Service Competition, Outsourcing and Co-Production in a Queuing Game*

Gérard P. Cachon
The Fuqua School of Business
Duke University
Durham, NC 27708
gpc@mail.duke.edu
www.duke.edu/~gpc

Patrick T. Harker
The Wharton School of Business
University of Pennsylvania
Philadelphia, PA 19104
harker@opim.wharton.upenn.edu

March 1999

Abstract

This paper studies competition between two firms that service time sensitive customers. Customers choose firms based on the firms' prices, the firms' expected waiting and service times, and the firms' brands. The firms may choose diverse strategies: one could choose a high price, but serve customers quickly, whereas the other could choose a low price with slow service. We consider three "service supply chain" design configurations: the firms perform the service, the firms outsource the service to a contractor, or the firms outsource the service to their customers; i.e., customers become co-producers of the service. With the first, there exists only one Nash equilibrium in prices and capacities (processing rates) in which both firms have positive market share. If the firms' costs are identical, the firms adopt identical strategies. If one firm has a lower cost, then that firm serves customers more quickly and has a higher market share, even though it might have a higher price. The outsourcing configurations might raise the firms' costs: outsourcing to a contractor requires coordination and monitoring; customers agree to co-production only if they receive a price break, but customers might be less efficient than the firms. Nevertheless, even if either outsouring configuration raises the firms' costs, we show that the firms may be strictly better off because outsourcing raises equilibrium prices. This benefit is particularly valuable in highly competitive markets.

*Thanks is extended to the seminar participants at the University of North Carolina's Operations Research Department and to Marty Lariviere for their many helpful comments.

A distinctive feature of many services is the close involvement of customers in the service delivery process. This involvement might be entirely passive: customers do nothing more than wait for their service to begin and then wait for their service to be completed. Or, customers might actively engage in the production of their service; i.e., they might be co-producers along with the firm. In either case, the inevitable temporary imbalances between supply and demand create congestion; some lucky customers complete their service quickly, whereas other customers experience the frustration of a long service process.

This paper studies competition between two firms that service time-sensitive customers. Each customer decides which firm to patronize (if any) based on the firms' prices, the expected amount of time the customer will spend at either firm (which depends on the firms' capacities), and the customer's *a priori* preference for the firms' brands (some customers have an *a priori* preference for the first firm, while others prefer the second firm).

We begin with the assumption that customers are passive participants in the service process. The firms simultaneously choose prices and capacities. Hence, the firms have the flexibility to choose very different strategies; one firm can charge a high price and serve customers quickly, whereas the other firm might choose a low price but serve customers more slowly. Each firm attempts to maximize its profit per unit time, knowing that the other firm has the same objective. We find that there exists a unique Nash equilibrium in which both firms are willing to serve customers. When the firms have identical costs, they choose identical strategies; firms do not attempt to mitigate competition by adopting diverse strategies. When the firms' costs are different, the low cost firm will have a higher market share and serve customers more quickly. Further, the low cost firm might even have a higher price! As Stalk and Hout (1990) conjecture, a "time-based competitor" may indeed enjoy a price premium in the market.

We next turn to the firms' "service supply chain" design decision. Instead of doing the service themselves, each firm might outsource the service production to another firm, which we call a contractor. Outsourcing to a contractor creates some additional costs; in particular, the firm will need to monitor the contractor to ensure that it is providing a desirable service quality. Further, a contractor may require a price premium so that it can earn a positive profit. Therefore, intuition suggests that outsourcing is a viable option only if the contractor is more efficient than the firm.

Outsourcing to a contractor is not the only option. Alternatively, a service firm could also outsource to its customers; an option which is generally not available to manufacturing firms. Indeed, all service firms, whether they realize it or not, must decide the degree to which their customers will engage in co-production: a retailer can either assemble a product for a customer, or let the customer perform the assembly; a bank teller can provide a customer with account information, or the bank can let customers access the information on their own; a grocery retailer lets customers transport items from the retailer's shelves to the customers' homes, or home delivery is provided. In an influential article on re-engineering, Hammer (1990) argues that firms should exploit co-production as much as possible because it accelerates response times and reduces overhead. But there are risks associated with a co-production design. To increase its customer's co-production, a firm will presumably need to decrease its price since it is shifting some work onto its customers. Lower prices can lead to lower revenues. Further, a firm that relies on co-production might lose customers to a competitor that is ready to pamper customers with full service.

Despite its risks, we find that outsourcing can be quite beneficial to firms in a competitive market. In particular, we find that outsourcing (either to contractors or to customers) may make the firms better off even if it increases their costs. The reason is rooted in the nature of the firms' production technology. When a firm outsources its service delivery, it switches from a production technology that exhibits increasing returns to scale to a production technology that has constant returns to scale. Cost per customer declines with the number of customers served in a queuing process, so it is an increasing returns to scale technology. With a typical outsourcing contract a firm will pay a constant fee per customer served, i.e., a constant returns to scale technology. If outsourcing is done via co-production the firm will need to give a constant price discount to each customer, so again, a firm's cost is linear in the number of customers it serves. The nature of the production technology matters because it influences the firms' incentive to cut prices. The marginal incentive to cut price is greater for a firm with an increasing returns to scale technology than for a firm with a constant returns to scale technology: the former gains a cost reduction per customer due to increased demand, whereas price changes have no impact on the latter's cost per customer. Hence, through outsourcing the firms become less price competitive and equilibrium prices are higher than they would be without outsourcing. A cost increase due to outsourcing (e.g., coordination costs or inefficient consumers) may be a secondary concern relative to the benefit of strictly higher equilibrium prices, especially in price competitive markets.

The next section summarizes the literature which is related to this work. §2 details the queuing game. §3 describes the deterministic version of the queuing game and summarizes its key results. §4 analyzes the queuing game and §5 discusses outsourcing. The final section discusses the managerial implications of these results.

1. Literature review

There are many papers that investigate competition when customers are sensitive to time: Armory and Haviv (1998), Davidson (1988), De Vany (1976), De Vany and Saving (1983), Gilbert and Weng (1997), Kalai, Kamien and Rubinovitch (1992), Lederer and Li (1997), Li (1992), Li and Lee (1994), and Loch (1994). In most of these models, either the firms' prices are exogenous or their processing rates are exogenous. Hence, the firms compete either with prices or with processing rates, but not both.[1] Allowing for both decisions creates significant analytical complications; in particular, the players' profit functions are not well behaved (unimodal). A second distinction between this work and the previous models is that they assume the firms are *a priori* homogenous; i.e., if the firms charge the same prices and have the same expected time costs, then all customers are indifferent between which firm they visit.[2] Finally, in many of those models customers wait in a single queue.[3] In many service settings, customers must choose which firm to visit and each firm maintains

---

[1] Li and Lee (1992) analyze a model with fixed processing rates and then discuss how the model could be expanded to allow the firms to choose prices as well. However, they emphasize that the lack of pure strategy equilibria in that game imposes a significant challenge to the analysis of the expanded game. In Lederer and Li (1997), the firms have fixed overall production capacity, but they decide how to allocate that capacity across multiple customer classes. In the single class version of their model, the firms only compete on price.

[2] Li and Lee (1992) do consider the possibility that the firms's services have different quality levels. That is a different form of differentiation than what we consider: in their model all customers agree as to which firm is *a priori* better (the one with the higher quality).

[3] Gilbert and Weng (1997) do consider a model with separate queues, however the arrival process to each queue is set so that each firm has the same expected waiting time.

a separate queue; i.e., customers are not able to jockey between the competing firms while they are waiting for service. Further, with a single queue framework total market demand is constant (i.e., all customers join the queue and are eventually served), but in our queuing game customers may choose to forgo service, so total market demand is not constant.

Deneckere and Peck (1995) do consider a model in which firms simultaneously choose prices and processing rates, and customers choose firms based on expected utility maximization. However, the firms in their model are *a priori* homogenous.

Hotelling (1929) introduced the linear city game, which is the standard game for spatially differentiated competition. (This form of differentiation is also referred to as horizontal differentiation or brand differentiation. See Anderson, de Palma and Thisse, 1992, for other models of differentiation.) In the linear city game, there are two firms located at either end of a unit interval.[4] Customers are uniformly distributed over the interval (the linear city) and incur a cost to visit either firm that is proportional to the distance from the customer's location to the firm. That cost could be interpreted literally as a transportation cost, or it could be interpreted as a cost for purchasing an item that does not meet a customer's ideal specifications along some abstract dimension (e.g., brand preference, ambiance, etc.) Costs and customer arrivals in the linear city game are deterministic. The queuing game that we study is a stochastic version of the linear city game.

Several recent papers consider service competition when customers are unable to fully anticipate the service they will receive, and therefore implement a heuristic to choose between providers: Gans (1999) and Hall and Porteus (1998).[5] In our queuing game customers correctly anticipate the service they will receive.

Mendelson and Whang (1990) consider a game between a service provider and its customers. Customers have heterogenous service priorities, but the service provider cannot distinguish between customers with different priorities; i.e., holding price constant, all cus-

---

[4] There is also an extensive literature that investigates the firm's location decision. We do not consider that issue.

[5] In Gans (1999) customers do not know the firms' service qualities for sure, so they must search for a good provider over multiple service occasions. In Hall and Porteus (1998) customers are assumed to behave with a simple, yet realistic, heuristic: a fixed fraction of the customers that encounter a service failure switch service providers.

tomers would request the highest priority. They establish a pricing policy that induces customers to truthfully announce their priority. Their pricing scheme also maximizes total system value. In our model customers are indistinguishable as well, but once a customer arrives at a firm, the customer has the same priority valuation as the other customers. Hence, firms operate with a single price and a first-come-first serve queue discipline.

Chase (1978) and Karmarkar and Pitbladdo (1995) recognized that the degree to which customer engage in co-production is an important design decision for a service firm. Ha (1998) considers the interaction between pricing and co-production. In his model the amount of effort the service facility must expend is inversely related to the amount of effort its customers exert, but the service facility cannot dictate its customers' effort level. Instead, it can only influence their efforts via its price schedule. We assume a firm does not influence its customer service effort, but that the firm does influence whether the customer will be willing to perform the service.

Several papers consider pricing and capacity decisions for a single server: Dewan and Mendelson (1990), Stidham (1992), Stidham and Rump (1998), and So and Song (1998). (The first three papers seek to maximize system value, while the last maximizes a firm's profit.) In fact, the queuing game in this paper is a competitive extension of Stidham's model; when there is a single firm in the queueing game, that firm faces the same problem that a monopolist would face in Stidham (1992). Stidham (1992) and Stidham and Rump (1998) also provide an extensive discussion on the stability of the firm's pricing and capacity decisions. Given the formulation of our queuing game, equilibrium stability is not an issue.

Many papers investigate queue joining behavior in which customers compete for fast service, but the service provider is not a game participant: Bell and Stidham (1985), Kulkarni (1983), Lippman and Stidham (1977), Mendelson (1985) and Naor (1969).

There is an extensive literature on outsourcing/vertical integration. The primary theory is that integration should be observed when market mediated transactions are too costly (see Grossman and Hart, 1986; Klein, Crawford and Alchian, 1978; and Williamson, 1979.) If market based transactions are not problematic (i.e., contracts are specific, enforceable and easy to negotiate), the literature suggests outsourcing if outsourcing lowers production costs. The literature then proceeds to determine when one could expect costs to decline with outsourcing (see McMillan,1990; Venkatesan, 1992; van Miegham, 1997). Those papers do not consider the impact of outsourcing on equilibrium prices.

The marketing literature considers the issue of whether competing suppliers prefer to outsource the retailing function to independent retailers or to perform their own retailing (see McGuire and Staelin, 1983; and Moorthy, 1988). McGuire and Staelin (1983) demonstrate that outsourcing the retailing function can be better for the suppliers when retail price competition is high; outsourcing introduces double marginalization between the supplier and the retailer, which in turn mitigates price competition between the two suppliers. Outsourcing mitigates price competition in the queuing game as well, but for different reasons. In particular, outsourcing reduces the firms' incentive to build market share so as to improve their production efficiency, whereas in McGuire and Staelin (1983), market size has no qualitative impact. (In a queuing system, cost per customer declines in market share; i.e., there exists economies of scale.) Further, in our queuing game, firms outsource the production function (not the retailing function) and outsourcing in the queuing game is beneficial for all levels of price competition (assuming the firms actually compete for customers).

2. The Queuing Game Model

Two firms provide a single service; call them firm i and firm j. Customers evaluate these services along several dimensions. Let p_i be the price of service at firm i. Customers are impatient, so let $h > 0$ be the cost a customer incurs per unit of time spent at a firm (waiting or in service). Let g_i be a customer's expected time cost for visiting firm i; i.e., g_i/h is the expected amount of time a customer spends at firm i, including time in queue and time in service. Let f_i be the full price for a customer to visit firm i, $f_i = p_i + g_i$. (Firm i's full price does not include transportation costs because the customers visiting firm i travel different distances, but they all pay the same price and time cost.)

The firms' services differ along some quantifiable attribute that maps into a unit interval. For each customer, there is a most preferred value for this attribute, and every customer incurs a cost if they obtain a service that deviates from their ideal value. For convenience, assume this attribute is a physical location, but more conceptual interpretations are possible (e.g., quality, ease-of-use, ambiance, etc.) To be specific, one firm is located at position 0, the other is located at position 1, and customers incur a t transportation cost per unit of distance traveled to a firm.

Independent of the firm a customer visits, let a be the value a customer receives from the firm's service. The constant a represents the *base utility* a customer receives from the service, before price, time, and travel costs are included. Overall, a customer located y units from firm i receives utility $u_i(y, f_i)$ from firm i's service,

$$u_i(y, f_i) = a - f_i - ty.$$

The arrival of potential customers to this market is a Poisson process with rate Λ. Once a customer arrives at this market, the customer immediately decides whether or not to seek service from one of the firms. The customer visits the firm that maximizes utility as long as that firm provides a positive utility, otherwise the customer leaves the market without seeking service. Each customer decides which firm to visit based on the customer's location, the firms' prices, and their expected time costs to customers. Customers do not base their decisions on queue lengths, since they are not observed. If a customer chooses a firm, the customer immediately visits the firm, waits in a first-come-first serve queue, and leaves once service is completed. Customer never balk upon arrival at a queue.

For any set of full prices, some fraction of potential consumers will visit firm i. Therefore, the arrival process to firm i will be Poisson. Let λ_i be the mean arrival rate to firm i. Assume customer processing times are exponentially distributed with μ_i as firm i's average processing rate. Based on these conditions, each firm operates an $M/M/1$ queue. If $\mu_i > \lambda_i$, $(\mu_i - \lambda_i)^{-1}$ is the expected time a customer spends at firm i, so

$$g_i = h/\left(\mu_i - \lambda_i\right). \quad (1)$$

It might seem natural to assume that in this game each firm chooses a fixed processing rate. However, the stability of the queues is an issue under that construction. A customer's expected time cost is not finite when $\lambda_i > \mu_i$, but Stidham (1992) and Stidham and Rump (1998) demonstrate that even if $\lambda_i < \mu_i$ occurs in equilibrium, the equilibrium might not be stable: after a small perturbation in the arrival rate the system might not return to the equilibrium. In fact, they demonstrate that chaotic system behavior is possible. To avoid that issue, in the queuing game each firm commits to an expected time cost for its customers, g_i. This implies that the firm is able to adjust its processing rate μ_i in response to any perturbation in its arrival rate so that a customer's expected time cost remains g_i. One interpretation is that time is divided into periods in which the arrival rate changes only at the beginning of each period, the firm observes the arrival rate for each period and the firms chooses μ_i so that g_i will be achieved in that period. This ensures a stable equilibria because customers always receive what they expect to receive, independent of the arrival rate to the queue. It is worth noting that many service operations are managed by expected time rather than expected rate. For example, consider a call center operation, where one typically manages to achieve a certain average speed of answer, i.e., queue time. Employees are brought in at the last minute, or asked to work overtime, to achieve this expected queue time.

Capacity is not free, so let b_i be firm i's cost per unit of average processing rate per unit time. Rearranging (1), firm i's expected capacity cost per unit time is

$$b_i\left(\lambda_i + h_i/g_i\right). \tag{2}$$

Firm i incurs a higher capacity cost whenever it commits to a lower time cost for its customers; hence, the time cost commitment is akin to a capacity commitment.

There are two stages in this queuing game. In the first stage, the firms simultaneously select p_i and g_i. In the second stage, the firms observe their choices and customers arrive to the market over an infinite horizon. Firms maximize expected profits per unit time.

3. Linear City Game

This sections describes the well known (and analytically much simpler) linear city game, which is the stochastic version of the queuing game. Several results from that model are presented so that those results can be contrasted with the results from the queueing game.

In the linear city game each firm processes customers at an infinite rate, so customers never wait for service. Customers arrive at a constant rate Λ to the market. Each customer observes the firms' prices, p_i and p_j, and visits the firm that maximizes utility. Let $u_i^d(y, p_i)$ be the utility a customer located y units from firm i receives from visiting firm i,

$$u_i^d(y, p_i) = a - p_i - ty.$$

Firm i incurs a cost c_i to serve each customer. Assume $c_i < a$; it is possible for either firm to serve at least some customers with a positive price. Firms simultaneously choose their prices and attempt to maximize their profits per unit time. (The equivalent objective of maximizing profits over a finite horizon is typically assumed.)

A Nash equilibrium in this game is a pair of prices $\{p_i^*, p_j^*\}$, such that neither firm can increase its profit with a unilateral deviation from the equilibrium. Equilibria are divided into three types: *monopoly*, *attrition* and *competitive*. With either a monopoly or an attrition equilibrium, there does not exist a customer that could earn a positive utility from visiting either firm. In other words, in those equilibria, neither firm attempts to take customers away from the other firm. In a monopoly equilibrium each firm chooses the monopoly price and some customers are not served, whereas all customers are served in an attrition equilibrium. (In an attrition equilibrium, either firm could begin to capture some of the other firm's customers with just a marginal reduction in its price, but neither firm is willing to do so.) In a competitive equilibrium, there exists some customers that could receive a positive utility from visiting either firm, and every customer is served.

Table 1: Equilibrium results for the linear city game with symmetric costs

Equilibrium type	Existence condition	$\{p_i^*, p_j^*\}$	Each firm's equilibrium profits
competitive	$t < \frac{2}{3}(a-c)$	$\{t+c, t+c\}$	$\Lambda t/2$
attrition*	$\frac{2}{3}(a-c) \leq t \leq (a-c)$	$\{a-t/2, a-t/2\}$	$(\Lambda/2)(a-t/2-c)$
monopoly	$(a-c) < t$	$\{(a+c)/2, (a+c)/2\}$	$(\Lambda/2)(a-c)^2/4$

*A continuum of equilibria may exist. The reported one divides the market equally.

Table 1 summarizes the equilibrium results when the firms have symmetric costs, $c_i = c_j = c$; see Appendix A for the general cost results. There are several other results (see Appendix A for details): in all equilibria, both firms have positive market share and earn positive profits; if the firms have non-identical costs, the low cost firm has a higher market share and a lower price than its competitor in any competitive equilibrium.

4. Queuing Game Analysis

This section begins the analysis of the queuing game with the evaluation of how each firm responds to its opponent's decision. These player responses are used to construct the set of Nash equilibria. The section concludes by characterizing the firms' equilibrium choices.

4.1 Firm Responses

In equilibrium, each firm assumes the other firm will choose its equilibrium actions. Therefore, it is necessary to evaluate how a firm should respond to any set of actions that the other firm may select.

A player's response falls into one of two categories; a player can be either *passive* or *competitive.* A passive player does not steal any customer away from the other player, whereas a competitive player does. Whether a player is passive or competitive depends on the full prices. Let $\overline{y}(f_i)$ be the farthest distance a customer is willing to travel to firm i (a customer traveling that distance earns exactly zero utility),

$$\overline{y}(f_i) = (a - f_i)/t.$$

Firm i is passive (given f_j) when it chooses f_i such that

$$\overline{y}(f_i) + \overline{y}(f_j) \leq 1,$$

otherwise it is competitive. When firm i is competitive, there will be some customer that is indifferent between visiting either firm. Let $y(f_i, f_j)$ be the distance from firm i to that customer,

$$y(f_i, f_j) = (f_j - f_i)/2t + 1/2.$$

Note that the firms are either both passive or both competitive (i.e., it is not possible for one player to be passive while the other is competitive).

To evaluate a player's optimal response, first determine the player's optimal passive response, then determine the player's optimal competitive response and choose the best among those responses.

4.1.1 Passive Response

Define

$$\widetilde{\pi}_i^p(p_i, g_i) = p_i \overline{y}(p_i + g_i)\Lambda - b_i\left(\overline{y}(p_i + g_i)\Lambda + h/g_i\right).$$

$\widetilde{\pi}_i^p(p_i, g_i)$ is player i's profit when none of its customers receives a positive utility from visiting firm j; i.e., when

$$0 \leq \overline{y}(p_i + g_i) \leq 1 - \overline{y}(f_j).$$

Since any customer closer than $\overline{y}(p_i + g_i)$ will visit firm i, and customers are uniformly distributed over the market, the arrival process to firm i is Poisson with rate $\overline{y}(p_i + g_i)\Lambda$. The first term in $\widetilde{\pi}_i^p(p_i, g_i)$ is the firm's revenue per unit time and the final term is the firm's cost per unit time.

Player i's optimal passive response problem is

$$\begin{array}{cc} \max_{p_i,g_i} & \widetilde{\pi}_i^p(p_i,g_i) \\ \text{s.t.} & 0 \le \overline{y}(p_i+g_i) \le 1-\overline{y}(f_j) \end{array}.$$

Define $\pi_i^p(f_i,g_i) = \widetilde{\pi}_i^p(f_i-g_i,g_i)$,

$$\pi_i^p(f_i,g_i) = (\Lambda/t)\,(f_i-g_i-b_i)(a-f_i) - b_ih/g_i.$$

The firm's problem is now expressed with f_i and g_i as the choice variables,

$$\begin{array}{cc} \max\limits_{g_i,\ f_i} & \pi_i^p(f_i,g_i) \\ \text{s.t.} & 2a-t-f_j \le f_i \le a \end{array}.$$

In the queuing game $a > t$ is always assumed; a customer located at one end of the interval could potentially receive positive utility from visiting the firm located at the other end of the interval. Due to this assumption, $2a-t-f_j$ is non-negative, so there is no need to also impose a $f_i \ge 0$ constraint. When $f_i = 2a-t-f_j$, firm i serves all of the uncontested market: the portion of the market that firm j does not serve, $1-\overline{y}(f_j)$.

For a fixed f_i, $\pi_i^p(f_i,g_i)$ is strictly concave in g_i. Let $g_i^p(f_i)$ be the firm's optimal customer time cost for a fixed full price:

$$g_i^p(f_i) = \left(\frac{thb_i}{\Lambda(a-f_i)}\right)^{1/2}.$$

Let $\pi_i^p(f_i) = \pi_i^p(f_i, g_i^p(f_i))$,

$$\pi_i^p(f_i) = \frac{\Lambda}{t}\left((f_i-b_i)(a-f_i) - 2\left(\frac{thb_i(a-f_i)}{\Lambda}\right)^{1/2}\right).$$

The firm's problem is now expressed in terms of its full price, constrained by the full price of the other firm,

$$\begin{array}{cc} \max\limits_{f_i} & \pi_i^p(f_i) \\ \text{s.t.} & 2a-t-f_j \le f_i \le a \end{array}. \tag{3}$$

For a sufficiently high full price, $f_i = a$, the firm earns a zero profit because it attracts no customers. This may indeed be the firm's optimal solution to (3), since $\pi_i^p(f_i)$ is not necessarily positive for any $f_i < a$. In particular, since $(f_i-b_i)(a-f_i)$ is strictly concave in f_i, the objective function will not have a positive solution whenever $\sqrt{thb_i/\Lambda}$ is too large; a passive firm can earn a positive profit only if there is a sufficient scale to the market (i.e., Λ is sufficiently large).

Let $\widehat{f}_i^p$ be an optimal solution to (3). As already mentioned, it may be optimal to serve none of the market, but it is also might be optimal to serve all of the uncontested market.

To determine if an interior $\widehat{f}_i^p$ exists, begin with some derivatives,

$$\frac{\partial \pi_i^p(f_i)}{\partial f_i} = \frac{\Lambda}{t}\left(a - 2f_i + b_i + \left(\frac{thb_i}{\Lambda}\right)^{1/2}(a - f_i)^{-1/2}\right)$$

and

$$\frac{\partial^2 \pi_i^p(f_i)}{\partial f_i^2} = \frac{\Lambda}{t}\left(-2 + \frac{1}{2}\left(\frac{thb_i}{\Lambda}\right)^{1/2}(a - f_i)^{-3/2}\right).$$

The above demonstrate that $\pi_i^p(f_i)$ is concave-convex in f_i, and concave in the interval,

$$\left[2a - t - f_j, a - (thb_i/(16\Lambda))^{1/3}\right]. \tag{4}$$

Hence, if an interior $\widehat{f}_i^p$ exists, then it is unique. A closed form solution to the first order condition, which would provide the candidate for the interior solution, is not practical (it is the solution of a cubic equation). Nevertheless, since the concave interval is identified, it is simple to evaluate $\widehat{f}_i^p$ numerically.

The passive response problem can also be applied if one of the firms decides to not participate in the market. For example, suppose firm j attracts no customers; i.e., $f_j = a$. In that case firm i's passive response problem is equivalent to a monopolist's problem with a maximum demand rate. Indeed, when $f_j = a$, it can be shown that (3) is the same as the $M/M/1$ queue problem studied by Stidham (1992) (see Appendix B for details).

Let $\widehat{f}^m$ be the optimal full price for a monopolist. Since this research focuses on competition, it is natural to assume that both firms cannot simultaneously enjoy monopoly profits, which occurs if both firms wish to serve more than half of the market. The following Lemma outlines sufficient conditions for each firm. They are assumed throughout the paper.

Lemma 1 *If firm i is a monopolist, then it serves more than half of the market whenever*

$$t < \frac{1}{2}(a - b_i) - \frac{1}{4}\gamma_i \tag{5}$$

and

$$\gamma_i \leq (2/5)(a - b_i), \tag{6}$$

where

$$\gamma_i = (2hb_i/\Lambda)^{1/2}.$$

All proofs are in Appendix C.

The firm's cost per customer when it serves exactly half of the market is $b_i + \gamma_i$, so the constant γ_i is the cost per customer of idle capacity. For symmetric costs, $b_i = b_j = b$, let $\gamma = \gamma_i = \gamma_j$. The first condition in Lemma 1 ensures that transportation costs are sufficiently low that each firm ideally wants to attract at least half of the market, and the second condition ensures that there is enough market demand for a monopolist to earn a positive profit even if t approaches the upper bound.

4.1.2 Competitive Response

As with the passive firm, the competitive firm's profit function in first developed in terms of p_i and g_i, and then it is expressed in terms of the full prices.

When firm i chooses a competitive full price, the customer located $y(f_i, f_j)$ from firm i receives a positive utility from visiting firm i, but also receives the same utility from visiting firm j. Hence, firm i's arrive rate is Poisson with rate $\Lambda y(f_i, f_j)$. Define

$$\widetilde{\pi}_i^c(p_i, g_i, f_j) = p_i \Lambda y(p_i + g_i, f_j) - b_i \left(\Lambda y(p_i + g_i, f_j) + h/g_i\right) :$$

$\widetilde{\pi}_i^c(p_i, g_i, f_j)$ is firm i's profit assuming it chooses a competitive full price.

Firm i's optimal competitive response problem is

$$\begin{array}{ll} \max\limits_{p_i,\ g_i} & \widetilde{\pi}_i^c(p_i, g_i, f_j) \\ \text{s.t.} & (1 - \overline{y}(f_j))^+ < y(p_i + g_i, f_j) \leq 1 \end{array} .$$

The constraint ensures that a competitive full price is chosen.

Let $\pi_i^c(f_i, g_i, f_j) = \widetilde{\pi}_i^c(f_i - g_i, g_i, f_j)$, so the firm's problem can be expressed such that f_i and g_i are the choice variables

$$\begin{array}{ll} \max\limits_{g_i,\ f_i} & \pi_i^c(f_i, g_i, f_j) = \frac{\Lambda}{2t}\left(f_i - g_i - b_i\right)\left(f_j + t - f_i\right) - \frac{b_i h}{g_i} \\ \text{s.t.} & (f_j - t)^+ \leq f_i < \min\left\{2a - f_j - t, f_j + t\right\} \end{array} .$$

The upper constraint on f_i ensures that firm i takes more than $(1 - \overline{y}(f_j))^+$ of the market: when $2a - f_j - t > f_j + t,$ all customers receive positive utility from visiting firm j, so firm i can charge up to $f_j + t$ and still capture none of the market. To take the entire market, firm i need only charge $f_j - t$; but, of course, a negative full price is never optimal.

For fixed f_i and f_j, $\pi_i^c(f_i, g_i, f_j)$ is strictly concave in g_i. Let $g_i^c(f_i, f_j)$ be the firm's optimal customer time cost for fixed full prices,

$$g_i^c(f_i, f_j) = \left(\frac{2thb_i}{\Lambda(f_j + t - f_i)}\right)^{1/2} .$$

Let $\pi_i^c(f_i, f_j) = \pi_i^c(f_i, g_i^c(f_i, f_j), f_j)$, so the firm's problem is now

$$\begin{array}{ll} \max\limits_{f_i} & \pi_i^c(f_i, f_j) = \frac{\Lambda}{2t}\left((f_i - b_i)(f_j + t - f_i) - 2\left(\frac{2thb_i(f_j + t - f_i)}{\Lambda}\right)^{1/2}\right) \\ \text{s.t.} & (f_j - t)^+ \leq f_i < \min\left\{2a - f_j - t, f_j + t\right\} \end{array} .$$

Let $\widehat{f_i^c}$ be an optimal competitive full price for the firm. (Since the feasible interval is not closed, $\widehat{f}_i^c$ may not exist.) The lower constraint might bind, or there may exist an interior $\widehat{f}_i^c$. Since the profit function is concave-convex,

$$\frac{\partial \pi_i^c(f_i, f_j)}{\partial f_i} = \frac{\Lambda}{2t}\left(f_j + t - 2f_i + b_i + \left(\frac{2thb_i}{\Lambda}\right)^{1/2}(f_j + t - f_i)^{-1/2}\right),$$

$$\frac{\partial^2 \pi_i^c(f_i, f_j)}{\partial f_i^2} = \frac{\Lambda}{2t}\left(-2 + \left(\frac{thb_i}{2\Lambda}\right)^{1/2}(f_j + t - f_i)^{-3/2}\right),$$

if there is an interior optimal full price, it is unique. As with the passive response, a closed form solution is not practical, but a numerical search over the concave interval,

$$\left[(f_j - t)^+, \min\left\{2a - f_j - t, f_j + t - (1/2)\left(thb_i/\Lambda\right)^{1/3}\right\}\right),$$

quickly yields the optimal interior solution.

4.2 Equilibrium Analysis

A Nash equilibrium in the queuing game is a set of strategies, $\{\{p_i^*, g_i^*\}, \{p_j^*, g_j^*\}\}$, such that each firm chooses an optimal strategy given the strategy of the other player. These strategies imply a set of full prices, $\{f_i^*, f_j^*\}$, such that $f_i^* = p_i^* + g_i^*$. In fact, a set of full prices $\{f_i^*, f_j^*\}$ is sufficient to characterize a Nash equilibrium; the previous section demonstrated that for fixed full prices, there exists a unique optimal time cost for each player and each player's optimal price and time cost choice problem depends only on the other player's full price (and not on how the other player's full price is allocated between price and time costs). That insight simplifies the subsequent equilibrium analysis: even though the actual queuing game has multi-dimensional strategies (prices and time costs), it can be analyzed as a queuing game with single dimensional strategies (full prices). The remainder of this paper only considers equilibria expressed in full prices.[6]

As in the linear city game, there are three types of Nash equilibria: monopoly, attrition,

[6] The analysis of the queuing game is significantly easier if either prices or time costs (capacities) are held constant. Firm i's profit function is strictly concave in p_i or g_i, holding the other decisions fixed. Hence, while in this queuing game each player has a single strategy choice (full price), the non-concavity of the players' profit functions results in non-continuous best responses. See Fudenberg and Tirole (1991) for a general discussion on the complexity of games with non-continuous best responses. In particular, it is difficult even to demonstrate the existence of a pure-strategy Nash equilibrium.

and competitive. With either a monopoly or attrition equilibrium, the firms choose passive full prices. With a monopoly equilibrium, some customers are not served, whereas with an attrition equilibrium all customers are served. All customers are also served in a competitive equilibrium, but now the firms choose competitive prices.

The following theorem removes the attrition equilibria from consideration, even though they may exist in the linear city game when each firm would want to serve at least half of the market.

Theorem 2 *Given the conditions in Lemma 1, there does not exist an attrition Nash equilibrium in the queuing game.*

It follows from Lemma 1 that in any monopoly equilibrium, only one firm serves customers: each firm wishes to serve at least half of the market, so it is not possible for a monopoly equilibrium to exist in which some of the market is left unserved. In fact, monopoly equilibria, $\{\widehat{f}^m, a\}$ and $\{a, \widehat{f}^m\}$, may exist. If monopoly equilibria exist and one of the firms were allowed to move before the other, then that firm would certainly enjoy a strong first-mover advantage: the first mover earns monopoly profits without facing the threat of competitive entry. It is the need to operate at a sufficient scale to achieve positive profits that allows the monopoly equilibria: once one firm takes a large share of the market, the second firm is neither willing to serve the remaining uncontested market nor willing to price sufficiently low to capture some of the large share firm's market. Comparable equilibria do not exist in the linear city game: in the linear city game, each firm always serves some portion of the market.

Since there are no attrition equilibria, and only one firm serves the market in any monopoly equilibrium, only in a competitive equilibrium will both firms have positive market share. Therefore, from the next theorem it follows immediately that there is at most one equilibrium in which both firms serve customers.

Theorem 3 *If a competitive Nash equilibrium exists, then it is the unique competitive Nash equilibrium.*

While the competitive equilibrium may be unique, it is also possible that the competitive equilibrium coexists with the two monopoly equilibria.

Although there are no simple expressions for the firms' optimal responses, when the firms are symmetric (identical processing costs) the competitive equilibrium in full prices nevertheless takes a simple form.

Theorem 4 *When the firms' processing costs are identical,* $b_0 = b_1 = b$, *and*

$$\gamma \le t, \tag{7}$$

then there is a unique competitive Nash equilibrium, $\{f^c, f^c\}$, *where*

$$f^c = t + b + \gamma. \tag{8}$$

Equilibrium profits are

$$\pi_i^c(f^c, f^c) = (\Lambda/2)\,(t - \gamma),$$

so (7) ensures the firms earn a positive profit in equilibrium. The above implies that t is a good proxy for the level of price competition in the market: as t declines, so do prices and profits. Indeed, when $t < \gamma$, price competition is so fierce that the market becomes unstable: there is no pure-strategy Nash equilibrium. Also note that as either $h \to 0$ or as $\Lambda \to \Lambda$, i.e., as queueing effects disappear either because consumers are indifferent to time or if the market's scale is quite large, $\pi_i^c(f^c, f^c) \to \Lambda/2$, which is the competitive equilibrium profit in the linear city game.

Identical processing costs leads to an equilibrium in which the firms choose identical prices and time costs for their customers. Hence, even though the firms could choose to compete along different dimensions, price or time cost, they nevertheless choose the same strategy. This occurs because consumers choose which firm to visit based only on the firms' full prices, and not based on how the full prices are allocated between price and time cost. Interestingly, there are other models with multiple competitive dimensions in which firms do choose diverse strategies (for one example, see Shaked and Sutton, 1983).

Note that γ is a customer's time cost in the symmetric equilibrium, $g_i^c(f^c, f^c) = \gamma$. Using (2), each firm's expected processing cost per unit time is $(b+\gamma)(\Lambda/2)$. Let c^c be the expected processing cost per customer in the competitive equilibrium. Since each firm's arrival rate is $\Lambda/2$, $c^c = b + \gamma$. Thus, in equilibrium, each firm's *full price* is $t + c^c$, which is analogous to each firm's *price* in the linear city's competitive equilibrium, $t + c$.

It also holds that γ is a firm's cost per customer of idle capacity in equilibrium. Therefore, in equilibrium, the processing cost the firm incurs while there are no customers to serve equals the time costs customers incur to receive service.

If the firms have different processing costs, then in the competitive equilibrium, the firm with the lower processing cost will have higher market share and serve customers more quickly (lower time costs).

Theorem 5 *When firm i's processing costs are lower than firm j's, $b_i < b_j$, in the competitive Nash equilibrium firm i serves more than half of the market and firm i's time cost is lower than firm j's.*

Absent from Theorem 5 is a statement regarding price. In fact, the low cost firm can either have a lower or higher price. The latter is certainly an enviable situation; a higher market share with a higher price! To investigate when that is likely, define $p_i^c(f_i^c, f_j^c)$ as firm i's price when $\{f_i^c, f_j^c\}$ is the competitive equilibrium,

$$p_i^c(f_i^c, f_j^c) = f_i^c - g_i^c(f_i^c, f_j^c).$$

Let $f_i^c(b_j)$ be firm i's competitive equilibrium full price as a function of b_j, so

$$\frac{dp_i^c(f_i^c, f_j^c)}{db_j} = \frac{\partial p_i^c(f_i^c, f_j^c)}{\partial b_j} + \frac{\partial p_i^c(f_i^c, f_j^c)}{\partial f_i^c}\frac{\partial f_i^c(b_j)}{\partial b_j} + \frac{\partial p_i^c(f_i^c, f_j^c)}{\partial f_j^c}\frac{\partial f_j^c(b_j)}{\partial b_j}.$$

When firm j's processing cost increases, firm i's equilibrium price becomes higher relative to firm j's equilibrium price when

$$\frac{dp_i^c(f_i^c, f_j^c)}{db_j} - \frac{dp_j^c(f_i^c, f_j^c)}{db_j} > 0. \tag{9}$$

To provide some tractability, assume $b_i = b_j = b$; i.e., before firm j's marginal cost increase the firms had identical processing costs. Given that assumption, (9) becomes

$$\left(\frac{h}{2b\Lambda}\right)^{1/2} > \frac{1}{t}\left[t - \gamma\right]\left(\frac{\partial f_j^c(b_j)}{\partial b_j} - \frac{\partial f_i^c(b_j)}{\partial b_j}\right). \tag{10}$$

The right hand parentheses is evaluated by using the implicit function theorem on the pair of first-order-conditions for competitive full prices. With that result the above can be further simplified (after a significant amount of algebra):

$$(1/t + 1/b)\,\gamma > 1. \tag{11}$$

Several observations are made from (10) and (11) regarding when the low cost firm is likely to enjoy higher prices in equilibrium than the higher cost firm. First, it becomes more likely as t decreases; i.e., as consumers become more price sensitive and thus, as price competition reduces profits. Indeed, as $t \to \gamma$ the competitive firms' profits approach zero and (10) clearly holds; in a competitive environment a firm takes advantage of its lower costs by raising prices. Second, (11) is less likely as Λ increases; as the market's scale increases it behaves like the linear city model in which the low cost firm prices lower than its competitor. Finally, the relationship between b and (11) is intricate (recall that γ depends on b). The left hand side of (11) is quasi-convex in b. As b increases to its maximum value for the existence of a competitive equilibrium, $\Lambda t^2/(2h)$, the competitive firms' profits decline to zero and, as in the case with the transportation costs, the low cost firm takes advantage of its position by increasing its price to gain profits. However, (11) also holds as b decreases to zero. To explain that effect, note that

$$-\frac{\partial p_j^c(f_i^c, f_j^c)}{\partial b_j} = \frac{\partial g_j^c(f_i^c, f_j^c)}{\partial b_j} = \left(\frac{h}{2b\Lambda}\right)^{1/2},$$

which is the left hand side of (10). When processing cost is free, $b = 0$, the firms choose an infinite processing rate and customers experience zero time cost. But it is no longer possible to support that large processing rate if processing cost increases even a tiny bit; holding the firm's full price constant, the firm's time cost soars and its price plummets, leaving the lower cost firm with a higher price.

Does (11) hold for intermediate b values? In fact, it need not. The left hand side of (11) is minimized when $b = t$, in which case (11) does not hold if $t \geq 8h/\Lambda$. Thus, the low cost firm will have higher prices either for high or low processing cost values, but not necessarily for intermediate values; low equilibrium profits cause the former, while a sensitive optimal time cost causes the latter.

In contrast, Lederer and Li (1997) observed that, when there is a single customer class (as in this game) and the firms have access to the same processing technology (in terms of the mean and variance of the service time distribution), the low cost firm always has a higher price and a lower time cost (faster service). However, there are some significant structural differences between their model and this game: they study perfect competition (prices are set to clear markets, rather than as the result of endogenous profit maximizing behavior), and the firms' processing rates are fixed.

5. The Outsourcing Game

This section explores the motivation for outsourcing in the queueing game. Two outsourcing techniques are considered: either the firms can outsource to another firm, which will be called the contractor, or they can outsource to their customers. To provide analytical tractability, assume the firms have identical processing costs, b, and the competitive Nash equilibrium is the expected outcome of the queuing game.

Suppose the firms let a contractor operate each of their queueing processes. The contractor and the firms have the same processing technology and costs: the contractor operates an $M/M/1$ queue, chooses a time cost for customers and incurs a b cost per unit of expected processing capacity. Consider the following contract: the contractor charges each firm c^c per customer served, the contractor guarantees a customer time cost no greater than γ and each firm guarantees a $\Lambda/2$ arrival rate. With this contract, the customers' time costs are the same as in the queuing game equilibrium. While it is assumed that there is one contractor, the analysis remains unchanged if the firms contract with distinct contractors.

After signing those contracts, the firms play a linear city game (i.e., they simultaneously choose their prices) in which their marginal costs are c^c per customer and $a - \gamma$ is each customer's base utility. (From the base utility the customers deduct price and travel costs to obtain their final utility.)

The contractor just breaks even with those contracts: the contractor's revenue per customer is the same as the firms' cost per customer in the queuing game. But both of the firms are strictly better off than in the queuing game.

Theorem 6 *Given the outsourcing contracts, the firms' equilibrium profits and prices are strictly higher than their profits and prices in the queuing game's competitive Nash equilibrium.*

The value of the outsourcing contract is particularly high in competitive markets: as $t \to \gamma$, competitive profits approach zero if the firms maintain control of the queues, but with the outsource contract the firms earn a positive profit even if $t = \gamma$.

The firms are better off because outsourcing replaces their increasing returns to scale technology (a queue) with a constant returns to scale technology (a constant fee per customer). With the latter, the firms are less price competitive than they are with the former. With an increasing returns to scale technology, a marginal decrease in price not only leads to a marginal increase in the arrival rate, it also causes a marginal decrease in cost per customer. That "build scale" incentive does not exist with the constant returns technology, so firms have a lower incentive to cut price; they are less price competitive. (We strongly suspect that this impact of production technology on equilibrium prices is general, and not limited to just queuing games.)

When $t < (2/3)(a - \gamma - c^c)$, the firms do not have to guarantee a $\Lambda/2$ arrival rate to the contractor because the contractor can anticipate that the firms will choose prices that result in that arrival rate. For higher transportation costs, the firms have less incentive than the contractor to expand their arrival rate. Thus, without the guarantee, the contractor correctly anticipates that the firms would choose prices that result in a lower arrival rate.[7]

Since the contractor breaks even with the outsourcing contract and the firms strictly increase their profits, there exists outsourcing contracts that earn the contractor a positive profit and also increase the firms' profits. In other words, the firms would even be willing to consider contracts that have higher costs per customer than c^c. This "win-win" opportunity does not exist in the linear city game: outsourcing in the linear city game, holding the firms' costs constant, has no impact on equilibrium prices, and so the contractor's service creates no additional rents for the firms.

The firms can even be better off if they outsource the service to their customers. Suppose customers choose their own processing rate, μ_s, and incur a b_s cost per unit of μ_s; working faster is costly. Assume there are no congestion effects with self service, so a customer expects to complete the service in $1/\mu_s$ time, which is independent of the number of customers performing self service. In some settings that is reasonable, e.g., self serve assembly. But there are self services in which that is an approximation, e.g., self serve gas pumping is restricted by the number of gas pumps at a service station. Nevertheless, congestion effects are probably significantly lower with self service than if the firm performs the service: in full service gas stations capacity is usually more constrained by the attendants than by the number of gas pumps. We return to this assumption at the end of this section.

Consumers also incur an h_s cost per unit time spent performing the service, so working faster reduces the consumer's time cost. Considering the sum of processing and time costs, $(h_s/\mu_s) + b_s\mu_s$, the consumer's optimal processing rate is $\mu_s^* = \sqrt{h_s/b_s}$ and the resulting cost for the consumer is $2\sqrt{h_s b_s}$ (not including price and travel costs). Finally, without loss of generality, assume each firm has a zero marginal cost per customer when the consumers perform self service.

Suppose the sum of the consumers' processing and time costs with self service equals the processing cost per customer in the queuing game, $b + \gamma$, plus the customer's time cost in

[7] When $t \leq (a - \gamma - c^c)$, each firm serving half the market is an equilibrium, but not the only equilibrium. If the contractor were certain that equilibrium would prevail, then the volume guarantee is not needed. However, the volume guarantee is certainly needed for $t > (a - \gamma - c^c)$.

the queuing game, γ; i.e., $2\sqrt{h_s b_s} = b + 2\gamma$. In that case the firms are strictly better off with self serving customers than they are in the queueing game.

Theorem 7 *When $2\sqrt{h_s b_s} = b + 2\gamma$, each the firms' profit in the outsourcing game with self serving customers is strictly higher than in the queuing game.*

Note that outsourcing to the contractor is not exactly the same as outsourcing to customers, even if costs are identical across all scenarios. A contractor must operate a queue, and since it is an increasing returns to scale technology, the contractor needs a volume guarantee. When the process is outsourced to customers there no longer exists an increasing returns to scale technology in the system, so volume can decline.

Since outsourcing to the customers strictly increases the firms' profits when the customers are as efficient as the firms (the sum of their processing and time costs in the self service game equals the sum of the firms' processing cost and their time cost in the queueing game), the firms can even benefit from outsourcing to the customers if the customers are less efficient than the firms (to some degree). Further, all else being equal, outsourcing to customers becomes more attractive as the market becomes smaller, since then $2\sqrt{h_s b_s} < b+2\gamma$ becomes more likely.

Two issues regarding outsourcing are worth discussing. First, will one firm outsource to a contractor if the other firm keeps the process internal? It is straightforward to show that the outsourcing firm will raise its full price (because it does not need to build scale), and therefore its market share will decline. The contractor will raise its fee because the firm's lower market share makes for a less attractive customer. So the firm will experience lower market share and higher costs, thereby leading to potentially lower profits. In that case, the firms could be playing a game that resembles the famous Prisoner's Dilemma: each firm is better off if they both outsource, but if one firm outsources, then the second firm will be tempted to keep the process internal. In the end, both firms might keep the process internal, and vicious price competition destroys profits. We conjecture that outsourcing will still be observed because over a long time horizon either the firms or the contractors will force sanity onto the players, but the details of that dynamic are left to future research.

The second issue to discuss is the possibility of co-production with congestion. For example, suppose a financial institution establishes a computer network to allow customers to perform co-production. While congestion on the computer network may not be as significant an issue as congestion in a call center (if computer capacity is cheaper than call center capacity, then the firm will provide better service with the former) it nevertheless could remain a concern. We conjecture that our qualitative results continue to hold, i.e., that adopting a more constant returns to scale technology raises equilibrium prices and profits: since the equilibrium full price in the queuing game equals the equilibrium prices in the linear city game, we suspect that the equilibrium full price in a hybrid game (one in which there is a constant cost per customer plus queuing costs) would also be the same, and since the queuing component in the hybrid game would be lower than in the queuing game, the firms' equilibrium prices would be higher, thereby leading to higher profits. (Recall that full price is time cost plus price, so lowering time cost with a constant full price means that price must increase.) Again, confirmation of this intuition is left to future research.

6. Practical Implications and Conclusions

The results of this paper imply that one should see more and more outsourcing of service capacity as markets become more competitive. To illustrate that this phenomenon is occurring quite frequently, consider two examples from the financial services industry.

Pushed by growing consumer demand and the fear of losing market share, banks are investing heavily in PC banking technology (Frei and Kalakota, 1997). Collaborating with hardware, software, telecommunications and other companies, banks are introducing new ways for consumers to access their account balances, transfer funds, pay bills, and buy goods and services without using cash, mailing a check, or leaving home. A key component of this new service is the bill payment function; i.e., writing checks. Typically, banks provided their own fulfillment processing (i.e., handling the check). With the advent of electronic bill payment, the question of who provides this back-office function has become quite important. Banks could invest in this capability themselves, but the transaction volumes required to make this economically viable are quite high. In addition, the peaks in this process do indeed cause the type of scale economies described herein.

All banks have been forced, due to competition, to drive the prices they charge for electronic banking to essentially zero, thereby erasing the profits they assumed would be generated by this service (Harker, Hitt and Frei 1999). How can one restore profitability to this service (or in the case of Citibank, which offered to *pay* customers $50 to adopt PC banking, at least break even?) Enter firms such as Checkfree, which provide this fulfillment service on an outsourced basis (Dalton, 1998). Checkfree will provide the bill payment service to a bank, and charges a flat fee per customer. Our results suggest that this outsourcing is possibly the only way to make such a service viable. Each bank, with its unique brand identity, provides a location in the financial service industry's attribute space. By outsourcing this processing step, they are able to profit from this new distribution channel.

What of outsourcing to the customer? One need only look at online brokerage firms such as Charles Schwab and E*Trade to see the rapid increase in customer involvement in the service delivery process (Schonfeld, 1998). Increasingly, brokers no longer enter orders, customers do. While driving prices down, online trading fees have not uniformly dropped to the lowest cost in the market. Rather, Schwab still commands a premium over its competitors and yet maintains a very healthy market share[8]. This outcome is fully consistent with the predictions of our model.

Thus, our model is only the first attempt to provide an analytical understanding for outsourcing service delivery operations, both to a third-party and, more interestingly, to the consumer. Moreover, the paper provides a novel analysis of a queuing game which also contributes to the theoretical literature in competitive queuing models.

The two examples given above illustrate the fact that the outsourcing phenomenon is quite real and is forcing managers to confront the conceptual issues discussed in this paper. Questions of stability of the equilibria, extensions to multiple firms, empirical analysis of relative profitability based on the hypotheses one can generate from Theorems 6 and 7, etc. must be addressed to fully understand the implications of these outsourcing issues.

---

[8] Note that just 30% share of daily trading volume at Schwab roughly equals that of its next three competitors combined (Schonfeld 1998).

References

Anderson, S., A. de Palma, and J.F. Thisse. 1992. *Discrete Choice Theory of Product Differentiation.* Cambridge, MIT Press.

Armony, M. and M. Haviv. 1998. Nash Equilibrium in a Duopoly Model with Waiting Lines. Hebrew University of Jerusalem working paper.

Bell, C. E. and S. Stidham. 1983. Individual versus Social Optimization in the Allocation of Customers to Alternative Servers. *Management Sci.* 29, 831-839.

Chase, R. 1978. Where Does the Customer Fit in a Service Operation? *Harvard Business Review.* November-December, 138-139.

Dalton, G. 1998. Online Billing Partners. *Informationweek.* 701 30.

Davidson, C. 1988. Equilibrium in Servicing Industries: An Economic Application of Queuing Theory. *J. of Business.* 61, 347-367.

De Vany, A. 1976. Uncertainty, Waiting Time, and Capacity Utilization: A Stochastic Theory of Product Quality. *J. of Political Economy.* 84(3), 523-541.

De Vany, A. and T. Saving. 1983. The Economics of Quality. *J. of Political Economy.* 91(6), 979-1000.

Deneckere, R. and J. Peck. 1995. Competition Over Price and Service Rate When Demand is Stochastic: A Strategic Analysis. *RAND Journal of Economics.* 26(1), 148-162.

Dewan, S. and H. Mendelson. 1990. User Delay Costs and Internal Pricing for a Service Facility. *Management Sci.* 36, 1502-1517.

Frei, F.X. and R. Kalakota. 1997. Frontiers of Online Financial Services. In M.J. Cronin (ed.), *Banking and Finance on the Internet.* New York: Van Nostrand Reinhold Press.

Friedman, J. 1986. *Game Theory with Applications to Economics.* Oxford, Oxford University Press.

Fudenberg, D. and J. Tirole. 1991. *Game Theory.* Cambridge, MIT Press.

Gans, N. 1999. Customer Loyalty and Supplier Strategies for Quality Competition. Wharton School working paper.

Gilbert, S. and Z. K. Weng. 1997. Incentive Effects Favor Non-Competing Queues in a Service System: The Principal-Agent Perspective. *Management Sci.* 44(12), 1662-1669.

Grossman, S. and O. Hart. 1986. The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration. *J. of Political Economy.* 94, 691-719.

Hall, J. and E. Porteus. 1998. Customer Service Competition in Capacitated Systems. Stanford University working paper.

Ha, A. 1998. Incentive-Compatible Pricing for a Service Facility with Joint Production and Congestion Externalities. *Management Sci.* 44(12), 1623-1636.

Hammer, M. 1990. Reengineering Work: Don't Automate, Obliterate. *Harvard Business Review.* 68(4), 104-112.

Harker, P.T., L.M. Hitt, and F.X. Frei, 1999. How Financial Service Firms Decide on Technology. *Brookings-Wharton Papers on Financial Services.* Washington, DC: The Brookings Institution, forthcoming.

Hotelling, H. 1929. Stability in Competition. *Economic Journal.* 39, 41-57.

Kalai, E., M. Kamien, and M. Rubinovitch. 1992. Optimal Service Speeds in a Competitive Environment. *Management Sci.* 38(8), 1154-1163.

Karmarkar, U. and R. Pitbladdo. 1995. Service Markets and Competition. *J. of Operations Management.* 12(3-4), 397-411.

Klein, B., R. Crawford, and A. Alchian. 1978. Vertical Integration, Appropriable Rents, and the Competitive Contracting Process. *J. of Law and Economics.* 21, 297-326.

Kulkarni, V. 1983. A Game Theory Model for Two Types of Customers Competing for |Service. *Operations Research Letters.* 2(3), 119-122.

Lederer, P. J. and L. Li. 1997. Pricing, Production, Scheduling, and Delivery-Time Competition. *Operations Research.* 45(3), 407-420.

Li, L. 1992. The Role of Inventory in Delivery Time-Competition. *Management Sci.* 38(2), 182-197.

Li, L. and Y. S. Lee. 1994. Pricing and Delivery-time Performance in a Competitive Environment. *Management Sci.* 40(5), 633-646.

Lippman, S. and S. Stidham. 1977. Individual Versus Social Optimization in Exponential Congestion Systems. *Operations Research.* 25 233-247.

Loch, C. 1994. Incentive Compatible Equilibria in Markets with Time Competition. INSEAD working paper.

McGuire, T. and R. Staelin. 1983. An Industry Equilibrium Analysis of Downstream Vertical Integration. Marketing Sci. 2, 161-192.

McMillan, J. 1990. Managing Suppliers: Incentive Systems in Japanese and U.S. Industry. *California Management Review*. 32(4), 38-55.

Mendelson, H. 1985. Pricing Computer Services; Queuing Effects. *Comm ACM*. 28, 312-321.

Mendelson, H. and S. Whang. 1990. Optimal Incentive-Compatible Priority Pricing for the M/M/1 Queue. *Operations Research*. 38, 870-883.

Moorthy, S. 1988. Product and Price Competition in a Duopoly. *Marketing Sci.* 7 (2) 141-168.

Naor, P. 1969. On the Regulation of Queue Size by Levying Tolls. *Econometrica*. **37** 15-24.

Rosenthal, R. 1973. A Class of Games Possessing Pure-Strategy Nash Equilibria. *International J. of Game Theory*. 2, 65-67.

Schonfeld, E. 1998. Schwab Puts it All Online. *Fortune*. 138(11), 94-100.

Shaked, A. and J. Sutton. 1983. Natural Oligopolies. *Econometrica.* 51, 1469-1484.

So, K. and J.-S. Song. 1998. Price, Delivery Time Guarantees and Capacity Selection. *European J. of Operations Research*. 111, 28-49.

Stalk, G. Jr. and T.M. Hout. 1990. Competing Against Time. *Research-Technology Management*. 33(2), 19-24.

Stidham, S. 1992. Pricing and Capacity Decisions for a Service Facility: Stability and Multiple Local Optima. *Management Sci.* 38 (8), 1121-1139.

Stidham, S. and C. Rump. 1998. Stability and Chaos in Input Pricing for a Service Facility with Adaptive Consumer Response to Congestion. *Management Sci.* 44(2), 246-261.

van Mieghem, J. 1997. Capacity Investment Under Demand Uncertainty: The Option Value of Subcontracting. Northwestern University working paper.

Venkatesan, R. 1992. Strategic Sourcing: To Make or Not to Make. *Havard Business Review*. 70(6), 98-107.

Williamson, O. 1979. Transaction-Cost Economics: The Governance of Contractual Relations. *J. of Law and Economics*. 22, 233-261.

Appendix A. Linear City Game Analysis

This section summarizes the equilibrium analysis of the linear city game. Let $y_d(p_i)$ be the farthest distance a customer is willing to travel to firm i when it charges p_i, $y_d(p_i) = (a - p_i)/t$, and let $y_d(p_i, p_j)$ be the distance from firm i of the customer that is indifferent between the firms, $y_d(p_i, p_j) = (p_j - p_i + t)/2t$. Let $\pi_i^d(p_i, p_j)$ be firm i's profit,

$$\pi_i^d(p_i, p_j) = \begin{cases} \Lambda y_d(p_i)(p_i - c_i) & \text{if } y_d(p_i) + y_d(p_j) \leq 1 \\ \Lambda y_d(p_i, p_j)(p_i - c_i) & \text{otherwise} \end{cases},$$

where $(p_j - t)^+ \leq p_i \leq a$ is assumed. When $y_d(p_i) + y_d(p_j) \leq 1$, the firms choose passive prices, otherwise, the firms choose competitive prices. Let $\widehat{p}_m^d$ be a firm's optimal monopoly price,

$$\widehat{p}_i^m = \arg\max_{p_i} \pi_i^d(p_i, a) = (a + c_i)/2$$

Since $c_i < a$, the monopolist always serves some customers.

A monopoly Nash equilibrium, $\{\widehat{p}_i^m, \widehat{p}_j^m\}$, is a set of prices such that $y_d(\widehat{p}_i^m) + y_d(\widehat{p}_j^m) < 1$, which simplifies to

$$a - (c_i + c_j)/2 < t. \tag{1}$$

Let $\{\widehat{p}_i^c, \widehat{p}_j^c\}$ be the solution to the first-order-conditions, assuming competitive prices are chosen, $y_d(p_i) + y_d(p_j) > 1$, $\widehat{p}_i^c = t + (2c_i + c_j)/3$. The pair $\{\widehat{p}_i^c, \widehat{p}_j^c\}$ is a competitive Nash equilibrium if $y_d(\widehat{p}_i^c) + y_d(\widehat{p}_j^c) > 1$, which simplifies to

$$t < (2/3)(a - (c_i + c_j)/2). \tag{2}$$

When neither (A-1) nor (A-2) are satisfied, there exists a continuum of attrition Nash equilibria, $\{\widehat{p}_i^a, \widehat{p}_j^a\}$, such that $y_d(\widehat{p}_i^a) + y_d(\widehat{p}_j^a) = 1$. For $y_d(\widehat{p}_i^a) = 1/2$ to be an equilibrium, it must be that each firm wants at least half the market and neither is willing to cut price further to take more than half of the market. For firm i, that condition is met when, $(2/3)(a - c_i) \leq t \leq (a - c_i)$, and the analogous condition must hold for firm j.

Appendix B. Relationship to Stidham (1992)

This section demonstrates that a monopolist in the queuing game faces the same $M/M/1$ queueing problem studied in Stidham (1992) (S92). (S92 also considers the more general $GI/GI/1$ queue.) Unless otherwise noted, the notation in this paper matches the notation in S92. In S92, the monopolist chooses a processing rate μ and incurs a per unit time cost

b per unit of processing rate. Customers incur time costs at rate h. Hence, $h/(\mu - \lambda)$ is a customer's expected time cost, assuming $\mu > \lambda$. While in the queuing game, the firm chooses an expected time cost and the average processing rate is inferred; in S92, the firm chooses a processing rate and the expected time cost is inferred. Since there is a one-to-one relationship between the time cost and the processing rate, those decisions are equivalent. (Although, they are not the same when considering equilibrium stability.)

It remains to confirm that the demand models are the same. In S92, when λ is the arrival rate to the firm, $V'(\lambda)$, is the value the marginal customer receives from the service, $V'(\lambda) = \widehat{a} - (\widehat{a} - d)(\lambda/\Lambda)$,where "$\widehat{a}$" replaces "a" used by S92, since "a" has a slightly different meaning in this paper, and d is an exogenous parameter. $V'(\lambda)$ determines the firm's demand rate because the marginal consumer will visit the firm only if $V'(\lambda)$ is no less than the sum of the firm's price and time cost. In the queuing game $\widetilde{V}(\lambda) = (\lambda/2)(a - t\lambda/\Lambda)$ is the total value per unit time received by customers when the firm's demand rate is $\lambda = y\Lambda$, where y is the farthest distance a customer will travel to firm i. Differentiating, $V'(\lambda) = (a/2) - t(\lambda/\Lambda)$. When $a/2 = \widehat{a}$ and $t = \widehat{a} - d$, $V'(\lambda) = \widetilde{V}'(\lambda)$, and thus, the demand models are the same: both are downward sloping and linear in λ, and both have two degrees of freedom to adjust the intercept and slope of the demand curve.

Appendix C. Proofs

Proof of Lemma 1. Since $\pi_i^p(f_i)$ is continuous in f_i, the monopolist's optimal full price is either the solution to the first-order-condition, or one of the constraints binds. Let $f_i^m(t)$ be the solution to the first-order-condition as a function of the transportation cost parameter. From the implicit function theorem, $f_i^m(t)$ is non-decreasing in t. Therefore, there exists some transportation cost such that $f_i^m(t) = a - t/2$ ($\Lambda/2$ is the monopolist's optimal demand rate) and for all lower transportation costs the monopolist's optimal demand rate is more than $\Lambda/2$. The full price $a - t/2$ solves the first-order-condition when

$$t = (a - b_i)/2 - \gamma_i/4, \tag{1}$$

hence the first condition in the lemma. The monopolist must also earn non-negative profits, $\pi_i^p(a - t/2) \geq 0$, which simplifies to

$$t \leq 2(a - b_i) - 4\gamma_i. \tag{2}$$

Given (C-1), condition (C-2) holds when $\gamma_i \leq (2/5)(a - b_i)$, thereby confirming the second condition. □

Proof of Theorem 2. Let $\{f_i^a, f_j^a\}$ be an attrition equilibrium. By definition, $\overline{y}(f_i^a) + \overline{y}(f_j^a) = 1$, which simplifies to

$$f_i^a = 2a - t - f_j^a. \tag{3}$$

The market is not necessarily divided equally between the two firms, even if they have identical processing costs. Without loss of generality, assume firm j serves at least half of the market, $f_j^a \leq a - t/2$. If $\{f_i^a, f_j^a\}$ is indeed an equilibrium, then firm i must not wish to steal any of firm j's customers,

$$\frac{\partial \pi_i^c(f_i^a, f_j^a)}{\partial f_i} \geq 0, \tag{4}$$

and firm i must earn a non-negative profit,

$$\pi_i^p(f_i^a) \geq 0. \tag{5}$$

Using (C-3), (C-4) simplifies to

$$(thb_i/\Lambda)^{1/2} \left(f_j^a + t - a\right)^{-1/2} \geq 4a - 3t - b_i - 3f_j^a. \tag{6}$$

The left hand side of (C-6) is strictly convex and decreasing in f_j^a over the interval of interest, $[a - t, a - t/2]$. (When $f_j^a = a - t$, firm j takes the entire market and thus a lower price is never optimal.) If $f_j^a = a - t/2$ (i.e., the firms divide the market equally), (C-6) reduces to $t \geq (2/3)(a - b_i - \gamma_i)$, which cannot be satisfied given the conditions in Lemma 1. Define $\overline{f}$ such that $\partial \pi_i^c(\overline{f}, f_j^a)/\partial f_i = 0$ and $\overline{f} < a - t/2$; when $f_j^a = \overline{f}$, (C-6) holds with equality. Since the right hand side of (C-6) is a decreasing linear function of f_j^a, there is a unique $\overline{f}$. Further, (C-6) holds for any $f_j^a \in [a - t, \overline{f}]$.

The positive profit condition, (C-5), simplifies to

$$\left(2a - t - f_j^a - b_i\right)/2 \geq (thb_i/\Lambda)^{1/2} \left(f_j^a + t - a\right)^{-1/2}. \tag{7}$$

Comparison of (C-6) and (C-7) demonstrates that a feasible f_j^a exists only if

$$\left(2a - t - f_j^a - b_i\right)/2 \geq (thb_i/\Lambda)^{1/2} \left(f_j^a + t - a\right)^{-1/2} \geq 4a - 3t - b_i - 3f_j^a. \tag{8}$$

Define $\widetilde{f}$ such that $\left(2a - t - \widetilde{f} - b_i\right)/2 = 4a - 3t - b_i - 3\widetilde{f}$, i.e., $\widetilde{f} = a - t + \frac{1}{5}(a - b_i)$. Since both the left and right hand terms in (C-8) are decreasing linear functions, but the right hand term decreases at a faster rate, there exists a feasible f_j^a only if $\widetilde{f} \leq \overline{f}$. Alternatively, there exists a feasible f_j^a only if $\partial\pi_i^c\left(2a - t - \widetilde{f}, \widetilde{f}\right)/\partial f_i \geq 0$, which simplifies to

$$(2/5)\left(2(a-b_i)/5t\right)^{1/2}(a-b_i) \leq \gamma_i. \tag{9}$$

Since the left hand side of (C-9) is decreasing in t, and (5) bounds t from above, if (C-9) holds strictly for

$$t = 1\left(a-b_i\right)/2 - \gamma_i/4, \tag{10}$$

then it holds for all feasible t. Substitute (C-10) into (C-9) and simplify,

$$\left((2/5)\left(a-b_i\right)\right)^3 < \gamma_i^2\left((a-b_i)/2 - \gamma_i/4\right), \tag{11}$$

where one notes that strict inequality is required. Given (6), the right hand side of the above condition is strictly increasing in γ_i. However, (C-11) does not hold even if γ_i equals its upper bound, $(2/5)(a-b_i)$. Therefore, there does not exist parameter values such that (C-9) is satisfied. Hence, $\widetilde{f} > \overline{f}$ and there does not exist an $f_j^a \in [a-t, a-t/2]$ such that (C-3), (C-4) and (C-5) are satisfied. □

Proof of Theorem 3. Let $\widehat{f_i^c}(f_j)$ be firm i's optimal competitive full price given firm j's full price. In a competitive equilibrium, each firm must have some market share. Hence, when $\widehat{f_i^c}(f_j)$ exists, it is the solution to a first-order-condition, it is continuous and differentiable. From the implicit function theorem

$$\frac{\partial\widehat{f_i^c}(f_j)}{\partial f_j} = \frac{1 - \frac{1}{2}\gamma_i t^{1/2}\left(f_j + t - f_i\right)^{-3/2}}{2 - \frac{1}{2}\gamma_i t^{1/2}\left(f_j + t - f_i\right)^{-3/2}}. \tag{12}$$

From the first-order-condition, $(f_i - b_i) - (f_j + t - f_i) = \gamma_i t^{1/2}\left(f_j + t - f_i\right)^{-1/2}$, which can be written as

$$(1/2)\left((f_i - b_i)/(f_j + t - f_i\right) - 1/2 = (1/2)\,\gamma_i t^{1/2}\left(f_j + t - f_i\right)^{-3/2}.$$

Substituting the above into (C-12) yields

$$\frac{\partial\widehat{f_i^c}(f_j)}{\partial f_j} = \frac{3 - (f_i - b_i)/(f_j + t - f_i)}{5 - (f_i - b_i)/(f_j + t - f_i)}$$

Since profits must be non-negative in the competitive equilibrium,

$$(f_i - b_i) - 2\gamma_i t^{1/2}\left(f_j + t - f_i\right)^{-1/2} > 0.$$

Combining the above with the first-order-condition yields

$$\left(f_i - b_i\right)/\left(f_j + t - f_i\right) < 2.$$

Therefore,

$$0 < \frac{\partial \widehat{f}_i^c(f_j)}{\partial f_j} < 1.$$

The above implies that the firms' competitive response functions are contraction mappings, which means that any competitive equilibrium is unique (see Friedman, 1986).□

Proof of Theorem 4. From Theorem 3, if there is a competitive equilibrium, then it is unique. When the firm's processing costs are identical, $\{f^c, f^c\}$ is the only solution to the first-order-conditions. Some additional conditions must be met before $\{f^c, f^c\}$ is confirmed as an equilibrium. First, each firm must make a non-negative profit in equilibrium (because the best passive response guarantees at least a zero profit). Firm profits in equilibrium are

$$\pi_i^c(f^c, f^c) = (\Lambda/2)\,(t-\gamma)\,, \tag{13}$$

which is non-negative when (7) holds. Second, each firm's competitive response must be better than its best passive response; i.e., $\pi_i^c(f^c, f^c) \geq \pi_i^p(\widehat{f}_i^p, f^c)$. From Lemma 1, each firm wants to serve at least half of the market, so the optimal passive response is one of the interval boundaries: either the firm serves none of the market or the firm captures the entire uncontested market, $(1-\overline{y}(f^c))^+$. If there is no uncontested market, $1-\overline{y}(f^c) < 0$, then there is no passive response, so the condition is trivially met. If there is an uncontested market, let $f_i^p = 2a - t - f^c$; f_i^p is the full price that lets the firm serve the entire uncontested market. However, in that case $\lim_{f \to f_i^p} \pi_i^c(f, f^c) = \pi_i^p(f_i^p)$. Thus, if there is an optimal interior competitive response, $\widehat{f}_i^c < f_i^p$, then it is better than the best passive response. f^c is the optimal interior competitive response if f^c is a local optimum and the customer located $y(f^c, f^c)$ from firm i receives a positive utility. The local optimum condition is $\partial^2 \pi_i^c(f^c, f^c)/\partial f_i^2 \leq 0$, which simplifies to $\gamma/4 \leq t$, which is less restrictive than the non-negative profit condition. That consumer is willing to purchase from either firm when $t < (2/3)\,((a-b)-\gamma)$, which is no more restrictive than (5) whenever (6) holds. □

Proof of Theorem 5: From the implicit function theorem, $\partial \widehat{f}_i^c(f_j)/\partial b_j = 0$ and $\partial \widehat{f}_j^c(f_i)/\partial b_j > 0$. From Theorem 3, $0 < \partial \widehat{f}_i^c(f_j)/\partial f_j < 1$. So an increase in b_j increases both f_i and f_j in equilibrium, but f_j increases more than f_i. Thus, firm i has a larger market share and $g_i^c(f_i, f_j) < g_j^c(f_i, f_j)$.□

Proof of Theorem 6. For notational convenience, let $\overline{u} = a - \gamma$. When $t < (2/3)(\overline{u} - c^c)$, equilibrium profits in the linear city game per firm are $t\Lambda/2$, which is clearly greater than equilibrium profits in the queueing game, $\Lambda(t-\gamma)/2$. Their equilibrium prices are $t + c^c$, which is greater than prices in the queueing game, $t+b$. When $(2/3)\left(\overline{u}-c^c\right) \leq t \leq \left(\overline{u}-c^c\right)$, there exists a continuum of equilibria in the linear city game, but only one in which each firm guarantees its contractor a $\Lambda/2$ arrival rate. Each firm's equilibrium profit is

$$(\Lambda/2)\left(a-b-t/2-2\gamma\right), \tag{14}$$

which is greater than the equilibrium profit in the queueing game if $(2/3)\left(a-b-\gamma\right) > t$, which hold because (5) and (6) are assumed. Their prices in that equilibrium are $a-t/2$, which is higher than $t+b$, given (5). When $\left(\overline{u}-c^c\right) \leq t$, each firm wants to serve less than half of the market, but given their commitment to a $\Lambda/2$ arrival rate, (C-14) is each firm's profit. So the above results hold in that case as well. □

Proof of Theorem 7. The outsourcing game with self serving customers is a linear city game in which each customer's base utility is $\overline{u} = a - 2\sqrt{h_s b_s}$ $(\overline{u} = a - b + 2\gamma)$ before including price and travel costs. We check each of the three equilibrium types. When $t < (2/3)\overline{u}$ there is a competitive equilibrium, with prices equal to t. Profits are $\Lambda t/2$, which are strictly greater than in the queuing game, $(\Lambda/2)(t-\gamma)$. When $(2/3)\overline{u} \leq t \leq \overline{u}$ there are a continuum of attrition equilibria in which all customers are served. Assume the equilibrium in which the market is evenly divided between the firms. Each firm's profit is $(\Lambda/2)(\overline{u}-t/2)$, which is strictly greater than in the queuing game given (5) and (6). When $t > \overline{u}$, there is a unique equilibrium in which each firm chooses its monopoly price and some customers are not served. Each firm's profits is $(\Lambda/t)(\overline{u}/2)^2$, which is greater than in the queuing game when

$$(a-b-2\gamma)^2 > 2t(t-\gamma). \tag{15}$$

The right hand side is strictly convex in t and increasing in t when $2t(t-\gamma) > 0$. Thus, given (5), (C-15) holds for all t if it holds for $t = (a-b)/2 - \gamma/4$. At that t value, (C-15) simplifies to

$$(a-b)^2 - 5\left(a-b\right)\gamma + (27/4)\,\gamma^2 > 0. \tag{16}$$

The left hand side of (C-16) is convex in γ, and decreasing when $\gamma < (20/27)(a-b)$. From (6), $\gamma \leq (2/5)(a-b)$, so (C-16) holds for all γ if it holds for $\gamma = (2/5)(a-b)$. At that γ, (C-16) simplifies to $(2/25)(a-b)^2 > 0$, so profits increase.□