We consider a single buyer who wishes to outsource a fixed demand for a manufactured good or service at a fixed price to a set of potential suppliers. We examine the value of competition as a mechanism for the buyer to elicit service quality from the suppliers. We compare two approaches the buyer could use to orchestrate this competition: (1) a supplier-allocation (SA) approach, which allocates a proportion of demand to each supplier with the proportion allocated to a supplier increasing in the quality of service the supplier promises to offer, and (2) a supplier-selection (SS) approach, which allocates all demand to one supplier with the probability that a particular supplier is selected increasing in the quality of service to which the supplier commits. In both cases, suppliers incur a cost whenever they receive a positive portion of demand, with this cost increasing in the quality of service they offer and the demand they receive. The analysis reveals that (a) a buyer could indeed orchestrate a competition among potential suppliers to promote service quality, (b) under identical allocation functions, the existence of a demand-independent service cost gives a distinct advantage to SS-type competitions, in terms of higher service quality for the buyer and higher expected profit for the supplier, (c) the relative advantage of SS versus SA depends on the magnitude of demand-independent versus demand-dependent service costs, (d) in the presence of a demand-independent service cost, a buyer should limit the number of competing suppliers under SA competition but impose no such limits under SS competition, and (e) a buyer can induce suppliers to provide higher service levels by selecting an appropriate allocation function. We illustrate the impact of these results through three example applications.

Keywords: outsourcing; supplier competition; service quality; inventory systems; queueing analysis

History: Accepted by Paul H. Zipkin, operations and supply chain management; received July 24, 2004. This paper was with the authors 8 months for 4 revisions.
requested. For physical goods, typical measures of service quality, or service levels, include fill rate, expected order delay, the probability that order delay does not exceed a quoted lead time, and the percentage of orders fulfilled accurately. For services, measures of service quality include expected customer waiting time, the probability that the customer receives service within a specified time window, and the probability that a customer does not leave (renge) before being served. Selecting suppliers who are able to consistently deliver on one or more of these service measures is particularly important when the buyer envisions a long-term relationship with her suppliers.

In this paper, we consider a single buyer who wishes to outsource a fixed demand for a manufactured good or service at a fixed price to a set of N suppliers. We examine the value of competition as a mechanism for the buyer to elicit good service quality from her suppliers. We consider two plausible schemes the buyer could use to set up a competition. In the first, the buyer allocates a proportion of demand to each supplier, with the proportion a supplier receives increasing in the service level she offers. In the second, the buyer selects a single supplier with the probability that a particular supplier is selected increasing in the service level the supplier offers. In both types of competition, we show that expected service quality is sensitive to the allocation function and the actual demand the supplier receives. The demand-independent component includes costs that depend on the service level promised, but are independent of the demand actually allocated. However, when a demand-independent service cost exists, SS competition leads to higher service levels than those obtained under SA (assuming that the same demand allocation function is used for both competitions). In this case, the two types of competition also differ in the effect of the number of suppliers on service quality. In particular, service levels always increase in the number of suppliers under SS, but may initially increase and then decrease under SA (implying a finite optimal number of suppliers).

In both types of competition, we show that expected service quality is sensitive to the allocation function of the buyer used to translate service level into expected market share. We show that with a properly designed allocation function, the buyer can in some cases maximize service quality and extract all supplier profits.

We illustrate our results with three example applications. The first example involves competition in a make-to-order environment where service quality is measured by response time and suppliers affect their service offering by investing in capacity. The second example looks at a make-to-stock environment where service level is measured by fill rate and is determined by the supplier’s chosen base-stock level. The third example considers a single-period problem where competing suppliers decide on order quantities prior to demand realization and service level is determined by the ability of a supplier to fulfill allocated demand immediately.

The remainder of this paper is organized as follows. In §2, we provide a brief review of related literature. In §3, we describe our problem formulation and the two types of competition. In §4, we study the effect of allocation functions. In §5, we describe a model for supplier selection under SS competition. In §6, we discuss the example applications. In §7, we summarize the main results and comment on possible extensions.
2. Related Literature

The competition described in this paper can be viewed as a form of a rent-seeking game (Tullock 1980). In a rent-seeking game, there are $N$ contestants who compete for a prize. The probability that a contestant wins the prize (the rent) increases with her expenditures and decreases in the expenditures of other contestants. In the rent-seeking literature, the probability of winning is typically assumed to have the form $e_i^\gamma / \sum_{i=1}^N e_i^\gamma$, where $e_i$ is the expenditure of contestant $i$, $N > 1$ is the number of contestants, and $\gamma \geq 0$ is a parameter denoting the ease with which expenditures affect outcome. A focus of this literature has been documenting the inefficiency of rent-seeking games. Rent seeking is viewed as wasteful because depending on the value of $N$ and $\gamma$, the total expenditures by the contestants can equal the value of the prize itself, a phenomenon called rent dissipation. Recent papers from the rent-seeking literature include Niti (1997), Konrad and Schlesinger (1997), Skaperdas (1996), and Perez-Castrillo and Verdier (1992). A review can be found in Nitzan (1994). Related literature on other forms of contests include Lazear and Rosen (1981), Green and Stokey (1983), Dixit (1987), and Kalra and Shi (2001).

There are important differences between the supplier competition we consider in this paper and rent-seeking contests. In our models, we explicitly model two parties: a buyer and her suppliers, with the buyer orchestrating the contest. We introduce the notion of service quality, absent in rent-seeking contests, which is used by the buyer to measure the efficiency of the contest. Consequently, the buyer does not necessarily value the cumulative effort over all suppliers because contests that yield higher levels of cumulative effort do not necessarily yield higher average service levels. Furthermore, in our models, expenditures by contestants occur only after a contestant has been declared a winner and is allocated a fraction of demand. We also allow for general definitions of effort cost and demand allocation.

Our supplier competition is also related to competition among multiple firms for market share, where the share realized by one firm depends on its own effort (e.g., its advertising budget) as well as the effort of other competing firms. The market share captured by firm $i$ is commonly modeled via a market attraction function of the form $a_i e_i^\gamma / \sum_{i=1}^N a_i e_i^\gamma$, where $a_i \geq 0$ represents the effectiveness of effort expended by firm $i$ (alternatively, a measure of customer bias toward firm $i$) and $\gamma > 0$ the attraction elasticity of effort of firm $i$ (see, for example, Moorthy 1993, §5.1; Cooper 1993, p. 262; Monahan 1987; Monahan and Sobel 1997; and the references therein). Bell et al. (1975) identify attributes that lead to market share functions having this form. Kotler (1984, p. 231) refers to such a market-share allocation as the “Fundamental Theorem of Market Share.” Demand allocations with a market-attraction form can also arise as the equilibrium of a Markovian consumer choice process (Mahajan and van Ryzin 2001a).

Wang and Gerchak (2001) use a market-attraction function to model marketing effort in the form of inventory displayed on a retailer’s shelf space. They consider a setting with two competing retailers, $\gamma = 1$, and no supplier bias ($a_i = 1$ for $i = 1, 2$) but with total demand increasing concave in the cumulative effort of the competing retailers. Boyaci and Gallego (2004) also use a market-attraction function to model competition between two supply chains, where effort is measured by fill rate. Bernstein and Federgruen (2004) consider a more general form of competition involving both price and fill rate. They also consider a general allocation function, of which market-attraction functions are special cases.

The above papers are related to a growing literature in operations management on inventory competition among multiple firms. Each firm is typically modeled by a newsvendor that decides on an order quantity prior to observing demand. Two different approaches for allocating demand have been considered (Cachon 2003). Under the first approach, total market demand $D$ is allocated to firms proportionally to their order quantities, with retailer $i$ receiving $D_i = q_i D/(q_1 + \cdots + q_N)$, where $q_i$ is the quantity ordered by retailer $i$ ($i = 1, \ldots, N$). In this case, the demands realized by the firms are perfectly correlated, with either each firm having excess demand (when $D > q_1 + \cdots + q_N$) or each firm experiencing shortages (when $D < q_1 + \cdots + q_N$). Under the second approach, each retailer faces an independent demand $D_i$ and only excess demand from firm $i$ can be reallocated among the other retailers according to some fixed reallocation rule.

An important insight from this literature is that retailers tend to overstock, choosing order quantities that are higher than those observed in the absence of competition. Examples of papers that study inventory competition include Lippman and McCardle (1997), Parlar (1988), Karjalainen (1992), Netessine and Rudy (2003), Li and Ha (2003), and Mahajan and van Ryzin (2001a, b). A review and discussion of this literature can be found in Cachon (2003). Inventory competition can be viewed as a variation on a rent-seeking contest where, instead of a single winner, the prize (total demand) is shared among the contestants according to an allocation rule. There is also a growing literature on market-share competition based on service quality, where service quality is a function of effort parameters other than inventory (e.g., delivery lead times). Recent examples include Hall and Porteus (2000), Gans (2002), Ha et al. (2003), Allon and Federgruen (2005), Bernstein and Federgruen (2004), Boyaci and Ray (2003), and the references therein.
This literature does not consider settings where a buyer is orchestrating the competition and specifying the allocation function. Instead, the allocation emerges endogenously from the competition of independent firms. Consequently, most of this literature is not concerned with identifying forms of competition that maximize supplier effort. However, notable exceptions include recent papers by Elahi et al. (2003) and Cachon and Zhang (2006a). Elahi et al. (2003) consider a system with a single buyer and multiple suppliers. The buyer allocates demand among the suppliers based on their fill rates. The suppliers are modeled as make-to-stock queues who affect their fill rates by increasing their inventory base-stock levels. This model is revisited in §6.2 and is shown to be a special case of the general model we describe in this paper. Cachon and Zhang (2006a) consider a problem with a single buyer and multiple suppliers, where the buyer uses suppliers’ delivery lead times to allocate demand. The suppliers are modeled as single-server queueing systems who affect their lead time performance by exerting effort in the form of capacity. The objective of the buyer is to induce suppliers to invest sufficient capacity to meet a target average lead time. The authors evaluate several allocation functions and show that not all allocation functions induce the desired capacity investments. The model described in Cachon and Zhang (2006a) extends previous models by Gilbert and Weng (1997) and Kalai et al. (1992).

Much of the literature dealing with firms competing for market share does not consider forms of competition where a single firm is allocated the entire market, except as a result of extreme asymmetry among the firms (e.g., the existence of a firm with a zero cost of effort). However, there is extensive literature dealing with supplier selection when there is a single buyer making the procurement decision. In most of this literature, the mechanism by which suppliers are selected is an auction where price is the selection criterion. The literature on procurement auctions is vast and spans both the fields of operations management and economics. Reviews can be found in Klemperer (1999), McAfee and McMillan (1987), Laffont and Tirole (1994), and Elmaghraby (2000). Some of this literature involves auctions with multiple sourcing as in Laffont and Tirole (1987), Anton and Yao (1989), and Seshadri (1995).

Recently, there has been renewed interest in the operations management literature in supplier selection and the allocation of supply contracts via auction mechanisms. For example, Cachon and Zhang (2006b) consider a buyer that selects one out of $N$ potential suppliers with the objective of minimizing the sum of procurement, inventory, and backordering costs. A supplier is selected using a scoring-rule auction based on price and lead time. This creates a price and capacity competition among the suppliers where in this case each supplier’s capacity cost is private information. The authors analyze the relative performance of a number of scoring rules including total cost, lead time only (with fixed price), and price only (with a fixed lead-time target). Other examples of auction-based supplier selection include Chen (2004) and Seshadri and Zemel (2003), who use supplier competition to determine both price and order quantity.

Finally, there is an extensive literature dealing with inventory replenishment policies when there are multiple suppliers or multiple supply modes. In this literature, the characteristics of the suppliers are exogenous and are not affected by the amount of demand that each supplier receives. Examples include Whittimore and Saunders (1977), Moinzadeh and Nahmias (1988), Ramasesh et al. (1991), Anupindi and Akella (1993), Rosenblatt et al. (1998), Swaminathan and Shanthikumar (1999), Chen et al. (2001), Fong et al. (2001), and the references therein.

3. Competition Formulation and Nash Equilibrium
We consider a system with a single buyer that seeks to outsource the provisioning of a product with an expected demand quantity $\lambda$ to $N$ identical potential suppliers. The price of the product, $p$, is fixed and identical across all suppliers. The supplier realizes a revenue $r = p - c$ per unit sold, where $c$ is the unit production cost. Let $s_i \geq 0$ denote the service level offered by supplier $i$ and $\alpha_i = a_i \lambda$ the amount of demand allocated to supplier $i$, $0 \leq \alpha_i \leq 1$ for all $i = 1, \ldots, N$. Also, let $f(s_i, \lambda)$ denote the cost supplier $i$ incurs in providing service level $s_i$ ($s_i \geq 0$) if given demand allocation $\lambda_i$ with $f(s_i, \lambda_i)$ nondecreasing in both $s_i$ and $\lambda_i$. We choose to separate production costs from service level costs because we assume that unit production costs remain the same regardless of the service level offered. We assume that each supplier commits to fulfilling the amount of demand allocated while maintaining the service level promised.

We focus on a particular class of plausible cost functions of the form

$$f(s_i, \lambda_i) = \lambda_i u(s_i) + v(s_i), \quad (1)$$

where $u(s_i)$ and $v(s_i)$ are nondecreasing convex functions in $s_i$, with either $u(s_i)$ or $v(s_i)$ increasing in $s_i$, and $v(0) = 0$ for $i = 1, \ldots, N$. The first term, $\lambda_i u(s_i)$, captures service-related costs that increase linearly with the amount of demand allocated. We refer to this as a demand-dependent cost because it varies with the demand allocated to the supplier. The second term, $v(s_i)$, captures cost that increases only with the service
We consider SA and SS competitions as two plausible strategies the buyer might use to induce service-based competition across the $N$ potential suppliers. Under SA competition, the buyer announces a criterion for allocating demand among the suppliers with the understanding that a supplier can increase her fraction of demand by increasing the service level she promises to offer the buyer. This does not prevent the buyer from taking into account factors other than service level in making the allocation decision.

Under SS competition, the buyer selects a single supplier to whom the entire demand is allocated. The probability that a particular supplier is selected is increasing in the service level the supplier promises to offer. Of course, this does not exclude settings where the supplier with the highest service level is always selected. SS competition is different from SA competition in that a winner takes all under SS (the supplier commits a-priori to sole sourcing), while more than one supplier may be awarded a share of the demand under SA (the supplier does not preclude a-priori the possibility of multisourcing from all suppliers). Under SS, only the selected supplier incurs a cost, while under SA, all suppliers that promise a positive service level eventually do. The probabilistic selection implies that quality of service alone may not guarantee that a supplier would be selected or that there is inherent randomness in the buyer’s decision-making process. A supplier only increases her chances of being selected by offering a higher service level. An alternative interpretation of SA and SS competitions, for which the analysis remains the same, is one where the allocation functions are estimated by the suppliers rather than explicitly announced by the buyer. In fact, in the case of SS, it is unlikely that the buyer would explicitly announce the selection probability function. Instead, the buyer may announce a decision-making process through which the probability function is inferred by the suppliers (see §5 for further discussion).

We assume that, once promised, service levels offered by the suppliers are enforceable. In practice, this would occur if the cost or, more likely, the associated effort expended by each supplier after the buyer allocates demand, is observable. The buyer can then ascertain whether or not a supplier has exerted sufficient effort (expended sufficient cost) to meet the promised service level. For instance, the buyer may observe the amount of capacity invested by the supplier after the demand was allocated and determines whether or not it is sufficient to meet the expected lead time that was initially promised by the supplier. Of course, there can also be settings where suppliers voluntarily deliver on promised service levels (regardless of observability of cost or effort) because they worry about their reputation or expect repeated interactions with the buyer in the future.

For SA competition, demand allocation is carried out via a demand allocation function vector $\alpha^{SA} = (\alpha_1^{SA}, \alpha_2^{SA}, \ldots, \alpha_N^{SA})$, where $\alpha_i^{SA}(s_i, s_{-i})$ specifies the fraction of demand allocated to supplier $i$ given the supplier’s own service level $s_i$ as well as the service levels $s_{-i} = (s_1, \ldots, s_{i-1}, s_{i+1}, \ldots, s_N)$ offered by her competitors with $0 \leq \alpha_i^{SA}(s_i, s_{-i}) \leq 1$. The function $\alpha_i^{SA}(s_i, s_{-i})$ is nondecreasing concave in $s_i$ and is equal to zero when $s_i = 0$ for $i = 1, \ldots, N$. By offering a certain service level, the supplier commits to exerting the necessary effort (and incurring the associated cost) to maintain this service level regardless of the demand it may eventually receive. However, because $\alpha_i^{SA}(s_i, s_{-i})$ is nondecreasing concave in $s_i$, $\alpha_i^{SA}(s_i, s_{-i}) > 0$ if $s_i > 0$, and $\alpha_i^{SA}(s_i, s_{-i}) = 0$ if and only if $s_i = 0$.

For SS competition, the demand allocation is carried out via a selection probability function vector $\alpha = (\alpha_1^{SS}, \alpha_2^{SS}, \ldots, \alpha_N^{SS})$, where $\alpha_i^{SS}(s_i, s_{-i})$ denotes the probability that supplier $i$ is selected. The probability $\alpha_i^{SS}(s_i, s_{-i})$ is nondecreasing concave in $s_i$ and is equal to zero when $s_i = 0$ for $i = 1, \ldots, N$ with $0 \leq \alpha_i^{SS}(s_i, s_{-i}) \leq 1$. Section 5 provides a discussion of how a probabilistic selection might arise and how selection probability functions might be specified.

Let $C$ denote the type of competition chosen by the buyer, with $C = SA$ or SS. The expected quality of service received by the buyer under competition of type $C$ is then

$$q^C(s) = \sum_{i=1}^{N} \alpha_i^C(s_i, s_{-i}) s_i, \quad (2)$$

where $s = (s_1, \ldots, s_N)$. The buyer chooses a structure for $\alpha^C$ to induce high quality of service by rewarding better performing suppliers with either higher market share (under SA) or a higher probability of selection (under SS). Given the buyer’s choice of $C$ and $\alpha^C$, the suppliers respond by competing against each other for the buyer’s fixed demand.

Each supplier competes by choosing a service level $s_i$ that maximizes her own expected profit, subject to the behavior of other suppliers. Under SA competition, this implies that supplier $i$ will choose $s_i$ to maximize

$$\pi_i^{SA}(s_i, s_{-i}) = \alpha_i^{SA}(s_i, s_{-i}) \lambda r - f(s_i, \alpha_i^{SA}(s_i, s_{-i}) \lambda)$$

$$= \alpha_i^{SA}(s_i, s_{-i}) \lambda [r - u(s_i)] - v(s_i), \quad (3)$$
while under SS competition, supplier $i$ will choose $s_i$ to maximize

$$
\pi_i^{SS}(s_i, s_{-i}) = \alpha_i^{SS}(s_i, S_s)(\lambda r - f(s_i, \lambda)) = \alpha_i^{SS}(s_i, S_s)\lambda[r - u(s_i)] - \alpha_i^{SS}(s_i, S_s)v(s_i).
$$

Note that in both cases, supplier $i$’s expected revenue and expected cost depend on her own service level $s_i$ as well as the service level profile $s_{-i}$ of her competitors. We assume that all parties have full access to information about each other’s costs. Also, in systems where some of the parameters are random variables, we assume all suppliers to be risk neutral and to be profit maximizers. Under both forms of competition, costs are incurred by a supplier only after demand allocations are announced by the buyer and only if the supplier receives a positive portion of demand. In the SS case, only the winner will incur costs, with non-winners walking away with no cost (and no revenue). The case where some cost may be incurred prior to demand allocation is discussed at the end of this section.

It is difficult to show the existence and uniqueness of a Nash equilibrium without further specifying the allocation functions (we use the term “allocation function” in the rest of the paper to refer to both $\alpha^{SA}$ and $\alpha^{SS}$, even though $\alpha^{SS}$ is a selection rather than an allocation function). Therefore, in our analysis we focus on a particular class of allocation functions of the form

$$
\alpha^c_i(s_i, s_{-i}) = \frac{g(s_i)}{\sum_{j \neq i} g(s_j) + g(s_i)},
$$

where $g(s_i)$ is a nondecreasing concave function of $s_i$ with $g(0) = 0$ for $i = 1, \ldots, N$ and $g$ is twice differentiable. In its simplest form, the function $g(s_i)$ could represent the service level a supplier chooses to offer, i.e., $g(s_i) = s_i$. This leads to a service-proportional allocation function of the form $\alpha^c_i(s_i, s_{-i}) = s_i / \sum_{j=1}^N s_j$. A more general proportional allocation may take the form of a market attraction function $\alpha^c_i(s_i, s_{-i}) = s_i / \sum_{j=1}^n s_j$ where $0 < \gamma_i \leq 1$. We choose to focus on proportional allocation functions because of their simplicity, mathematical tractability, and their wide use in the literature. Proportional allocation functions arise naturally in some cases through buyer decision processes (see §5 for an example). Although we do not pursue it in this paper, we expect that many of our results would continue to hold for other allocation functions including nonconcave functions (see Cachon and Zhang 2006a for examples of other allocation functions). In §4, we show how the analysis can be extended in some cases to proportional but nonconcave allocation functions.

In the following theorem, we show that the supplier competition defined by either SA or SS admits a unique symmetric Nash equilibrium. The appendix provides a proof for this and all subsequent results.

**Theorem 1.** A Nash equilibrium for SA and SS competitions exists and is unique with equilibrium service levels $s^c_i = s^c$ where $s^c > 0$ for $i = 1, \ldots, N$ and $C = SA, SS$.

Although we restrict our discussion in this paper to symmetric suppliers, it can be shown that a Nash equilibrium continues to exist for nonidentical suppliers and for more general service cost and allocation functions (see Elahi 2007 for details).

To gain some insight into the differences in the equilibrium service levels obtained under SA and SS, it is useful to examine the suppliers’ profit functions (Equations (3) and (4)) more closely. First, note that in the absence of a demand-independent cost component, i.e., when $v(s_i) = 0$ for all $s_i$, SA and SS have the same profit structure. Consequently, if $\alpha^{SA}(s_i, s_{-i}) = \alpha^{SS}(s_i, s_{-i})$, both SA and SS lead suppliers to choose the same service levels and consequently lead to the same expected quality of service for the buyer. In contrast, when $v(s_i) > 0$, SA and SS behave quite differently. For SA, the fraction of demand allocated to each supplier decreases with $N$, which reduces the expected revenue each supplier receives. Although the demand-dependent cost also diminishes, the demand-independent cost is unaffected. Consequently, depending on the relative strength of $u$ and $v$, the incentive for a supplier to offer a high service level could diminish with $N$. In contrast, under SS, a selected supplier is awarded the entire demand and incurs the demand-independent cost only after the supplier is indeed selected. Although the expected revenue diminishes with $N$, so does the expected cost. Hence, an increase in $N$ could in fact intensify competition, forcing suppliers to increase their service levels.

Theorems 2 and 3 confirm this intuition and offer further comparisons between SA and SS competitions.

**Theorem 2.** The following holds for all suppliers $i = 1, \ldots, N$:

1. If $v(s_i) = 0$, then $s^{SA} = s^{SS} = s^c$ and $q^{SA} = q^{SS}$.
2. If $v(s_i) > 0$ for $s_i$, then $s^{SA} < s^{SS}$ and $q^{SA} < q^{SS}$.

Furthermore, if $v(s_i) > 0$, $g(s_i) = s^c_i$, where $0 < \gamma_i \leq 1$ and both $u$ and $v$ are linear in $s_i$, then

3. $\sum_{i=1}^N f(s_i^{SA}, s_{-i}) = Nf(s^c) > f(s^c, s^c)$ and

4. $\pi_i^{SA} = \pi_i^{SS} = \pi^c$, where $\pi_i^c$ refers to the equilibrium expected profit for $C = SA, SS$.

Theorem 2 implies that for a given allocation function, SA and SS are equivalent when there are no demand-independent service costs. However, if there are demand-independent costs, the service levels offered by the suppliers under SS are higher than
those offered under SA. Consequently the average quality of service received by the buyer is also higher under SS. Interestingly, the cumulative cost incurred by all the suppliers under SA (which is indicative of the cumulative effort being exerted by the suppliers) can actually be higher than the cost incurred by the single selected supplier under SS. In other words, under SA competition, the buyer is able to get suppliers to invest a greater proportion of their revenues into effort, which explains the lower supplier profits under SA. Hence, somewhat paradoxically, although the suppliers cumulatively spend more on service under SA competition, both buyers and suppliers are worse off.

Theorem 3 describes the impact of the number of participating suppliers (N) on service level, quality of service, and supplier profits.

THEOREM 3. The following holds for all suppliers \( i = 1, \ldots, N \):

1. For SS, \( s^{SS} \) and \( q^{SS} \) are increasing in \( N \) with \( s^{SS} \rightarrow s^{SS} \) and \( q^{SS} \rightarrow q^{SS} \) as \( N \rightarrow \infty \), where \( s^{SS} \) and \( q^{SS} \) are positive values.

2. For SA, we distinguish three cases:
   (a) if \( v(s_i) = 0 \) for \( s_i > 0 \), then \( s^{SA} \) and \( q^{SA} \) are increasing in \( N \) with \( s^{SA} \rightarrow s^{SA} \) and \( q^{SA} \rightarrow q^{SA} \) as \( N \rightarrow \infty \), where \( s^{SA} \) and \( q^{SA} \) are positive values.
   (b) if \( u(s_i) = 0 \) for \( s_i > 0 \), then \( s^{SA} \) and \( q^{SA} \) are decreasing in \( N \) with \( s^{SA} \rightarrow 0 \) and \( q^{SA} \rightarrow 0 \) as \( N \rightarrow \infty \), and
   (c) if \( u(s_i) = 0 \) for \( s_i > 0 \), then \( s^{SA} \) and \( q^{SA} \) are decreasing in \( N \) with \( s^{SA} \rightarrow 0 \) and \( q^{SA} \rightarrow 0 \) as \( N \rightarrow \infty \).

3. \( \pi^C \) is decreasing in \( N \) with \( \pi^C \rightarrow 0 \) as \( N \rightarrow \infty \) for \( i = 1, \ldots, N \) and \( C = SA, SS \).

Theorem 3 shows that the effect of increased competition can be different under the SA and SS schemes and is sensitive to the form of the cost function. Under SS competition, larger \( N \) always leads to higher service levels. Here, the buyer favors having a large number of suppliers participate in the selection process. This also holds true for SA when there is no demand-independent service cost. However, when service cost only contains a demand-independent component, SA yields the opposite effect with larger \( N \) always leading to lower service levels. When both types of costs exist with SA, the effect of \( N \) is generally not monotonic. An increase in \( N \) can lead to an initial increase in service levels, but further increases in \( N \) eventually lead to a decrease in service levels, with service levels approaching zero in the limit case.

Supplier profits under both SA and SS decrease in \( N \) regardless of the cost function and vanish in the limiting case of perfect competition (i.e., \( N \rightarrow \infty \)). However subtle differences in supplier profits exist. While under SA, the actual profit of each supplier approaches zero as \( N \) becomes large; only expected supplier profit may under SS competition (expected supplier profit approaches zero because the probability of being selected approaches zero). The actual profit of a selected supplier (i.e., a supplier’s profit given that the supplier is selected) can be strictly positive. This means that the buyer may not be able, even under perfect competition, to extract all the post-selection profit from the selected supplier.

So far in our analysis, we have assumed that all costs are incurred by the suppliers once the allocations are made. However, in some applications, some demand-independent costs could occur before the allocations are announced. For example, to qualify as potential suppliers, the buyer might require some initial investment from the suppliers for them to qualify for the competition. The timing of when these demand-independent costs occur does not impact the structure of the supplier profit function under SA, but it does change the supplier profit function under SS. For example, if supplier \( i \) incurs the entire demand-independent component \( v(s_i) \) before supplier selection takes place, expected profit for supplier \( i \) becomes \( \pi_i^{SS}(s_i, s_{-i}) = \alpha_i^{SS}(s_i, s_{-i}) \lambda[r - u(s_i)] - v(s_i) \). This function is identical to the profit function under SA. Consequently, if the same allocation function is used for both SA and SS, both forms of competition are equivalent and yield the same expected service level. SS would retain some advantage over SA if suppliers incur only a portion of the demand-independent cost prior to supplier selection, with this advantage diminishing as the pre-selection portion increases. This insight also implies that the comparison results between SA and SS described in this paper could be recast as a comparison between two forms of SS competition, one with demand-independent costs incurred prior to supplier selection and one with demand-independent costs incurred post selection.

4. The Effect of Allocation Functions

In this section, we explore how the form of the allocation function impacts the competition outcomes. In general, the Nash equilibrium service levels are sensitive to the functional form of the allocation function. That is, different allocation functions can induce different service levels. This can be verified, for example, by observing that the Nash equilibrium service levels are solutions to the following sets of equations:

\[
\frac{\partial \pi_i^{SA}}{\partial s_i} = \frac{\partial \pi_i^{SA}(s_i, s_{-i})}{\partial s_i} \lambda r - \frac{\partial f(s_i, s_{-i})}{\partial s_i} = 0, \quad \text{and (6)} \\
\frac{\partial \pi_i^{SS}}{\partial s_i} = \frac{\partial \pi_i^{SS}(s_i, s_{-i})}{\partial s_i} \lambda r - \frac{\partial f(s_i, s_{-i})}{\partial s_i} - \alpha_i^{SS}(s_i, s_{-i}) \frac{\partial f(s_i, s_{-i})}{\partial s_i} = 0 \quad \text{(7)}
\]
for \( i = 1, \ldots, N \). The solutions appear to depend on \( \partial \alpha_i^C/\partial s_i \) \((C=SA, SS)\), the rate at which market share increases with increases in \( s_i \). Intuitively, we expect that if the rate \( \partial \alpha_i^C/\partial s_i \) decreases slowly (recall that \( \alpha_i^C \) is concave), then the Nash equilibrium would occur at higher values of service than if \( \partial \alpha_i^C/\partial s_i \) decreased abruptly. In other words, the Nash equilibrium service levels appear to depend on the second derivative of the allocation function, which can be viewed as a measure of the intensity of the competition. This is easily verified for proportional allocation functions of the form \( \alpha_i^C(s_i, s_{-i}) = s_i^\gamma/\sum_{i=1}^N s_i^\gamma \), where \( 0 \leq \gamma \leq 1 \). Here, \( \partial^2 \alpha_i^C/\partial s_i^2 \) is decreasing in \( \gamma \), \( s_i^SS \) and \( s_i^SS \) are increasing in \( \gamma \), with \( \gamma = 1 \) maximizing service for the buyer.

To apply Theorem 1, we require that \( \gamma \leq 1 \) so that the allocation function is concave. However, concavity is sufficient but not a necessary condition. This leads to the question as to what would happen if we allowed \( \gamma \) to be greater than one. Would we still have an equilibrium and would it lead to an even higher service level for the buyer? If so, could the buyer choose a high enough \( \gamma \) to force suppliers to offer the maximum feasible service level and realize zero profits? In the following propositions, we examine the special case of linear cost functions and show that a Nash equilibrium may exist for \( \gamma > 1 \) and that a buyer can indeed induce suppliers in some cases to provide the maximum feasible service level.

**Proposition 1.** Under SS competition with \( \alpha_i^SS(s_i, s_{-i}) = s_i^\gamma/\sum_{i=1}^N s_i^\gamma \), \( u(s_i) = k_1 s_i \), and \( v(s_i) = k_2 s_i \), a unique Nash equilibrium exists for any \( \gamma > 0 \). The equilibrium service levels are increasing in \( \gamma \) while expected supplier profit is decreasing in \( \gamma \) with \( \lim_{\gamma \to \infty} s_i^SS = \lambda r/(\lambda k_1 + k_2) \) and \( \lim_{\gamma \to \infty} s_i^SS = 0 \).

**Proposition 2.** Under SA competition with \( \alpha_i^SA(s_i, s_{-i}) = s_i^\gamma/\sum_{i=1}^N s_i^\gamma \), a symmetric Nash equilibrium exists for \( \gamma > 1 \) if one of the following conditions holds.

1. \( u(s_i) = k_1 s_i \), \( v(s_i) = k_2 s_i \), and \( \gamma \leq \gamma_{\text{max}} = N/(N-1) \); the corresponding equilibrium service levels are increasing in \( \gamma \), with \( s_i^SA = \lambda(r - k_1)/Nk_2 \) when \( \gamma = \gamma_{\text{max}} \), while the corresponding expected supplier profits are decreasing in \( \gamma \) with \( \pi_i^SA = 0 \) when \( \gamma = \gamma_{\text{max}} \).

2. \( u(s_i) = k_1 s_i \) and \( v(s_i) = 0 \); the corresponding equilibrium service levels are increasing in \( \gamma \), with \( \lim_{\gamma \to \infty} s_i^SA = r/k_1 \) and \( \lim_{\gamma \to \infty} \pi_i^SA = 0 \).

In the case of (1), the symmetric Nash equilibrium is the unique equilibrium if \( N = 2 \). In the case of (2), the symmetric Nash equilibrium is always the unique equilibrium.

These observations highlight the important role allocation functions play in determining the level of service suppliers provide. Using a service proportional allocation rule, a buyer may be able to extract all the profit from the suppliers and induce them to provide the maximum feasible service level. These results appear consistent with those in Cachon and Zhang (2006a) who consider an application with competition similar to SA with \( u(s_i) = k_1 \) and \( v(s_i) = k_2 s_i \). It is interesting to note that for SA the maximum feasible service level under condition (1) is decreasing in \( N \) while for SS it is always independent of \( N \). This means that for SA, the buyer can maximize her expected service levels by setting \( N = 2 \) and choosing \( \gamma = 2 \), which leads to \( s_i^SA = \lambda(r - k_1)/2k_2 \). For SS, the maximum feasible service level \( s_i^SS = \lambda r/(\lambda k_1 + k_2) \) is achievable with any \( N \) by letting \( \gamma \to \infty \). The latter is not surprising. When \( \gamma \to \infty \), SS becomes equivalent to an auction where the supplier with the highest service level is selected with probability one.

The above analysis raises the question as to whether it is possible for every service level achievable with SS to choose an allocation function that makes that service level achievable with SA. In other words, is it possible for the buyer to specify a service level and then choose allocation functions, with the one for SA possibly different from the one for SS, to obtain the specified service level from either type of competition? It turns out that this is not always possible. For example, under demand-independent service costs, the maximum feasible service level under SS is always strictly greater than the one achieved under SA. Therefore, there may be a range of service levels (between the maximum feasible service level for SA and the maximum feasible service level for SS) achievable by SS but not by SA regardless of what allocation function is used for SA.

We end this section by discussing a useful reformulation that allows us in certain cases to extend results to more general service cost functions with the only requirement that the function is increasing in service level. For these cases, we show that it is always possible for the buyer to orchestrate a competition that produces a Nash equilibrium, maximizes service level, and results in zero expected supplier profits. Consider first SS competition. Recognizing that as long as \( f(s_i, \lambda) \) is increasing in \( s_i \), there is a one-to-one correspondence between the service level \( s_i \) and \( f(s_i, \lambda) \), and suppliers can be viewed as competing on cost expenditures. We refer to supplier cost expenditures as **effort** and denote it by \( e_i \), where \( e_i \equiv f(s_i, \lambda) \). The buyer could reformulate the competition, including the allocation function, in terms of this effort. This leads to expected supplier profit functions given by \( \pi_i^SS(e_i, e_{-i}) = \alpha_i^SS(e_i, e_{-i})(\lambda r - e_i) \) for \( i = 1, \ldots, N \). If the buyer uses a proportional allocation function \( \alpha_i^SS(e_i, e_{-i}) = e_i/\sum_{i=1}^N e_i^\gamma \), then the buyer could induce the suppliers to exert maximum feasible effort by letting \( \gamma \to \infty \). This maximum feasible effort is given by \( e_i^SS = \lambda r \) and the corresponding maximum feasible service level \( s_i^SS \) is the unique solution to \( f(s_i, \lambda) = \lambda r \).
For SA, a similar reformulation is not always possible because there is not a one-to-one correspondence between service level and cost expenditures. Cost expenditures depend on both the service level and the amount of demand allocated. However, a reformulation is feasible for the following two important cases (see §6 for example applications): (1) \( f(s_i, \alpha(s_i, s_{-i}) \lambda) = \alpha(s_i, s_{-i}) \lambda k_1 + v(s_i) \), and (2) \( f(s_i, \alpha(s_i, s_{-i}) \lambda) = \alpha(s_i, s_{-i}) \lambda u(s_i) \), where the only requirement on \( u \) and \( v \) is that they are increasing in \( s_i \). For case (1), when \( v(s_i) \) is increasing in \( s_i \), there is a one-to-one correspondence between the service level \( s_i \) and the demand-independent cost \( v(s_i) \), and so the buyer could reformulate the supplier competition in terms of these expenditures. Letting \( \epsilon_i \equiv v(s_i) \), expected supplier profits can be rewritten as \( \pi^{SA}_i(e_i, \epsilon_{-i}) = \alpha^{SA}_i(e_i, \epsilon_{-i}) \lambda (r - k_i) - \epsilon_i \). If the buyer uses again an allocation function of the form \( \alpha^{SA}_i = \epsilon_i^\gamma / \sum_{i=1}^{N} \epsilon_i^\gamma \), he would maximize service quality by choosing \( \gamma = 2 \) and \( N = 2 \), which leads to the maximum feasible demand-independent cost expenditure \( e^{SA} = \lambda (r - k_i) / 2 \) and corresponding service level \( s^{SA} \) given by the unique solution to \( v(s) = \lambda (r - k_i) / 2 \). A similar treatment can be carried out for case 2. Letting \( \epsilon_i \equiv u(s_i) \), the buyer would induce the suppliers to exert maximum feasible effort \( e^{SA} = r \) by letting \( \gamma \to \infty \), with the corresponding service level being the unique solution to \( u(s) = r \).

Finally, we should note that staging a competition in terms of effort can lead to different equilibrium service levels than a competition based on service levels, depending on the cost and allocation functions. The main advantage of an effort-based competition is that simple allocation functions can be designed to induce suppliers to provide maximum service level. However, clearly the usefulness of the transformation depends on whether or not effort is observable (see §6 for examples where this might be plausible).

5. A Model for Supplier Selection

Choosing an allocation function and announcing it to the suppliers is straightforward under SA. An allocation function in this case is a verifiable formula for how demand is allocated once service levels are announced. For SS, specifying an allocation function (which corresponds to a selection probability) is less obvious. Typically, a selection probability is implied by the buyer’s past behavior in choosing suppliers. The selection probability is often learned by the suppliers through repeated interactions between buyer and suppliers, rather than being explicitly announced by the buyer. In this section, we describe an example supplier selection process through which a probabilistic selection naturally arises. We show how under some conditions the resulting selection probability fits our assumptions.

Consider a setting where suppliers announce their service levels and the buyer responds by assigning each supplier a score \( w_i(s_i) = g(s_i) + \epsilon_i \), where \( \epsilon_i \) is a random variable denoting an error term with mean zero and standard deviation \( \sigma \). The random variables \( \epsilon_i \) are independent and identically distributed (i.i.d.). The functional form of \( g(s_i) \) is announced by the buyer to the suppliers before they commit to \( s_i \). However, the value of \( \epsilon_i \) is revealed only after the suppliers announce their service levels. The buyer then chooses the supplier with the highest score \( w_i(s_i) \). The term \( \epsilon_i \) reflects inherent and unbiased randomness in the selection process. For example, it could denote the outcome of an opinion poll of the buyer’s purchasing managers or the outcome of an audit of the suppliers after the service levels have been announced. Alternatively, it could result from a multiplicity of decision makers at the buyer’s firm (Ha 2004). The variance of \( \epsilon_i \) reflects the amount of uncertainty associated with the selection process. When variance is low, the outcome of the selection procedure is primarily determined by the service level \( s_i \). When variance is high, the outcome of the selection is mostly random and service level is not the main determinant of the selection decision.

The probability that supplier \( i \) is selected can now be stated as

\[
\alpha^{SS}_i(s_i, s_{-i}) = \prod_{j \neq i} \Pr[g(s_j) + \epsilon_j \geq g(s_i) + \epsilon_i], \quad (8)
\]

or equivalently,

\[
\alpha^{SS}_i(s_i, s_{-i}) = \prod_{j \neq i} \Pr[\epsilon_j \leq \epsilon_i - g(s_j) - g(s_i)]
= \prod_{j \neq i} F_{\epsilon_j}[g(s_i) - g(s_j)]. \quad (9)
\]

where \( F_{\epsilon_j} \) is the distribution of the difference \( \zeta = \epsilon_j - \epsilon_i \). Obtaining closed-form expressions for \( \alpha^{SS}_i \) is difficult in general. However, in the case where the \( \epsilon_j \)'s are Gumbel distributed random variables (i.e., \( F_{\epsilon_j}(x) = e^{-e^{-\theta x}} \)), where \( \theta > 0 \) is a scale parameter such that \( \text{Var}(\epsilon_j) = \mu^2 \pi^2 / 6 \) and \( \kappa = 0.5772 \ldots \) is Euler’s constant, we have (see, for example, Talluri and van Ryzin 2004, Chapter 7)

\[
\alpha^{SS}_i(s_i, s_{-i}) = \frac{e^{g(s_i)/\mu}}{\sum_{i=1}^{N} e^{g(s_i)/\mu}}. \quad (10)
\]

As we can see, the selection probability has the form of a proportional allocation function. Furthermore, by choosing \( g(s_i) = \mu \ln[h(s_i)] \), the buyer can reduce it to

\[
\alpha^{SS}_i(s_i, s_{-i}) = \frac{h(s_i)}{\sum_{i=1}^{N} h(s_i)}. \quad (11)
\]

Therefore, all the analysis and results of the previous sections apply.
The above supplier selection process is one example of how a probabilistic selection might arise. In practice, it is not uncommon for some uncertainty to surround the supplier selection process whenever sole sourcing is involved, even when the declared primary selection criterion is service level. Sole sourcing poses greater risks to the buyer and the final selection typically involves deliberations whose outcome can be uncertain. Of course, it is possible to consider settings where decisions are based only on service level. This corresponds in our model to the case where \( \epsilon_i \to 0 \).

6. Example Applications

In this section, we illustrate the general framework and results of the previous sections with three example applications. The first example views suppliers as make-to-order service providers who influence service through capacity investments. The second example views suppliers as make-to-stock manufacturers having fixed utilization targets who influence service levels through inventory investments. The third example views suppliers as newsvendors who make a single-period decision about capacity which then determines service levels. The examples illustrate different types of service levels, different forms of effort, and different cost functions.

6.1. Competition with Make-to-Order Suppliers

Consider a system of \( N \) potential suppliers who operate in a make-to-order fashion, provisioning services in response to real-time requests. A buyer, in outsourcing her service requests to this supply pool, is interested in inducing high time-based service performance, using measures such as expected fulfillment time of requests or the probability of fulfilling requests within a quoted lead time. Because time-based performance is driven primarily by the capacity of the suppliers, we assume that suppliers commit to investing in capacity sufficient to meet the service levels they promise to offer. Hence, the service level costs incurred by the supplier are capacity investment costs.

We assume that service requests from the buyer to the suppliers occur continuously over time according to a renewal process with rate \( \lambda \). Each potential supplier \( i \) can be viewed as a service facility with service rate \( \mu_i \) and i.i.d. service times with mean \( 1/\mu_i \), for \( i = 1, \ldots, N \). Under SA competition, demand is partitioned among the \( N \) potential suppliers with supplier \( i \) receiving a long run fraction \( \alpha_i^{SA}(s_i, s_{-i}) \) of total demand. Hence, the demand rate that supplier \( i \) sees is \( \lambda_i = \alpha_i^{SA}(s_i, s_{-i}) \lambda \). Because service requests arrive dynamically over time, \( \alpha_i^{SA} \) in fact specifies the probability that an incoming service request is assigned to supplier \( i \). Although a truly probabilistic allocation is unlikely in practice, it is useful in approximating the behavior of a central dispatcher that attempts to adhere to a specified allocation for each supplier. It is also useful in modeling settings where demand arises from a sufficiently large number of sources. The parameter \( \alpha_i^{SA}(s_i, s_{-i}) \) would then correspond to the fraction of demand sources (e.g., geographical locations) for type \( i \) that is always satisfied by supplier \( i \).

Service level in a service system can be defined in a variety of ways. For the purpose of illustration, we consider the probability of fulfilling a service request within a quoted lead time. For ease of exposition and to allow for closed-form expressions for service levels, we assume that demand occurs according to a Poisson process and service times are exponentially distributed. This is consistent with assumptions in Cachon and Zhang (2006a), Gilbert and Weng (1997), and Kalai et al. (1992). Because the probabilistic splitting of a Poisson process is itself Poisson, the demand process each supplier sees is also Poisson. Thus, each supplier behaves like an M/M/1 queue. Given these assumptions, the probability of meeting a quoted lead time \( \tau \) is

\[
\Pr(W_i \leq \tau) = 1 - e^{-(\mu_i - \alpha_i \lambda) \tau},
\]

where \( W_i \) is a random variable denoting fulfillment time by supplier \( i \) given service rate \( \mu_i \).

Under SA competition, if a supplier offers service level \( s_i = \Pr(W_i \leq \tau) \), then she commits to acquiring an amount of capacity (in the form of a service rate) equal to \( \alpha_i^{SA}(s_i, s_{-i}) \lambda + \ln[1/(1-s_i)]/\tau \). The fraction of demand \( \alpha_i^{SA}(s_i, s_{-i}) \) allocated to supplier \( i \) is increasing in \( s_i \) with \( \sum_{i=1}^{N} \alpha_i^{SA}(s_i, s_{-i}) \leq 1 \). Under SS competition, supplier \( i \) commits to acquiring an amount of capacity equal to \( \lambda + \ln[1/(1-s_i)]/\tau \) if she wins the business, with the probability \( \alpha_i^{SS}(s_i, s_{-i}) \) that supplier \( i \) is selected as the sole service provider increasing in \( s_i \) and \( \sum_{i=1}^{N} \alpha_i^{SS}(s_i, s_{-i}) \leq 1 \). Each supplier incurs a variable production cost \( c \) per unit produced and an amortized capacity cost of \( k \) per unit of service rate (the treatment can be extended to a general increasing capacity cost function). Expected supplier profit can then be written as

\[
\pi_i^{SA}(s_i, s_{-i}) = \alpha_i^{SA}(s_i, s_{-i}) \lambda [(p - c) - k] - k \ln[1/(1-s_i)]/\tau
\]

(12)

and

\[
\pi_i^{SS}(s_i, s_{-i}) = \alpha_i^{SS}(s_i, s_{-i}) \lambda [(p - c) - k] - k \ln[1/(1-s_i)]/\tau.
\]

(13)

Letting \( u(s_i) = k \) and \( v(s_i) = k \ln[1/(1-s_i)]/\tau \) and noting that \( v(s_i) \) is increasing convex in \( s_i \), we can see that the profit functions have the same form as (3) and (4). Therefore, all the associated results of §§3 and 4 immediately apply.
The example illustrates a case where the demand-dependent cost is linear in the allocated demand but independent of the service level. Hence, the service level is solely determined by the demand-independent cost. Because costs correspond to capacity investment levels, this means that the total capacity invested by a supplier is always equal to the amount of demand allocated (the minimum capacity needed to guarantee finite fulfillment time) plus a fixed amount that depends only on service level. Comparing the resulting capacity utilizations, we can see that under SA, supplier \( i \) has an average utilization
\[
\rho_i^{SA}(s_i, s_{-i}) = \frac{\alpha_i^{SA}(s_i, s_{-i}) \lambda}{(s_i - s_i) \lambda + \ln[1/(1 - s_i)]} \tau
\]
while under SS,
\[
\rho_i^{SS}(s_i, s_{-i}) = \frac{\lambda}{(s_i - s_i) \lambda + \ln[1/(1 - s_i)]} \tau.
\]
It is not difficult to verify that \( \rho_i^{SS}(s_i, s_{-i}) \geq \rho_i^{SA}(s_i, s_{-i}) \). This implies that under SS, the supplier is able to maintain a higher utilization (i.e., invest in less capacity relative to the allocated demand) than SA while providing the same service level to the buyer. This is consistent with results about the benefit of pooling in queueing systems where it is known that less capacity is needed to meet a target service level in a system with a single server and a single queue than in a system with multiple servers and independent queues (see, for example, Benjaafar et al. 2005).

As described in §4, the buyer could reformulate the competition in terms of the demand-independent costs or, equivalently, the extra capacity beyond the minimum required. The results of §4 could then be used to show that if the buyer chooses a proportional allocation function \( \alpha_i^C = c_i^C / \sum_{i=1}^N c_i^C \), where \( c_i = v(s_i) \), he would be able to maximize his expected service quality by choosing \( \gamma = 2 \) and \( N = 2 \) under SA and by letting \( \gamma \to \infty \) under SS. The corresponding maximum feasible service levels would be given by \( S^{SA} = 1 - e^{-\lambda \tau}/2k \) and \( S^{SS} = 1 - e^{-\lambda \tau}/k \).

### 6.2. Competition with Make-to-Stock Suppliers

Consider a buyer who seeks to outsource the manufacturing of a physical good among a set of \( N \) potential suppliers. The problem is similar to the one described in the previous section, except that now suppliers are able to produce goods ahead of demand in a make-to-stock fashion. By holding finished goods inventory, each supplier is able to improve the quality of service (in terms of order fulfillment delay) she offers the buyer.

As in the previous example, we assume that the buyer faces demand that takes place continuously over time. We assume again that this demand forms a renewal process with rate \( \lambda \) and is allocated to suppliers according to either the SA or SS competition scheme. Each supplier has a finite production rate \( \mu_i \). Each supplier may hold a buffer of finished goods inventory. If so, we assume that this buffer is managed according to a base-stock policy with base-stock level \( b_i \). This means that a replenishment order is placed with the production system each time inventory drops below the base-stock level. This also means that each demand arrival triggers a replenishment order.

We assume that a supplier scales her capacity proportionally to the demand she receives so that she always maintains a fixed target utilization level \( \rho = \lambda_i / \mu_i \). That is, supplier \( i \) sets her capacity to \( \mu_i = \alpha_i \lambda_i / \rho_i \). This assumption, in addition to being plausible in many settings, allows us to focus on only one form of effort—inventory level. Various measures of service quality could be used. For illustration, we consider fill rate, the probability of fulfilling an order from on-hand inventory, which is a commonly used measure of service level in inventory management.

There are three types of cost incurred by a supplier: (1) a unit variable production cost \( c \), (2) an amortized capacity cost \( k \) per unit of capacity per unit time, and (3) a holding cost \( h \) per unit of finished goods inventory held per unit of time. Expected supplier profit can be written as
\[
\pi_i^{SA}(s_i, s_{-i}) = \alpha_i^{SA}(s_i, s_{-i}) \lambda[p - c - k/\rho] - hE[I_i(s_i)]
\]
and
\[
\pi_i^{SS}(s_i, s_{-i}) = \alpha_i^{SS}(s_i, s_{-i}) [\lambda(p - c - k/\rho) - hE[I_i(s_i)]],
\]
where \( E[I_i(s_i)] \) denotes expected inventory level for supplier \( i \) given a choice of service level \( s_i \). If we assume that the demand occurs according to a Poisson process and production times are i.i.d. and exponentially distributed, then given a base-stock level \( b_i \), the fill rate of supplier \( i \) is given by (see, for example, Buzacott and Shanthikumar 1993)
\[
Pr[I_i > 0] = 1 - \rho^b.
\]
If supplier \( i \) commits to service level (fill rate) \( s_i \), supplier \( i \) commits to choosing base-stock level \( b_i(s_i) = \ln(1 - s_i)/\ln(\rho) \). If we treat base-stock levels as continuous, a common assumption in inventory theory (see Zipkin 2000, Benjaafar et al. 2004, and Buzacott and Shanthikumar 1993), then expected inventory can be obtained as (Buzacott and Shanthikumar 1993)
\[
E[I_i(s_i)] = b_i(s_i) - \frac{\rho}{1 - \rho} (1 - \rho^{b_i(s_i)}),
\]
or equivalently, 
\[ E[I_i(s_i)] = \frac{\ln(1 - s_i)}{\ln(p)} - \frac{\rho}{1 - \rho} s_i, \]
which is increasing convex in \( s_i \). Expected supplier profits can be rewritten as
\[ \pi_i^{SA}(s_i, s_{-i}) = \alpha_i^{SA}(s_i, s_{-i}) \left[ p - c - \frac{k}{\rho} \right] - h_i \left( \frac{\ln(1 - s_i)}{\ln(p)} - \frac{\rho}{1 - \rho} s_i \right) \]
and
\[ \pi_i^{SS}(s_i, s_{-i}) = \alpha_i^{SS}(s_i, s_{-i}) \left\{ \lambda \left( p - c - \frac{k}{\rho} \right) - h_i \left( \frac{\ln(1 - s_i)}{\ln(p)} - \frac{\rho}{1 - \rho} s_i \right) \right\}. \]

Letting \( u(s_i) = k/\rho \) and \( v(s_i) = h(\ln(1 - s_i)/\ln(p) - \rho/(1 - \rho)s_i) \) and noting that \( v(s_i) \) is increasing convex in \( s_i \), we can see that the profit functions have the same form as in the previous example of §6.1. Hence, similar analysis and insights apply. In particular, because there is a one-to-one correspondence between service levels and the demand-independent inventory costs \( v(s_i) \), the buyer could reformulate the competition in terms of these cost expenditures. If the buyer then chooses a proportional allocation function \( \alpha^e_i = e_i^r \sum_{i=1}^N e_i^r \), where \( e_i \equiv v(s_i) \), she would be able to maximize her expected service quality by setting \( \gamma = 2 \) and \( N = 2 \) under SA and by letting \( \gamma \to \infty \) under SS. The corresponding maximum feasible service levels would be given by the unique solution to \( h(\ln(1 - s_i)/\ln(p) - \rho/(1 - \rho)s_i) = \lambda r/2k \) under SA and by the unique solution to \( h(\ln(1 - s_i)/\ln(p) - \rho/(1 - \rho)s_i) = \lambda r/k \) under SS.

6.3. Competition with Newsvendor Suppliers

Now consider a setting with a single period. The demand of the buyer during this period is stochastic and is described by a random variable \( D \) with distribution \( F_D \). The buyer wishes to outsource the fulfillment of this demand to one or more outside suppliers from a set of \( N \) potential suppliers. The selection of suppliers and allocation of demand among suppliers takes place prior to demand realization. Prior to demand realization suppliers also choose capacity levels (e.g., the quantity of the product to produce or purchase). Once demand is realized, these capacity levels determine how much of the demand is satisfied using existing capacity. The buyer is interested in inducing the suppliers to invest in as much capacity as possible so as to maximize the service level she receives from these suppliers, as measured by the probability that all demand allocated to a supplier is fulfilled immediately from available capacity.

Prior to demand realization, the buyer shares information about the demand with the suppliers, specified in the form of the distribution \( F_D \). The buyer also announces the allocation function. The suppliers respond by specifying the service levels they each commit to offer. In turn, the buyer responds with a specification of the fraction of the demand that will be allocated to each supplier after demand is realized, consistent with the announced allocation function. Once the suppliers know the fraction of demand they will be allocated, they each proceed with acquiring the capacity necessary to provide the promised service level. When demand is realized, the actual allocations are carried out. Depending on the realized demand, a supplier may not be able to fulfill all of her allocated demand. This excess demand is backlogged and fulfilled once the supplier is able to acquire the additional capacity (e.g., produce additional units). Depending on the realized demand, a supplier may also be left with excess capacity whose salvage price we assume is zero. Although all the demand that is allocated to a supplier is eventually satisfied, the buyer is interested in minimizing the delays that result from backlogging. Therefore, the buyer is interested in inducing the suppliers to invest in as much initial capacity as possible. In contrast, suppliers are interested in minimizing their risk and would prefer to invest as little initial capacity as possible.

By committing to a service level \( s_i \), supplier \( i \) commits to investing in capacity \( q_i \) such that
\[ \Pr(\alpha_i^{SA}(s_i, s_{-i})D \leq q_i) = s_i \]
under SA and \( \Pr(D \leq q_i) = s_i \) under SS (in fact, \( q_i \) could be substituted for service level in this case because there is a one-to-one correspondence between \( q_i \) and \( s_i \)). Let \( c \) denote the cost incurred by a supplier for one unit of capacity and \( psi \) the price paid by the supplier for each unit of allocated demand. Also, let \( \theta(s_i) \) be the unique solution to \( F_D(x) = s_i \). Expected supplier profit can then be written as
\[ \pi_i^{SA}(s_i, s_{-i}) = E[p\alpha_i^{SA}(s_i, s_{-i})D - c \max(\alpha_i^{SA}(s_i, s_{-i})\theta(s_i), \alpha_i^{SA}(s_i, s_{-i})D)] \]
and
\[ \pi_i^{SS}(s_i, s_{-i}) = \alpha_i^{SS}(s_i, s_{-i})E[pD - c \max(\theta(s_i), D)], \] (17)
and
\[ \pi_i^{SS}(s_i, s_{-i}) = \alpha_i^{SS}(s_i, s_{-i})E[pD - c \max(\theta(s_i), D)], \] (18)
Expressions (17) and (18) can be rewritten as
\[ \pi_i^{c}(s_i, s_{-i}) = \alpha_i^{c}(s_i, s_{-i})[rE(D) - u(s_i)], \] (19)
where \( C = SA, SS, r = p - c, v(s_i) = 0, u(s_i) = cE([\theta(s_i) - D]^+), \) and \( u(s_i) \) is increasing in \( s_i \).
The above illustrates an example where there is only a demand-dependent service cost that increases linearly in the allocated demand. We know from Theorem 2 that SA and SS lead to identical service levels in this case. Because there is a one-to-one correspondence between \( u(s_i) \) and the service level \( s_i \), here too it is possible for the buyer to reformulate the competition in terms of \( u(s_i) \). The buyer would then realize the maximum feasible service level under either SA or SS by choosing a proportional allocation function \( \alpha_i^e = e_i / \sum_{i=1}^N e_i^e \), where \( e_i = u(s_i) \) and letting \( \gamma \to \infty \). The maximum service level would be given by the unique solution of

\[
E[(\theta y^e_i - D)^+] = r E(D)/c \quad \text{for} \quad \mathcal{C} = \text{SA, SS.}
\]

In addition to inducing the maximum feasible service levels, the advantage of this reformulation is that a unique equilibrium is always guaranteed regardless of the functional properties of \( u \) (e.g., concave, convex, or neither), as long as \( u \) is increasing in \( s_i \).

It may appear surprising that there is no pooling benefit associated here with SS the way there was in the previous two examples. This is because the amounts of demand allocated to the suppliers under SA are now perfectly correlated due to the proportional allocation, with all the suppliers either allocated amounts in excess of their capacity or allocated amounts less than their capacity. Therefore, consistent with known results from inventory theory (see, for example, Eppen 1979), there is no benefit to pooling under these conditions. This was not the case in the examples of §§6.1 and 6.2, where the demand streams seen by the different suppliers under SA consist of independent Poisson processes.

Finally, we note that one might have expected the buyer to apply a backorder penalty to suppliers whenever shortages occur. However, such penalties are unnecessary here. Competition alone is sufficient to guarantee that suppliers provide good service and, with the appropriate allocation function, the maximum feasible service level.

7. Conclusion

In this paper, we have studied the value of competition as a mechanism for the buyer to elicit service quality from a set of suppliers. We examined two approaches the buyer could use to orchestrate this competition. The first is a supplier-allocation (SA) approach in which a portion of demand is allocated to each supplier based on the supplier’s promised service level. The second is a supplier-selection (SS) approach where a single supplier is selected with the probability of being selected depending on the supplier’s promised service level. The analysis reveals that (a) a buyer could indeed orchestrate a competition among potential suppliers to promote service quality, (b) under identical allocation functions, the existence of a demand-independent service cost gives a distinct advantage to SS-type competitions, in terms of higher service quality for the buyer and higher expected profit for the supplier, (c) the relative advantage of SS versus SA depends on the magnitude of demand-independent versus demand-dependent service costs, (d) in the presence of a demand-independent service cost, a buyer should limit the number of competing suppliers under SA competition but impose no such limits under SS competition, and (e) a buyer can induce suppliers to provide higher service levels by selecting an appropriate allocation function and number of suppliers.

Our results suggest that, given similar allocation functions, SS competition is preferable to SA from a quality of service perspective when there is a demand-independent cost incurred after demand is allocated. Under this condition, competition that leads to single sourcing is preferable to competition that leads to multisourcing, and if multisourcing is implemented, then dual sourcing is optimal. However, these results do not take into account additional factors that may favor one form of sourcing versus the other. For example, in environments where there are few buyers (e.g., defense industries), a supplier that does not receive an allocation could go out of business. In that case, the buyer needs to support multiple suppliers to ensure continued competition in the future. A decision on the part of the buyer to single, dual, or multisource should trade off service quality benefits with these other tangible and less tangible benefits. In fact, our results are most useful for separately documenting the impact of different parameters on the performance of each type of competition. Finally, our results show that the differences between SA and SS depend on the timing of when the demand-independent costs are incurred. In particular, if the independent costs are incurred prior to the demand allocation under SA and prior to supplier selection under SS, SA and SS are equivalent in terms of the expected service level they yield to the buyer.

Our results comparing SA and SS can be recast as a comparison between two forms of SS competition—one with demand-independent costs incurred prior to demand allocation and one with demand-independent costs incurred post demand allocation. These results could be extended to examine settings where demand-independent costs are incurred in two stages: one portion occurring pre-allocation and the other post allocation. In practice, pre-allocation costs may be desirable because they could reduce the risk of a supplier reneging on the promised service levels. However, this risk-mitigation benefit needs to be balanced against the service level reduction it induces.
There are several possible avenues for future research. Our analysis currently relies on the assumption of identical service cost functions among suppliers. Dropping this assumption would allow us to consider situations where some suppliers are more cost efficient than others. The degree of cost asymmetry could affect the behavior and performance of our two types of competition, as well as the type of allocation functions that maximize service quality. Allocation functions that intensify competition could lead under SA to a small number of suppliers capturing most of the demand. With asymmetry, it is also not clear if there would always be allocation functions that lead to zero supplier profits. In highly asymmetric settings, the most cost-effective supplier could capture most of the demand while expending only a fraction of her revenue on service cost.

We have also assumed that the set of participating suppliers is exogenously determined. However, one could consider the joint decision of choosing \( M \) suppliers out of the pool of \( N \) and then allocating demand among these \( M \) winners. SA and SS are actually special cases of this more general problem, with SA implying \( M = N \) and SS implying \( M = 1 \). Under this generalized scheme, the buyer decides on both a selection function to determine the \( M \) winners and an allocation function to determine how demand is divided. This generalized form of the competition could capture benefits of both SA and SS. For example, by choosing a large value for \( N \), the buyer (with appropriate choice of selection and allocation functions) may be able to intensify the competition and extract high service levels while still maintaining multiple suppliers, which might be desirable to manage for reasons other than service quality. However, we suspect that the effect of \( M \), the number of suppliers that are eventually selected, would remain the same. In particular, smaller \( M \) leads to higher service quality with the highest service quality realized when \( M = 1 \).

In certain settings, the buyer may not be interested in maximizing the average service quality received from his suppliers. Instead, the buyer may be interested in measures of service that depend in different ways on the effort profile of the various suppliers. For example, the buyer could be interested in cumulative expenditures by the suppliers (e.g., total capacity investments in the supply chain). Alternatively, the buyer could be interested in reducing the variance of service levels across different suppliers (e.g., minimizing maximum delay over all suppliers). We expect different service measures to favor different types of competition. For example, we know that SA induces a lower average service level but a higher total cost expenditures on service. Therefore, when buyers care about cumulative expenditures, SA becomes superior and multisourcing more desirable.

Finally, our analysis could be extended to settings where the buyer may choose to outsource only a fraction of her demand under SA or reserve the right not to select any suppliers under SS. This could be implemented by the buyer by choosing, for example, an allocation function of the form \( \alpha_i^A(s_i, s_{-i}) = s_i/(\kappa + \sum_{i=1}^N s_i) \), where \( \kappa > 0 \). Under SA, the fraction \( \kappa/(\kappa + \sum_{i=1}^N s_i) \) corresponds to the fraction of demand that is not allocated to any supplier, while under SS the same fraction corresponds to the probability that no supplier is selected. The parameter \( \kappa \) could correspond to a known service level offered by an incumbent (current) supplier or to the service level realized if the buyer decides to produce in-house. We expect the threat of partial outsourcing to affect the outcome of the competition. In cases where the buyer can choose the parameter \( \kappa \), we also suspect that there may be values of this parameter that maximize the service levels received from the suppliers.

**Acknowledgments**

The authors are grateful to Paul Zipkin, the associate editor, and two anonymous reviewers for many helpful comments and suggestions. They also thank seminar participants at the following universities: Chicago, Cornell, Hong-Kong, Koç, Purdue, and USC for many insightful questions and comments that led to improvements in their analysis. This research was supported by the National Science Foundation through Grant DMI-0500381.

**Appendix**

**Proof of Theorem 1**

We first note that the decision space for each supplier \( i = 1, \ldots, N \) is given by \( [0, s^i] \), where \( s^i \) is the unique solution of \( \lambda [r - u(s)] = v(s) = 0 \). The service level \( s^i \) is finite since \( u \) and \( v \) are nondecreasing convex.

**SA Competition.** It can be easily shown that the profit function of supplier \( i \) is concave with respect to \( s_i \), \( i = 1, \ldots, N \). Therefore, a Nash equilibrium can be obtained as the solution to the following system of \( N \) first order optimality condition equations:

\[
\frac{\partial \pi_i^{SA}(s_i, s_{-i})}{\partial s_i} = \frac{\partial \alpha_i^{SA}(s_i, s_{-i})}{\partial s_i} \lambda (r - u(s_i)) \\
- \left[ \frac{s_i}{s_i} \lambda \frac{\partial u(s_i)}{\partial s_i} + \frac{\partial v(s_i)}{\partial s_i} \right] = 0
\]

for \( i = 1, \ldots, N \), \hspace{1cm} (20)

or equivalently

\[
\sum_{i=1}^N g(s_i) = g(s_i) \lambda (r - u(s_i)) \\
- \left[ \frac{g(s_i)}{\sum_{i=1}^N g(s_i)} \lambda u^\prime(s_i) + v^\prime(s_i) \right] = 0
\]

for \( i = 1, \ldots, N \), \hspace{1cm} (21)
where \( g'(s_i) = \partial g(s_i)/\partial s_i \), \( u'(s_i) = \partial u(s_i)/\partial s_i \), and \( v'(s_i) = \partial v(s_i)/\partial s_i \).

Let \( g^A > 0 \) be the unique solution to the following equation:

\[
\frac{N-1}{N^2 g(s)} g'(s) \lambda(r - u(s)) - \left[ \frac{1}{N} \lambda u'(s) + v'(s) \right] = 0. \tag{22}
\]

Also, let

\[
A^{SA}(s) = (N-1)g'(s)\lambda(r - u(s))/N^2 g(s)
\]

and

\[
B^{SA}(s) = \lambda u'(s)/N + v'(s).
\]

Therefore, (22) can be written as \( A^{SA}(s) - B^{SA}(s) = 0 \). This equation has a unique strictly positive solution since

(a) \( \lim_{r \to 0} A^{SA}(r) = +\infty \),

(b) \( A^{SA}(s_0) = 0 \), where \( s_0 \) is the solution to \( r = u(s_0) \); furthermore, \( A^{SA}(s) = 0 \) does not admit any other solution,

(c) the positive part of \( A^{SA}(s) \) is decreasing in \( s \),

(d) \( B^{SA}(0) \geq 0 \),

(e) \( B^{SA}(s) > 0 \) for \( s > 0 \).

It is easy to check that \( s_i = s_i^{SA} \), \( i = 1, \ldots, N \) is a solution to the system of Equations (21). Hence, to complete the proof of the theorem, it only remains to show that there cannot be another solution to (A2), which implies that \( s_i = s_i^{SA} \), \( i = 1, \ldots, N \), is the unique Nash equilibrium.

The following system of \( N + 1 \) equations with unknowns \( s_i \) and \( G \) is equivalent to the system of Equations (21):

\[
\frac{g(s_i) - g(s_i)}{G^2} g'(s_i) \lambda(r - u(s_i)) - \left[ \frac{g(s_i)}{G} \lambda u'(s_i) + v'(s_i) \right] = 0,
\]

for \( i = 1, \ldots, N \) \( \tag{23} \)

and

\[
\sum_{i=1}^{N} \tilde{g}(s_i) - G = 0. \tag{24}
\]

By virtue of Lemma A1 below, each equation in (23) admits at most one positive solution in the decision space of each supplier. Therefore, \( s_i = s_i^{SA} \) for \( i = 1, \ldots, N \) and \( G = N\tilde{g}(s^{SA}) \) is the only solution for the system of Equations (23) and (24).

**Lemma A1.** Let \( \phi^{SA}_i(s_i, G) = [(G - g(s_i))/G^2]g'(s_i)\lambda(r - u(s_i)) - [(g(s_i)/G)\lambda u'(s_i) + v'(s_i)] \) where \( G > 0 \) is a constant. Then equation \( \phi^{SA}_i(s_i, G) = 0 \) admits at most one strictly positive solution in the decision space of supplier \( i \).

**Proof.** Let \( \phi^{SA}_i(s_i, G) = \phi_{i,1}(s_i, G) - \phi_{i,2}(s_i, G) \), where

\[
\phi_{i,1}(s_i, G) = [(G - g(s_i))/G^2]g'(s_i)\lambda(r - u(s_i))
\]

and

\[
\phi_{i,2}(s_i, G) = [(g(s_i)/G)\lambda u'(s_i) + v'(s_i)]
\]

We know that

(a) since \( u(s_i) \) is an increasing convex function, there is at least one infinite solution to equation \( \phi_{i,1}(s_i, G) = 0 \); let \( s_{i,0} \) be the smallest solution to this equation,

(b) \( \phi_{i,1}(0, G) > 0 \),

(c) \( \phi_{i,1}(s_i, G) \) is decreasing in \( s_i \in [0, s_{i,0}] \), and

(d) \( \phi_{i,2}(s_i, G) \) is nonnegative and increasing in \( s_i \).

Consequently, equation \( \phi^{SA}_i(s_i, G) = 0 \) admits a unique positive solution in \( [0, s_{i,0}] \) if \( \phi^{SA}_i(0, G) > 0 \). Otherwise, it admits no solution. We also notice that the term \( (r - u(s_i)) \) in \( \phi_{i,1} \) is always nonnegative for \( s_i \in [0, s^*] \). Therefore, \( \phi_{i,1} = 0 \) cannot admit more than one solution in this interval. As a result, \( \phi^{SA}_i(s_i, G) = 0 \) cannot have more than one solution in \( [0, s^*] \). This completes the proof of the lemma.

**SS Competition.** It can be easily shown that the profit function of supplier \( i \) is concave with respect to \( s_i \), \( i = 1, \ldots, N \). Therefore, a Nash equilibrium can be obtained as the solution to the following system of \( N \) first order optimality condition equations:

\[
\frac{\partial \pi^{SS}_i(s_i, s_{-i})}{\partial s_i} = \frac{\partial q^{SS}_i(s_i, s_{-i})}{\partial s_i} (\lambda r - \lambda u(s_i) - v(s_i)) - \alpha_i^{SS} \left( \frac{\lambda u'(s_i) + \partial v(s_i)}{\partial s_i} \right) = 0
\]

for \( i = 1, \ldots, N \). \( \tag{25} \)

or equivalently

\[
\sum_{i=1}^{N} \frac{g(s_i)}{N^2 g(s)} g'(s_i) (\lambda r - \lambda u(s_i) - v(s_i)) - \frac{g(s_i)}{N^2 g(s)} \lambda u'(s_i) + v'(s_i) = 0
\]

for \( i = 1, \ldots, N \). \( \tag{26} \)

Let \( s^{SS} > 0 \) be the unique solution to the following equation:

\[
\frac{N-1}{N^2 g(s)} g'(s) (\lambda r - \lambda u(s) - v(s)) - \frac{1}{N} (\lambda u'(s) + v'(s)) = 0. \tag{27}
\]

Also, let \( A^{SS}(s) = (N-1)(\lambda r - \lambda u(s) - v(s))/N^2 g(s) \) and \( B^{SS}(s) = (\lambda u'(s) + v'(s))/N \). Therefore, (27) can be written as \( A^{SS}(s) - B^{SS}(s) = 0 \). This equation has a unique strictly positive solution since

(a) \( \lim_{s \to 0} A^{SS}(s) = +\infty \),

(b) \( A^{SS}(s_0) = 0 \), where \( s_0 \) is the solution to \( \lambda r - \lambda u(s) - v(s) = 0 \). Furthermore, \( A^{SS}(s) = 0 \) does not admit any other solution,

(c) the positive part of \( A^{SS}(s) \) is decreasing in \( s \),

(d) \( B^{SS}(0) \geq 0 \),

(e) \( B^{SS}(s) \) is nondecreasing in \( s \), and

(f) \( B^{SS}(s) > 0 \) for \( s > 0 \).

It is easy to check that \( s_i = s_i^{SS} \), \( i = 1, \ldots, N \) is a solution to the system of Equations (26). Hence, to complete the proof of the theorem, we only need to show that there cannot be another solution to (26).

The following system of \( N + 1 \) equations with unknowns \( s_i \) and \( G \) is equivalent to the system of Equations (26):

\[
\frac{G - g(s_i)}{G^2} g'(s_i) (\lambda r - \lambda u(s_i) - v(s_i)) - \frac{g(s_i)}{G} (\lambda u'(s_i) + v'(s_i)) = 0
\]

for \( i = 1, \ldots, N \). \( \tag{28} \)

and

\[
\sum_{i=1}^{N} g(s_i) - G = 0. \tag{29}
\]
By virtue of Lemma A2 below, each equation in (28) admits at most one positive solution in the decision space of each supplier. Therefore, \( s_i = s_i^{SS}, i = 1, \ldots, N \) and \( G = N g(s^{SS}) \) is the only solution for the system of Equations (28) and (29).

**Lemma A2.** Let

\[
\phi^{SS}_i(s_i, G) = \left[ (G - g(s_i))/G \right] g'(s_i)(\lambda r - \nu(s_i) - v(s_i)) - \left[ (g(s_i)/G) (\lambda u'(s_i) + v'(s_i)) \right],
\]

where \( G > 0 \) is a constant. Then equation \( \phi^{SS}_i(s_i, G) = 0 \) admits at most one strictly positive solution in the decision space of supplier \( i \).

**Proof.** Let \( \phi^{SS}_i(s_i, G) = \phi_{i,1}(s_i, G) - \phi_{i,2}(s_i, G) \), where

\[
\phi_{i,1}(s_i, G) = \left[ (G - g(s_i))/G \right] g'(s_i)(\lambda r - \nu(s_i) - v(s_i))
\]

and

\[
\phi_{i,2}(s_i, G) = \left[ (g(s_i)/G)(\lambda u'(s_i) + v'(s_i)) \right].
\]

We know that

(a) since \( u(s_i) \) and \( v(s_i) \) are increasing convex functions, there is at least one finite solution to equation \( \phi_{i,1}(s_i, G) = 0 \); let \( s_i^0 \) be the smallest solution to this equation,

(b) \( \phi_{i,1}(0, G) > 0 \),

(c) \( \phi_{i,1}(s_i, G) \) is decreasing in \( s_i \in [0, s^0] \),

(d) \( \phi_{i,2}(s_i, G) \) is nonnegative and increasing in \( s_i \) with \( \phi_{i,2}(0, G) = 0 \).

Consequently, equation \( \phi^{SS}_i(s_i, G) = 0 \) admits a unique positive solution in \( [0, s^0] \). Note also that the term \( (\lambda r - \nu(s_i) - v(s_i)) \) in \( \phi_{i,1} \) is always nonnegative in the decision space of supplier \( i \), \( [0, s^0] \). Therefore, \( \phi_{i,1} = 0 \) cannot admit more than one solution in this interval. As a result, \( \phi_i(s_i, G) = 0 \) cannot have more than one solution in \( [0, s^0] \). \( \square \)

**Proof of Theorem 2**

Result 1. The result is obvious because the profit functions of SA and SS competitions have the same form.

Result 2. First recall that the Nash equilibria for SA and SS competitions are, respectively, the solutions to the following equations:

**SA:**

\[
\frac{N - 1}{N} \frac{g'(s)}{g(s)} \lambda (r - u(s)) = \frac{1}{N} \lambda u'(s) + v'(s),
\]

and

**SS:**

\[
\frac{N - 1}{Ng(s)} \frac{g'(s)}{g(s)} \left[ \lambda (r - u(s)) - v(s) \right] = \lambda u'(s) + v'(s),
\]

which can also be rewritten as

**SA:**

\[
\lambda (r - u(s)) - N \frac{g(s)}{g'(s)} v'(s) = \frac{N}{N - 1} \frac{g(s)}{g'(s)} \left[ \lambda u'(s) + v'(s) \right],
\]

and

**SS:**

\[
\lambda (r - u(s)) - v(s) = \frac{N}{N - 1} \frac{g(s)}{g'(s)} \left[ \lambda u'(s) + v'(s) \right].
\]

In order to show that the solution to Equation (32), \( s^{SS} \), is always less than the solution to Equation (33), \( s^{SS} \), we state and prove the following set of claims.

**Claim 1.** For the functions \( g, u, \) and \( v \), \( g(s)/g'(s) \geq s, u(s)/u'(s) \leq s, \) and \( v(s)/v'(s) \leq s \).

To verify Claim 1, let \( \theta_s(s) = g(s)/g'(s) - s \), then \( \theta_s(0) = 0 \) and

\[
\frac{d\theta_s(s)}{ds} = \frac{g'(s)g''(s) - g(s)g''(s)}{g'(s)^2} - 1 = -\frac{g(s)g''(s)}{g'(s)^2} \geq 0.
\]

Therefore, \( \theta_s(s) = g(s)/g'(s) - s \geq 0 \). Next, let \( \theta_u(s) = u(s)/u'(s) - s \), then \( \theta_u(0) = 0 \) and

\[
\frac{d\theta_u(s)}{ds} = \frac{u'(s)u''(s) - u(s)u''(s)}{u'(s)^2} - 1 = -\frac{u(s)u''(s)}{u'(s)^2} \leq 0.
\]

Therefore, \( \theta_u(s) = u(s)/u'(s) - s \leq 0 \). Similarly, we can show that \( \theta_v(s) = v(s)/v'(s) - s \leq 0 \).

**Claim 2.** Let \( \eta_1(s) = \lambda (r - u(s)) - N g(s)/g'(s) v'(s) \) and \( \eta_2(s) = \lambda (r - u(s)) - v(s) \). For any \( s > 0 \), we have \( \eta_2(s) > \eta_1(s) \).

**Claim 2 follows by noting that, by virtue of Claim 1, we have**

\[
\eta_2(s) - \eta_1(s) = N g(s)/g'(s) v'(s) - v(s) \\
\geq N s v'(s) - v(s) \geq N v(s) - v(s) > 0.
\]

**Claim 3.** Let \( \eta_3(s) = N/[(N - 1)g(s)/g'(s)](\lambda u'(s) + v'(s)), \)

\( s^1 \) be the unique solution of \( \eta_3(s) = \eta_2(s) \), and \( s^2 \) be the unique solution of \( \eta_3(s) = \eta_1(s) \), then \( s^2 < s^1 \).

Claim 3 can be shown by arguing that, since \( \eta_3(s) \) is increasing in \( s \) and \( \eta_3(s) > \eta_1(s) \), we have \( s^2 < s^1 \). Recognizing that \( s^1 = s^{SS} \) and \( s^2 = s^{SA} \), proves the result. Figure A1 offers a graphical illustration of this argument.

Results 3 and 4. Substituting \( g(s) = as^\gamma \) and \( u(s) = k_1 s \), and \( v(s) = k_2 s \) in the supplier expected profit functions, we can rewrite Equations (32) and (33) as follows:

**SA:**

\[
\frac{\lambda (r - k_1 s)}{N} = \frac{N}{N - 1} \gamma s \frac{1}{\lambda k_1 + k_2},
\]

**SS:**

\[
\lambda (r - k_1 s) - k_2 s = \frac{N}{N - 1} \gamma s (\lambda k_1 + k_2),
\]

from which we obtain

\[
s^{SA} = \frac{\lambda \gamma (N - 1)}{N[\lambda k_1 + N k_2] + \lambda k_1 \gamma (N - 1)},
\]

and

\[
s^{SS} = \frac{\lambda \gamma (N - 1)}{(\lambda k_1 + k_2)(\gamma (N - 1) + N)}.
\]

It is easy to verify that \( N s^{SS} > s^{SS} \). Furthermore, we have

\[
f(s^{SA}, \lambda_i) = \lambda k_1 s^{SA}/N + k_2 s^{SA}
\]
and

\[ f(s^{SS}, \lambda) = \lambda k_1 s^{SS} + k_2 s^{SS}. \]

Hence,

\[
Nf(s^{SA}, \lambda) - f(s^{SS}, \lambda) = (\lambda k_1 + Nk_2)s^{SS} - (\lambda k_1 + k_2)s^{SS} = \lambda \gamma(N-1) \left( \frac{\gamma N(N-1)k_2}{(N[\lambda k_1 + Nk_2] + k_2)(N-1)(\gamma(N-1)+N)} \right) > 0.
\]

Rewriting the profit functions as

\[
\pi^{SA} = \frac{\lambda \gamma (N-1) \gamma \lambda \gamma}{N(N + \gamma(N-1) k_1)/(\lambda k_1 + N k_2)}
\]

and

\[
\pi^{SS} = \frac{\lambda \gamma (N-1) \gamma \lambda \gamma}{N(N + \gamma(N-1))},
\]

we can also easily see that \( \pi^{SA} < \pi^{SS} \).

**Proof of Theorem 3**

**Result 1.** The Nash equilibrium service level is the solution to the following equation:

\[
\frac{N-1}{N} \frac{g'(s)}{g(s)} (\lambda r - \lambda u(s) - v(s)) = \lambda u'(s) + v'(s).
\]

Since, the right-hand side is independent of \( N \) and increasing in \( s \) and the left-hand side is increasing in \( N \), the Nash equilibrium service level \( s^{SS} \) is also increasing in \( N \). Furthermore, since

\[
\lim_{N \to \infty} \left[ \frac{N-1}{N} \frac{g'(s)}{g(s)} (\lambda r - \lambda u(s) - v(s)) \right] = \frac{g'(s)}{g(s)} (\lambda r - \lambda u(s) - v(s)),
\]

the Nash equilibrium service level approaches \( s^{SS} \). the unique solution to

\[
g'(s)(\lambda r - \lambda u(s) - v(s))/g(s) = \lambda u'(s) + v'(s).
\]

Since the solution is symmetric we have \( q^{SS} = s^{SS} \).

**Result 2.** When \( v(s) = 0 \), the proof is similar to that of the result (1), and \( s^{SA} = q^{SA} \) is the solution to the following equation:

\[
g'(s)(\lambda r - \lambda u(s))/g(s) = \lambda u'(s).
\]

When \( u(s) = 0 \), the Nash equilibrium service level solves the following equation:

\[
\frac{N-1}{N^2} \frac{g'(s)}{g(s)} \frac{\lambda r - \lambda u(s)}{\lambda u'(s)} = \frac{\lambda r - \lambda u(s)}{\lambda u'(s)}.
\]

The right-hand side is independent of \( N \) and increasing in \( s \), while the left-hand side is decreasing in \( N \). Consequently, the equilibrium point \( s^{SA} \) is decreasing in \( N \). Finally, since \( \alpha^{SA} = 1/N \) and \( v(0) = 0 \), \( \lim_{N \to \infty} \pi^{SA} = 0 \). For part (c), when both \( u(s) \) and \( v(s) \) are positive for \( s > 0 \), the profit function is given by

\[
\pi^{SA}(s, s) = \alpha^{SA}(s, s) \lambda (r - u(s)) - v(s).
\]

Since \( \alpha^{SA} \) approaches zero as \( N \) goes to infinity, the only service level which leads to a nonnegative profit is \( s^{SA} = 0 \). In this case, the Nash equilibrium is the solution to

\[
\frac{N-1}{N^2} \frac{g'(s)}{g(s)} \lambda (r - u(s)) = \frac{\lambda}{N} u'(s) + v'(s).
\]

Noting that both sides of the above equation are decreasing in \( N \), the right-hand side is increasing in \( N \) and the left-hand side is decreasing in \( s \), the solution to this equation is not necessarily increasing or decreasing in \( N \). However, since the solution approaches zero as \( N \) goes to infinity it should be decreasing for large values of \( N \).

**Result 3.** Result 3 follows immediately from Results 1 and 2.

**Proof of Proposition 1**

We can rewrite the profit function as follows:

\[
\pi^{SA}_i(s_i, s_{-i}) = \frac{\lambda}{\Sigma} \left( \lambda r - \lambda k_1 s_i - k_2 s_i \right)
\]

where \( \Sigma = \sum_{i=1}^{N} s_i \). The derivative of the profit function of supplier \( i \) with respect to her service level is:

\[
\frac{\partial \pi^{SA}_i(s_i, s_{-i})}{\partial s_i} = \frac{\Sigma - s_i}{\Sigma} \gamma s_i \gamma^{-1} (\lambda r - \lambda k_1 s_i - k_2 s_i)
\]

\[-s_i \frac{\lambda}{\Sigma} (\lambda k_1 + k_2).
\]

The Nash equilibrium is therefore the solution to the following set of equations.

\[
(\Sigma - s_i) \gamma (\lambda r - \lambda k_1 s_i - k_2 s_i) = \Sigma (\lambda k_1 + k_2) s_i \quad \text{for } i = 1, \ldots, N.
\]

Using an approach similar to the one used in the proof of Theorem 1, one can show that there is a unique symmetric solution to this set of equations, which is the Nash equilibrium. Because the Nash equilibrium is symmetric, it is given by

\[
s^{SS} = \frac{(N-1) \gamma \lambda r}{(\lambda k_1 + k_2)(N + \gamma(N-1))}.
\]

To see if \( s^{SS} \) is increasing in \( \gamma \), note that its derivative, with respect to \( \gamma \), is positive:

\[
\frac{\partial s^{SS}}{\partial \gamma} = \frac{(N-1) \lambda r N}{\lambda k_1 + k_2 N(N-1)} > 0.
\]

It is easy to verify that \( \lim_{\gamma \to \infty} s^{SS} = \lambda r / (\lambda k_1 + k_2) \). Expected supplier profit is given by \( \pi^{SS} = \lambda r / (N + \gamma(N-1)) \), which is decreasing in \( \gamma \) and approaches zero as \( \gamma \to \infty \).

**Proof of Proposition 2**

**Result 1.** The expected supplier profit functions can be written as

\[
\pi^{SA}_i(s_i, s_{-i}) = \frac{s_i}{\Sigma} \lambda (r - k_1) - k_2 s_i
\]

where

\[
\Sigma = \sum_{i=1}^{N} s_i \quad \text{for } i = 1, \ldots, N,
\]

which leads to

\[
\frac{\partial \pi^{SA}_i(s_i, s_{-i})}{\partial s_i} = \frac{\Sigma - s_i}{\Sigma^2} \gamma s_i \gamma^{-1} (\lambda r - k_1) - k_2.
\]
The Nash equilibrium could be the solution to the following set of equations:

\[(\Sigma - s_i')s_i^{\gamma - 1}\lambda (r - k_i) = \Sigma^2 k_2, \text{ for } i = 1, \ldots, N. \quad (35)\]

Because all \(N\) equations have the same form, there exists a symmetric solution \(s_1 = \cdots = s_i = s\) that solves (35). To show that \(s\) is indeed a Nash equilibrium service level, we need to show that the profit function of supplier \(i\) has a unique maximum at \(s_i = s\) when all other suppliers choose \(s\). Given that all other suppliers choose service level \(s\), the expected profit function of supplier \(i\) is given by:

\[\pi^{SA}_i(s_i, s_{-i}) = \frac{s_i^\gamma}{(N-1)s^\gamma + s_i^\gamma} \lambda (r - k_i) - k_2 s_i.\]

This leads to:

\[\frac{\partial \pi^{SA}_i(s_i, s_{-i})}{\partial s_i} = \frac{(N-1)s^\gamma}{(N-1)s^\gamma + s_i^\gamma} s_i^{\gamma - 1} \lambda (r - k_i) - k_2 = 0,\]

or equivalently

\[\frac{(N-1)s^\gamma}{(N-1)s^\gamma + s_i^\gamma} s_i^{\gamma - 1} \lambda (r - k_i) = k_2. \quad (36)\]

It is easy to see that (36) could admit more than one solution. To show that the expected supplier profit function cannot have more than one maximum, we check the behavior of the second derivative of the profit function with respect to \(s_i:\)

\[\frac{\partial^2 \pi^{SA}_i(s_i, s_{-i})}{\partial s_i^2} = \frac{(N-1)s^\gamma s_i^{\gamma - 2}((\gamma - 1)(N-1)s^{\gamma - (\gamma + 1)} + (\gamma + 1)s_i^\gamma)}{(N-1)s^\gamma + s_i^\gamma)^2} \gamma \lambda (r - k_i). \quad (37)\]

If a function with continuous first and second derivative has more than one local maximum, then the sign of the second derivative of the function must change more than once. Since the second derivative of the profit function of supplier \(i\) is positive for \(s_i' < (\gamma - 1)(N-1)s^{\gamma - (\gamma + 1)} + (\gamma + 1)s_i^\gamma\) and remains negative for \(s_i' > (\gamma - 1)(N-1)s^{\gamma - (\gamma + 1)} + (\gamma + 1)s_i^\gamma\), this profit function cannot have more than one maximum. We know when \(s_i = s\) the first order optimality condition is satisfied. A condition for the profit function to admit its maximum at \(s_i = s\) is for the second derivative to be negative at \(s_i = s\). This condition is satisfied if \(\gamma < N/(N-2)\). Hence, the solution to Equation (35) is a Nash equilibrium if \(\gamma < N/(N-2)\) and the resulting Nash equilibrium service levels and profit function are given by:

\[s^{SA}_i = \frac{(N-1)\gamma \lambda (r - k_i)}{N_3 k_2}\]

and

\[\pi^{SA}_i = \frac{(N-1)\gamma \lambda (r - k_i)}{N^2 k_2}.

To ensure a nonnegative profit we need the condition \(\gamma \leq \gamma_{\max} = N/(N-1)\), which is more restrictive than \(\gamma < N/(N-2)\).

It is straightforward to verify that the equilibrium service level is increasing in \(\gamma\) and the equilibrium expected profit is decreasing in \(\gamma\). For \(\gamma = \gamma_{\max}\) we have \(s^{SA}_i = \lambda (r - k_i)/Nk_2\) and \(\pi^{SA}_i = 0\). Furthermore, when \(N = 2\), Equation (36) simplifies to

\[s_i = \frac{k_2}{\gamma \lambda} (s_i^\gamma + s_i^\gamma)^\gamma - 1, \text{ for } i = 1 \text{ and } 2.\]

It is easy to check that the solution to the above system of equations is unique.

**Result 2**. The proof is similar to that of Proposition 1 with \(k_2 = 0\).

**References**


Li, Q., A. Y. Ha. 2003. Accurate response, reactive capacity and inventory competition. Working paper, Hong Kong University of Science and Technology, Hong Kong.


