Procuring Fast Delivery: Sole Sourcing with Information Asymmetry

Gérard P. Cachon
Operations and Information Management, The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104, cachon@wharton.upenn.edu

Fuqiang Zhang
Paul Merage School of Business, University of California, Irvine, California 92697-3125, fzhang@uci.edu

This paper studies a queuing model in which a buyer sources a good or service from a single supplier chosen from a pool of suppliers. The buyer seeks to minimize the sum of her procurement and operating costs, the latter of which depends on the supplier’s lead time. The selected supplier can regulate his lead time, but faster lead times are costly. Although the buyer selects the supplier to source from (possibly via an auction) and dictates the contractual terms, the buyer’s bargaining power is limited by asymmetric information: The buyer only has an estimate of the suppliers’ costs, while the suppliers know their costs precisely. We identify a procurement mechanism that minimizes the buyer’s total cost (procurement plus operating). This mechanism is not simple: It is a numerically derived nonlinear menu of contracts. Therefore, we study several simpler mechanisms: e.g., one that charges a late fee and one that specifies a fixed lead-time requirement (no menus, no nonlinear functions). We find that simple mechanisms are nearly optimal (generally within 1% of optimal) because asymmetric information conveys significant protection to the supplier, i.e., the supplier is able to retain most of the benefit of having a lower cost. Renegotiation is another concern with the optimal mechanism: Because it does not minimize the supply chain’s cost, the firms can be both better off if they throw away the contract and start over. Interestingly, we find that the potential gain from renegotiation is relatively small with either the optimal or our simple mechanisms. We conclude that our simple mechanisms are quite attractive along all relevant dimensions: buyer’s performance, supply chain performance, simplicity, and robustness to renegotiation.

Key words: mechanism design; reverse auctions; supply chain coordination; game theory; renegotiation

History: Accepted by William Lovejoy, operations and supply chain management; received December 27, 2003.

This paper was with the authors 8 months for 4 revisions.

1. Introduction

In the sourcing of a product or service, a buyer should consider both procurement price and delivery lead time. The faster a supplier’s delivery lead time, the lower a buyer’s operating costs (e.g., inventory holding and backorder penalty costs). A supplier’s delivery lead time depends on the supplier’s capacity, but capacity is costly, and so there is a classic incentive conflict within the supply chain: The supplier incurs the direct cost of capacity but the buyer enjoys its benefit. To complicate matters, the buyer often has only an estimate of the supplier’s capacity cost, while the supplier knows it precisely.

While practitioners and academics surely understand the importance of lead times in the procurement process (see Burt 1989, Pyke and Johnson 2003, McNealy 2001, Wise and Morrison 2000), and the advent of the Internet has created an explosion of new marketplaces in the business-to-business arena (Pinker et al. 2003), there has been surprisingly little research on how a buyer should design her procurement process to achieve minimum total cost through an effective balance of price and delivery lead time. That is the focus of this paper. Our main research questions are summarized as follows.

What is an optimal procurement mechanism for the buyer? A mechanism is any process that takes information the suppliers announce (e.g., their bids, their costs, etc.) and outputs the buyer’s decisions: which supplier is chosen, what actions the suppliers must take, and how much they are paid. The optimal mechanism minimizes the buyer’s total cost (procurement plus operating), and it is the benchmark to assess all other mechanisms.

Do simple procurement mechanisms exist that give the buyer near-optimal performance? The optimal mechanisms are complex along several dimensions: They may be hard to evaluate, or they may involve nonlinear functions or a complex menu of functions. While we admit that there is no definitive way to measure how much simpler one mechanism is over another,
this ambiguity should not cause research to focus exclusively on optimal mechanisms. We believe simple mechanisms are worth studying because they are more likely to be implemented in practice. Beil and Wein (2003) make a similar observation based on their discussions with industry practitioners. Assuming we can identify effective simple mechanisms, we then wish to provide some insight into why they are effective.

To what extent is supply chain efficiency reduced by the buyer’s desire to minimize her own total cost? The literature on supply chain coordination, which generally does not consider asymmetric information, suggests the buyer offer the supplier a coordinating contract (one that induces the supplier to choose the supply chain optimal capacity) and then negotiate for as large a share of the supply chain’s profit as possible (e.g., Caldentey and Wein 2003). However, implementing a coordinating contract is difficult with asymmetric information: The coordinating contract parameters may depend on the unknown information, thereby creating doubt with at least one firm as to what the proper contract parameters are. In addition, it is well known (see Laffont and Martimort 2002) that ex post efficiency (i.e., maximizing supply chain performance) is at odds with the buyer’s ex ante desire to maximize her own profit. This creates a renegotiation opportunity: After the optimal mechanism is implemented, the firms have an incentive to scrap it to capture the lost efficiency. We wish to determine the magnitude of this trade-off in the context of our model.

The next section describes the model and §3 relates our work to the literature. Section 4 minimizes the supply chain’s total cost. Section 5 covers procurement strategies with one potential supplier, and §6 covers competitive bidding among multiple potential suppliers. Section 7 provides numerical results, and §8 discusses two extensions to the model. Section 9 discusses our results.

2. The Model
A buyer must acquire a component from one of \( n \geq 1 \) potential suppliers. The buyer uses this component in the assembly of a product sold to consumers. (In §8.2 we assume the buyer is unable to hold component inventory, so in that case it is possible to interpret the model in terms of a buyer procuring a service rather than a physical product.) Customer demand arrives at the buyer according to a Poisson process with rate \( \lambda \).

The suppliers are make-to-order manufacturers. Let \( \mu \) be a supplier’s production rate, which we generally refer to as the supplier’s capacity. A supplier’s interproduction times are exponentially distributed with mean \( 1/\mu \), and the supplier incurs a capacity cost at rate \( b \mu \) (\( b > 0 \)) to maintain its capacity. The suppliers’ capacity costs are independent draws from a continuous random variable, \( b \), where \( b \in [b_1, b_2] \) and \( 0 < b_1 \leq b_2 \). Let \( F \) and \( f \) be the cdf and pdf, respectively. The variable production cost is normalized to zero. Once the production of a unit is completed, it is immediately delivered to the buyer.

The buyer incurs inventory holding costs at rate \( h \) per unit. A constant holding cost is reasonable if the physical holding cost plus the financial holding cost on the variable production cost dominates the financial holding cost due to the supplier’s capacity cost and margin. Alternatively, a constant \( h \) can be considered as an approximation for the holding cost given the possible range of procurement costs. Section 8 discusses the implications of a holding cost that varies with the procurement cost.

Unsatisfied demand is backlogged, and the buyer incurs a goodwill cost at rate \( p \) per backorder. The sum of the holding and backorder costs is referred to as the operating cost. To control her operating cost, the buyer uses a base-stock policy with base-stock level \( s \). The supplier does not carry inventory.1

The buyer’s procurement strategy includes two tasks—supplier selection (which supplier to source from) and contract design (the details of the transfer payment between the buyer and the supplier). We consider several procurement strategies within two distinct scenarios. The first scenario is sole sourcing with one potential supplier (\( n = 1 \)): The buyer only offers a procurement contract to a single potential supplier, possibly because there is only one supplier with the necessary technology, or the buyer has a long-run relationship with the supplier, or because the buyer wishes to develop the component quickly. The next scenario involves competitive bidding among at least two potential suppliers (\( n \geq 2 \), i.e., the buyer selects her supplier via some auction mechanism. (These are often called reverse auctions because the suppliers are bidding for the right to sell to the buyer, but we shall just refer to them as auctions.) With either scenario the buyer knows the distribution function of the suppliers’ capacity costs, but the buyer does not observe each supplier’s cost realization. All other rules and parameters in the game are common knowledge.

The sequence of events is as follows: The buyer announces her supplier selection process (some
auction mechanism, if \( n \geq 2 \) and her transfer payment contract; assuming the supplier accepts the contract, the supplier chooses his capacity \( \mu \); the buyer observes the supplier’s lead times and chooses \( s \); the buyer incurs costs (procurement and operating) and the supplier earns a profit (transfer payment minus capacity costs) over an infinite horizon. The buyer minimizes the sum of her procurement and operating costs per unit of time. The suppliers maximize their own expected profit per unit of time. All firms are risk neutral.

Although we did not design this model with a specific industry in mind, the model is most representative of the contract manufacturing industry, in which firms assemble specialized components on a make-to-order basis (see Thurm 1998, Bulkeley 2003).

### 3. Literature Review

Our model studies procurement strategies in a queuing framework with asymmetric information. There is much related work, the closest of which is Cachon and Zhang (2003). As in this paper, there is a single buyer with Poisson demand, the suppliers are make-to-order producers that choose capacity, and the buyer is concerned with procurement and operating costs. However, this paper considers sole-sourcing strategies, whereas Cachon and Zhang (2003) work with multisourcing strategies and do not have asymmetric information. (For additional work on dual sourcing see Gilbert and Weng 1998, Ha et al. 2003, and Kalai et al. 1992.)

Caldentey and Wein (2003) study a model similar to ours, but they have only one potential supplier and do not have asymmetric information. They focus on coordination strategies, whereas we consider the buyer’s optimal mechanism. Benjaafar et al. (2004) study multisourcing versus sole-sourcing strategies with several potential suppliers, but the buyer’s price is fixed, and they also do not have asymmetric information.

The following papers study a supply chain with two firms and asymmetric information in nonqueuing models: Corbett and de Groot (2000), Corbett (2001), Corbett and Tang (1998), Corbett et al. (2004), and Ha (2001). As in this paper, those papers design an optimal menu of contracts, but we also consider a broader set of procurement strategies. There is literature on quality contracting with asymmetric information (e.g., Baijman et al. 2000 and Lim 2001), but those models focus on the buyer’s inspection decisions and the ability to contract on the outcome of inspections, neither of which is present in our model with lead times. There is work on supply chain signaling (e.g., Cachon and Lariviere 2001 and Özer and Wei 2002), in which the firm that designs the contract has private information, but in our model the contract designer lacks information.

See Elmaghraby (2000) for a survey of the procurement literature, and see Klemperer (1999) and McAfee and McMillan (1987a) for surveys of the auction literature. Most closely related to our work are McAfee and McMillan (1987b), Laffont and Tirole (1987), and Che (1993). As in our paper, the first two articles study adverse selection (suppliers vary in their costs) with moral hazard (suppliers exert costly effort that benefits the buyer, where effort is analogous to capacity). Although there are some differences, we show that their results can be used to evaluate the optimal mechanism in our model. However, they do not study the effectiveness of simple mechanisms. Che (1993) implements the optimal mechanism in McAfee and McMillan (1987b) and Laffont and Tirole (1987) via a scoring-rule auction in which the suppliers bid on both price and quality. Bushnell and Oren (1995) study scoring-rule auctions for an organization in which internal suppliers have private cost information, but they do not consider the supplier’s delivery performance. See Asker and Cantillon (2004) for additional discussion on scoring-rule auctions.

There are several other papers that study multiattribute procurement. Chen et al. (2003) study procurement over price and transportation costs, but take the perspective of a third-party auctioneer rather than the buyer. Manelli and Vincent (1995) consider (in effect) a multiattribute situation in which the buyer’s value is correlated with the suppliers’ costs, i.e., the additional attribute is the supplier’s identity. In our model the buyer is indifferent between any two suppliers as long as the suppliers have the same delivery time. Beil and Wein (2003) study multiattribute auctions that occur over multiple rounds so that the buyer learns information regarding the suppliers in each round. We have a single-round auction, so learning is not possible. They do not consider sole sourcing with only one potential supplier.

Dasgupta and Spulber (1990), Chen (2001), Hansen (1988), Jin and Wu (2002), and Seshadri and Zemel (2003) study procurement with competitive bidding and variable quantity. In our model, the buyer’s expected purchase quantity is fixed per unit time. There is literature on lead-time competition through operational strategies (e.g., Li 1992, Cachon and Harker 2002, So 2000), but in those papers the competitive structure is exogenous, whereas in our model

---

2 The organization’s profit depends on which supplier is chosen and the purchase quantity, but not on the actions of the selected supplier. In our model, the selected supplier chooses how much capacity to build, which affects the buyer’s operating cost. However, their model is more general than ours in the sense that their suppliers differ both in their fixed production cost and their nonlinear marginal production cost.
it is endogenous. Ramasesh et al. (1991), Anupindi and Akella (1993), Sedarage et al. (1999), and Li and Kouvelis (1999) are representative studies investigating a buyer’s procurement strategy given exogenous characteristics for each supplier (such as delivery time and price). There are a number of papers (for surveys see Cachon 1998, 2003) that study supply chain lead-time coordination in a multi-echelon inventory setting, but those papers do not have asymmetric information, nor do they consider procurement costs. We touch upon the issue of renegotiation. See Plambeck and Zenios (2000) and Plambeck and Taylor (2002, 2004) for other papers that discuss renegotiation, but in settings quite different than ours.

4. Centralized Management

This section defines and derives several useful functions and presents the optimal policy for the supply chain. It is optimal for the supply chain to have one supplier (because the capacity cost is linear in \( \mu \)), and it is optimal to use a base-stock policy. Let \( N \) be the number of outstanding orders at the supplier in steady state. \( N \) is geometrically distributed. The buyer’s operating cost is

\[
c(\mu, s) = E[h(s-N)^{+} + p(N-s)^{+}]
\]

\[
= h\left(s - \frac{\mu}{\mu - \lambda}\right) + (h + p)\left(\frac{\mu}{\mu - \lambda}\right)^{s} \frac{\lambda}{\mu - \lambda},
\]

the supplier’s cost is \( b\mu \), and the supply chain’s total cost is \( C(\mu, s, b) = c(\mu, s) + b\mu \), where \( (x)^{+} = \max(0, x) \), \( \mu \geq 0 \) and \( s \in \{0, 1, 2, \ldots\} \). Because \( s \) is restricted to the set of non-negative integers, it is not possible to provide a closed-form solution for the minimum cost. Therefore, in the remainder of this paper we treat \( s \) as a continuous variable.

\( C(\mu, s, b) \) is convex in \( s \) and let \( s^{*}(\mu) \) be the optimal base-stock level,

\[
s^{*}(\mu) = -\ln\left(\frac{h}{h + p} \frac{\mu/\lambda - 1}{\ln(\mu/\lambda)}\right)/\ln(\mu/\lambda).
\]

The buyer’s operating cost with \( s^{*}(\mu) \) is

\[
c(\mu) = c(\mu, s^{*}(\mu))
\]

\[
= h\left[\frac{1}{\ln(\mu/\lambda)} - 1\right] \left(1 - \ln\left(\frac{h}{h + p} \frac{\mu/\lambda - 1}{\ln(\mu/\lambda)}\right)\right),
\]

and the supply chain’s total cost is \( C(\mu, b) = c(\mu) + b\mu \). According to the next theorem, \( c(\mu) \) is convex in \( \mu \). Consequently, \( C(\mu, b) \) is convex in \( \mu \). Let \( \mu^{*}(b) \) be the supply chain’s optimal capacity,

\[
\mu^{*}(b) = \arg\min_{\mu} C(\mu, b),
\]

and let \( C^{*}(b) = C(\mu^{*}(b), b) \) be the supply chain minimum cost.

**Theorem 1.** The buyer’s operating cost, \( c(\mu) \), is convex in \( \mu \geq \lambda \). (All proofs are in the appendix.)

We later take advantage of the following approximations. The exponential distribution is the continuous counterpart to the geometric distribution, so an exponential distribution with mean \( E[N] \) approximates the geometric distribution for \( N \). This tends to underestimate the average waiting time, but it is justified in a heavy-traffic analysis (see Caldentey and Wein 2003). Let \( \hat{C}(\mu, s, b) = \hat{c}(\mu, s) + b\mu \) be the supply chain’s cost function according to the exponential approximation, where \( \hat{c}(\mu, s) \) is the buyer’s operating cost,

\[
\hat{c}(\mu, s) = hs + \frac{(h + p)e^{-s(\mu/\lambda - 1)} - h}{\mu/\lambda - 1}.
\]

(We use “\( \sim \)” throughout to indicate a function derived from this approximation.) From Caldentey and Wein (2003), the unique global minimizers of \( \hat{C}(\mu, s, b) \) are

\[
\hat{\mu}(b) = \lambda + \sqrt{a/b} \quad \text{and} \quad \hat{s}(b) = \sqrt{b\alpha/h^2},
\]

where \( \alpha = h\lambda \ln((h + p)/h) \).

The operating cost and the supply chain’s optimal cost (i.e., the buyer’s operating cost plus the supplier’s capacity cost) are then

\[
\hat{c}(\mu) = \hat{c}(\hat{\mu}(b), \hat{s}(b), b) = b\lambda + 2\sqrt{ab}.
\]

Zhang (2004) finds that the supply chain’s cost is nearly minimized with capacity \( \hat{\mu}(b) \) as long as utilization is reasonably high (say, more than 0.17).

5. One Potential Supplier (\( n = 1 \))

In this section we consider the case in which there is only one potential supplier (or the buyer has already selected her supplier), so the buyer only needs to set the transfer payment. We begin with the optimal mechanism, then consider supply chain coordination and two simpler mechanisms.

5.1. Buyer’s Optimal Mechanism (OM)

A mechanism consists of a feasible message space for the supplier and a mapping from the supplier’s message space to the space of feasible payment and action schedules. We consider direct mechanisms in which the supplier’s message space is identical to his set of private cost values. The mechanism is truth inducing if the supplier finds it optimal to reveal his cost to the buyer. Although the space of possible mechanisms is quite large, according to the revelation principle (see Myerson 1981, 1983; Salanie 1998), an optimal mechanism for the buyer is a menu of contracts that satisfies two constraints. The menu is a pair of functions,
\[ \{ \mu_0(x), R_0(x) \} \], such that the supplier chooses from this menu by announcing his cost to be \( x \), then he builds capacity \( \mu_0(x) \) and the buyer pays him \( R_0(x) \) per unit produced. One constraint imposed on this menu is the incentive compatibility constraint:

\[ b = \arg \max_x \pi(x) = R_0(x) \lambda - b \mu_0(x), \]

i.e., the supplier’s true cost maximizes his profit, therefore, he builds capacity \( \mu_0(b) \) and receives \( R_0(b) \) per unit delivered. The second is an individual rationality constraint:

\[ \pi(b) \geq 0 \quad \text{for all } b \in [b_l, b_h], \]

i.e., the supplier participates only if his profit is non-negative (we assume zero profit is the supplier’s best outside alternative). According to (3), the buyer designs a menu that even the highest-cost supplier accepts, which implicitly assumes there is a severe penalty for failing to make an agreement with the supplier. (Corbett et al. 2004 relax this assumption in a different model.)

The buyer’s total cost (procurement and operating) is \( R_0 \lambda + c(\mu_0) \), and the buyer’s optimal menu is the solution to the following problem:

\[ \min_{\mu_0} \int_{b_l}^{b_h} (R_0(x) \lambda + c(\mu_0(x))) f(x) \, dx \]

s.t. (2), (3).

Theorem 2. If \( F(x) \) is log-concave, then the buyer’s optimal menu of contracts to offer the supplier (i.e., the solution to (4)) is characterized by

\[ c'(\mu_0(x)) = -x - F(x)/f(x) \]

\[ R_0(x) \lambda = x \mu_0(x) + \int_x^{b_h} \mu_0(y) \, dy. \]

The log-concave requirement on \( F(x) \) is sufficient (but not necessary) for the second-order condition on each buyer’s incentive compatibility constraint, (2). It is a mild restriction, satisfied by many commonly used distributions (see Bagnoli and Bergstrom 1989 for details). Because \( c(\mu) \) is complex, we do not have a closed-form solution for \( \mu_0(x) \) and \( R_0(x) \), but it is possible to numerically evaluate the optimal menu and the buyer’s expected cost.

With the buyer’s optimal mechanism the supplier builds less than the supply chain optimal capacity, \( \mu^\star(b) \) (the optimal capacity satisfies the first-order condition, \( c'(\mu^\star) = -b \), because \( c(\mu) \) is convex), hence, the buyer sacrifices some ex post efficiency to increase his own profit. This is why the optimal mechanism is vulnerable to renegotiation: After the supplier announces his capacity, both the buyer and the supplier can be better off if they renegotiate (choose \( \mu^\star(b) \) and a Pareto division of the supply chain’s profit).

Although in Theorem 2 the optimal mechanism specifies the capacity the supplier must build for each announced cost, this menu can also be expressed in terms of the supplier’s lead time (i.e., the set of lead times the supplier can choose from): Because \( F(x) \) is log-concave, \( \mu_0(x) \) is decreasing in \( x \), which means the required lead time is also decreasing in \( x \). Nevertheless, the optimal contract implicitly assumes the parties can enforce a contract based on capacity or lead time. In some settings it appears that this issue is not a major concern because contracts like this are observed in practice (see §5.4). However, we admit that there could be settings in which contracting on capacity (or lead time) is challenging or maybe even infeasible. For example, due to the stochastic nature of lead times, it is not always possible for the buyer to know whether or not the supplier cheated on his capacity commitment, and convincing a third party of an infraction may be even more difficult. One solution in those settings would be to use a contract that is not based on capacity or lead-time commitments, such as the late-fee contract we describe in §5.3. That contract is not optimal, but we show that it performs quite well for the buyer. Another possible solution is to construct a relational contract that is enforced by the threat of breaking off all future relations. For example, the buyer could pay the supplier a bonus at the end of a period if a certain lead time were achieved in that period, and if the buyer refused to pay that bonus (or if the supplier refused to accept the bonus), then the parties never interact again. It is also possible to combine both explicit contracts with relational contracts (see Gibbons 2005). A full investigation of the relational contract approach is beyond the scope of this paper, and hence, is left to future research.

5.2. Supply Chain Coordination (CC)

Coordination requires that the supplier builds capacity \( \mu^\star(b) \), the supplier earns a nonnegative profit, and the chosen base-stock level is \( s^\star(\mu^\star(b)) \). This can be done with the following arrangement: Charge the supplier \( h \) per unit in the buyer’s inventory and \( p \) per unit in the buyer’s backorder; the supplier chooses \( s \) and the unit price is

\[ R_0 = C(\mu^\star(b_h), s^\star(b_h), b_h)/\lambda. \]

This works because the supplier incurs all supply chain costs, so the supplier has an incentive to choose \( \mu^\star(b) \) and \( s^\star(\mu^\star(b)) \), and even the highest-cost supplier earns a nonnegative profit. The buyer’s total cost is then \( C(\mu^\star(b_h), s^\star(b_h), b_h) \), and the supplier’s profit is

\[ \pi = C(\mu^\star(b_h), s^\star(b_h), b_h) - C(\mu^\star(b), s^\star(b), b). \]

This resembles vendor-managed inventory (because the supplier chooses \( s \)) with consignment and service
penalties. Supply chain coordination is not achievable with a simpler mechanism: Because only the supplier knows $b$, only a full transfer of the buyer’s operating cost to the supplier results in the supplier choosing $\mu^*(b)$, and due to the full transfer of costs, the supplier must also choose the buyer’s base-stock level.

5.3. Late-Fee Mechanism (LF)
With a late-fee mechanism the buyer pays the supplier $R_f$ per unit and charges the buyer $\eta_f$ per outstanding order per unit time. This mechanism is simple to explain (just two parameters, no menu), easy to implement (it is based on data verifiable by both parties, the number of outstanding orders), and it is observed in practice (e.g., Beth et al. 2003). Although we would ideally like to find the optimal pair $[R_f, \eta_f]$, the complexity of $c(\mu, b)$ precludes a closed-form solution. As an alternative, we take advantage of the exponential approximation for $N$ to derive closed-form solutions for $R_f$ and $\eta_f$. We later show that this approximation yields excellent results.

Let $\mu_f(b)$ be the supplier’s optimal capacity given the late fee:

$$\mu_f(b) = \arg\min_{\mu} (b\mu + \eta_f \lambda/(\mu - \lambda)) = \lambda + \sqrt{\eta_f \lambda/b}.$$ 

Recall that $\hat{\mu}(b) = \lambda + \sqrt{\alpha/b}$ minimizes $\hat{C}(\mu, \hat{s}(b), b)$. Matching $\mu_f(b)$ with $\hat{\mu}(b)$ yields

$$\eta_f = \alpha/\lambda:$$

(5)

If the late fee is $\eta_f$, then the supplier minimizes his cost with capacity $\hat{\mu}(b)$, which also happens to be the capacity that minimizes $\hat{C}(\mu, \hat{s}(b), b)$, the supply chain’s cost function based on the exponential approximation. Hence, $\eta_f$ minimizes the approximate cost function, and while it does not minimize the actual cost function, as already mentioned, it nearly does so when the optimal utilization is not too low.

To ensure participation, the buyer should pay $R_f$ per delivered unit, such that

$$\pi(b_h) = R_f \lambda - b_h \hat{\mu}(b_h) - \eta_f \left(\frac{\lambda}{\hat{\mu}(b_h)} - \lambda\right) = 0,$$

which yields

$$R_f = b_h + 2\sqrt{\alpha b_h/\lambda}.$$ (6)

$N$ does not depend on $s$, so the buyer’s optimal base-stock level is $s^* (\hat{\mu}(b))$. However, a base-stock policy may no longer be optimal for the buyer; unlike a base-stock policy for which the sum of on-hand and on-order inventory is sufficient for implementation, the buyer’s optimal policy probably depends on both on-hand and on-order inventory in a complex way. Thus, with this late-fee mechanism the buyer must commit to implementing a base-stock policy.\(^3\)

5.4. Lead-Time Mechanism (LT)
With the lead-time mechanism the buyer tells the supplier the lead time that must be delivered and how much the buyer pays for each unit. Due to the one-to-one relationship between the delivered lead time and the supplier’s capacity, we can think of this mechanism in terms of two parameters, $\mu_s$ and $R_s$, the supplier’s required capacity and the buyer’s price per unit, respectively. We assume that the supplier must deliver the lead time $(\mu_s - \lambda)^{-1}$ if the supplier accepts the contract, i.e., there is a substantial penalty for failing to adhere to the agreement. In other words, we assume the buyer can enforce this contract, which, as discussed with the optimal mechanism, is not assured. However, FreeMarkets (see Rangan 1998) implements this mechanism with an auction (multiple potential suppliers), so we suspect that at least in some cases the compliance issue can be sufficiently managed.\(^4\)

With the lead-time mechanism the supplier’s expected profit is $\pi = \lambda R_s - b \mu_s$. To ensure participation, the unit price must be $R_s(\mu_s) = b_h \mu_s/\lambda$. The buyer’s cost is then

$$c(\mu_s) + \lambda R_s(\mu_s) = c(\mu_s^*) + b_h \mu_s^*,$$

which is the supply chain’s cost with the highest capacity cost, $C(\mu_s, b_h)$. Hence, the buyer’s optimal lead-time requirement is $(\mu_s - \lambda)^{-1}$, where $\mu_s = \mu^*(b_h)$, and the buyer pays the supplier $R_s(\mu^*(b_h))$ per unit. Interestingly, from the buyer’s perspective this mechanism is equivalent to the supply chain coordination mechanism. However, now the supply chain optimal capacity is chosen only when $b = b_h$.

the contract after the supplier has revealed his private information. Furthermore, the pattern of orders submitted by the buyer should provide an indication (albeit not perfect) to the supplier as to whether the buyer is indeed implementing a base-stock policy. If the buyer is indeed unable to commit to implementing a base-stock policy, then the buyer can implement a late-fee mechanism in which the late fee is only collected on the first outstanding order, not all outstanding orders. A base-stock policy is indeed optimal with that approach to the late fee. While this modified late-fee mechanism provides good results for the buyer, it is not as effective as the original late-fee mechanism. Therefore, it is in the buyer’s interest to commit to implementing a base-stock policy. Details are available from the authors.

\(^4\) Enforcement may also be easier if the firms agree to payments that include a nonlinear menu based on each realized lead time delivered. For example, there would be no penalty if the lead time is $(\mu_s - \lambda)^{-1}$ or earlier, but a convex and increasing penalty is imposed for longer lead times (the supplier is then compensated so that the expected profit remains unchanged). With this scheme, the buyer need only verify to courts the duration of each realized lead time, which seems easier than verifying capacity. Assuming the penalty scheme is appropriately designed, the supplier will not deviate from the specified capacity level.

\(^3\) We note that the optimal mechanism also requires a commitment on the part of the buyer, a commitment not to renegotiate
6. Competitive Bidding \( (n \geq 2) \)

Now suppose there are at least two potential suppliers, so competitive bidding is possible. We evaluate an optimal mechanism and several simpler mechanisms that implement second-bid auctions, e.g., a sealed-bid auction in which the winner pays the second-highest bid or an open outcry ascending-bid auction (an English auction).

6.1. Optimal Mechanism (OM)

Similar to the case of \( n = 1 \), when \( n \geq 2 \) the buyer offers to the suppliers a menu, \( \{q^i_r(\cdot), \mu^i_r(\cdot), R^i_r(\cdot)\} \), where \( i \in [1, n] \): Supplier \( i \) is the winner with probability \( q^i_r(b) \geq 0 \), where \( b = (\hat{b}^1, \ldots, \hat{b}^n) \) is the vector of announced costs and \( \sum q^i_r(b) = 1 \); supplier \( i \) receives a unit price \( R^i_r(b) \) from the buyer; the winner builds capacity \( \mu^i_r(\hat{b}) \); and the losers do nothing but enjoy their payment.

Consider the suppliers’ bidding behavior. Supplier \( i \) maximizes her own expected profit:

\[
\max_{\hat{b}^i} \pi^i = E_{\hat{b}^i} \left[ R^i_r(b) - q^i_r(b) \hat{b}^i \mu^i_r(\hat{b}) \right].
\]

According to the revelation principle, we need only consider truth-telling mechanisms,

\[
b^i = \arg \max_{\hat{b}^i} \pi^i(\hat{b}^i). \tag{7}
\]

The individual rationality constraints are

\[
\pi^i(b^i) \geq 0. \tag{8}
\]

Let \( b = (\hat{b}^1, \ldots, \hat{b}^n) \) be the true cost vector. The buyer’s problem is

\[
\min_{\{\varphi_i(\cdot), \mu_r(\cdot), R_r(\cdot)\}} E_{\hat{b}} \left\{ \sum_i R^i_r(b) - \sum_i q^i_r(b) c(\mu^i_r(b)) \right\} \tag{9}
\]

s.t. (7) and (8).

The following theorem gives the solution to (9).

**Theorem 3.** If \( F(\cdot) \) is log-concave, then in the optimal mechanism for \( n \geq 2 \) heterogeneous suppliers, the suppliers announce their true costs and the most efficient supplier is chosen. The same menu is offered to the suppliers with functions given by

\[
q_r(\hat{b}) = \begin{cases} 1 & \text{if } \hat{b} = \min(\hat{b}^1, \ldots, \hat{b}^n) \\ 0 & \text{otherwise} \end{cases}, \quad c'(\mu_r(x)) = -F(x)/f(x),
\]

\[
R_r(x) = (1 - F(x))^{n-1} x \mu_r(x) + \int_x^{\hat{b}} (1 - F(y))^{n-1} \mu_r(y) \, dy.
\]

From Theorems (2) and (3) we see that the incentive scheme (i.e., the capacity function \( \mu_r(x) \)) applies for all \( n \), so, again the optimal mechanism results in less capacity than optimal for the supply chain. We numerically evaluate the functions in Theorem 3.

This optimal mechanism is strange because the losers receive a payment even though they do not build any capacity. However, it is possible to show that the optimal mechanism can be implemented so that only the winner receives a payment (see Zhang 2004).

6.2. A Scoring-Rule Auction (SA)

With a scoring-rule auction, suppliers submit bids that contain a price and a lead time, and the buyer evaluates these bids by assigning each bid a value via a publicly announced function (i.e., the scoring rule). Because from the buyer’s perspective there is a one-to-one relationship between lead time and capacity, we shall assume, without loss of generality, that the suppliers submit \( \{\mu, R\} \) bids, i.e., a capacity and a unit price. Let \( Y(\mu, R) \) be the buyer’s scoring rule and let \( Y_i \) be the \( i \)th highest score. The winner is the supplier whose bid has the highest score and the winner chooses any \( \{\mu, R\} \) pair such that \( Y_2 = Y(\mu, R) \) (i.e., the winner does not have to exactly match the second-best bid, he matches the second-best bid’s score). There are many possible scoring rules, but we work with an intuitive one: Let the buyer’s scoring rule be the buyer’s total cost,

\[
Y(\mu, R) = c(\mu) + R\lambda.
\]

Therefore, the highest score refers to the lowest total cost. Due to the next lemma, we can think of the suppliers as if they are bidding on \( b \mu^*(b) + \pi \), the supply chain’s optimal capacity cost plus a profit.

**Lemma 1.** In a scoring-rule auction with the buyer’s total cost as the scoring rule, \( Y(\mu, R) \), the dominant strategy for a supplier with cost \( b \) is to bid the supply chain optimal capacity \( \mu^*(b) \).

**Theorem 4.** Consider the total cost, \( Y(\mu, R) \), scoring-rule auction. For each supplier it is a dominant strategy to bid according to \( \mu_r(x) = \mu^*(x) \) and \( R_r(x) = x \mu^*(x)/\lambda \).
Che (1993) shows that this scoring rule is not optimal for the buyer (the buyer is better off distorting the supplier to a lower-than-optimal capacity). Nevertheless, we present this scoring rule as a simple and intuitive alternative to the optimal mechanism.

6.3. Lead-Time Mechanism (LT) with a Price Auction
One idea to further simplify the scoring-rule auction is to reduce its dimensionality: Fix one of the dimensions and have the suppliers bid on the other dimension. In the lead-time mechanism with a price auction, the buyer announces the lead time the selected supplier must deliver, and the selected supplier is the winner of a price auction. (This is the natural extension of the lead-time mechanism with one potential supplier to \( n \geq 2 \) potential suppliers.) As already mentioned, FreeMarkets runs auctions like this (Rangan 1998). However, Asker and Cantillon (2004) demonstrate that a scoring-rule auction always performs better for the buyer than a mechanism like this one, in which a subset of the parameters is fixed (here, lead time) and the bidders bid on the remaining dimensions (here, price). As before, we analyze this mechanism as if the buyer announces a required capacity, \( \rho \), instead of a lead time.

**Theorem 5.** Consider the lead-time mechanism with price auction. The weakly dominant strategy for a supplier is to bid \( R_r(x) = \mu x / \lambda \). The buyer’s expected total cost is convex in \( \rho \).

In the price auction, because the required capacity is given, the suppliers essentially bid on their profit, so in a second-bid auction they bid their break-even price. Given that the buyer’s total cost is convex in \( \rho \), a numerical search finds the optimal lead time.

6.4. Late-Fee Mechanism (LF) with a Price Auction
In the late-feee mechanism with a price auction, the buyer charges the winner of the price auction the late fee \( \eta_f \) per outstanding order per unit time. This is similar to the lead-time mechanism in that the selection of the supplier is based only on the suppliers’ price bids, but it is different in that now the winning supplier is free to choose his capacity/lead time to minimize his own costs. Because the winner’s price bid does not influence his capacity choice, the winner chooses capacity \( \bar{\mu}(b) \). As a result, the suppliers effectively bid their capacity cost, \( b \bar{\mu}(b) \), plus a profit. As with one potential supplier, \( \eta_f \) is not the buyer’s optimal late fee, but we show in §7 that it is quite good. The results for this mechanism are summarized in the following theorem.

**Theorem 6.** Consider the late-fee mechanism with a price auction and the late fee \( \eta_f \). The suppliers’ dominant strategy is to bid \( R_r(x) = \bar{C}(x) / \lambda \), and the winner chooses capacity \( \mu_*(x) = \bar{\mu}(x) \).

The results of Asker and Cantillon (2004) do not apply with this mechanism because the late fee does not restrict the supplier’s capacity choice. Hence, it is possible that the late-fee mechanism outperforms the scoring-rule auction, and this is confirmed in our numerical study, §7.

6.5. The Optimal Mechanism Compared to the Late-Fee Mechanism and the Supply Chain Optimal Solution
Although the optimal mechanism in Theorem 3 can be evaluated numerically, the theorem provides little insight into its qualitative features. In particular, we are interested in how the optimal mechanism compares relative to the late-fee mechanism (i.e., what is the potential gain of using a complex mechanism relative to a simpler mechanism) and how it compares to the supply chain optimal solution (i.e., how much does it increase the supply chain’s total cost). We are unable to answer such questions with the actual cost functions and general distributions for \( b \). However, in this section we obtain insights by using the exponential approximation of the cost functions and assuming \( b \) is uniformly distributed on the interval \( [b_1, b_2] \). Our numerical study confirms that these insights apply more generally.

Let \( b_1 = \theta(1 - \delta) \) and \( b_2 = \theta(1 + \delta) \), so \( \theta \) is the expected capacity cost and \( \delta \) is the maximum percentage variation about that mean. It follows that \( F(x) = (x - b_1) / 2\theta \) and \( f(x) = 1 / 2\delta \). From Theorem 3, with the optimal mechanism

\[
\hat{c}'(\mu_*) = -b - F(b) / f(b) = -2b + b_1.
\]

Replacing \( c(\mu_*) \) with \( \hat{c}(\mu_*) \), (10) yields the supplier’s capacity in the optimal mechanism

\[
\hat{\mu}_*(b) = \lambda + \sqrt{\frac{\alpha}{2b - b_1}}.
\]

The buyer’s operating cost is then

\[
\hat{c}(b) = \hat{c}(\hat{\mu}_*(b), b) = \sqrt{\alpha(2b - b_2)}.
\]

In this section we work with the implementation of the optimal mechanism described in Theorem 3, in which only the winner gets paid (see Zhang 2004 for details). Let \( b_1 \) and \( b_2 \) (\( b_1 \leq b_2 \)) be the lowest and second-lowest costs, respectively. The lowest-cost supplier is chosen, builds capacity \( \mu_*(b_1) \), and receives a payment

\[
b_1 \mu_*(b_1) + \int_{b_1}^{b_2} \mu_*(y) dy,
\]

which is similar to the single-supplier case, but now the upper limit of the integral is truncated at \( b_2 \) instead of \( b_2 \). This is expected because the single-supplier case is equivalent to the multiple-supplier case with \( b_2 = b_2 \).
Based on the exponential approximation, replace $\mu_i$ with $\hat{\mu}_i$ in (11). The buyer’s total cost with the optimal mechanism is then

$$\hat{C}_o(b_1, b_2) = \hat{c}(b_1) + b_1\hat{\mu}_i(b_1) + \int_{b_1}^{b_2} \hat{\mu}_i(y) dy$$

$$= \lambda b_2 + \sqrt{\alpha \left( \frac{b_1}{\sqrt{2b_1 - b_1}} + \sqrt{2b_2 - b_1} \right)}.$$  

With the late-fee mechanism, the buyer’s payment to the supplier is $R_{k\lambda} = \hat{C}(b_1) = b_2\lambda + 2\sqrt{\alpha b_2}$, the buyer’s operating cost is approximately $\hat{c}(\mu_i(b_1), b_1) = \sqrt{\alpha b_1}$, but the late fees collected are also $\sqrt{\alpha b_1}$, so the buyer’s total cost is

$$\hat{C}_f(b_2) = \hat{C}(b_2) = b_2\lambda + 2\sqrt{\alpha b_2};$$

the buyer’s total cost is independent of $b_1$, which implies that with the late-fee mechanism the buyer is unable to extract any rents from the most efficient supplier relative to those of the second-most efficient supplier.

Now let us compare the optimal and late-fee mechanisms. It is straightforward to show that $\hat{C}_o(b_1, b_2)$ is increasing in $b_1$. Hence, for any given $b_2$, $\hat{C}_o(b_1, b_2)$ ranges from a minimum of $\hat{C}_o(b_1, b_2)$ to a maximum of $\hat{C}_o(b_2, b_2)$. It is also straightforward to show that $\hat{C}_f(b_1, b_2) < \hat{C}(b_1) < \hat{C}_o(b_1, b_2)$. Therefore, if $\hat{C}_o(b_1, b_2)$ is relatively flat in $b_1$ for any given $b_2$, then the buyer’s expected cost with the optimal mechanism, $E[\hat{C}_o(b_1, b_2)]$, is approximately equal to the buyer’s expected cost with the late-fee mechanism, $E[\hat{C}_f(b_2)]$. To see that $\hat{C}_o(b_1, b_2)$ is flat, consider the following ratio:

$$\frac{\hat{C}_o(b_2, b_2)}{\hat{C}_o(b_1, b_2)} = \frac{b_2\lambda + \sqrt{\alpha b_2(\sqrt{2b_2 - b_1} + \sqrt{2b_2 - b_1})}}{b_1\lambda + \sqrt{\alpha (\sqrt{b_1} + \sqrt{2b_2 - b_1})}}$$

$$< \frac{b_2/\sqrt{2b_2 - b_1} + \sqrt{2b_2 - b_1}}{\sqrt{b_1} + \sqrt{2b_2 - b_1}},$$  

where the inequality follows because $b_2 > \sqrt{b_1}$.

$\sqrt{2b_2 - b_1}$ implies the left-hand side is decreasing in $\lambda$ (so let $\lambda \to 0$). The right-hand side of (12) is increasing in $b_2$, so given that $b_2 \leq b_h$,

$$\frac{\hat{C}_f(b_2, b_2)}{\hat{C}_o(b_1, b_2)} < \frac{b_h/\sqrt{2b_h - b_1} + \sqrt{2b_h - b_1}}{\sqrt{b_1} + \sqrt{2b_h - b_1}}.$$  

Use $b_1 = \theta(1 - \delta)$ and $b_h = \theta(1 + \delta)$ to obtain

$$\frac{\hat{C}_o(b_2, b_2)}{\hat{C}_o(b_1, b_2)} < \frac{2 + 4\delta}{\sqrt{(1 + 3\delta)(1 - \delta) + (1 + 3\delta)}}.$$  

The right-hand side of (13) equals 1.025 and 1.075, with $\delta = 0.2$ and $\delta = 0.4$, respectively. Therefore, even if the supplier’s cost can vary up to 40% around its mean ($\delta = 0.4$) and $\lambda \approx 0$, the buyer’s cost with the late-fee mechanism cannot be more than 7.5% higher than with the optimal mechanism for any value of $b_2$. Note that the bound in (13) is evaluated by assuming $b_2 = b_h$. Therefore, the buyer’s total cost in the optimal mechanism tends to be flatter as the number of suppliers increases because $b_2$ is more likely to be small.

Why is the buyer’s total cost relatively insensitive to the supplier’s capacity cost with the optimal mechanism? It is not because the supply chain’s cost function is flat. With the supply chain’s optimal cost we obtain the following ratio:

$$\frac{\hat{C}(b_2)}{\hat{C}(b_1)} = \frac{(1 + \delta)\lambda + 2\sqrt{\alpha(1 + \delta)}}{(1 - \delta)\lambda + 2\sqrt{\alpha(1 - \delta)}} > \sqrt{\frac{1 + \delta}{1 - \delta}}.$$  

The right-hand side in (14) is 1.22 and 1.53 with $\delta = 0.2$ and $\delta = 0.4$, respectively. Therefore, when there is a 40% variation in cost, the supply chain’s cost with the least efficient supplier is at least 53% higher than with the most efficient supplier. Given that the buyer is unable to extract rents from low-cost suppliers, it must be that the supplier earns most of the efficiency rents. In other words, asymmetric information conveys substantial protection to suppliers: Even when the buyer implements the optimal mechanism, a low-cost supplier is able to keep most of the benefit from being a low-cost supplier (relative to the second-best supplier in an auction and relative to the worst-possible supplier, $b_h$, when there is no bidding).

This implies that the late-fee mechanism is effective, because the optimal mechanism is not very effective: The late-fee mechanism extracts no rents from a low-cost supplier and the optimal mechanism extracts not much more.

Now consider the relationship between the supply chain’s total cost with the buyer’s optimal mechanism as compared to the supply chain minimum total cost. Note that the magnitude of the gap $\hat{\mu}_i(b) - \hat{\mu}(b)$ increases with $b$, so the greatest supply chain inefficiency occurs with $b = b_h$. The following ratio indicates the magnitude of loss in supply chain performance due to the implementation of the optimal mechanism:

$$\frac{\hat{C}(\hat{\mu}_i(b_h))}{\hat{C}(\hat{\mu}(b_h))} = \frac{b_h\lambda + \sqrt{\alpha(\sqrt{2b_h - b_1} + \sqrt{2b_h - b_1})}}{b_2\lambda + \sqrt{\alpha(2\sqrt{b_h})}}$$

$$< \frac{1 + 2\delta}{\sqrt{1 + \delta} \sqrt{1 + 3\delta}}.$$  

For $\delta = 0.2$ and $\delta = 0.4$, the right-hand side in (15) is 1.010 and 1.026, respectively. Given that those are upper bounds, in expectation the optimal mechanism nearly yields the supply chain minimum cost. This
implied that even with the optimal mechanism there is little value to renegotiation.

Figure 1 displays these results graphically for the single-supplier case by setting $\lambda \approx 0$ and normalizing $\sqrt{a\theta} = 1$. The buyer’s operating cost with the optimal mechanism decreases significantly as the supplier becomes more efficient (as $b$ decreases), but the buyer’s procurement cost increases nearly as much, hence, the buyer’s total cost is nearly flat. Thus, as $b$ decreases, most of the supply chain’s efficiency gains accrue to the supplier, as can be seen by the supplier’s increasing profit.

7. Numerical Study
This section reports on a numerical study of the procurement strategies analyzed in the previous two sections. We first constructed 144 scenarios from all combinations of the following parameters: $h = 1$, $\lambda \in \{0.1, 1, 10, 100\}$, $p \in \{3, 40, 200\}$, $b$ is uniformly distributed on the interval $[b_l, b_u]$ where $b_l = \theta(1 - \delta)$ and $b_u = \theta(1 + \delta)$, $\theta \in \{0.5, 5, 50, 200\}$, and $\delta \in \{0.05, 0.1, 0.2\}$. We take the scenarios with $\delta = 0.05$ to represent reasonably low uncertainty with respect to the suppliers’ cost (within 5% of forecast), and the scenarios with $\delta = 0.20$ to represent high uncertainty (it is unlikely that qualified suppliers would have costs that range more than 20% from the buyer’s forecast).

We fix $h$ to a single value, because it is easy to show that the buyer’s cost depends on the ratios $p/h$ and $b/(p/h)$, so it is sufficient to vary $p$ and $b$ and hold $h$ fixed. Because backorder penalty costs are generally higher than holding costs, we allow $p$ to range from a low value of three times $h$ to a high value of two hundred times $h$. Similarly, because of economies of scale in queuing systems, we range the demand rate from a low of 0.1 to a high of 100. Capacity costs range from very low, $\theta = 0.5$, which generally results in low utilizations, to very high, $\theta = 200$, which generally results in high utilizations. Table 1 contains a list of the strategies we evaluate and the mnemonics we use to identify each one in the subsequent tables. With the competitive bidding strategies, we assume there are two potential suppliers.

Table 2 provides data on the performances of each strategy relative to the OM. Provided are the average, 90th percentile, and maximum cost increase of each mechanism relative to the OM across all scenarios. For example, with the LF, for 90% of the scenarios the buyer’s percentage cost increase relative to the OM is only 0.19%. Table 2 shows that with a single potential supplier, both the CC and the LF are nearly optimal. In fact, LF even performs slightly better than CC at the 90th percentile because (we conjecture) it makes

Table 1  Procurement Strategies Evaluated in the Numerical Study

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>OM</td>
<td>OM (minimizes the buyer’s cost)</td>
</tr>
<tr>
<td>CC</td>
<td>Supply chain coordination mechanism (minimizes total supply chain cost)</td>
</tr>
<tr>
<td>LF</td>
<td>Late-fee mechanism: The buyer pays $R_i$ per unit but charges the supplier the late fee $p_i$ per outstanding order per unit time. With one potential supplier, $R_i$ is chosen by the buyer, otherwise it is chosen via a price auction.</td>
</tr>
<tr>
<td>LT</td>
<td>Lead-time mechanism: The buyer pays $R_i$ per unit and requires the supplier to achieve the lead time $(\mu_i - \lambda)^{-1}$. With one potential supplier, $R_i$ is chosen by the buyer, otherwise it is chosen via a price auction.</td>
</tr>
<tr>
<td>SA</td>
<td>Scoring-rule auction: The scoring-rule is $Y_f(\mu, R)$, which is the buyer’s cost with unit price $R$ and capacity $\mu$; the winning supplier bids the lowest score and then must achieve his bid.</td>
</tr>
</tbody>
</table>

Table 2  Percentage Cost Increase Relative to the OM in 144 Scenarios

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Average (%)</th>
<th>90th percentile (%)</th>
<th>Maximum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single supplier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC and LT</td>
<td>0.09</td>
<td>0.24</td>
<td>0.56</td>
</tr>
<tr>
<td>LF</td>
<td>0.20</td>
<td>0.19</td>
<td>2.85</td>
</tr>
</tbody>
</table>

Multiple suppliers

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Average (%)</th>
<th>90th percentile (%)</th>
<th>Maximum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>0.16</td>
<td>0.26</td>
<td>0.47</td>
</tr>
<tr>
<td>LT</td>
<td>0.18</td>
<td>0.31</td>
<td>0.59</td>
</tr>
<tr>
<td>LF</td>
<td>0.32</td>
<td>0.31</td>
<td>3.27</td>
</tr>
</tbody>
</table>

Notes. $h = 1$, $\lambda \in \{0.1, 1, 10, 100\}$, $p \in \{3, 40, 200\}$, $b$ is uniformly chosen via a price auction, $b_l = \theta(1 - \delta), b_u = \theta(1 + \delta)$, $\theta \in \{0.5, 5, 50, 200\}$, and $\delta \in \{0.05, 0.1, 0.2\}$. 

Figure 1 With One Potential Supplier, $\lambda \approx 0$ and $\delta = 40\%$, the Optimal Mechanism (OM) Relative to the Late-Fee Contract (LF) and the Supply Chain’s Minimum Cost

Note. All functions are based on the exponential approximation and normalized so that $(a\theta)^{1/2} = 1$. 

With One Potential Supplier, $\lambda \approx 0$ and $\delta = 40\%$, the Optimal Mechanism (OM) Relative to the Late-Fee Contract (LF) and the Supply Chain’s Minimum Cost
the supplier build less capacity than optimal, just like the OM. We found that \( \eta_f \) is nearly the optimal late fee, so there is little value in numerically searching for the optimal LF. Figure 2 illustrates this result for a sample of the scenarios. It also illustrates that it is possible to increase costs substantially with a poorly chosen late fee, especially a late fee that is too low.

With two potential suppliers, the SA is no longer optimal, but it still generates total costs that are close to the OM. Both the LTs and the LFs generate good results for the buyer. LF is least effective when both the capacity cost and the demand rate are very low (e.g., \( \theta = 0.5, \lambda = 0.1 \)) because then the exponential approximation is not accurate due to very low system utilization. We find that the buyer is much better off with two potential suppliers than only one potential supplier (when OMs are used): On average, the buyer’s cost is 5.3% lower with two heterogeneous suppliers relative to just one potential supplier.

Overall, we see from Table 2 that the LFs and LTs perform quite well. To test whether this continues to hold under extreme conditions, we repeated our previous numerical study but only with \( \delta = 0.4 \). With these additional 48 scenarios Table 3 reports that our simple mechanisms are still close to optimal. The LF performs even better than the SA in this set of scenarios.\(^6\) We also evaluated two ratios in these scenarios. The first is the (expected) ratio between the supply chain’s capacity cost and the supply chain’s total cost in the OM with a single supplier: This ratio ranges from 27% to 99%, with mean 73%. As a result, the strong performance of the simple mechanisms is not limited to situations in which the capacity cost represents a negligible share of the supply chain’s total cost. The second ratio is the supplier’s capacities with the lowest and highest costs in the OM, i.e., \( \mu_s(b_l)/\mu_s(b_h) \). We find that the capacity ratios in our numerical experiments are much smaller than the ratio \( b_h/b_l = 1.4\theta/0.6\theta = 2.33 \): On average, \( \mu_s(b_l) \) is only about 25% higher than \( \mu_s(b_h) \), despite the fact that \( b_h \) is 133% higher than \( b_l \). In other words, with the OM the capacity function is much less variable than the range of capacity costs, which is why the LT, which specifies a single capacity, works well.

To provide another robustness test, we constructed another set of 144 scenarios that are identical to the first except that in each scenario the capacity cost distribution is changed from a uniform distribution with mean \( \theta \) and range \([\theta(1-\delta), \theta(1+\delta)]\) to a normal distribution with mean \( \theta \) and standard deviation \( \delta \theta/4 \). Table 4 summarizes those results. In short, both mechanisms continue to perform well.

Table 5 provides data on the incentive to renegotiate with the mechanisms that do not coordinate the supply chain. We see that the OMs do create some opportunity for renegotiation, but that opportunity is generally relatively small (less than 1% for all scenarios). The LF also presents a small opportunity in most of the scenarios, except if the system utilization is very low (e.g., when the capacity cost and demand rate are very low). Again, this is because \( \eta_f \) is derived from the exponential approximation, which is less accurate for systems with very low utilization. The renegotiation opportunity with the LTs is comparable to the LFs.

\(^6\) Recall that the results in Asker and Cantillon (2004) demonstrate that the LT is always worse than the SA mechanism, but LF can be better because it does not fix the supplier’s capacity decision.
To summarize, we observe in an extensive numerical study that the LTs and LFs perform for the buyer nearly as well as the OMs’ and they generally create a relatively small renegotiation opportunity (i.e., they nearly coordinate the supply chain). These results are entirely consistent with the analytical observations obtained in §6.5 even though those results were derived from approximations.

8. Two Extensions

This section discusses two extensions to the model: The buyer’s holding or backorder cost is allowed to vary linearly with the buyer’s procurement cost or the buyer operates in a make-to-order fashion, so the buyer does not hold inventory. The latter case is useful for the analysis of a service that has been outsourced to the supplier.

8.1. Generalized Costs

So far in our analysis we have assumed that the buyer’s backorder and holding costs are fixed at rate \( p \) and \( h \), respectively. As already mentioned, this is reasonable if the physical costs of holding inventory dominate the financial holding costs and if the buyer’s margin, which presumably influences the backorder cost in some manner, is independent of the purchase price per unit. Our assumption is also correct in an important special case: The backorder cost is proportional to the buyer’s markup, and the holding cost is proportional to the buyer’s procurement cost. To explain, suppose \( p = (m - 1)v \), where \( v \) is the buyer’s procurement cost for each unit and \( m \) is the markup the buyer uses to set the price of the good, i.e., the buyer’s price to its customers is \( mv \). Furthermore, let \( r \) be the buyer’s return on capital and \( h = rv \), i.e., the holding cost is composed only of the financial opportunity cost of the capital tied up in inventory. Then the \( p/h \) ratio is \( (m - 1)v/rv = (m - 1)/r \), which is independent of the actual procurement price. Recall that the ratio \( p/h \) is sufficient for our analysis, so we need not be concerned with how \( p \) and \( h \) vary independently.

Although there are situations in which it is reasonable to assume constant \( p \) and \( h \), there are also cases in which the \( p/h \) ratio varies with the procurement price (such as when there are fixed goodwill penalties associated with the backorder cost or when there are fixed holding costs per unit that are independent of the purchase cost). To test the robustness of our results, we considered a model in which \( p \) is held constant and \( h = h_0 + rv \), where \( h_0 \) is a constant representing the physical holding cost, \( r \) is the interest rate, and \( v \) is the buyer’s unit cost, which may differ from the unit price \( R \). For example, the buyer’s unit cost with a late fee is \( R \) minus the late fee per unit. Unfortunately, the evaluation of the OM with this new holding cost structure is quite difficult. There are several complications. First, the buyer’s operating cost is not always jointly convex in \( \mu \) and \( v \), nor everywhere differentiable, which prevents finding solutions via first-order conditions. Second, the transfer payment and the operating cost are no longer separable (i.e., the transfer payment can no longer be chosen arbitrarily for a given capacity and operating cost), which significantly complicates the evaluation of the optimal transfer payment and capacity function. As a result, full enumeration over the contract space is required to evaluate an OM. Hence, we can only determine the OM when the suppliers’ costs are drawn from a discrete distribution and the suppliers are only allowed to choose capacities from a discrete set. Under these conditions we considered the 144 scenarios described in the previous section, each with three different interest rates. Our numerical results generally yielded the same insights as we found with the fixed holding cost, so we conclude that our findings are robust even to generalized holding cost structures. Specific details on those results are available from the authors.

8.2. Make-to-Order Buyer

If the buyer is a make-to-order manufacturer or a service provider, then the buyer is unable to hold buffer inventory to mitigate the consequence of slow delivery. Hence, we investigate whether the LTs and LFs perform well in this setting. The buyer’s operating cost is \( c(\mu) = \lambda p/(\mu - \lambda) \). Theorem 3 still applies because \( c(\mu) \) is convex, so we can evaluate the OM for this case. It is easy to show that the buyer’s optimal LT has \( \mu_b = \lambda + \sqrt{p\lambda/b_h} \) and the unit price \( R_b(\mu_b) = b_h + \sqrt{p b_h/\lambda} \).
To choose a late fee we find the buyer’s optimal late fee with one potential supplier. With the supplier’s optimal capacity, $\mu_f(b)$, the supplier’s profit is

$$\pi(\mu_f, b) = R\lambda - \left(b\lambda + 2\sqrt{\eta_f\lambda b}\right).$$

Setting $\pi(\mu_f, b_0) = 0$ gives the optimal transfer price:

$$R_f = b_0 + 2\sqrt{\eta_f b_0 / \lambda}.$$

The buyer’s expected cost with $\{\eta_f, R_f\}$ is

$$C_f = b_0 \lambda + 2\sqrt{\eta_f \lambda b_0} + (p - \eta_f)\sqrt{\lambda / \eta_f}E(\sqrt{b}),$$

which is convex in $\eta_f$ and minimized by

$$\eta_f = \frac{E(\sqrt{b})}{2\sqrt{b_0} - E(\sqrt{b})}p. \tag{16}$$

With multiple suppliers the unit price is determined via an auction.

Table 6 reports that both the LTs and the LFs perform well relative to the OM with the scenarios defined in §7, and Table 7 reports that the renegotiation opportunities are generally small with either mechanism. We conclude that the simple mechanisms perform well even if the buyer is unable to use inventory to buffer the supplier’s lead time performance.

9. Discussion

A buyer procures a component from a single supplier whose capacity cost is unknown to the buyer. There are two tasks in the buyer’s procurement strategy, supplier selection (which supplier to source from) and contract terms (how much to pay the supplier). Two situations are considered: With one potential supplier, the buyer need only choose contract terms, whereas with two or more potential suppliers the two procurement tasks (selection and contract terms) are bundled.

We identify optimal procurement strategies for the buyer and provide alternative strategies as well—in particular, simple mechanisms with a few fixed parameters rather than menus of nonlinear functions. We judge each mechanism along two key dimensions: how well it minimizes the buyer’s total cost (procurement plus operating) and how well it minimizes the supply chain’s cost. If a mechanism does not score well on the latter dimension, then the mechanism may not be implementable due to the threat of renegotiation. Our main finding is that there exist simple mechanisms that are effective along both dimensions. One is a late-fee mechanism: The buyer charges the supplier a fixed late fee for on-order units and either sets the unit price (with one potential supplier) or conducts an auction to set the unit price. The other is a lead-time mechanism: The buyer sets a fixed lead-time requirement and uses the same procedure as the late-fee mechanism to set the unit price.

We show that with the simple late-fee and lead-time mechanisms the buyer’s total cost is completely insensitive to the supplier’s capacity cost, i.e., a low-cost supplier is able to retain all of the rents from being a low-cost supplier. In other words, asymmetric information provides significant protection to the supplier. The optimal mechanism indeed reduces that protection: With the optimal mechanism the buyer is able to capture some of the rents from a low-cost supplier. Unfortunately, the optimal mechanism captures very little: We show that the buyer’s total cost is relatively insensitive to the supplier’s capacity cost information even in the optimal mechanism. Hence, we conclude that simple mechanisms are effective because the optimal mechanism is almost ineffective.

If a buyer is not able to overcome an information handicap over a supplier by implementing a sophisticated contracting mechanism, what is a buyer to do? Our results indicate that the buyer’s procurement cost is sensitive to the capacity cost of the least efficient supplier the buyer is willing to procure from, $b_h$. Hence, the buyer should be able to earn significant returns from any activity that reduces uncertainty in the supplier’s cost, e.g., visiting the potential supplier to fully understand the supplier’s operating process, technology, and cost structure. Alternatively, the

Table 6 With $s = 0$, Percentage Cost Increase Relative to the OM in 144 Scenarios

<table>
<thead>
<tr>
<th></th>
<th>Average (%)</th>
<th>90th percentile (%)</th>
<th>Maximum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single supplier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT</td>
<td>0.13</td>
<td>0.40</td>
<td>0.55</td>
</tr>
<tr>
<td>LF</td>
<td>0.03</td>
<td>0.08</td>
<td>0.12</td>
</tr>
<tr>
<td>Multiple suppliers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT</td>
<td>0.21</td>
<td>0.43</td>
<td>0.51</td>
</tr>
<tr>
<td>LF</td>
<td>0.14</td>
<td>0.20</td>
<td>0.24</td>
</tr>
</tbody>
</table>

Table 7 With $s = 0$, the Percentage Cost Increase of the Supply Chain’s Total Cost over the Supply Chain’s Minimum Cost, i.e., the Value of Renegotiation

<table>
<thead>
<tr>
<th></th>
<th>Average (%)</th>
<th>90th percentile (%)</th>
<th>Maximum (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single supplier</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT</td>
<td>0.15</td>
<td>0.49</td>
<td>0.58</td>
</tr>
<tr>
<td>LT max</td>
<td>0.53</td>
<td>1.74</td>
<td>2.05</td>
</tr>
<tr>
<td>LF</td>
<td>0.10</td>
<td>0.33</td>
<td>0.39</td>
</tr>
<tr>
<td>LF max</td>
<td>0.11</td>
<td>0.33</td>
<td>0.39</td>
</tr>
<tr>
<td>Multiple suppliers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LT</td>
<td>0.09</td>
<td>0.29</td>
<td>0.35</td>
</tr>
<tr>
<td>LT max</td>
<td>0.26</td>
<td>0.87</td>
<td>1.03</td>
</tr>
<tr>
<td>LF</td>
<td>0.10</td>
<td>0.33</td>
<td>0.39</td>
</tr>
<tr>
<td>LF max</td>
<td>0.11</td>
<td>0.33</td>
<td>0.39</td>
</tr>
</tbody>
</table>

Notes. The subscript $e$ denotes the expected supply chain inefficiency ex ante, the subscript max denotes the maximum possible supply chain inefficiency ex post.
buyer’s information disadvantage is reduced if the buyer is willing and able to have multiple potential suppliers because auctions effectively reduce the buyer’s information handicap.

We recognize that none of the mechanisms we present are perfect. For example, with the late-fee mechanism the buyer must commit to implementing a base-stock policy, and with the lead-time mechanism the firms must be reasonably confident that the supplier does not significantly distort the supplier’s capacity decision, so there is little renegotiation opportunity (the supplier’s capacity cost). Given that the buyer is unable to extract rents from low-cost suppliers, the buyer does not significantly distort the supplier’s capacity decision, so there is little renegotiation opportunity.

To summarize, this research is about how a buyer should procure when both procurement and operating costs are important. It has been frequently articulated in the procurement literature that a buyer should not focus on just the purchase price, but rather on the total procurement cost. Unfortunately, there has been no rigorous analysis of how a buyer should go about balancing price with operating costs. The mechanism design literature suggests an approach that uses a menu of contracts to minimize the buyer’s total cost, albeit at the expense of supply chain inefficiency. The supply chain coordination literature seeks to maximize the supply chain’s efficiency, but ignores the likely possibility of asymmetric information. Neither approach (mechanism design or supply chain coordination) values a simple design explicitly. Our practical approach is a blend of all three. For both simplicity and outstanding performance (the buyer’s and the supply chain’s), we recommend either the lead-time or the late-fee mechanisms.

Acknowledgments
The authors would like to thank for their helpful comments Martin Lariviere, William Lovejoy, Serguei Netessine, Erica Plambeck, Tunay Tunca, the associate editor, the anonymous reviewers, and seminar participants at Columbia University, Stanford University, University of Pennsylvania, and the Second MIT Symposium in Operations Research: Procurement and Pricing Strategies to Improve Supply Chain Performance. The first version of this paper was titled "Procuring fast delivery, part II: Sole sourcing with information asymmetry."

Appendix
Proof of Theorem 1. $c(\mu)$ is convex in $\mu$ if $c'(\phi) \geq 0$, where $\phi = \mu/\lambda$,

$$c'(\phi) = h \left[ -\frac{1}{\phi} + \frac{2}{\phi} \ln \phi + \frac{1}{\phi} \ln(\phi - 1) - \frac{2}{\phi - 1} \ln(\phi) \right]$$

where $h = h(h + p)$. $c'(\phi) \geq 0$ if $g(\phi) \geq 0$ for $\phi > 0$, where

$$g(\phi) = 2\phi^2 \ln \phi - 2\phi^3 - 2\phi^2 \ln(\phi - 1) + 2\phi^3 \ln(\phi - 1) - 2\phi^3 \ln(\phi - 1)$$

A simple plot reveals $g(\phi) \geq 0$ for all $\phi > 0$. A more rigorous proof, based on Taylor series expansions, is provided by Zhang (2004).

Proof of Theorem 2. This is a special case of the proof for Theorem 3.

Proof of Theorem 3. The proof is adapted from Laffont and Tirole (1987). We provide a sketch of the proof; Zhang (2004) provides a complete proof. A necessary condition for truth telling is

$$\frac{\partial}{\partial \hat{b}^i} E_{\hat{b}^i}(\hat{b}) = 0, \quad \hat{b}^i = \hat{b}^i \text{ for all } i.$$  \hspace{1cm} (17)

We now assume that $q_i(\cdot)$ and $\mu_i(\cdot)$ are nonincreasing functions in $\hat{b}^i$, and check later that they are indeed nonincreasing in the optimal mechanism. It follows that the first-order condition (17) is sufficient for truth telling (see Zhang 2004).

Define $U^i(\hat{b}^i)$ to be the expected profit for supplier $i$ under truth telling:

$$U^i(\hat{b}^i) = E_{\hat{b}^i} \left[ R_i(\hat{b}) - q_i(\hat{b}) \mu_i(\hat{b}) \right].$$  \hspace{1cm} (18)

From (17) and (18) we have

$$U^i(\hat{b}^i) = E_{\hat{b}^i} \left[ R_i(\hat{b}) - q_i(\hat{b}) \mu_i(\hat{b}) \right].$$  \hspace{1cm} (19)

We can see that $U^i$ is nonincreasing in $\hat{b}^i$, so we can set

$$U^i(\hat{b}_i) = 0, \quad \text{all } i.$$  \hspace{1cm} (20)

The buyer’s problem now is

$$\min_{[q_i(\cdot), \mu_i(\cdot), U^i(\cdot)]} \left\{ E_{\hat{b}} \left[ \sum_i U^i(\hat{b}) + \sum_i [q_i(\hat{b}) \mu_i(\hat{b}) + c(\mu_i(\hat{b}))] \right] \right\}$$

s.t. (19) and (20).
According to Zhang (2004), letting \( \mu^*_i(b) \) be dependent on \( b_i \) \((i \neq i) \) is not optimal. Therefore, the above program can be simplified by only considering functions \( \mu^*_i(b) \) that are functions of \( b' \) only. Once the optimal \( q^*_i(\cdot) \) is given, so that \( Q'(b') = E_{i\rightarrow i} q^*_i(b') \) is given, the optimization with respect to \( \mu^*_i(b') \) can be decomposed into \( n \) programs as follows:

\[
\min \int_{b_1}^{b_n} \left[ U_i'(b') + Q'(b') \left[ b'_i \mu^*_i(b') + c(\mu^*_i(b')) \right] f(b') \right] \, db'
\]

\[
s.t. \quad U_i'(b') = -Q'(b') \mu^*_i(b'),
\]

\[
U_i(b_0) = 0.
\]  

(22)

(23)

(24)

This is a dynamic control problem with \( U_i' \) as the state variable and \( \mu^*_i \) as the control variable. Solving this problem gives

\[
c(\mu^*_i) = -b^i - F(b') / f(b').
\]  

(25)

From the above equation we can derive \( \mu^*_i \) as a function of \( b' \). Because \( \mu^*_i \) is the same for all \( i \), we can drop the superscript. From (23), we have

\[
\int_{b_1}^{b_n} U_i'(b') f(b') \, db' = U_i'(b') F(b') \bigg|_{b_1}^{b_n} - \int_{b_1}^{b_n} F(b') \, dU_i'(b')
\]

\[
= \int_{b_1}^{b_n} \left[ F(b') \frac{Q'(b') \mu^*_i(b')}{b'} \right] f(b') \, db'.
\]

Therefore, the cost function in (22) can be written as

\[
\int_{b_1}^{b_n} Q'(b') \left[ \frac{F(b')}{f(b')} \mu^*_i(b') + b_i \mu^*_i(b') + c(\mu^*_i(b')) \right] f(b') \, db'.
\]

Let \( A'(b') = \frac{F(b') / f(b')}{{\mu^*_i(b')} + b_i \mu^*_i(b') + c(\mu^*_i(b'))} \). From (25), we have

\[
dA' = \mu^*_i(b') \left( 1 + \frac{d}{db'} \frac{F(b')}{f(b')} \right) > 0,
\]

so \( A'(b') \) is increasing in \( b' \). Hence, we should give more weight to \( Q'(b') \) when \( b' \) is small. Because there are \( n \) symmetric suppliers, the optimal \( q^*_i(\cdot) \) must be \( q^*_i(\cdot) = 1 \) if \( b' < \min_{i=1}^n b_i \) and \( q^*_i(\cdot) = 0 \) otherwise. That is, in the optimal mechanism, the most efficient supplier is chosen with probability one. As a result, \( Q'(b') = (1 - F(b'))^{n-1} \)

We can derive the profit function \( U_i' \) from (23) and (24):

\[
U_i'(b') = \int_{b_1}^{b_n} \left[ (1 - F(x))^{n-1} \mu^*_i(x) \right] \, dx.
\]

Again, we can drop the superscript for \( U_i' \). The transfer payment function is therefore given by

\[
R(b') \lambda = (1 - F(b'))^{n-1} b_i \mu^*_i(b')
\]

\[
+ \int_{b_1}^{b_n} \left[ (1 - F(y))^{n-1} \mu_o(y) \right] \, dy.
\]  

\( \Box \)

Proof of Lemma 1. For a fixed score \( y \), the following program determines a supplier’s bid, \( (\mu, R) \), because the supplier’s probability of winning depends only on \( y \):

\[
\max_{\mu, R} \quad \pi = R\lambda - b\mu
\]

\[
s.t. \quad c(\mu) + R\lambda = y.
\]

Substitute the constraint into the profit function shows that the supplier chooses \( \mu \) to minimize the system’s total cost:

\[
\max_{\mu} \pi = y - c(\mu) - b\mu \quad \Rightarrow \quad \min_{\mu} C(\mu, b). \]  

\( \Box \)

Proof of Theorem 4. From Lemma 1, \( \mu^*(x) \) is a supplier’s dominant strategy, so the supplier’s profit is \( RA - b\mu^*(b) \). With second bid it is a dominant strategy to bid the minimum price the supplier is willing to receive, \( R(x) = b\mu^*(x)/\lambda \). \( \Box \)

Proof of Theorem 6. This proof follows the proof of Theorem 4. See Zhang (2004) for details. \( \Box \)

References


