) and firms obtain zero profits (Tirole 1988).[8] Therefore, in the symmetric NE for the SLG game the total demand served is the unique solution Q^{SLG} to $P(Q^{SLG}) = t'(Q^{SLG}/N)$, and the NE full price is $t'(Q^{SLG}/N)$. We refer to this marginal cost as the "perfectly competitive full price" (by analogy with standard pricing theory).

Further, because price equals marginal cost, this solution also maximizes social surplus, i.e., the sum of consumer surplus and producer surplus, over all demand vectors $\boldsymbol{q}$. Additionally, recall that in the NE all firms invest efficiently, that is, for all j, $I_j^{SLG} = I\left(Q^{SLG}/N\right)$ (see discussion in Section 4.1).

We note that if all firms charge full price $t'(Q^{SLG}/N)$, then the NE price is equal to the *congestion externality*, i.e., the marginal congestion effect of one consumer on the others at the same service provider. This is expected as the congestion externality corresponds to the "socially optimal" price in this context (Pigou 1920). Indeed, using Lemma 2 in the Appendix, $t'(q) = \ell\left(q, I(q)\right) + q\partial\ell\left(q, I(q)\right)/\partial q$. Recall that the price charged by firms in a NE for the SLG game, p^{SLG}, is given by the difference between the full price and the congestion cost. We conclude that:

$$p^{SLG} = \frac{Q^{SLG}}{N}\frac{\partial \ell}{\partial q}\left(\frac{Q^{SLG}}{N}, I\left(\frac{Q^{SLG}}{N}\right)\right),$$

which corresponds to the congestion externality. We also refer to this quantity as the "perfectly competitive price."

In the NE for the BE game, firms invest efficiently, that is, for all j, $I_j^{BE} = I\left(Q^{BE}/N\right)$. We know that for constant returns to investment, $\ell(q, I(q))$ is constant for all $q \geq 0$ (cf. the proof of Proposition 3). Therefore, the congestion cost perceived by consumers is the same in both the SLG and BE game. In the BE game, however, by equation (9), the NE price exhibits a markup above

[8] Note that if t is linear and $N \geq 3$, there are asymmetric NE. However, in all of these equilibria, the unique selling price equals marginal cost and all firms obtain zero profits.

the congestion externality. Using this expression, WJV show that Q^{BE} is strictly lower than the social surplus maximizing quantity, which is equivalent to Q^{SLG} in the case of constant returns to investment. Note that this also establishes that the full price in the NE for the BE game is strictly higher than in the SLG game. We have proved the following key result.

Proposition 4. *Suppose Assumptions 5, 6, and 9 hold. Suppose that demand systems for the SLG game and BE games are given by Assumptions 7 and 8, respectively. Let* $\ell^{SLG}, p^{SLG}, f^{SLG}, Q^{SLG}, \pi^{SLG}$, *and* $\ell^{BE}, p^{BE}, f^{BE}, Q^{BE}, \pi^{BE}$ *be the consumers' congestion costs, prices, full prices, total demand quantities, and firms' profits in the unique symmetric NE for the SLG and BE game, respectively. Then,* $\ell^{SLG} = \ell^{BE}$, $p^{SLG} < p^{BE}$, $f^{SLG} < f^{BE}$, $Q^{SLG} > Q^{BE}$, *and* $\pi^{BE} > \pi^{SLG} = 0$. *Moreover,* p^{SLG} *and* f^{SLG} *correspond to the perfectly competitive price and full price, respectively. Further,* Q^{SLG} *corresponds to the socially optimal quantity (that maximizes the sum of consumer and producer surplus). Finally, firms invest efficiently in both games given the mass of consumers they serve.*

The preceding comparison highlights that the choice of strategic variables has a significant impact on the resulting Nash equilibria for congestion games. In industries with constant returns to investment, competition is stronger in a game where firms choose prices and SLGs than in a game where firms choose prices and investment levels. In the former, Bertrand price competition results and firms cannot exercise market power. In the SLG game firms face an infinitely elastic demand at the full price charged by the competitors. On the other hand, in the BE game, firms face a smooth demand function at the price and investment levels chose by the competitors. Hence, competition is less intense and firms are able to sustain markups in equilibrium. Our comparisons are similar in spirit to the well known difference between price competition and quantity competition in standard oligopoly games (see Vives 2001, Farahat and Perakis 2006, and references therein). We emphasize, however, that while the SLG game can be reduced to an equivalent pricing game, the BE game *cannot* be reduced to an equivalent quantity competition game.

5.4 Comparison: Convex Costs

In this section we compare the NE for the BE and SLG game when Assumptions 4 and 6 hold; we assume the total cost function t is convex. In this case, we cannot provide a sharp comparison as in the case of constant returns to investment; we show that the NE full price for the BE game can be above or below a NE full price for the SLG game.

To simplify the presentation, we assume there is a perfectly inelastic demand of size M. Therefore, the demand system for the SLG game in this section is given by the following assumption.

Assumption 10. *In the SLG game, suppose Assumption 7 holds with $Q(f) = M$, for all $f \geq 0$.*

The definition of WE is changed to accommodate a perfectly inelastic demand. The demand system in this section for the BE game is given by the following assumption.

Assumption 11. *In the BE game, for given price and investment vectors $\mathbf{p}$ and $\boldsymbol{I}$, the vector of demand quantities $\boldsymbol{q} \geq 0$ is given by a* Wardrop equilibrium*, that is,*

$$\begin{aligned} &p_j + \ell(q_j, I_j) = \min_k\{p_k + \ell(q_k, I_k)\}, \qquad \textit{for all } j \textit{ with } q_j > 0, \textit{ and} \\ &\textstyle\sum_{j=1}^N q_j = M. \end{aligned}$$

First, we characterize the symmetric NE for the SLG game played by firms under the previous assumptions. By Proposition 1 it is enough to consider the pricing game with profit functions $\pi_j(f_j, \boldsymbol{f}_{-j})$. Vives (2001) and Dastidar (1995) provide a characterization of NE for this game when demand is downward sloping. A similar argument can be applied in the case of perfectly inelastic demand to find symmetric equilibria. Since we are interested in symmetric NE, we will say *the full price f is a NE* if all firms setting full price f is a NE. Define:

$$\underline{f}^{SLG} = \frac{t\,(M/N)}{(M/N)};\ \ \overline{f}^{SLG} = \frac{(t(M) - t\,(M/N))}{(M(N-1)/N)}\,. \tag{10}$$

Suppose Assumptions 4, 6, and 10 hold. Then, one can show that f can be sustained as a NE for the SLG game if and only if $f \in [\underline{f}^{SLG}, \overline{f}^{SLG}]$. Further, $\underline{f}^{SLG} \leq \overline{f}^{SLG}$, so at least one symmetric NE exists. Moreover, the perfectly competitive full price $t'\,(M/N)$ is an element of $[\underline{f}^{SLG}, \overline{f}^{SLG}]$. If, in addition, t is strictly convex, there are no asymmetric NE, and $\overline{f}^{SLG} - \underline{f}^{SLG} > 0$. The proof of these results can be found in the Appendix.

Therefore, there exists a continuum of full prices (hence, of prices) that can be sustained as NE for the SLG game. In any of these equilibria, all firms invest efficiently, that is, they invest $I(M/N)$ (see discussion in Section 4.1).

Now we study the BE game. The analysis in Section 5.2 can be specialized to a setting with a perfectly inelastic demand. In this case, for all firms j, the demand quantities and investment levels in the unique and symmetric NE for the BE game are given by $q_j^{BE} = M/N$, and $I_j^{BE} = I\,(M/N)$, while prices are given by $p_j^{BE} = M/(N-1)\partial\ell\left(q_j^{BE}, I_j^{BE}\right)/\partial q$. Assumption 9 is modified accordingly for this section. Note that in the BE game, p^{BE} exceeds the congestion externality; thus $f^{BE} = \ell\,(M/N, I(M/N)) + p^{BE} > \underline{f}^{SLG}$. Therefore, in the SLG game, it is possible to sustain lower full prices in equilibrium; indeed, the perfectly competitive full price can be sustained as a NE. On the other hand, in general, f^{BE} can be smaller or larger than $\overline{f}^{SLG}$ depending on the problem instance. Therefore, in some cases, the NE full price for the BE game is larger than all full prices that can be sustained as symmetric NE for the SLG game. However, in other cases, in the pricing and SLG game higher full prices can be sustained in equilibrium. We provide a condition and an example below. First, we formalize our discussion in the following proposition.

Proposition 5. *Suppose Assumptions 4, 6, and 9 hold. Suppose that demand systems for the SLG game and BE games are given by Assumptions 10 and 11, respectively. In the SLG game, a continuum of symmetric NE can be sustained. In the pricing and investment game, there is a unique NE. In all of these NE, each firm attracts a mass M/N of customers and invests the efficient level, $I(M/N)$. Moreover, we have that $f^{BE} > \underline{f}^{SLG}$. However, f^{BE} can be smaller or larger than $\overline{f}^{SLG}$ depending on the problem instance.*

We study the difference between f^{BE} and $\overline{f}^{SLG}$ in the following lemma; the proof can be found in the Appendix.

Lemma 1. *Suppose Assumption 4 holds. The expression $f^{BE} - \overline{f}^{SLG}$ has the same sign as $Mt'(M/N) + I(M/N) - t(M)$.*

Under Assumption 4, the function t is convex with $t(0) = 0$ (since $t = v$, this follows by Proposition 2). Hence, $t'(x) \geq t(x)/x,\ \forall x \geq 0$. However, if t is sufficiently "steep," it will be the case that $t'(M/N) < t(M)/M - I(M/N)/M$, hence $f^{BE} - \overline{f}^{SLG} < 0$. For the first inequality to hold we need t to be "steep" enough (relative to N). We provide an example.

Example 1. *Suppose that $N \geq 2$ and that the congestion cost function is given by $\ell(q, I) = q^r/I$, where $r \geq 1$.*[9] *It is simple to check that the efficient investment level and the total cost function are given by $I(q) = q^{(1+r)/2}$ and $t(q) = v(q) = 2q^{(1+r)/2}$, respectively. Substituting, we obtain that $Mt'(M/N) + I(M/N) - t(M)$ is positive (hence, $f^{BE} - \overline{f}^{SLG} > 0$) if and only if $1 + N + rN - 2N^{(1+r)/2}$ is positive.*

For example, suppose $N = 4$. If $r = 1$, $f^{BE} - \overline{f}^{SLG} > 0$, hence, the NE full price for the BE game is larger than all full prices that can be sustained as NE for the SLG game. This is not surprising, since for $r = 1$ the congestion cost function exhibits constant returns to investment.

[9]WJV show that ℓ satisfies Assumption 4

Moreover, the difference $f^{BE} - \overline{f}^{SLG}$ is positive for $r \leq 1.413$. However, if r increases over this value and ℓ and t become more "steep," then $f^{BE} - \overline{f}^{SLG}$ becomes negative.

We conclude by noting that WJV show that NE for the BE game may fail to exist, even under Assumptions 4 and 6, and provide conditions that guarantee its existence. By contrast, the previous analysis states that Assumptions 4 and 6 (and inelasticity of demand) suffice to ensure NE for the SLG game always exist.

6. Firms' Choice of Contractual Agreement

Thus far, we have assumed that all firms share the same exogenously given strategy space; i.e., they decide prices and SLGs, or prices and investments. However, in many cases, one of the most important strategic decisions made by firms is *which* contractual agreement they will offer to customers: a SLG or BE service. Motivated by this fact, in this section we wish to study whether an industry would naturally result in competition in prices and investment levels, or in prices and SLGs. To investigate this question, we must allow the possibility that firms can *choose* the strategy space they will use in competition. We thus introduce a multistage game where at the first stage, each firm decides whether it will set a price and investment level, or a price and an SLG at the second stage. We refer to the former action in the first stage as "P-I," and the latter action as "P-SLG." In the second stage, each firm chooses its actions: either a price and investment level (if P-I was chosen in the first stage), or a price and SLG (if P-SLG was chosen in the first stage). As in most of Section 5, services are perfect substitutes and consumers generate a downward sloping demand function. We analyze the pure strategy subgame perfect equilibria of this multi-stage game.

Formally, let S_j denote the decision of firm j in stage 1: either $S_j =$P-I, or $S_j =$P-SLG. Let $A = \{j : S_j = \text{P-I}\}$, and $B = \{j : S_j = \text{P-SLG}\}$. For firms $j \in A$, let p_j and I_j denote the price and investment level chosen at the second stage, respectively; and for firms $j \in B$, let p_j and $\overline{\ell}_j$ denote the price and SLG chosen at the second stage, respectively.

Given p_j and $\overline{\ell}_j$ for firms $j \in B$, and given p_j and I_j for firms $j \in A$, demand allocation is as follows. Essentially, if any firm chose P-SLG at the first stage, and serves a positive mass of customers at the final stage, then the full price in the market must be the full price set by that firm. Formally, define $f^B = \min_{j \in B}[p_j + \overline{\ell}_j]$, and let f^A be the full price at a WE if only the firms in A compete for consumers. If $f^B \leq p_j + \ell(0, I_j)$ for all $j \in A$, or $A = \emptyset$, then all firms in B with full price equal to f^B equally split the demand $Q(f^B)$ (Assumption 7), and no firms in A serve any customers. On the other hand, if $f^B \geq f^A$, or $B = \emptyset$, then no firms in B serve any customers, and firms in A split the demand according to the WE conditions (Assumption 8). Finally, if $\min_{j \in A}[p_j + \ell(0, I_j)] < f^B < f^A$, then the demand allocation is the unique solution to the following equations:

$$p_j + \ell(q_j, I_j) = f^B, \text{ if } j \in A, \; q_j > 0;$$
$$p_j + \ell(q_j, I_j) \geq f^B, \text{ if } j \in A, \; q_j = 0;$$
$$q_j = (Q(f^B) - Q^A)/n, \text{ if } j \in B, \; p_j + \bar{\ell}_j = f^B;$$
$$q_j = 0, \text{ if } j \in B, \; p_j + \bar{\ell}_j > f^B,$$

where $Q^A = \sum_{j \in A} q_j$ and $n = |\{j \in B : p_j + \bar{\ell}_j = f^B\}|$. In other words, the full price in the market is fixed at f^B, and demand allocates accordingly: firms in A receive customers until their full price reaches f^B, and all remaining demand is split equally among firms in B with full price equal to f^B. It is straightforward to verify that these equations have a unique solution as long as $\ell(q, I)$ is strictly increasing in q for any $I > 0$.

For simplicity, we restrict attention to a setting where ℓ exhibits constant returns; i.e., we assume that Assumption 5 holds, and that $\beta_j = 0$ for all j. Our key insight is the following.

Proposition 6. *Suppose Assumptions 5 and 6 hold, and that demand is allocated as previously described. Consider any subgame at the second stage where at least one firm has chosen P-SLG in the first stage. Then there exists a NE of the resulting second stage game; and further, in any such NE, all firms obtain zero profits.*

Proof. Since Assumption 5 holds and $\beta_j = 0$, we know from Proposition 3 that $t(q) = v(q) = cq$ for some constant $c > 0$. We first prove that all firms make zero profits in the second stage equilibrium; we then establish existence of an equilibrium.

Given a NE of the second stage game, define A, B, f^A, f^B, Q^A, and n as in the demand allocation discussion above. Let $f = \min\{f^A, f^B\}$; then the demand allocation implies that for any firm j with $q_j > 0$, we have:

$$p_j + \ell(q_j, I_j) = f.$$

The key observation is that regardless of whether such a firm has chosen P-SLG or P-I, in equilibrium, it will invest efficiently given its demand allocation; i.e., $I_j = I(q_j)$. (WJV establish this insight for the BE game.) Thus, for equilibrium profits we have:

$$p_j q_j - I_j = f q_j - v(q_j) = (f - c) q_j.$$

We conclude that in equilibrium, all firms j make profit $(f - c)q_j$. Thus to establish the result, it suffices to show that if at least one firm has chosen P-SLG in the first stage, then $f = c$ in the second stage NE. We assume to the contrary that $f > c$, and establish a contradiction.

First suppose that all firms have chosen P-SLG in the first stage. In this case, the resulting game at the second stage is a standard Bertrand pricing game with constant marginal costs, and thus we must have $f = c$. Thus if $f > c$ in equilibrium, at least one firm must have chosen P-I in the first stage, so both A and B are nonempty.

We distinguish two cases. First suppose that all firms $i \in B$ earn zero profits; i.e., any firm that has chosen P-SLG in the first stage has $q_j = 0$. This implies all customers are served by firms in A. We consider a deviation where a firm $i \in B$ acquires ϵ of the demand to serve (where ϵ is small). Let f^ϵ be the unique full price that results at a Wardrop equilibrium if only the firms in A compete to serve a market with demand function $\hat{Q}(\hat{f}) = Q(\hat{f}) - \epsilon$; in other words, f^ϵ is the unique solution to:

$$\begin{aligned} p_j + \ell(q_j, I_j) &= f^\epsilon, \qquad \text{if } q_j > 0,\ j \in A; \\ p_j + \ell(q_j, I_j) &\geq f^\epsilon, \qquad \text{if } q_j = 0,\ j \in A; \\ \sum_{j \in A} q_j &= Q(f^\epsilon) - \epsilon. \end{aligned}$$

It is straightforward to check that as $\epsilon \to 0$, we have $f^\epsilon \to f$. Therefore for sufficiently small ϵ we have $f^\epsilon > c$ (by continuity). We conclude that for any firm $i \in B$, there exists a deviation where firm i sets SLG $\overline{\ell}'_i = \ell(\epsilon, I(\epsilon))$, and price $p'_i = f^\epsilon - \overline{\ell}'_i$. For sufficiently small $\epsilon > 0$, firm i will serve ϵ consumers at the resulting demand allocation; further, firm i earns positive profits, since $p'_i\epsilon - I(\epsilon) = (f^\epsilon - c)\epsilon > 0$. This is a profitable deviation for firm i, so the original configuration could not have been a NE of the second stage game.

The second possibility is that at least one firm $i \in B$ earns positive profits, i.e., there exists $i \in B$ with $q_i > 0$; let $Q^B = \sum_{i \in B} q_i$, so $Q^B > 0$. Let $j \in A$ be any firm in A. We consider the following deviation for firm j. Let $q'_j = q_j + Q^B$, $I'_j = I(q_j + Q^B)$, and $p'_j = f - \ell(q'_j, I(q'_j))$. Note that $q'_j\ell(q'_j, I(q'_j)) \leq v(q'_j) = cq'_j < fq'_j$, so $p'_j > 0$. Under this deviation, firm j captures all the customers currently being served by firms in B, and invests (efficiently) to keep the resulting equilibrium full price unchanged. Thus in the resulting demand allocation, firm j attracts exactly q'_j customers, and earns profit $(f - c)q'_j$. Since $Q^B > 0$, we have $q'_j > q_j$, and thus we have found a profitable deviation for firm j, contradicting the assumption of NE.

Thus we cannot have $f > c$ in the second stage NE. Thus $f = c$, and all firms make zero profits in any second stage equilibrium, as required.

It remains to establish that a second stage NE exists if at least one firm has chosen P-SLG. Again let A be the set of firms that chose P-I at the first stage, and let B be the set of firms that chose P-SLG at the first stage. If A is empty, then all firms play a Bertrand pricing game at the second stage with constant marginal costs, for which all firms setting a full price $f = c$ is a NE.

Suppose instead that $A \neq \emptyset$, and consider the following strategy profile. Let $K = |A|$, and suppose that all firms $j \in A$ choose $I_j = I(Q^A/K)$ and $p_j = c - \ell(Q^A/K, I(Q^A/K))$ where $Q^A = Q(c)$. (As above, it is straightforward to check that $p_j \geq 0$.) Suppose further that all firms $j \in B$ choose $\overline{\ell}_j$ and p_j so that $p_j + \overline{\ell}_j = c$. In the resulting demand allocation, observe that all demand is equally divided among the firms in A, and firms in B serve no customers; in particular, all firms in B have no investment expenditure. Further, the full price at the resulting WE is exactly c, and since all firms have invested efficiently, it is straightforward to check that all firms obtain zero profits.

We claim the proposed strategy profile is a NE. Since all firms in B have set price plus SLG equal to c, even if any firm in A changes its price and/or investment decision, the full price in the resulting demand allocation cannot be any greater than c. Note that because all firms with $q_j > 0$ invest efficiently, all firms j make profits $(f - c)q_j$ in a resulting demand allocation with full price f. Thus a firm $j \in A$ cannot earn positive profits by deviating. On the other hand, consider a firm $j \in B$. Raising the full price (i.e., the sum $p_j + \overline{\ell}_j$) has no effect on this firm's profit, since it is not serving any demand; and while lowering the full price below c attracts demand, the resulting profit earned will be negative. Thus no firm $j \in B$ has a profitable deviation either, establishing that the proposed profile is a NE, as required. □

Proposition 6 leads to the strong conclusion that if even one firm chooses to offer SLGs to customers, then all firms "suffer" as a result: intense price competition immediately results and all firms make zero profits. Thus the only subgame perfect equilibrium where all firms make positive profits is the one where all firms choose P-I at the first stage, i.e., where firms compete only using prices and investment levels (see Proposition 4).

In particular, our result suggests that if all other firms have chosen to compete by offering prices and investment levels, a deviation to offer SLGs is not profitable. We state the following corollary.

Corollary 1. *Suppose Assumptions 5, 6 and 9 hold, and that demand is allocated as previously described. Then, the unique pure strategy subgame perfect equilibrium of the two stage game where all firms make positive profits is obtained when all firms choose P-I in the first stage, and play the unique and symmetric NE of the pricing and investment game in the second stage. Conversely,* *in any subgame perfect equilibrium where one firm has chosen P-SLG in the first stage, all firms earn zero profits. As a consequence, assuming firms play according to NE in the second stage, choosing P-I in the first stage is a (weakly) dominant action.*

We believe the correct way to interpret the result is that in the absence of customer differentiation, setting SLGs leads to marginal cost pricing. This viewpoint implies that for constant returns to investment, SLGs are likely to be useful only in scenarios where market segmentation is possible; in such a setting, a firm might offer a high quality SLG, and charge a high price for the service. Only consumers who are willing to pay a large amount for such quality would subscribe to this firm. In our model, by contrast, all consumers trade off money and congestion identically; thus such market segmentation is not possible.

7. Conclusions

This paper compares market outcomes in service industries with two different contractual models: best effort service (BE), and service level guarantees (SLGs). The comparison is made through the Nash equilibria of the games played by firms. On the path to this comparison, we first characterized equilibrium behavior in a game where firms simultaneously set prices and SLGs (the SLG game). Our key insight is that such a game is equivalent to a standard pricing game. Moreover, for an important class of congestion models, the pricing game yields convex or linear costs.

We used these results to compare the SLG game with a game where firms choose prices and investment levels (the BE game), yielding insight into policy and business considerations. The analysis crucially depends on the cost structure. In the case of constant returns to investment the comparison is sharp; competition is more intense in the SLG game. In this case, markups are higher and social welfare is lower in the BE game than in the SLG game. For convex costs, the analysis is more subtle; depending on the shape of the congestion cost function, the NE full price for the BE game can be above or below a NE full price for the SLG game. The results highlight how the choice of strategic variables can dramatically affect equilibrium behavior in the provision of congestible services.

Finally, in a service industry where competition takes place by simultaneously choosing prices and investment levels, we ask whether a firm can obtain a competitive advantage by offering SLGs instead. Our results suggest that, in the absence of customer heterogeneity and with constant returns to investment, deviating to offer SLGs is not profitable. In future work, we plan to study the impact of SLGs on markets with several customer segments with different sensitivities for congestion cost. In this case, SLGs may help to segment the market and increase profits. As with the remainder of our work, the cost structure will of course have a first order impact on the firms' decision of strategy space as well.

Appendix: Proofs

Proposition 2. *Suppose that Assumptions 3 and 4 hold. For all j, the function t_j is strictly increasing, convex, and continuous, with $t_j(0) = 0$. Additionally, $t_j(q)$ is twice differentiable for $q > 0$.*

Proof. The proposition is a direct consequence of the following lemma that is proved by WJV.

Lemma 2. *Suppose that Assumption 4 holds. Then, for all j:*

1. *The function v_j is strictly increasing, convex, and continuous, with $v_j(0) = 0$;*
2. *$v_j(q)$ is twice differentiable for $q > 0$; and*
3. *$v_j'(q) = \ell_j(q, I_j(q)) + q\partial\ell_j(q, I_j(q))/\partial q$.*

□

Proposition 3. *Suppose Assumption 5 holds. Then, for all j, there exists $c_j > 0$ such that $t_j(q) = c_j q$.*

Proof. Fix $q_j > 0$. By Assumption 4.1, the minimization problem (2) admits an optimal solution. For $q_j = 0$, the optimal solution is $I_j = 0$. Moreover, for fixed $q_j > 0$, $\ell_j(q_j, I_j)$ is coercive in I_j, so any optimal solution to (2) must be interior. Let I_j be an optimal solution. Then I_j must satisfy the following first order condition:

$$-\frac{q_j^2}{I_j^2} h_j'\left(\frac{q_j}{I_j}\right) + 1 = 0 \,. \tag{11}$$

Consider the equation:

$$y^2 h_j'(y) = 1. \tag{12}$$

By Assumption 4, $qh_j(q/I)$ is strictly convex in I, hence, $y^2 h_j'(y)$ is strictly increasing. Therefore, equation (12) has a unique solution, say y_j^*. We conclude from (11) that $q_j/I_j = y_j^*$. This implies that for fixed q_j, the problem (2) has a unique optimal solution, given by $I_j = q_j/y_j^*$; i.e., the efficient investment level is *linear* in q_j. Substituting, we conclude that $v_j(q_j) = q_j h_j(y_j^*) + q_j/y_j^*$, which is linear in q_j. The result follows because by Assumption 5, we also have $b_j(q_j) = \beta_j q_j$. $\square$

Characterizing NE of the SLG Game with Convex Costs (Section 5.4). *Suppose Assumptions 4, 6, and 10 hold. Then, f can be sustained as a NE for the SLG game if and only if $f \in [\underline{f}^{SLG}, \overline{f}^{SLG}]$. Further, $\underline{f}^{SLG} \leq \overline{f}^{SLG}$, so at least one symmetric NE exists. If, in addition, t is strictly convex, there are no asymmetric NE, and $\overline{f}^{SLG} - \underline{f}^{SLG} > 0$.*

Proof. Through an abuse of notation, let $\pi(f;n)$ be the profit received by a firm when n firms set full price f and all other firms set a full price higher than f. Hence, $\pi(f;n) = fM/n - t\,(M/n)$.

We first note that f is a Nash equilibrium if and only if (1) $\pi(f;1) \leq \pi(f;N)$ and (2) $\pi(f;N) \geq 0$. If a firm undercuts such a full price by $\epsilon > 0$, the resulting profits are $\pi(f-\epsilon;1) < \pi(f;N)$. On the other hand, if a firm raises its full price, then it receives zero profit, so such a deviation is not profitable (since condition (2) guarantees that the full price f is individually rational).

Assume that $N > 1$, and define $\overline{f}^{SLG}$ as the unique solution to $\pi(\overline{f}^{SLG};1) = \pi(\overline{f}^{SLG};N)$; note that the solution is unique since $\pi(f;n)$ is linear in f for fixed n. Similarly, define $\underline{f}^{SLG}$ as the unique solution to $\pi(\underline{f}^{SLG};N) = 0$. Since $\pi(f;1)$ - $\pi(f;N)$ is strictly increasing in f, it follows that if $f > \overline{f}^{SLG}$ then $\pi(f;1) > \pi(f;N)$, and if $f < \overline{f}^{SLG}$, then $\pi(f;1) < \pi(f;N)$. Since $\pi(f;n)$ is strictly increasing in f for fixed n, if $f < \underline{f}^{SLG}$, then $\pi(f;N) < 0$, and, if $f > \underline{f}^{SLG}$, then $\pi(f;N) > 0$. We conclude that a full price f is a Nash equilibrium if and only if it lies in the interval $[\underline{f}^{SLG}, \overline{f}^{SLG}]$. (Note that this interval may be empty, e.g., if $\overline{f}^{SLG}$ is negative.)

We now show that if Assumption 4 holds, then $0 \leq \underline{f}^{SLG} \leq \overline{f}^{SLG}$. By the definitions:

$$\underline{f}^{SLG} = \frac{t\,(M/N)}{(M/N)}; \;\; \overline{f}^{SLG} = \frac{(t(M) - t\,(M/N))}{(M(N-1)/N)}, \tag{13}$$

Note that both $\underline{f}^{SLG}$ and $\overline{f}^{SLG}$ are nonnegative, the latter because t is strictly increasing (Lemma 2). Algebraic simplification yields:

$$\overline{f}^{SLG} - \underline{f}^{SLG} \;=\; \frac{t(M) - Nt\,(M/N)}{M(N-1)/N} \geq 0\,, \tag{14}$$

where the inequality follows because t is convex with $t(0) = 0$ (Lemma 2). This establishes that when Assumption 4 holds, then all symmetric pure strategy Nash equilibria of the simultaneous pricing and SLG game are given by full prices in the nonempty interval $[\underline{f}^{SLG}, \overline{f}^{SLG}]$. Moreover, if t is strictly convex, $\overline{f}^{SLG} - \underline{f}^{SLG} > 0$.

Now, we show that the competitive full price will be sustainable as a Nash equilibrium, that is, $t'(M/N) \in [\underline{f}^{SLG}, \overline{f}^{SLG}]$. From (10) we have $t'(M/N) \geq \underline{f}^{SLG}$, since t is convex and $t(0) = 0$. For the upper bound, note that by convexity:

$$t(M) \geq t\left(\frac{M}{N}\right) + t'\left(\frac{M}{N}\right)\left(M - \frac{M}{N}\right).$$

Rearranging terms in the preceding expression, and referring to (10), we find $t'(M/N) \leq \overline{f}^{SLG}$. Hence, in SLG game, there exists a NE where all firms charge the perfectly competitive full price $t'(M/N)$.

Furthermore, if t is strictly convex, we can establish that no asymmetric equilibria exist. To see this, note that if $\pi(f;n) \geq 0$, then $f \geq t(M/n)/(M/n)$. Thus if $p > n$:

$$\pi(f;p) \geq \left(\frac{n}{p}\right) t\left(\frac{M}{n}\right) - t\left(\frac{M}{p}\right).$$

Since t is strictly convex with $t(0) = 0$, it follows the right hand side of the preceding inequality is positive; thus $\pi(f;p) > 0$. We conclude that if t is strictly convex, there cannot be any asymmetric equilibria; at least one firm that is currently not attracting any demand would strictly prefer to set full price f and capture a share of the market. □

Lemma 1. *Suppose Assumption 4 holds. The expression* $f^{BE} - \overline{f}^{SLG}$ *has the same sign as* $Mt'(M/N) + I(M/N) - t(M)$.

Proof. Using the expression for f^{BE}, $\overline{f}^{SLG}$, and the definition of t, and simplifying, we obtain,

$$f^{BE} - \overline{f}^{SLG} = \frac{1}{M(N-1)/N}\big(M\ell(M/N, I(M/N)) + M^2/N\partial\ell(M/N, I(M/N))/\partial q + I(M/N) - t(M)\big).$$

By Lemma 2, $t'(q) = \ell(q, I(q)) + q\partial\ell_j(q, I(q))/\partial q$. Hence,

$$f^{BE} - \overline{f}^{SLG} = \frac{1}{M(N-1)/N}\left(Mt'(M/N) + I(M/N) - t(M)\right).$$

The result follows. □

Acknowledgments

We are grateful for helpful conversations with Gad Allon, Awi Federgruen, Sunil Kumar, Nolan Miller, Marcelo Olivares, Benjamin Van Roy, and seminar participants at the INFORMS Revenue Management and Pricing Conference 2007 and the INFORMS Annual Meeting 2007. We thank the NET Institute www.NETinst.org for financial support, as well as the National Science Foundation under grants DMI-0620811 and CNS-0644114.

References

Acemoglu, D., A. Ozdaglar. 2006. Competition and efficiency in congested markets. To appear, Mathematics of Operations Research.

Allon, G., A. Federgruen. 2005a. Competition in service industries. To appear, Operations Research.

Allon, G., A. Federgruen. 2005b. Competition in service industries with segmented markets. Working Paper, Kellogg.

Allon, G., A. Federgruen. 2006. Service competition with general queueing faciltites. Working Paper, Kellogg.

Cachon, G. P., P. T. Harker. 2002. Competition and outsourcing with scale economies. *Management Science* **48**(10) 1314–1333.

Campo-Rembado, M. A., A. Sundararajan. 2004. Competition in wireless telecommunications. Working Paper, New York University.

Cheng, L. 1985. Comparing bertrand and cournot equilibria: a geometric approach. *RAND Journal of Economics* **16**(1) 146–152.

Dastidar, K. G. 1995. On the existence of pure strategy Bertrand equilibrium. *Economic Theory* **5** 19–32.

Farahat, A., G. Perakis. 2006. A comparison of Bertrand and Cournot profits in oligopolies with gross substitutes. Working paper, MIT.

Hall, J., E. Porteus. 2000. Customer service competition in capacitated systems. *Manufacturing & Service Operations Management* **2**(2) 144–165.

Miller, N., A. Pazgal. 2006. Budget or target: The choice between input and output strategies. *RAND Journal of Economics* **37**(2) 391–415.

Ozdaglar, A. 2006. Price competition for elastic traffic. Working Paper, MIT.

Pigou, A. C. 1920. *The Economics of Welfare*. 1st ed. Macmillan.

Rosen, J. B. 1965. Existence and uniqueness of equilibrium points for concave n-person games. *Econometrica* **33**(3) 520–534.

Roughgarden, T. 2005. *Selfish Routing and the Price of Anarchy*. MIT Press, Cambridge, Massachusetts.

Scotchmer, S. 1985. Profit-maximizing clubs. *Journal of Public Economics* **27** 25–45.

Singh, N., X. Vives. 1984. Price and quantity competition in a differentiated duopoly. *RAND Journal of Economics* **15**(4) 546–554.

So, K. C. 2000. Price and time competition for service delivery. *Manufacturing & Service Operations Management* **2**(4) 392–409.

The Economist. 2006. The end of the line. Oct. 12.

Tirole, J. 1988. *The Theory of Industrial Organization*. 1st ed. MIT Press.

Vives, X. 2001. *Oligopoly Pricing*. 1st ed. MIT Press.

Wardrop, J. G. 1952. Some theoretical aspects of road traffic research. *Proc. of the Institute of Civil Engineers, Pt. II* **1** 325–378.

Weintraub, G. Y., R. Johari, B. Van Roy. 2007. Investment and market structure in industries with congestion. Working Paper, Stanford University.

Xiao, F., H. Yang, D. Han. 2007. Competition and efficiency of private toll roads. *Transportation Research Part B* **41** 292–308.