Procurement Mechanism Design in a Two-Echelon Inventory System with Price-Sensitive Demand

Fuqiang Zhang

Olin Business School, Washington University in St. Louis
zhang@olin.wustl.edu

This paper studies a buyer’s procurement strategies in a two-stage supply chain with price-sensitive demand. The buyer procures a product from a supplier and then sells to the marketplace. Market demand is stochastic and depends on the buyer’s selling price. The supplier’s production cost is private information, and the buyer only knows the distribution of the cost. Both the buyer and the supplier can hold inventories to improve service, and a periodic review inventory system is considered. The buyer takes two attributes into consideration when designing the procurement mechanism: quantity attribute (i.e., the total purchase quantity) and service-level attribute (i.e., the supplier’s delivery performance). We first identify the optimal procurement mechanism for the buyer, which consists of a nonlinear menu of contracts for each of the two attributes. It can be shown that the optimal mechanism induces both a lower market demand and a lower service level compared to the supply chain optimum. In view of the complexity of the optimal mechanism, we proceed to search for simpler mechanisms that perform well for the buyer. We find that the above two attributes have different implications for procurement mechanism design: The value of using complex contract terms is generally negligible for the service-level attribute, while it can be highly valuable for the quantity attribute. In particular, we demonstrate that a fixed service-level contract, which consists of a target service level and a price-quantity menu, yields nearly optimal profit for the buyer. Additionally, the price-quantity menu is essentially a quantity discount scheme widely observed in practice.

1. Introduction

Typical U.S. manufacturing firms spend half their revenue or more on goods and services procured from external suppliers (Rangan, 1998 and U.S. Department of Commerce, 2006). No matter how much revenue is pouring in, a firm is not likely to prosper if its spending is undisciplined. It is critical for companies to manage their procurement processes intelligently to maintain a competitive edge in the marketplace. Consider a buyer who obtains an input from an upstream supplier. The buyer can be either a manufacturer or retailer, and the input can be either a component or product. The supplier incurs a linear production cost for the input, which is private information.

Asymmetric information is common in practice when the firms are independent organizations. The

\( ^1 \)Forthcoming in *M&SOM*. The author would like to thank Gerard Cachon, Leon Chu, Marty Lariviere, Jeannette Song, Paul Zipkin, the Associate Editor, the reviewers, and seminar participants at UC Berkeley, UC Irvine, Duke University, Northwestern University, University of British Columbia, UT Austin, the 2008 OCSAMSE conference in Shanghai, China for their helpful comments.
buyer sells the product to customers by setting a market price. Both the buyer and the supplier can hold inventories to improve service performance. The buyer’s objective is to find a procurement mechanism that maximizes expected profit, i.e., the difference between revenues and costs. The revenues are determined by the market price set by the buyer and its associated selling quantity. The buyer’s costs consist of two parts: procurement price and operating cost. The former refers to the total amount the buyer pays the supplier for delivering the products, and the latter includes stock-holding cost and goodwill loss for backlogged demand.

The setting described above reflects a situation faced by many buyers who depend on their suppliers for the delivery of components or products. It is well known that a supplier’s delivery or service lead time should factor prominently in the procurement decision (Cavinato, 1994 and Laseter, 1998). The more responsive the supplier is, the lower the operating cost the buyer has to incur. For example, Sun Microsystems considers procurement cost and delivery performance as the two most important factors when choosing suppliers (Farlow, Schmidt, and Tsay, 1996). In a recent Aberdeen survey, about 70% of companies report that measuring supplier performance, including on-time delivery and fill-rate, is critical to their operations (Kay, 2005). The other major factor the buyer needs to consider is how much to order from the supplier. This decision is essentially the same as the market price chosen by the buyer, because it is the market price that determines the demand level. However, the optimal market price is dependent on the procurement price, which is linked to the supplier’s production cost.

In this paper, we study how the buyer should design procurement process in a supply chain while taking two primary attributes, order quantity and delivery performance, into account. The key to the above problem is the supplier’s private cost information. The supplier’s cost has a two-fold impact on the buyer’s profit: First, it determines the buyer’s procurement price paid to the supplier, which in turn may affect the buyer’s price in the marketplace. Second, it influences the supplier’s inventory holding cost, which in turn may affect the buyer’s operating policy. The buyer’s objective is to design a contract that maximizes her expected profit, under the constraint that the supplier is assured participation. We first derive the optimal procurement mechanism for the buyer. In the optimal mechanism, the supplier is asked to reveal his cost structure and, based on the announced information, the supplier provides certain delivery performance and receives a payment. Not surprisingly, the optimal mechanism is quite complex: It involves nonlinear functions for each of the attributes. Although it is less appealing from a practical point of view, the optimal mechanism serves as a benchmark to evaluate other simpler mechanisms. In addition, by comparing the optimal mechanism to the centralized optimal solution, we may derive some insight into the
impact of asymmetric information on supply chain behavior.

In view of the complexity of the optimal mechanism, we are curious about the performance of mechanisms that are intuitive and relatively easy to implement. We consider a so-called fixed service-level contract. In this contract, the buyer specifies a target level for delivery performance and provides a price-quantity menu from which the supplier can choose. In such a menu, the buyer would order more if the supplier chooses a low procurement price, so it represents a quantity discount scheme. This is also closely related to the service-level contracts in practice where buyers focus on price and quantity negotiations while taking care of the operational performance by specifying a minimum service level. Interestingly, we find that this simpler mechanism performs nearly as well as the optimal mechanism. The maximum profit loss from the simple mechanism relative to the optimal mechanism is 2.24% in an extensive numerical study. Such a finding implies that complex contracting terms on the delivery performance attribute are not necessary, since a fixed service-level target provides excellent results in general.

Given that specifying a fixed service level works well, a natural question arises: What happens if we use a fixed quantity? We study the performances of two additional contracts: The first is called fixed quantity contract, where the buyer specifies a fixed order quantity but offers a menu of service levels; the second is even simpler, where both quantity and service level are fixed (thus we call this a fixed quantity and service-level contract). Unfortunately, it has been found that the simplification of the quantity dimension may lead to remarkable profit losses for the buyer. Moreover, we show that the ability to use cost-contingent quantities is more valuable when market demand is highly sensitive to price. Therefore, the buyer should pay close attention to the quantity attribute when designing procurement mechanisms, especially when facing a price-sensitive market demand.

The basic model can be readily extended in a couple of directions, and we examine the robustness of the results in both extensions. In the first extension, there is a pool of potential suppliers, and the buyer can design a bidding mechanism to choose the most efficient one. In the second, there is only one potential supplier, but the buyer may use a cutoff-level policy in contract design, i.e., the buyer may purposely exclude some high-cost supplier types. Our numerical results confirm that the fixed service-level contract continues to perform very well. Further, we find that supplier competition and cutoff-level policies may significantly improve the performance of the fixed quantity and service-level mechanism. This is because both the bidding mechanism and the cutoff-level policy can be viewed as a screening process that selects certain supplier types for transaction.

This paper makes a two-fold contribution to the literature. First, we study a new contracting
problem in the classic two-echelon inventory system by introducing asymmetric information and price-dependent demand. We identify the buyer’s optimal procurement mechanism and investigate the impact of asymmetric information by comparing the optimal mechanism to the system’s optimal solution under centralized control. Second, we demonstrate that the buyer can achieve nearly optimal profit by using the fixed service-level contract, which specifies a target service level and offers a price-quantity menu to the supplier. This contract is relatively easy to implement as well, because the price-quantity menu is essentially a quantity discount scheme widely observed in practice.

The rest of the paper is organized as follows. The next section reviews the literature. Section 3 describes the model, and Section 4 presents several procurement mechanisms. The performances of these mechanisms are compared in Section 5. Section 6 extends the basic model and Section 7 concludes. All proofs are given in Appendix A of the Electronic Companion.

2. Literature Review

This paper builds upon the multi-echelon inventory management literature initiated by Clark and Scarf (1960). Their seminal work has been followed by a number of studies that focus on centralized multi-echelon inventory systems, the notable ones including Federgruen and Zipkin (1984) and Chen and Zheng (1994). Several papers on decentralized inventory systems have also been reported. Cachon and Zipkin (1999) analyze the competitive behavior of a two-stage serial system and propose a linear transfer payment contract to coordinate the supply chain. Ray, Li, and Song (2005) incorporate pricing decisions for both echelons and allow random delivery lead times. Our paper extends the traditional multi-echelon inventory models by introducing private information and price-dependent demand in a decentralized system. We study a procurement problem from the buyer’s perspective and show that in our model setting, the quantity and service-level attributes have different implications in designing procurement mechanisms.

There is an enormous amount of literature on procurement contract design in economics. Most studies in this literature focus on two issues. The first is how to select a cost-effective supplier from a pool of potential suppliers, and the second is how to induce the selected supplier to optimally invest in related activities such as R&D. See Klemperer (1999) and Laffont and Tirole (1987) for surveys of this literature. In general, papers in this category do not take operations factors such as delivery performance into consideration. There are numerous papers that study a buyer’s procurement or replenishment policy given exogenous suppliers characteristics (such as delivery lead time, quality,
and price). See Elmaghraby (2000) for a review. In our paper, the suppliers’ delivery performance is endogenous and can be influenced by the buyer’s procurement strategy.

A few papers study multi-attribute procurement, which is closely related to this paper. Che (1993) studies a two-attribute auction in which price and quality are the two attributes under consideration. Che focuses on the so-called scoring-rule auction with multiple suppliers whose costs are independent draws from a common distribution. Branco (1997) extends Che’s model to allow a correlation in the suppliers’ costs. We also consider a multi-attribute procurement problem (i.e., quantity and delivery performance), but we consider the single-supplier case and focus on the search for simple procurement mechanisms. It is worth noting that although these papers study procurement involving multiple attributes, there is only a single dimension of information asymmetry. The optimal mechanism design problem with multidimensional information asymmetry is challenging and has not been fully explored. See Zheng (2000) and the references therein for more discussions.

There are papers studying supply chain contracting with asymmetric cost information: Corbett and Tang (1998), Corbett and de Groote (2000), Corbett (2001), Ha (2001), Corbett, and Zhou and Tang (2004), to name a few. Lau, Lau and Wang (2008) considers a contracting problem with asymmetric cost information and price-dependent demand. There are also papers analyzing procurement auctions with endogenous quantity. Dasgupta and Spulber (1990), Chen (2007), and Li and Scheller-Wolf (2008) study optimal auction design when a buyer’s procurement quantity depends on the suppliers’ private cost information. Our research differs from these papers because we include both supplier delivery performance and price-sensitive market demand in our problem.

One of the main objectives of our paper is to identify simple mechanisms that perform well for the buyer. The importance of simple procurement contracts has long been emphasized in the economics literature. Indeed, most real-world incentive schemes seem to take less extreme forms than the sophisticated policies predicted by economic theory (Holmstrom and Milgrom, 1987; Bhattacharyya and Lafontaine, 1995). Under different model settings, Bower (1993) and Rogerson (2003) both demonstrate that a simple contract can capture a surprisingly large portion of the surplus delivered by the fully optimal menu of contracts. See Chu and Sappington (2005) and the references therein for more examples. Recently, there has been an emerging group of papers in operations management that are in a similar spirit. The most relevant papers are Cachon and Zhang (2006) and Feng and Zhang (2007). Cachon and Zhang (2006) study a supply chain where the supplier is a queueing server and possesses private capacity cost information. They demonstrate that a contract with a fixed procurement price and capacity level is nearly optimal
for the buyer. The main difference between our paper and theirs is that by incorporating price-dependent market demand, we show that fixing the price (or quantity) attribute may lead to significant profit losses, especially when the buyer’s market demand is highly price-sensitive. Feng and Zhang (2007) study supply contracts for a newsvendor retailer who does not have perfect information about the supplier’s capacity cost. They consider a simple contract similar to ours in the sense that the retailer specifies a fixed capacity level together with a fixed wholesale price. It has been shown that the performance of the simple contract depends critically on demand variability. In contrast, we find in our multi-period inventory model, fixing the service level is nearly optimal even when demand is highly variable.

Several papers investigate the performance of relatively simple procurement contracts in problem settings that are quite different from ours. Kayış, Erhun, and Plambeck (2007) examine the performance of a price-only contract relative to the complex, optimal procurement strategy for a buyer facing both first-tier and second-tier suppliers with private cost information. It has been shown that the performance of the price-only contract depends on model parameters. We only consider a single tier of suppliers in our paper. Taylor and Plambeck (2007) propose the use of simple relational contracts in a repeated procurement game for a buyer to induce capacity investment from a supplier. In a repeated relationship, the reputational effect plays an important role, which is absent in our problem. Tunca and Wu (2008) study a procurement process selection problem where multiple suppliers (i.e., multi-sourcing) can be used. Given the observation that theoretically optimal procurement formats are almost never employed in reality, they compare the performances of two relatively simple, but suboptimal procurement processes. All these studies emphasize the virtue of simplicity when designing procurement contracts. However, they do not consider suppliers’ delivery performance and therefore, their analyses and insights are quite different from ours (e.g., they do not study the effect of fixing the service level for suppliers).

3. Model Setting

A buyer wishes to procure a product from a supplier and sell to the marketplace within a given time horizon. In the basic model, we consider a situation where there is only one qualified supplier with whom the buyer can do business. The situation might be due to quality, technology, and strategic alliance considerations. Later we will show that key results will not change when there are multiple suppliers competing for the buyer’s business. The supplier has a marginal production cost $c$. The supplier knows his cost type $c$ before signing the contract, but the buyer does not. The buyer’s uncertainty about $c$ is represented by a distribution $F(c)$, which is differentiable on
a strictly positive support \([\underline{c}, \bar{c}]\) \((0 < \underline{c} < \bar{c})\). The corresponding density function is denoted \(f(c)\). Assume \(F(c)\) is log-concave, i.e., \(\frac{F(c)}{f(c)}\) is an increasing function of \(c\). The log-concave condition is satisfied by most commonly used distribution functions (see Bagnoli and Bergstrom, 2005).

In practice, many firms can hold inventories to improve service. We focus on products with relatively long life-cycles so multiple replenishments are possible. Thus we consider a two-echelon periodic-review inventory system. We divide the problem horizon into a number of periods and allow both firms to place replenishment orders in each period. For example, the contracting horizon is one year and each period is a day or a week. Demand for this inventory system is stochastic and price-sensitive. Specifically, the demand arriving at the buyer in each period is given by

\[
D(p) = D_0(p) + \varepsilon, \tag{1}
\]

where \(p\) is the market price chosen by the buyer and \(\varepsilon \geq 0\) is a random variable. In particular, \(D_0(p)\) is a downward sloping, concave, deterministic function of \(p\). Both the additive demand format and the assumptions about \(D_0(p)\) are commonly used in the literature (see, for example, Federgruen and Heching, 1999 and Ray, Li, and Song, 2005). The distribution of \(\varepsilon\) is i.i.d. across periods. To avoid unrealistic solutions, we assume that the range of feasible values for \(p\) is bounded from above by some large number \(\bar{p} > \bar{c}\), where \(D_0(\bar{p}) = 0\). Similar treatment can be found in Ha (2001) and Ray, Li, and Song (2005).

The inventory system consists of two stages: Stage 1 represents the buyer and stage 2 is the supplier. Throughout the paper, we will use subscripts 1 and 2 to refer to the buyer and the supplier, respectively. The transportation lead time from the supplier to the buyer is \(L_1\), which may also represent the assembly time at the buyer. The supplier either manufactures the products or obtains them from an outside source with ample capacity. There is a constant replenishment lead time \(L_2\) for the supplier. Both the buyer and the supplier adopt stationary base stock policies and incur linear inventory holding costs. Let \(\hat{s}\) and \(\hat{s}_1\) denote the (local) base stock levels chosen by the supplier and the buyer. (We omit the subscript of \(\hat{s}_2\) for concision due to its frequent use.) The buyer’s holding cost rate is \(h_1\) per period, and the holding cost rate for the supplier is \(h_2 = h_0 + rc\), where \(h_0\) is a constant and \(r\) is the interest rate. So a higher production cost \(c\) means a higher inventory holding cost for the supplier. We follow the literature to assume that \(h_1 > h_0 + rc\), i.e., the holding cost of the buyer is greater than that of the supplier. Unmet demand in each period is backlogged, and the buyer incurs a backorder cost \(b\) for each backlogged unit. More details of this classic inventory model can be found in Chen and Zheng (1994) and Cachon and Zipkin (1999).

All firms are risk-neutral and aim to maximize their expected profit per period.
needs to design a procurement mechanism to offer to the supplier. In this paper, we define delivery performance (or service level) as the expected fill-rate at the supplier, which is determined by the supplier’s stocking policy. Choi, Dai, and Song (2004) and Shang and Song (2006) provide detailed discussion about how to set internal fill rate between firms to improve the performance of decentralized inventory systems. Without loss of generality, assume the maximum value of outside options is zero for the supplier, i.e., the supplier requires a minimum of zero profit for participation. We restrict our attention to stationary contract formats in this paper. Specifically, contracting terms such as procurement price, once determined, remain constant throughout the entire horizon. This assumption implies that the buyer essentially faces a static procurement problem and provides tractability to model analysis. Stationary contract formats are of particular interest in our problem setting for several practical reasons as well. First, dynamic contracts with time-varying parameters would be unnecessarily complex, which makes them more difficult to implement in practice. Second, adjusting contracting terms across periods may be costly, since it may require re-negotiations between the two parties and cause planning problems. This is true especially in our setting where the firms can routinely replenish their inventories. Third, it is reasonable to focus on stationary formats as an initial step to understand the procurement mechanism design problem. The analysis may serve as a benchmark to evaluate the benefits of using other more sophisticated contracting processes in future research.

The sequence of events is as follows: Supplier cost \( c \) is realized (only observable to the supplier); the buyer offers a procurement contract to the supplier; the supplier either takes or leaves the contract; if the contract is taken, then the supplier’s base stock level \( \hat{s} \), the buyer’s market price \( p \), and the payment scheme are determined according to the contract; finally, demand is realized in each period and the firms’ profits are evaluated for the entire horizon. For notational parsimony, define \( (x)^+ = \max(x, 0) \), \( (x)^- = \max(-x, 0) \), and \( x \wedge y = \min(x, y) \). We use \( E \) for the expectation operation. Let \( D^\tau(p) = \tau D_0(p) + \varepsilon^\tau \) denote the demand over \( \tau \) periods given market price \( p \), where \( \varepsilon^\tau \) is the convolution of \( \tau \) i.i.d. random variables. Define \( \mu^\tau(p) = E(D^\tau(p)) \) as the mean demand over \( \tau \) periods. Let \( \Phi^\tau \) and \( \phi^\tau \) be the distribution and density functions of \( \varepsilon^\tau \). Assume \( \Phi^\tau \) is differentiable for all positive integers \( \tau \). The superscript \( \tau \) is omitted when \( \tau = 1 \), e.g., \( \Phi \) is the distribution function of \( \varepsilon \) and \( \mu \) is the mean demand per period.

As preparation for analysis, we first present the buyer’s operating cost function, i.e., the inventory holding cost plus the backorder cost. We adopt the cost accounting method proposed in Chen and Zheng (2004). Suppose the market price \( p \) and thus the per-period demand \( D(p) \) have been chosen. Define \( G_1(y) = h_1(y)^+ + b(y)^- \) to be the buyer’s operating costs in a period with inventory
level \( y \). Then the buyer’s expected per-period operating cost can be written as

\[
\hat{H}_1(\hat{s}, \hat{s}_1) = E[G_1(\hat{s}_1 \land (\hat{s}_1 + \hat{s} - D^{L_2}(p)) - D^{L_1+1}(p))],
\]  

(2)

where \( \hat{s}_1 + \hat{s} \) is the supplier’s echelon base stock level. The echelon inventory position for stage \( i \) is defined as the sum of the inventory positions at stage \( i \) and all downstream stages. See Axsäter and Rosling, 1993 for detailed discussion of local vs. echelon stock policies.

Notice that \( D^\tau(p) = \tau D_0(p) + \varepsilon^\tau \) contains two parts, where \( \tau D_0(p) \) is the deterministic part and \( \varepsilon^\tau \) is the random part. The deterministic part represents a predictable product flow in the supply chain and does not affect the buyer’s on-hand inventory and backorder levels in each period. Therefore, the buyer’s operating cost actually depends only on the random demand part. Define \( s = \hat{s} - L_2D_0(p) \) and \( s_1 = \hat{s}_1 - (L_1 + 1)D_0(p) \). Then we have

\[
\hat{H}_1(\hat{s}, \hat{s}_1) = E[G_1((s_1 - \varepsilon^{L_1+1}) \land (s_1 - \varepsilon^{L_1+1} + s - \varepsilon^{L_2}))] = E[G_1(s_1 \land (s_1 + s - \varepsilon^{L_2}) - \varepsilon^{L_1+1})],
\]

which is independent of the market price \( p \). Define \( H_1(s, s_1) = E[G_1(s_1 \land (s_1 + s - \varepsilon^{L_2}) - \varepsilon^{L_1+1})] \) as the buyer’s operating cost function. We will work with \( (s, s_1) \) because they exclude the deterministic demand part and thus separate the buyer’s operating cost from the market price decision. This treatment does not change the essence of the buyer’s problem, but can simplify the exposition as well as analysis. So, unless otherwise mentioned, hereafter by base stock level we refer to \( s \) or \( s_1 \).

Let \( s_1(s) \) be the buyer’s reaction function to the supplier’s base stock level \( s \). It is straightforward to show that \( H_1(s, s_1) \) is convex in \( s_1 \) by inspecting the derivatives and, hence, there is a unique \( s_1(s) \) for any given \( s \). Furthermore, we can characterize the properties of \( s_1(s) \) as follows.

**Lemma 1** (i) \( s_1(s) \) is decreasing in \( s \);

(ii) \( s_1(s) \to \underline{s}_1 \) as \( s \to \infty \), where \( \underline{s}_1 = (\Phi^{L_1+1})^{-1}(b/(h_1 + b)) \).

Based on Lemma 1, we also investigate how the buyer’s operating cost \( H_1(s, s_1(s)) \) depends on \( s \). It can be shown that \( H_1(s, s_1(s)) \) is not necessarily convex in \( s \). However, we can prove the following useful properties about \( H_1(s, s_1(s)) \). The first property is intuitive and states that the buyer’s operating cost decreases in \( s \). We will use the second property when introducing the buyer’s optimal mechanism in Section 4.

**Lemma 2**

(i) \( \frac{dH_1(s, s_1(s))}{ds} = 0 \) for \( s = 0 \) and \( \frac{dH_1(s, s_1(s))}{ds} < 0 \) for \( s > 0 \);

(ii) \( \left( \frac{dH_1(s, s_1(s))}{ds} \right) /\Phi^{L_2}(s) \) is increasing in \( s \).
Finally, the following lemma presents the optimal solution for the inventory system under centralized control, which will be used for comparison later. The performance of a centralized system is determined by three variables: market price $p$, supplier’s base stock level $s$, and the buyer’s base stock level $s_1$. Note that the per-period mean demand $\mu(p)$ is fixed once the market price $p$ is given, i.e., there is a one-to-one relationship between $\mu$ and $p$. Let $H_2(c, s) = (h_0 + rc) [E(s - \varepsilon L_2)^+ + L_1 \mu(p)]$ denote the operating cost incurred by the supplier. Then the profit for the inventory system can be written as

$$
\pi_{sc} = \mu(p)(p - c) - H_1(s, s_1) - H_2(c, s), \tag{3}
$$

where the subscript $sc$ stands for “supply chain”.

**Lemma 3** Suppose the supplier’s cost is $c$. Then the centralized optimal solution for the inventory system $(p^*, s^*, s_1^*)$ is characterized by

$$
\frac{d\mu(p)}{dp} [p - c - (h_s + rc)L_1] + \mu(p) = 0, \tag{4}
$$

$$
\Phi^{L_1+1}(s_1) = (h_2 + b)/(h_1 + b), \tag{5}
$$

$$
-b + (b + h_2)\Phi^{L_2}(s) + (b + h_1) \int_s^\infty \phi^{L_2}(y)\Phi^{L_1+1}(s_1 + s - y)dy = 0. \tag{6}
$$

4. Analysis of Mechanisms

4.1 Optimal Mechanism (OM)

What is the optimal profit for the buyer? First we derive the optimal procurement mechanism for the buyer and use it later as a benchmark to evaluate suboptimal, but simpler mechanisms. A mechanism is a mapping from the supplier’s message space to the supplier’s payment and action space. According to the Revelation Principle (Salanie, 1997), the buyer’s attention can be confined to mechanisms that are truth-inducing (i.e., it is in the supplier’s best interest to report the true cost). In our problem, the optimal mechanism (OM) involves the actions to be taken by the supplier (i.e., the base stock level and the mean delivery quantity per period) and a payment schedule (i.e., the transfer payment from the buyer to the supplier).

Suppose the buyer offers a menu of contracts $\{\mu(\cdot), s(\cdot), w(\cdot)\}$ to the supplier: If the supplier announces his cost to be $x$ ($x$ could be different from the true cost $c$), then the buyer purchases an average quantity $\mu(x)$ in each period, the supplier adopts a base stock level $s(x)$, and there is a unit procurement price $w(x)$ between the buyer and the supplier. To facilitate the derivation of the optimal menu of contracts, we introduce two auxiliary functions $T(\cdot)$ and $p(\cdot)$, where $T(\cdot) = w(\cdot)\mu(\cdot)$ represents the transfer payment and $p(\cdot)$ is the market price that satisfies $\mu(p(\cdot)) = \mu(\cdot)$. Recall
\[ H_2(c, s) = (h_0 + rc)[E(s - \varepsilon L^2)^+ + L^1\mu] \] is the supplier’s operating cost. Given the \( \{\mu(\cdot), s(\cdot), w(\cdot)\} \) contract, the supplier’s profit function is given by

\[ \pi_2(c, x) = T(x) - c\mu(p(x)) - H_2(c, s(x)), \]

where \( c \) is the supplier’s true cost and \( x \) is the announced cost. For any truth-inducing contract, we have the Incentive Compatibility (IC) constraint:

\[ c = \arg \max_x \pi_2(c, x) \text{ for all } c. \tag{7} \]

That is, the supplier will announce \( x = c \) in order to optimize his profit. Also, the buyer needs to make sure that the supplier will accept the contract even with the highest cost \( \bar{c} \). This is called the Individual Rationality (IR) constraint:

\[ \pi_2(c, c) \geq 0 \text{ for all } c. \tag{8} \]

Since there is a one-to-one relationship between the menus \( \{\mu(\cdot), s(\cdot), w(\cdot)\} \) and \( \{p(\cdot), s(\cdot), T(\cdot)\} \), the buyer’s problem can be written as

\[
\max_{\{p(\cdot), s(\cdot), T(\cdot)\}} \pi_1 = \int_{\mathcal{E}} [\mu(p(x))p(x) - T(x) - H_1(s, s_1(s))] f(x)dx
\]

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

Once the optimal menu \( \{p^\alpha(\cdot), s^\alpha(\cdot), T^\alpha(\cdot)\} \) is determined, we simply apply \( \mu^\alpha(x) = \mu(p^\alpha(x)) \) and \( w^\alpha(x) = \frac{T^\alpha(x)}{\mu^\alpha(x)} \) to obtain the optimal menu \( \{\mu^\alpha(\cdot), s^\alpha(\cdot), w^\alpha(\cdot)\} \).

**Theorem 1** The optimal menu of contracts \( \{p^\alpha(\cdot), s^\alpha(\cdot), T^\alpha(\cdot)\} \) (i.e., the solution to (9)) can be characterized by

\[
\frac{d\mu(p)}{dp} \left[ p - x - (h_0 + rx)L_1 - (1 + rL_1)\frac{F(x)}{f(x)} \right] + \mu(p) = 0, \tag{10}
\]

\[
\frac{dH_1(s, s_1(s))/ds}{\Phi^L(s)} + \left[ h_0 + rx + r\frac{F(x)}{f(x)} \right] = 0. \tag{11}
\]

The transfer payment \( T^\alpha(x) \) can be solved using (7) and (8).

Theorem 1 deserves some discussion. First, Equation (10) solves the market price function \( p^\alpha(x) \), which in turn determines the average quantity function \( \mu^\alpha(x) \). Note that \( \mu'(p) \leq 0, \mu''(p) \leq 0 \) and \( \frac{F(x)}{f(x)} \) is increasing in \( x \). Thus (10) holds only if \( p - x - (h_0 + rx)L_1 - (1 + rL_1)\frac{F(x)}{f(x)} \geq 0 \). For any given \( x \), the left-hand side of (10) is decreasing in \( p \), which means that there is a unique solution \( p^\alpha(x) \) to (10). Further, it can be shown that \( p^\alpha(x) \) must increase as \( x \) increases, since otherwise (10)
cannot hold. Second, the first term in (11) is increasing in $s$ (see Lemma 2) and the second term is increasing in $x$, so we know that there is a unique solution $s^0(x)$, which is a decreasing function of $x$. Finally, from the truth-telling condition (Equation (20) in the proof of Theorem 1)

$$T'(x) - x\mu'(p)p'(x) - (h_0 + rx)[\Phi^L(s)s'(x) + L_1\mu'(p)p'(x)] = 0,$$

we can see that $T'(x) \leq 0$, because $\mu'(p) \leq 0$, $p'(x) \geq 0$ and $s'(x) \leq 0$. Therefore, in the optimal mechanism, a less efficient supplier (i.e., a higher $x$) corresponds to a higher market price ($p$), a lower mean demand ($\mu$), a lower stocking level ($s$), and a smaller total payment ($T$). Although a closed-form expression of the optimal menu of contracts is not available, we are able to show in the next theorem that in the optimal mechanism, the buyer tends to set a market price that is higher than the centralized optimal solution and induce a service level that is lower than the centralized optimal solution. This is caused by the buyer’s intention to optimize her own profit under asymmetric cost information.

**Theorem 2** $p^0(x) \geq p^*(x)$ and $s^0(x) \leq s^*(x)$ for all $x$, where $p^*$ and $s^*$ are given in Lemma 3.

### 4.2 Fixed Service-Level Contract (FS)

From the optimal mechanism derived in the previous section, we can see that the buyer should offer a non-linear function to the supplier for each of the two attributes, i.e., order quantity and delivery performance. The form of the optimal mechanism is quite complex, and it may require a sophisticated negotiation process in implementation. See Tunca and Wu (2008) for more discussion about the difficulties to implement a theoretically optimal procurement mechanism. In practice, we observe service-level contracts under which the supply chain members reach an agreement on a target or minimum service level. The target service level could be either an industry standard or tailored to a particular buyer. For instance, in the consumer goods industry, retailers typically require a certain service level on delivery performance from the manufacturers (Thonemann et al., 2005). It has also been observed that a buyer may specify a target fill-rate over a finite contracting horizon, and the supplier is either rewarded with a bonus or penalized with a late fee depending on the realized delivery performance (Thomas, 2005). See Katok, Thomas and Davis (2008) for more examples of contracts involving inventory service-level agreements. Clearly, the procurement mechanism would be easier to implement if the buyer specifies a fixed service level rather than a nonlinear performance requirement contingent on the supplier’s realized cost. But how does this simplification affect the buyer’s profit? We study a so-called fixed service-level (FS) contract $\{s, \mu(\cdot), w(\cdot)\}$, where the interpretation of the contract terms are analogous to that of the optimal
mechanism, except that \( s \) is the target base stock level specified by the buyer. The buyer’s profit in this FS contract can be evaluated in two steps. First, for a given \( s \), we can solve the optimal functions \( \mu(\cdot) \) and \( w(\cdot) \). Second, we search over \( s \) to find the optimal \( s \) the buyer should specify. The analysis of this contract is similar to that of the optimal mechanism in Section 4.1, and the details can be found in the Supplemental Appendix.

Similar to the optimal mechanism presented in Theorem 1, it can be shown that \( \mu(x) \) is a decreasing function of \( x \). Further, \( w(x) \) is increasing in \( x \) in the optimal FS contract. The argument is as follows. With a fixed \( s \), the truth-telling condition in (12) becomes

\[
T'(x) - x\mu'(p)p'(x) - (h_0 + rx)L_1\mu'(p)p'(x) = 0.
\]

Plugging \( T(x) = \mu(p(x))w(x) \) into the above equation, we have

\[
\mu(p)w'(x) + \mu'(p)p'(x)[w(x) - x - (h_0 + rx)L_1] = 0.
\]

Recall that \( w(x) \) is the unit procurement price and \( x + (h_0 + rx)L_1 \) is the supplier’s production and inventory holding cost for delivering one unit product, so \( w(x) - x - (h_0 + rx)L_1 \geq 0 \). Together with \( \mu'(p)p'(x) \leq 0 \), we know there must be \( w'(x) \geq 0 \). Thus, under the FS contract, the buyer essentially offers a price-quantity menu from which the supplier can choose from: If the supplier picks a low procurement price in the menu, then the buyer will order more from the supplier. This resembles the widely studied quantity-discount contract, where a supplier offers a price-quantity menu to a buyer, and a large order quantity chosen by the buyer corresponds to a low wholesale price. See Jeuland and Shugan, 1983 and Lee and Rosenblatt, 1986 for representative studies of quantity discount pricing models. It is well known that the quantity discount contract can coordinate a channel or supply chain (Cachon, 2003). However, here we demonstrate that the quantity discount scheme is also useful in our procurement problem under the presence of asymmetric cost information. Given the prevalence of the quantity discount contract in practice, we believe that the FS contract is intuitive and relatively easy to implement. It will be shown later that the FS contract yields nearly optimal profit for the buyer as well.

### 4.3 Fixed Quantity Contract (FQ)

Instead of fixing the service-level attribute, we may fix the quantity attribute in the optimal mechanism. Suppose the buyer offers a so-called fixed-quantity (FQ) contract \( \{\mu, s(\cdot), w(\cdot)\} \) to the supplier. In this contract, the buyer specifies an order quantity \( \mu \), and sets the service level \( s(\cdot) \) and the procurement price \( w(\cdot) \) based on the supplier’s cost signal. The optimal FQ contract can
be derived as follows: For a given $\mu$, we can solve the optimal menus $s(\cdot)$ and $w(\cdot)$; then we find the optimal $\mu$ that maximizes the buyer’s expected profit. To save space, the detailed analysis about the FQ contract is given in the Supplemental Appendix.

It is straightforward to show that $s'(x) \leq 0$ in the FQ contract. That is, similar to the optimal mechanism in Theorem 1, the buyer would require a higher service level if the supplier is more efficient. Now the truth-telling condition in (12) becomes $T'(x) - (h_0 + rx)[\Phi^{L2}(s)s'(x)] = 0$. Together with $T(x) = \mu w(x)$, we have $\mu w'(x) = (h_0 + rx)[\Phi^{L2}(s)s'(x)] \leq 0$, i.e., $w(x)$ is a decreasing function of $x$. Thus, in the FQ contract, the buyer offers a price-service menu to the supplier; in particular, the buyer pays a low price if the supplier chooses a low service level due to a high cost realization.

We highlight the following differences between the FQ and FS contracts. First, there is $w'(x) \leq 0$ in the FQ contract but $w'(x) \geq 0$ in the FS contract. The explanation is as follows: In the FQ contract, the mean order quantity is fixed, thus the buyer’s price decreases in $x$ since a high cost corresponds to a low service level (i.e., $s'(x) \leq 0$). In the FS contract, however, the mean order quantity is decreasing in supplier’s cost (i.e., $\mu'(x) \leq 0$); so the buyer needs to pay a high-cost supplier a high unit price to induce truth-telling (otherwise the supplier will have an incentive to deflate his cost to achieve both a larger order quantity and a higher price). Second, due to the fixed mean order quantity, the buyer’s corresponding market price, $p$, as well as the revenue $\mu p$, is fixed in the FQ contract. In contrast, in the FS contract, it is the service level that is fixed, hence the buyer has the flexibility to exploit the market by setting a cost-contingent market price. For instance, the FS contract enables the buyer to extract more revenue from the market if the supplier’s cost realization is low.

Given the above differences, a natural question may arise: If only one of the two attributes (service level or order quantity) can be fixed to simplify the procurement process, which attribute should we choose? Later we will compare the performances of the FS and FQ contracts and provide an answer to this question.

### 4.4 Fixed Quantity and Service-Level Contract (FQS)

We may further reduce the complexity of our contracts by fixing both attributes that are of interest. Specifically, the buyer may offer the supplier a contract $\{s, \mu, w\}$ consisting of three numbers rather than three menus/functions. Call this a fixed quantity and service-level (FQS) contract. Under this contract, the supplier’s profit function is $\pi_2(c) = w\mu - c\mu - H_2(c, s)$. To ensure supplier participation (i.e., the supplier with cost $\tilde{c}$ will take the contract), we need $w\mu = \tilde{c}\mu + H_2(\tilde{c}, s)$. That is, once $s$ and
\( \mu \) are fixed, the buyer’s optimal procurement price \( w \) is also determined. Thus, the FQS contract essentially fixes the procurement price the buyer pays the supplier. Let \( p \) be the corresponding market price to induce the mean demand \( \mu \). Under this contract, the buyer’s profit can be written as

\[
\pi_1 = p\mu - H_1(s, s_1(s)) - w\mu = (p - \bar{c})\mu - H_1(s, s_1(s)) - H_2(\bar{c}, s).
\]

Note that this is the supply chain’s profit function given a supplier cost \( \bar{c} \). Therefore, the buyer’s optimal profit in this contract is

\[
\pi_1 = (p^*(\bar{c}) - \bar{c})\mu(p^*(\bar{c})) - H_1(s^*(\bar{c}), s_1^*(\bar{c})) - H_2(\bar{c}, s^*(\bar{c})).
\]

In other words, the buyer will choose a market price \( p^*(\bar{c}) \) and requires a service level \( s^*(\bar{c}) \) in the FQS contract. As a matter of fact, in this contract the buyer essentially sells the supply chain to the supplier at a price that equals the centralized supply chain’s optimal profit \( \pi_{sc}^*(\bar{c}) \).

5. Comparison of Mechanisms

This section compares the performances of the procurement mechanisms. An analytical comparison is difficult due to the complexity of the inventory system, so we compare the mechanisms using an extensive numerical study. In particular, we are interested in the performance of the simple mechanisms relative to the optimal mechanism. The numerical study shall cover a wide range of plausible situations that may arise in practice. This is done by assigning both high and low feasible values to each of the problem parameters. The supplier’s cost \( c \) follows a uniform distribution with a support \([\underline{c}, \bar{c}] = [\eta - \delta, \eta + \delta]\), where \( \eta = 1 \) and \( \delta = 0.4\eta \). Since the result of the numerical study depends only on the relative magnitude of the parameter values, we can fix the value of \( \eta \) to be 1 and vary other parameter values. Note that with \( \delta = 0.4\eta \), the ratio \( \bar{c}/\underline{c} \) is equal to 2.33, which represents an unusually high uncertainty in the supplier’s cost. It is clear that the simple mechanism approaches optimal as the cost uncertainty goes to zero, and therefore we only focus on situations where there is a relatively large cost uncertainty.\(^2\) As to the supplier’s holding cost, we set \( r \in \{1\%, 10\%\} \). In view of the 15% annual holding cost rate commonly used in textbooks, \( r = 10\% \) is quite large for inventory systems with review periods much shorter than a year. Recall that we have assumed \( h_1 > h_2 = h_0 + rc \). In the numerical study, we set \( h_0 = 0.01 \), \( h_1 = \alpha_h(h_0 + r\bar{c}) \) and choose \( \alpha_h \in \{1.1, 2\} \). In addition, we let \( b = \alpha_bh_1 \) and choose \( \alpha_b \in \{1, 10, 40\} \), where a larger \( \alpha_b \) means a more significant backorder cost compared to inventory holding cost. The random

\(^2\)We have tested values less than 0.4 for \( \delta \) and found that performance of the simple mechanisms is indeed better for smaller \( \delta \). See the Supplemental Appendix for details.
demand variable $\varepsilon$ follows a Gamma distribution with parameters $(k, \theta)$, where $k \in \{1, 4, 16\}$ is the scale parameter and $\theta = 10$ is the shape parameter. The mean and variance of $\varepsilon$ are given by $k\theta$ and $k\theta^2$, respectively, and the coefficient of variation $(1/\sqrt{k})$ takes values in $\{0.25, 0.5, 1\}$. A coefficient of variation of 1 is unusually high for a product item managed under a periodic review inventory system. The lead times for the two stages are $L_1 \in \{1, 4\}$ and $L_2 \in \{2, 4, 8\}$. Finally, we consider a quadratic demand function in this numerical study: $D_0(p) = \alpha - \beta p^2$, where $\alpha = 300$ and $\beta \in \{10, 20, 30\}$. The value of $\alpha$ guarantees an interior optimal solution for the market price, i.e., $p^* < \bar{p}$. The $\beta$ values represent different sensitivity levels of the market demand in response to the price. There are 648 combinations in total in this numerical study. Among the 648 parameter combinations, we check the following variables in the optimal mechanism: First, the expected fill-rate in the optimal mechanism ranges from 4% to 98%; second, the buyer’s expected operating cost as a fraction of the revenue ranges from 1% to 66%. So the numerical study covers situations with both low and high service levels, as well as with both relatively low and high operating costs.

Table 1. Percentage profit decrease of each mechanism relative to the optimal mechanism (OM) with a single supplier.

<table>
<thead>
<tr>
<th>cost distribution</th>
<th>mechanism</th>
<th>min 10th percentile</th>
<th>median</th>
<th>90th percentile</th>
<th>max</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>uniform</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.16</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td>FQ</td>
<td>0.45</td>
<td>0.69</td>
<td>2.34</td>
<td>8.89</td>
<td>51.23</td>
</tr>
<tr>
<td></td>
<td>FQS</td>
<td>0.45</td>
<td>0.69</td>
<td>2.37</td>
<td>8.98</td>
<td>52.07</td>
</tr>
<tr>
<td>normal</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0.02</td>
<td>0.37</td>
<td>2.24</td>
</tr>
<tr>
<td></td>
<td>FQ</td>
<td>1.81</td>
<td>3.30</td>
<td>7.43</td>
<td>19.28</td>
<td>63.56</td>
</tr>
<tr>
<td></td>
<td>FQS</td>
<td>1.91</td>
<td>3.42</td>
<td>7.77</td>
<td>19.38</td>
<td>65.68</td>
</tr>
</tbody>
</table>

Table 1 (the top part) summarizes the results from the numerical study. The performance of a mechanism is measured by the percentage profit decrease relative to the optimal mechanism (OM). For each mechanism, percentage profit decreases at several different percentiles are provided. For example, with the FS contract, the buyer’s profit decrease at the 90th percentile is 0.16%. In other words, the percentage profit decrease under the fixed service-level contract is no greater than 0.16% for 90% of all the scenarios. We emphasize the following observations. First, the FS contract is nearly optimal for all the tested scenarios: The profit decrease is 0.06% on average and the number is 0.84% at maximum. The difference between the FS contract and the OM lies in how we treat the service-level attribute, i.e., the former uses a fixed number while the latter includes a complex menu for the service-level attribute. Thus the numerical study suggests that using a complex menu on the service-level attribute adds little value for the buyer. Second and in contrast, Table 1 indicates that fixing the quantity attribute may remarkably decrease the buyer’s profit: The profit loss under the FQ contract is 4.44% on average and it can be as high as 51.23%. This
implies that the value of using a screening process for the quantity attribute could be significant. Third, the performance of the FQS contract is very close to that of the FQ contract. Note that the FQS contract differs from the FQ contract by further fixing the service-level attribute. Thus the result confirms that fixing the service level has a negligible impact on the buyer’s profit even when the quantity attribute has already been fixed. Finally, it is worth reporting the composition of the supplier’s profit (or equivalently, the information rent the supplier derives from his private cost information) in the optimal mechanism. Specifically, we partition the supplier’s information rent into two parts, i.e., the rent associated with the service-level attribute and the rent associated with the quantity attribute. The fraction of the rent associated with the service-level attribute has a maximum of 33% and an average of 11% in this numerical study. Therefore, the observation about the service-level attribute is not because its related information rent is so minuscule that can be ignored.

To test the robustness of the performance results, we construct another set of 648 scenarios that are identical to the above except that the uniform cost distribution is replaced by a normal distribution with mean \( \eta \) and standard deviation \( \delta/3 \) (the truncated probability on the two sides is redistributed evenly on the support). The performances of the simple mechanisms are presented in the bottom part of Table 1. Similarly, we can see that the FS contract performs very well relative to the OM (the profit decrease is 0.37% at the 90th percentile and merely 0.12% on average), while the FQ and FQS contracts perform much worse. For instance, the profit decrease for the FQ contract is 9.88% on average and 63.56% at maximum.

We also examine the impact of different parameters on the comparison of performances. From Table 1, the FS contract performs uniformly well in all tested scenarios. The performances of both the FQ and FQS contracts are highly volatile. In the rest of the paper, we will focus on the FQS contract rather than the FQ contract, because they exhibit similar performances but the former has a simpler format.\(^3\) A notable finding is that the price-sensitivity parameter plays a critical role in the performance of the FQS contract. Figure 1 displays the relationship between the performance of the FQS contract and the parameter \( \beta \) (with uniform cost distribution). We can see that as the demand becomes more sensitive in market price (i.e., as \( \beta \) increases), the performance of the FQS contract deteriorates accordingly.

In our procurement problem, the buyer and supplier interact along two dimensions, i.e., service level and order quantity. Next we study two special cases of the buyer’s problem. In the first case,
Figure 1: The impact of the price-sensitivity of market demand (measured by $\beta$) on the performance of the fixed quantity and service-level (FQS) contract (measured by the percentage profit decreases relative to the optimal mechanism at different percentiles).

we concentrate on the service level by assuming that the market price and demand are exogenously given; while in the second case, there is no demand uncertainty and we focus on the order quantity. This way we can separate the effects of these two dimensions in order to derive more intuition and insight into the above observations.

5.1 Exogenous Demand

We start with the observation about the service-level attribute, i.e., the buyer can achieve nearly optimal profit by simply posting a target service level. Here we consider a special case where the market price $p$, and thus the mean market demand $\mu(p) = D_0(p) + E(\varepsilon)$ are exogenously given. This may happen, for instance, when the buyer has little pricing power in the market, or the market price has been pre-determined. This treatment allows us to derive further insights about the service-level attribute by eliminating the effect of the quantity attribute. With an exogenous price $p$, we go back to the classic two-echelon inventory system (with asymmetric information). Does the FS contract continue to perform well? And if yes, why? Below we provide the answers to these questions.

Now the buyer’s optimal mechanism consists of two functions: $\{s(\cdot), w(\cdot)\}$. The analysis of this optimal mechanism is similar to that of Theorem 1 and thus given in the Supplemental Appendix. Without the presence of the market price, the FS contract studied in Section 4.2 reduces to $\{s, w(\cdot)\}$, where $s$ is the specified base stock level for the supplier and $w(x)$ is the procurement price function.
Under this contract, the actions taken by the supplier (i.e., the delivery quantities and the base stock level) are independent of his cost, so all supplier types will report a cost \( \bar{c} \) to receive the highest price \( w(\bar{c}) \). Let \( w = w(\bar{c}) \). Then we have \( w(x) = w \) for all \( x \) and the contract essentially consists of only two numbers \( \{ s, w \} \). In addition, to ensure supplier participation (i.e., the supplier with cost \( \bar{c} \) will accept the contract), we need \( \pi_2(\bar{c}) = \mu(p)w - \bar{c}\mu(p) - H_2(\bar{c}, s) \geq 0 \), or \( \mu(p)w \geq \bar{c}\mu(p) + H_2(\bar{c}, s) \).

Thus the buyer will choose \( \mu(p)w = \bar{c}\mu(p) + H_2(\bar{c}, s) \) and her profit is given by

\[
\pi_1 = \mu(p)p - H_1(s, s_1(s)) - \mu(p)w = \mu(p)(p - \bar{c}) - H_1(s, s_1(s)) - H_2(\bar{c}, s).
\]

Note that this is the supply chain's total profit given a supplier cost \( \bar{c} \). Therefore, the buyer will choose the supply chain’s optimal service level \( s^*(\bar{c}) \) (given by Lemma 3) in the FS contract.

Next we compare the FS contract to the buyer’s optimal mechanism. To this end, we modify the previous numerical study by assuming \( \mu(p) = E(\varepsilon) \). In other words, there is \( D_0(p) = 0 \) without losing generality. Since the revenue parts \( \mu(p)p \) are fixed and identical, it is equivalent to compare the cost parts in the two mechanisms. Let \( C^{OM}_1(c) = \mu(p)w(c) + H_1(s, s_1(s)) \) denote the buyer’s total costs in the OM with a supplier cost \( c \); and let \( C^{FS}_1 = \mu(p)\bar{c} + H_1(s, s_1(s)) + H_2(\bar{c}, s) \) be the buyer’s total costs in the FS contract (note that \( C^{FS}_1 \) is independent of \( c \)). Another advantage of comparing the cost parts is that we can isolate the effect of the revenue (e.g., a dominant revenue part in the profit function may significantly weaken the effect of the cost difference). The rest of the design of the numerical study is the same as before (there are 648 scenarios for uniform and normal cost distributions, respectively). The performance of the FS contract is summarized in Table 2, which presents the percentage cost increase of the simple mechanism. Again, we can see that the simple mechanism is nearly optimal.

<table>
<thead>
<tr>
<th>cost distribution</th>
<th>mechanism</th>
<th>min</th>
<th>10th pantile</th>
<th>median</th>
<th>90th pantile</th>
<th>max</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>uniform</td>
<td>FS</td>
<td>0</td>
<td>0.01</td>
<td>0.06</td>
<td>0.28</td>
<td>0.63</td>
<td>0.10</td>
</tr>
<tr>
<td>normal</td>
<td>FS</td>
<td>0</td>
<td>0.05</td>
<td>0.56</td>
<td>1.24</td>
<td>0.12</td>
<td></td>
</tr>
</tbody>
</table>

Why does the simple mechanism perform so well? To answer this important question, we proceed to calculate the ratio between the buyer’s costs in the OM when the supplier’s cost takes the two boundary values (i.e., \( C^{OM}_1(\bar{c})/C^{OM}_1(\underline{c}) \)). It has been found that among all the scenarios, the ratio ranges from 1.00 to 1.10, with an average of 1.01. Thus we know that despite significant cost uncertainty (\( \bar{c}/\underline{c} = 2.33 \)), there is a very small variation in the buyer’s total cost (or equivalently, profit) in the OM. This can be manifested with Figure 2, which demonstrates each party’s payoff as a function of the supplier cost \( c \) for one of the scenarios (with parameters \( r = 0.1, \alpha_h = 2, \alpha_b = 10, \)
We can see that the buyer’s total cost in the OM is quite flat (it almost coincides with the buyer’s cost in the FS contract), while the supplier’s profit increases considerably as the supplier becomes more efficient.\footnote{To confirm the generality of this observation, we have conducted additional numerical experiments with two-point cost distributions (rather than continuous uniform and normal distributions). It has been found that again the buyer’s cost is generally flat in the optimal mechanism. Details are presented in Appendix B of the Electronic Companion.}

The above observation suggests that along the service-level dimension, the supplier is able to keep essentially all the benefits of a low production cost due to the protection from information asymmetry. A similar finding has been reported in Cachon and Zhang (2006), where a buyer procures a product from a make-to-order supplier. It has been found that, in their model setting, fixing a capacity level (rather than constructing a complex menu) for the supplier works very well. We have extended their results to situations where firms can hold inventories to improve customer service. In both papers, market demand is exogenously given. Hence the revenue stream is fixed, and the supplier’s efficiency level only affects the supply chain’s total cost for satisfying the demand. Truth-telling is beneficial to the supply chain because the buyer could reduce the total cost by requesting cost-contingent service levels from the supplier. However, since it is the buyer who ultimately pays the total cost, the supplier has little incentive to report a low cost unless he will keep the majority of the benefits. In other words, inducing truth-telling reduces the supply chain’s total cost, but this mainly benefits the supplier through information rent due to his superior bargaining position. Therefore, even the optimal menu on the service-level attribute is not effective in extracting surplus from a low-cost supplier. But such a result does not carry over to the quantity dimension when the market demand is endogenous, as we will explain in the next subsection.

### 5.2 Deterministic Demand

We would also like to shed some light on the observation about the quantity attribute, i.e., there can be substantial benefits for the buyer to use a complex menu on the order quantity. As a counterpart to the previous subsection, here we consider a special case where demand is endogenous but without uncertainty (i.e., $\varepsilon = 0$ with probability 1) and inventory holding costs (i.e., $h_1 = h_2 = 0$). This corresponds to practical situations where the impact of demand uncertainty is insignificant. The supplier’s cost is uniformly distributed on $[0, 1]$, and the buyer faces a linear demand function $D(p) = \alpha - \beta p$. We require $\alpha > \beta$ so that there is a positive demand when $p = 1$. Although this is a simple example, it provides tractability and clear insight. Closed-form solution of the optimal menu of contracts can be derived, and the buyer’s optimal profit is given by $\pi^{OM}_1 = \frac{\alpha^2}{4\beta} - \frac{\alpha}{2} + \frac{\alpha}{3}$.\footnote{To confirm the generality of this observation, we have conducted additional numerical experiments with two-point cost distributions (rather than continuous uniform and normal distributions). It has been found that again the buyer’s cost is generally flat in the optimal mechanism. Details are presented in Appendix B of the Electronic Companion.}
Figure 2: Firms’ ex post payoffs in the OM and the FS contract (with a single supplier and exogenous market price).

details in the Supplemental Appendix). Since there is no demand uncertainty, the FQS contract studied in Section 4.4 becomes a fixed-quantity contract, i.e., the buyer pays the supplier a unit procurement price that is equal to 1. It can be shown that the buyer’s profit is $\pi_{FS}^1 = \frac{\alpha^2}{4\beta} - \frac{\alpha}{4} - \frac{\beta}{4}$.

Let $\rho = \frac{\alpha}{\beta} > 1$. Then the percentage profit loss caused by using the FQS contract is given by

$$\frac{\pi_{OM}^1 - \pi_{FS}^1}{\pi_{OM}^1} = \frac{1}{3\rho^2 - 6\rho + 4},$$

which approaches 1 as $\rho \to 1$. This example delivers two useful messages. First, it provides a worst-case scenario under which the profit loss can go up to 100%. This means that the optimal menu of contracts on the quantity attribute can be highly valuable, which is in stark contrast to the case of service-level attribute. Second, it demonstrates that the profit loss tends to increase as $\rho$ decreases. Note that a smaller $\rho$ corresponds to either a lower base demand $\alpha$ or a higher price sensitivity factor $\beta$. Indeed, this confirms the pattern in Figure 1, where the buyer’s percentage profit decrease in the FQS is higher with a larger $\beta$. So a complex menu for the quantity attribute tends to be more valuable when the buyer faces a more price-sensitive demand.

The above numerical study and analysis give rise to an interesting question: Why does the quantity attribute have a different implication in procurement mechanism design than the service-level attribute? The following example may help us understand this question. Consider $\alpha = 3$ and $\beta = 1.2$ in the deterministic demand case. (In this example, the buyer’s profit with the OM is about 15% higher than that with the FQS contract, so the benefit of using the complex screening contract
is quite large.) For illustration, Figure 3 highlights each party’s profit as a function of the supplier cost $c$ (i.e., the ex post profit realization). Again we observe that the supplier’s profit rises quickly as the cost $c$ decreases. However, the buyer’s profit in the OM also exhibits a clear increasing trend as the cost declines. Actually, $\pi_{OM}^{1}(\bar{c})/\pi_{OM}^{1}(\bar{c}) = 2.6$ in this example, which is much higher than the ratios $C_{OM}^{1}(\bar{c})/C_{OM}^{1}(\bar{c})$ observed in Section 5.1, including Figure 2. This means that, unlike the case with exogenous market demand (Section 5.1), the buyer can enjoy a significant portion of the benefits from having a low-cost supplier. A plausible explanation is as follows: A low supplier cost enables the buyer to charge a low market price and induce additional demand, which benefits both the buyer and the supplier; hence the supplier is incentivized to report a low cost. As a result, the buyer does not have to leave all the benefits to the supplier as information rent when the two firms interact over the order-quantity dimension.}

We emphasize that the endogenous demand is critical to the findings along the order-quantity dimension. Also, the above explanation is consistent with the earlier observation that the screening contract tends to be more valuable when the market demand is more sensitive in price (see Figure 1). The reason is that all else being equal, a higher sensitivity means that market demand increases faster as the supplier cost declines, which makes the screening contract more effective in attracting market demand.
6. Extensions

This section presents two extensions of the basic model. The first extension in Section 6.1 involves multiple competing suppliers from which the buyer can choose. The second extension in Section 6.2 allows the buyer to use a cutoff-level policy in procurement. For these two extensions, we investigate whether the key result from the basic model still holds, i.e., the buyer can fix a service level and achieve nearly optimal profit in procurement mechanism design.

6.1 Multiple Potential Suppliers

In this subsection, we consider an extension where there are \( n \geq 2 \) potential suppliers and the buyer may use competitive bidding to select the most efficient one. The suppliers’ marginal costs \( c_i \in [c, \bar{c}] \) are independent draws from a common distribution \( F \) (with density \( f \)). The buyer does not know the exact costs, but she knows the common distribution \( F \). Let \( c = (c_1, \ldots, c_n) \) denote the true cost vector.

We first derive the optimal mechanism for the buyer. Consider the following menu of contracts:

\[
\{ q_i(\cdot), \mu_i(\cdot), s_i(\cdot), T_i(\cdot) \}, \; i = 1, 2, \ldots, n
\]

That is, if the suppliers announce their costs to be \( \hat{c} = (\hat{c}_1, \ldots, \hat{c}_n) \), then each supplier \( i \) receives a payment \( T_i(\hat{c}_i) \); supplier \( i \) wins with probability \( q_i(\hat{c}_i) \geq 0 \); finally, given that supplier \( i \) wins the contract, the buyer orders an average quantity \( \mu_i(\hat{c}_i) \) (or equivalently, sets a market price \( p_i(\hat{c}_i) \)), and supplier \( i \) adopts a base stock level \( s_i(\hat{c}_i) \).

The analysis starts with the supplier’s bidding behavior. Supplier \( i \) maximizes the following expected profit:

\[
\max_{\hat{c}_i} \pi_2^i(c^i, \hat{c}_i) = \mathbb{E}_{\hat{c}} \{ T_i(\hat{c}_i) - q_i(\hat{c}_i)[c_i \mu_i(p_i(\hat{c}_i))] + H_2(c_i, s_i(\hat{c}_i)) \},
\]

where \( c^i \) is the true cost and \( \hat{c}_i \) is the announced cost. We focus on truth-inducing mechanisms, i.e., for each supplier \( i \),

\[
\hat{c}_i = \arg \max_{\hat{c}_i} \pi_2^i(c^i, \hat{c}_i) \text{ for all } c^i.
\]

Similarly, we can write the participation constraint or individual rationality constraint as

\[
\pi_2^i(c^i, \hat{c}_i) \geq 0 \text{ for all } c^i.
\]

Then the buyer’s problem is

\[
\max_{\{ q(\cdot), p(\cdot), s(\cdot), T(\cdot) \}} \pi_1 = \mathbb{E}_e \left\{ \sum_i q_i(e) [\mu_i(p_i(c^i)) p(c^i) - H_1(s_i, s^i)] - \sum_i T_i(c^i) \right\}.
\]

s.t. (13) and (14)
Let \( c_l = \min(c_1, \ldots, c_n) \) be the lowest cost among \( n \) suppliers. The following theorem characterizes the optimal mechanism with multiple suppliers (i.e., the solution to (15)).

**Theorem 3** In the optimal mechanism with \( n \geq 2 \) suppliers, the buyer offers the same menu of contracts \( \{q^o(\cdot), p^o(\cdot), s^o(\cdot), T^o(\cdot)\} \) to all suppliers. The suppliers announce their true costs and the most efficient supplier wins, i.e., \( q^o(c^i) = 1 \) if \( c^i = c_l \) and \( q^o(c^i) = 0 \) otherwise. The \( p^o(\cdot) \) and \( s^o(\cdot) \) functions are characterized by

\[
\frac{d\mu(p)}{dp} \left[ p - c^i - (h_0 + rc^i) L_1 - (1 + r L_1) \frac{F(c^i)}{f(c^i)} \right] + \mu(p) = 0, \quad (16)
\]

\[
\frac{dH_1(s, s_1(s))}{ds} + \left[ h_0 + rc^i + r \frac{F(c^i)}{f(c^i)} \right] = 0. \quad (17)
\]

The \( T^o(\cdot) \) function can be derived from (13) and (14).

The above theorem states that the market price function \( p^o(\cdot) \) and the service-level function \( s^o(\cdot) \) are exactly the same as in the optimal mechanism with a single supplier (see Theorem 1). So the optimal mechanism with \( n \geq 2 \) suppliers also induces a higher market price and a lower service level than the centralized optimal solution. However, the lump-sum \( T^o(\cdot) \) is now dependent on the number of suppliers, \( n \).

We examine whether the buyer can achieve nearly optimal profit by using a simpler mechanism with a fixed service level. Specifically, the buyer announces a required service level and then asks the suppliers to reveal their costs. Similar mechanism has been observed in non-manufacturing sectors. For instance, Cripps and Ireland (1994) study how to design auctions to award television franchises to private operators. One of the auction formats they consider involves a minimum quality level to be provided by the winning bidder. We call this a fixed service-level (FS) mechanism (it differs from the FS contract with a single supplier only in that the selection process described in Theorem 3 is used to single out the most efficient supplier).

It has been shown in Section 5 that the FQS contract performs poorly with a single supplier. How does competition affect the performance of the FQS contract? Here we consider the following FQS mechanism under competition: The buyer specifies an (average) order quantity \( \mu \) and a target service level \( s \) (both are independent of the supplier’s cost information), and then asks the suppliers to bid on the unit procurement price. Without losing generality by the Revenue Equivalence Principle, suppose a second-price, sealed-bid auction is used to determine the winner of the auction. In this auction, each supplier will bid a price that just breaks even, and the buyer can infer the suppliers’ true costs from their bids. More detailed analyses of the FS and FQS mechanisms under competition are presented in the Supplemental Appendix.
We have conducted a numerical study with \( n \geq 2 \) suppliers to compare the performances of different mechanisms. The same parameter values in Section 5 have been used. The results of the comparison for \( n = 2 \) are presented in Table 3. (The observation is similar for \( n > 2 \).) We can see that fixing the service level performs nearly as well as the optimal mechanism with multiple potential suppliers. There is an intuitive explanation for why the multi-supplier model yields a qualitatively similar result as the single-supplier model. Let \( c_s \) denote the second-lowest cost in the vector \((c^1, \ldots, c^n)\). Then the multi-supplier model essentially resembles a single-supplier model where the supplier’s cost is bounded by \( c_s \). Although \( c_s \) is a random variable, it has a support on \([\underline{c}, \bar{c}]\) and any realization of \( c_s \) shall not change the qualitative result from the single-supplier model. Note that \( c_s \) decreases stochastically as the number of suppliers increases, i.e., the cost uncertainty can be reduced by inviting more competition. This also explains why the performance of the simple mechanism with two suppliers is even better than the performance with a single supplier.

Table 3. Percentage profit decrease of the simple mechanisms relative to the optimal mechanism (OM) with \( n = 2 \) suppliers.

<table>
<thead>
<tr>
<th>cost distribution</th>
<th>mechanism</th>
<th>min</th>
<th>10(^{th}) pctile</th>
<th>median</th>
<th>90(^{th}) pctile</th>
<th>max</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>uniform</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.08</td>
<td>0.24</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>FQS</td>
<td>0.57</td>
<td>0.77</td>
<td>1.86</td>
<td>5.06</td>
<td>14.85</td>
<td>2.68</td>
</tr>
<tr>
<td>normal</td>
<td>FS</td>
<td>0</td>
<td>0.01</td>
<td>0.02</td>
<td>0.06</td>
<td>0.17</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>FQS</td>
<td>0.35</td>
<td>0.45</td>
<td>0.94</td>
<td>2.25</td>
<td>5.95</td>
<td>1.26</td>
</tr>
</tbody>
</table>

Table 3 also presents the performance of the FQS mechanism under competition. Comparing to Table 1, we can see that the FQS mechanism performs much better with competitive suppliers, especially for the worse-case scenarios (i.e., maximum profit losses). With a single supplier, the buyer has to offer a relatively high procurement price in the FQS in order to ensure participation from the least efficient supplier with a cost \( \bar{c} \) (see Section 4.4). With competing suppliers, the procurement price is determined by the bidding process, which will be clearly lower than the price in the single-supplier setting. Note that a lower procurement price enables the buyer to charge a higher market price and attract more demand. As a result, the inefficiency from using a fixed order quantity would be reduced when there is supplier competition. In other words, competitive bidding has an equivalent effect of screening because the procurement price is contingent on the suppliers’ true costs. Nevertheless, Table 3 shows that there is still a gap between the FQS and FS mechanisms. For example, with uniform distribution, the maximum profit loss with the FQS mechanism is 14.85%, while the number is 0.24% with the FS mechanism.

The above result has useful practical implications. The advent of information technology has led to a boom in online markets. In some B2B industrial exchanges, buyers post product specifications
and logistics requirements and then ask suppliers to bid on contracts (Rangan, 1998). One of
the biggest challenges in designing an online procurement process is bringing into consideration
multiple factors that affect the buyer-supplier relationship (Elmaghraby, 2004). Our result shows
that when both procurement price and logistics performance are taken into consideration, buyers
can simply post a target service level and then ask suppliers to bid only on their costs. In addition,
such a strategy can be made even more effective by increasing the number of competing suppliers.

6.2 Cutoff-Level Policy

In the previous sections we have implicitly assumed that the buyer needs to contract with the
supplier regardless of his cost realization. This assumption is appropriate under situations where
the buyer has other strategic concerns in decision-making (e.g., pre-commitment to transaction and
long-term relationship considerations), or the supplier has passed a rigorous screening process so
that his cost is within a reasonable range. However, in some other situations, the buyer may wish to
do business with the supplier only if his cost is below a cutoff level. Given that a cutoff-level policy
is allowed, we would like to investigate whether the results from the previous sections continue to
hold.

We are interested to know the performances of the simple mechanisms relative to the optimal
one when an optimal cutoff level is used. For the optimal mechanism, we restrict our attention
to the following format: There is a single cutoff level \( \tilde{c} < \bar{c} \) such that the buyer sources from
the supplier if and only if \( c < \tilde{c} \). (Corbett and de Groote, 2000 study a different contracting
problem under asymmetric information and derive the optimal menu of contracts within this class
of cutoff-level policies. Ha, 2001 shows that such a cutoff-level policy is optimal in his problem
setting.) We search for the optimal cutoff level \( \tilde{c} \) that maximizes the buyer’s profit within this
contract format. Similarly, for the simpler mechanisms (FS and FQS), we identify their optimal
cutoff levels, respectively. Consider the FQS contract for illustration. Now, instead of making sure
that the least efficient supplier will participate, the buyer optimally chooses a cutoff level \( \hat{c} \) such
that only suppliers with a cost lower than \( \hat{c} \) will participate. Intuitively, by employing a cutoff-level
policy, the buyer would improve her profit in the FQS contract because she can avoid paying a high
transfer payment that is needed to guarantee the participation from all supplier types. One may
also interpret the cutoff level as a naive screening policy as opposed to the full screening in the
optimal mechanism. We use the same numerical study in Section 5 to evaluate the performances
of the FS and FQS contracts.

Table 4 summarizes the results from the numerical study. There are two noteworthy obser-
vations. First, the FS mechanism is again nearly optimal. Second, the performance of the FQS contract has been significantly improved by using the cutoff-level policy. In other words, the buyer’s ability to not transact with unattractive suppliers (through a cutoff-level policy) is critical for her profitability. Since the performances of the FS and FQS contracts are now closer (as compared to Table 1), we need to evaluate the trade-off carefully when choosing between these procurement contracts: While the FQS contract has a simpler format, the FS contract could yield higher profit for the buyer (for example, with uniform distributions, the maximum profit loss relative to the optimal mechanism is 8.51% with the FQS mechanism, while the number is 0.28% with the FS mechanism).

Table 4. Percentage cost increase of the simple mechanisms relative to the optimal mechanism (OM) with a single supplier and optimal cutoff level policies.

<table>
<thead>
<tr>
<th>cost distribution</th>
<th>mechanism</th>
<th>min</th>
<th>10th percentile</th>
<th>median</th>
<th>90th percentile</th>
<th>max</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>uniform</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.13</td>
<td>0.28</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>FQS</td>
<td>0.45</td>
<td>0.69</td>
<td>2.37</td>
<td>6.42</td>
<td>8.51</td>
<td>3.05</td>
</tr>
<tr>
<td>normal</td>
<td>FS</td>
<td>0</td>
<td>0</td>
<td>0.01</td>
<td>0.06</td>
<td>0.11</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>FQS</td>
<td>0.52</td>
<td>0.86</td>
<td>1.26</td>
<td>2.13</td>
<td>3.47</td>
<td>1.37</td>
</tr>
</tbody>
</table>

7. Conclusion

This paper studies the procurement strategy for a buyer who obtains a product from a supplier and in turn sells to customers. The buyer faces an uncertain demand that depends on the market price she sets. Since the buyer and the supplier are independent organizations, the supplier’s production cost is private information and the buyer only has an unbiased belief about the distribution of the cost. Both firms can hold inventories to improve customer service, and a two-echelon periodic-review inventory system is considered. The buyer wants to design a procurement process to maximize her expected profit, which is equal to the revenue minus the operating cost. The buyer’s revenue is a function of the market price (and the corresponding selling quantity) she chooses, which depends on the supplier’s production cost; the operating cost arises from the holding of safety stock and backlogged customer demand, which depends on the supplier’s delivery performance. Thus, the buyer needs to take care of both the quantity (i.e., how much to order) and the service-level attributes (i.e., how to deliver the products) in this procurement problem.

We first derive the optimal mechanism for the buyer. In the optimal mechanism, the buyer offers the supplier a nonlinear menu of contracts on each of the two attributes. With information asymmetry, the buyer would prefer to distort the centralized optimal solution in order to optimize her own objective function. For any cost realization, the optimal mechanism will induce an order
quantity and a service level that are both lower than the supply chain optimal solution.

We also include simplicity as a major consideration in the design of procurement mechanism. To simplify the nonlinear menu of contracts in the optimal mechanism, we use a fixed number for each of the two attributes. Through both numerical studies and analysis, we find that the quantity and service-level attributes have different implications for procurement mechanism design. For the service-level attribute, a fixed number works almost as well as the nonlinear menu in the optimal mechanism. On the other hand, using a fixed procurement quantity may lead to significant profit losses, especially when the buyer faces price-sensitive market demand. These results seem to be consistent with the practical observation that buyers tend to focus more on price-quantity negotiation while taking care of the supplier’s operational performance by specifying a minimum service level. In particular, we find that the fixed service-level contract, which consists of a target service level and a price-quantity menu, can achieve nearly optimal profit for the buyer. Furthermore, the fixed service-level contract resembles the quantity discount contract commonly used in practice, except that it is the buyer, rather than the supplier, who offers the contract. The quantity-discount contract has been widely recognized as a channel/supply chain coordination device in the marketing and operations literature. Here we demonstrate that it is also a useful tool in procurement contract design under information asymmetry.

An implicit assumption throughout the paper is that both firms will faithfully execute the contract terms. The buyer needs to commit to an average order quantity, and the supplier should adopt a contract-specifed service level. This assumption is reasonable in situations where firms care about long-term relationships, or firms’ actions are verifiable with a positive probability and there is a severe deviation penalty. However, there are also situations where firms’ actions are difficult to verify. In this case, the above contracts need to be modified to induce self-enforcing actions. A detailed discussion of incentive schemes that may achieve this goal is provided in Appendix C of the Electronic Companion.

Finally, this research can be extended in several directions. First, an alternative to model price-sensitive demand is to use the multiplicative form (i.e., $D(p) = D_0(p)\varepsilon$). With a multiplicative demand function, the analysis is less transparent because now the pricing decision affects the demand variance in each period, which further complicates the convoluted distribution for the lead time demand. Second, stationary contract formats have been assumed in this paper. Given the multi-period setting, the buyer may design a dynamic procurement mechanism in which contracting terms vary over time. For example, the procurement price may be a function of the supplier’s delivery performance in the previous periods. Next, one may incorporate information-updating into
the procurement problem. That is, the buyer might be able to learn about the supplier’s private cost information through the first few periods. Meanwhile, the supplier may react accordingly by not revealing the true information too early. All these extensions are analytically challenging. Nevertheless, they are quite interesting and deserve further research attention.

References


*Operations Research* 47(3) 454–475.


Electronic Companion

Appendix A: Proofs of Lemmas and Theorems

Proof of Lemma 1

(i) We can rewrite \( H_1(s, s_1) = \int_0^s \phi^{L_2}(y)G_1(s_1)dy + \int_s^\infty \phi^{L_2}(y)G_1(s_1 + s - y)dy \). Note that

\[
\frac{ds_1(s)}{ds} = -\frac{\partial^2 H_1(s, s_1)/\partial s_1 \partial s}{\partial^2 H_1(s, s_1)/\partial s_1^2} = -\frac{\int_s^\infty \phi^{L_2}(y)G_1''(s_1 + s - y)dy}{\Phi^{L_2}(s)G_1'(s_1) + \int_s^\infty \phi^{L_2}(y)G_1''(s_1 + s - y)dy}.
\]

We know \( s_1(s) \) is decreasing in \( s \) because \( G_1(\cdot) \) is convex and \( G_1''(\cdot) > 0 \).

(ii) Given \( s \), the first-order condition for the buyer is

\[
\frac{\partial H_1(s, s_1)}{\partial s_1} = \Phi^{L_2}(s)G_1'(s_1) + \int_s^\infty \phi^{L_2}(y)G_1'(s_1 + s - y)dy = 0.
\]

When \( s \to \infty \), we have \( G_1'(s_1) = 0 \) which simplifies to \( (h_1 + b)\Phi^{L_1+1}(s_1) - b = 0 \). Inverting the above equation gives the desired result. \( \square \)

Proof of Lemma 2

(i) By the Envelope Theorem

\[
\frac{dH_1(s, s_1(s))}{ds} = \frac{\partial H_1(s, s_1)}{\partial s} \bigg|_{s_1=s_1(s)} = \int_s^\infty \phi^{L_2}(y)G_1'(s_1 + s - y)dy.
\]

Since \( s_1 \) is optimally chosen given \( s \), the first-order condition \( \partial H_1(s, s_1)/\partial s_1 = 0 \) implies

\[
\Phi^{L_2}(s)G_1'(s_1) + \int_s^\infty \phi^{L_2}(y)G_1'(s_1 + s - y)dy = 0. \tag{18}
\]

Note also that \( G_1'(\cdot) = (h_1 + b)\Phi^{L_1+1}(\cdot) - b. \) Together we have

\[
\frac{dH_1(s, s_1(s))}{ds} = -\Phi^{L_2}(s)[(h_1 + b)\Phi^{L_1+1}(s_1) - b]. \tag{19}
\]

Since \( G_1(\cdot) \) is convex and hence \( G_1'(\cdot) \) is increasing, the first-order condition in (18) implies that \( G_1'(s_1) = (h_1 + b)\Phi^{L_1+1}(s_1) - b > 0 \). Otherwise there is \( G_1'(s_1 + s - y) < 0 \) for \( y > s \) and (18) cannot hold. This completes the proof for part (i).

(ii) Note

\[
\left( \frac{dH_1(s, s_1(s))}{ds} \right) /\Phi^{L_2}(s) = -[(h_1 + b)\Phi^{L_1+1}(s_1) - b],
\]

and the term in the bracket is positive. The proof is complete since \( s_1(s) \) is decreasing in \( s \) as shown in Lemma 1. \( \square \)

Proof of Lemma 3
Recall that $\mu(p)$ is concave by assumption. So the unique profit-maximizing $p^*$ can be characterized by

$$\frac{d\pi_{sc}}{dp} = \frac{d\mu(p)}{dp} [p - c - (h_0 + rc)L_1] + \mu(p) = 0.$$ 

Once the optimal market price $p^*$ and mean demand $\mu(p^*)$ are chosen, we have a standard two-stage serial inventory system. We know from the literature (e.g., Chen and Zheng, 1994) that the optimal inventory policy $(s^*, s^*_1)$ is characterized by the following equations:

$$\Phi^{L+1}(s_1) = (h_2 + b)/(h_1 + b),$$

$$0 = -b + (b + h_2)\Phi^{L}(s) + (b + h_1) \int_{s}^{\infty} \phi^{L}(y)\Phi^{L+1}(s_1 + s - y)dy.$$ 

\[\square\]

**Proof of Theorem 1**

The first-order condition for truth-telling is

$$T'(x) - c\mu'(p)p'(x) - (h_0 + rc)[\Phi^{L}(s)s'(x) + L_1\mu'(p)p'(x)] = 0$$

at $x = c$ for all $c$. Equivalently, replacing $c$ by $x$ gives

$$T'(x) - x\mu'(p)p'(x) - (h_0 + rx)[\Phi^{L}(s)s'(x) + L_1\mu'(p)p'(x)] = 0. \quad (20)$$

The second-order condition for truth-telling is

$$T''(x) - c(\mu'(p)p')' - (h_0 + rc)[\Phi^{L}(s)s''(x) + \phi^{L}(s)(s'(x))^2 + L_1(\mu'(p)p'(x))'] \leq 0 \quad (21)$$

at $x = c$ for all $c$. Differentiating (20) and plugging into (21) gives $\mu'(p)p' + r[\Phi^{L}(s)s' + L_1\mu'(p)p'] \leq 0$. So for the second-order condition to hold, it suffices to have

$$\mu'(p)p' \leq 0 \text{ and } s' \leq 0 \quad (22)$$

Let us assume (22) holds for now and later we will show that it is true.

Conditions (20) and (21) constitute the local incentive constraints, which ensure that the supplier will truthfully announce his cost locally. Now we check the global incentive constraint, that is, $\pi_2(c, c) \geq \pi_2(c, x)$ for all $x \neq c$. Note that

$$\pi_2(c, c) - \pi_2(c, x) = T(c) - T(x) - [c\mu(p(c)) - c\mu(p(x))]$$

$$-(h_0 + rc) \left\{ s(c)\Phi^{L}(s(c)) - s(x)\Phi^{L}(s(x)) - \int_{s(x)}^{s(c)} t\phi^{L}(y)dy + L_1[\mu(p(c)) - \mu(p(x))] \right\}.$$

34
From the truth-telling condition (20) we have

\[ T(c) - T(x) - [\epsilon(p(c)) - \epsilon(p(x))] - (h_0 + rc) L_1[\mu(p(c)) - \mu(p(x))] = \int_x^c (h_0 + r y) \Phi^{L^2}(s(y))s'(y)dy = \int_{s(x)}^{s(c)} (h_0 + rs^{-1}(z)) \Phi^{L^2}(z)dz, \]

where the second equality follows from transformation of variables \((z = s(y))\). Hence we need to show

\[ \int_{s(x)}^{s(c)} (h_0 + rs^{-1}(z)) \Phi^{L^2}(z)dz \geq (h_0 + rc) \int_{s(x)}^{s(c)} \Phi^{L^2}(z)dz \]

or \( \int_{s(x)}^{s(c)} (s^{-1}(z) - c) \Phi^{L^2}(z)dz \geq 0 \), which is straightforward for both \( x > c \) and \( x < c \) since \( s'(x) \leq 0 \).

Under truth-telling, we have \( \pi_2(x) = T(x) - x \mu(p) - (h_0 + rx) \left[ E(s - DL^2)^+ + L_1 \mu(p) \right] \). Taking derivative of \( \pi_2(x) \) and using (20), we have

\[ \pi_2'(x) = -\mu(p) - r \left[ \int_0^{s(x)} \phi^{L^2}(y)(s(x) - y)dy + L_1 \mu(p) \right]. \] (23)

If we use the profit function, \( \pi_2(x) \), to replace the lump-sum payment function \( T(x) \), then the buyer’s problem can be written as

\[
\max_{\{\mu(\cdot), s(\cdot), \pi_2(\cdot)\}} \pi_1 = \int_c^e [\mu(p)(p - x) - \pi_2(x) - H_1(s, s_1(s)) - H_2(x, s)] f(x)dx
\]

s.t. (23), (22) and (8).

This is an optimal control problem: \( \pi_2 \) is the state variable, and \( p \) and \( s \) are the control variables. Define

\[
J(p, s, \pi_2, \lambda, x) = [\mu(p)(p - x) - \pi_2(x) - H_1(s, s_1(s)) - H_2(x, s)] f(x) + \lambda \left[ -\mu(p) - r \int_0^{s(x)} \phi^{L^2}(y)(s(x) - y)dy - rL_1 \mu(p) \right]
\]

as the Hamiltonian. From the Pontryagin principle: \( \lambda' = -\partial J/\partial \pi_2 = f(x) \). From the transversality condition (since there is no constraint on \( \pi_2(\cdot) \) at \( e \)), we know \( \lambda = F(x) \). Optimizing with respect to \( s \) yields

\[
\frac{dH_1(s, s_1(s))/ds}{\Phi^{L^2}(s)} = - \left[ h_0 + rx + r \frac{F(x)}{f(x)} \right].
\]

Due to Lemma 2 and log-concavity of \( F \), there is a unique solution \( s \) to the above equation for given \( x \) and \( s(x) \) is a decreasing function. Optimizing with respect to \( p \) gives

\[
\left\{ \frac{d[\mu(p)(p - x)]}{dp} - (h_0 + rx)L_1 \frac{d\mu(p)}{dp} \right\} f(x) - \frac{d\mu(p)}{dp}(1 + rL_1)F(x) = 0,
\]

35
which in turns gives (10). Finally, the lump-sum payment function $T(x)$ can be solved using the IC and IR constraints in (20) and (8). □

**Proof of Theorem 2**

First we show $p^o(x) \geq p^*(x)$. Note that $\mu'(p) \leq 0$, $\mu''(p) \leq 0$ and $\frac{F(x)}{f(x)}$ is increasing in $x$. The result follows by comparing the two equations in (4) and (10). This further implies that $\mu^o(x) \leq \mu^*(x)$.

Second we show $s^o(x) \leq s^*(x)$. The optimal inventory policy $(s^*_1, s^*_o)$ for the centralized system are given by (5) and (6). From (19) in Lemma 2 and (11) in Theorem 1, we can characterize the base stock levels $(s_1, s^o)$ in the optimal mechanism for given supplier cost $x$:

$$
\Phi^{L1+1}(s_1) = \left[ h_2 + rF(x) / f(x) + b \right] / (h_1 + b) \quad (24)
$$

$$
0 = -b + (b + h_2 + rF(x) / f(x))\Phi^{L2}(s^o) + (b + h_1) \int_{s^o}^{\infty} \Phi^{L2}(y) \Phi^{L1+1}(s_1 + s^o - y) dy. \quad (25)
$$

From (5) and (24) we know $s_1 \geq s^*_1$. Consider the right hand side of Equation (25). Taking derivative with respect to $s^o$ and using Equation (24), we can show that it is increasing in $s^o$. Then from (6) and (25) we can infer $s^o \leq s^*$. □

**Proof of Theorem 3**

The proof is similar to that of Theorem 1. Consider the suppliers’ bidding behavior. Supplier $i$ maximizes her own expected profit:

$$
\max_{c^i} \pi^i_2 = E_{c^i} [T^i(c^i) - q^i(c^i)(c^i \mu + H_2(c^i, s^i))] .
$$

The first-order condition for truth-inducing is

$$
\frac{d}{dc^i} T^i(c^i) = \frac{d}{dc^i} E_{c^i} [q^i(c^i)(c^i \mu + H_2(c^i, s^i))] \quad \text{at } c^i = c^o \quad \text{for all } i. \quad (26)
$$

We now assume that $q^i(\cdot)$ and $s^i(\cdot)$ are nonincreasing functions in $c^i$, and check later that they are indeed nonincreasing in the optimal mechanism. It follows that the first-order condition in (26) is sufficient for truth telling (see the proof of Theorem 1).

Define $\pi^i_2(c^i)$ to be the expected profit for supplier $i$ under truth telling:

$$
\pi^i_2(c^i) = E_{c^i} [T^i(c^i) - q^i(c^i)(c^i \mu + H_2(c^i, s^i))]. \quad (27)
$$

From (26) and (27) we have

$$
\pi^i_2(c^i) = -E_{c^i} \left[ q^i(c^i) \left( \mu + \frac{\partial H_2(c^i, s^i)}{\partial c^i} \right) \right]. \quad (28)
$$
We can see that \( \pi_2^i \) is nonincreasing in \( c^i \) \((H_2(c^i, s^i(c^i))) \) is convex in \( c^i \), so we can set

\[
\pi_2^i(c) = 0, \text{ all } i. \quad (29)
\]

Once the optimal \( q^i(\cdot) \) is given, so that \( Q^i(c^i) = E_{c^i} q^i(c) \) is given (i.e., the expected probability for supplier \( i \) to win is given), the optimal control problem can be decomposed into \( n \) programs as follows:

\[
\max_{\{p^i(\cdot), s^i(\cdot), \pi_2^i(\cdot)\}} \pi_1 = \int_\mathcal{C} \{Q^i(c^i)[\mu(p^i)(p^i - c^i) - H_1(s^i, s_1(s^i)) - H_2(c^i, s^i(c^i))] - \pi_2^i(c^i)\} f(c^i) dc^i \tag{30}
\]

s.t. \( \pi_2^i(c^i) = -Q^i(c^i) \left( \mu + \frac{\partial H_2(c^i, s^i(c^i))}{\partial c^i} \right) \), \( \pi_2^i(c) = 0. \tag{31} \tag{32} \)

This is an optimal control problem with \( \pi_2^i \) as the state variable and \( p^i \) and \( s^i \) as the control variables. Solving this problem gives (see the proof of Theorem 1):

\[
\frac{d\mu(p^i)}{dp^i} \left[ p^i - c^i - (h_0 + rc^i)L_1 - (1 + rL_1)\frac{F(x)}{f(x)} \right] + \mu(p^i) = 0,
\]

\[
\frac{dH_1(s^i, s_1(s^i))/ds^i}{\Phi L_2(s^i)} + \left[ h_0 + rc^i + r\frac{F(c^i)}{f(c^i)} \right] = 0.
\]

The functions \( p^i \) and \( s^i \) are exactly the same as in Theorem 1. Thus \( p^i \) is nondecreasing in \( c^i \) and \( s^i \) is nonincreasing in \( c^i \). From (31), we have

\[
\int_\mathcal{C} \pi_2^i(c^i)f(c^i) dc^i = \pi_2^i(c^i)F(c^i)|_\mathcal{C} - \int_\mathcal{C} F(c^i) d\pi_2^i(c^i)
= \int_\mathcal{C} \left[ F(c^i)Q^i(c^i) \left( \mu + \frac{\partial H_2(c^i, s^i(c^i))}{\partial c^i} \right) \right] dc^i.
\]

Therefore, the buyer’s profit function in (30) can be written as

\[
\int_\mathcal{C} Q^i(c^i) \left[ \mu(p^i)(p^i - c^i) - H_1(s^i, s_1(s^i)) - H_2(c^i, s^i) - \frac{F(c^i)}{f(c^i)} \left( \mu + \frac{\partial H_2(c^i, s^i(c^i))}{\partial c^i} \right) \right] f(c^i) dc^i.
\]

It is straightforward to show that the terms inside the bracket is decreasing in \( c^i \). Hence we should give more weight to \( Q^i(c^i) \) when \( c^i \) is small. Since there are \( n \) symmetric suppliers, the optimal \( q^i(\cdot) \) must be \( q^i(c) = 1 \) if \( c^i < \min_{j \neq i} c^j \) and \( q^i(c) = 0 \) otherwise. That is, in the optimal mechanism, the most efficient supplier is chosen with probability one. As a result, \( Q^i(c^i) = (1 - F(c^i))^{n-1} \).

We can derive the profit function \( \pi_2^i \) from (31) and (32) and then the payment function \( T^i \) from (13) and (14). \( \Box \)
Appendix B: Numerical Experiments with Two-Point Cost Distributions

Section 5.1 demonstrates that in the numerical study with uniform and normal cost distributions, the buyer’s cost in the optimal mechanism is quite flat. To confirm the generality of this observation and obtain further intuition, we conduct additional numerical experiments with two-point cost distributions. With only two supplier types (efficient type and inefficient type), we can define information rent as the profit for the efficient supplier (the inefficient type will get zero profit). The purpose of the information rent is to induce truth telling from the efficient supplier.

We modify the numerical study in Section 5.1 as follows: The supplier’s cost can be either $c_L = 1$ with probability $v$ or $c_H = 1.4$ with probability $1 - v$, i.e., the inefficient supplier has a cost that is 40% higher than that of the efficient supplier. We consider $v = \{0.1, 0.5, 0.9\}$ and there are totally 648 scenarios. Note that $v = 0.1$ and 0.9 represent extreme cases (the cost distribution is highly skewed to one side). We assume that the buyer needs to contract with the supplier regardless of his cost type. First we compare the performance of the FS contract to that of the optimal mechanism. It has been found that the percentage cost increase in the FS contract has a maximum of 3.13% and an average of 0.18% among all scenarios. Thus the simple mechanism is nearly optimal even with extreme two-point cost distributions. Then we examine the buyer’s ex post cost in the optimal mechanism. In particular, we find that the ratio $C_{OM}^1(c_H)/C_{OM}^1(c_L)$ ranges from 1.00 to 1.23, with a value of 1.01 on average and 1.02 at the 90th percentile. So the buyer’s total cost in the optimal mechanism is again quite flat in most cases. (Note that the FS contract is close to optimal even when there is some variation in the buyer’s ex post cost in the optimal mechanism, e.g., when $C_{OM}^1(c_H)/C_{OM}^1(c_L) = 1.23$. It is because the buyer’s cost in the FS contract is a constant, which is higher than $C_{OM}^1(c_L)$ but lower than $C_{OM}^1(c_H)$.) This has confirmed our intuition described in Section 5.1 based on continuous cost distributions: Inducing truth telling reduces the supply chain’s total cost (since the buyer can request a cost-contingent service level), but such a cost reduction mainly benefits the supplier through information rent due to his superior bargaining position.

Appendix C: Unverifiable Firm Actions

This paper has implicitly assumed that supply chain firms’ actions are enforceable. This assumption is reasonable under certain situations and has been used in the literature (see Cachon and Lariviere, 2001 for a discussion of various enforcement schemes in supply chain contracts). However, there are also situations where firms’ actions are difficult to enforce. For example, it may not be feasible for the buyer to verify the supplier’s base stock level, especially with the presence of random demand noise. Similarly, the average purchase quantity in each period depends on the
buyer’s selling price, which may not be publicly observable when transacting with distributors and retailers. In this appendix we briefly discuss how to design incentive schemes to induce desirable firm actions.

First we focus on the supplier’s action. For illustration, suppose the buyer offers the supplier a fixed service-level (FS) contract \(\{\tilde{s}, \mu(\cdot), w(\cdot)\}\), where \(\tilde{s}\) is the required base stock level. Since the true base stock level is unverifiable, the supplier may have incentives to reduce his base stock level to \(s < \tilde{s}\). To deal with this moral hazard problem, the buyer may charge the supplier a penalty cost \(\omega\) per unit of backorder (see Cachon and Zipkin, 1999 for a similar scheme). Note that the buyer observes the quantity received in each period and knows the order history as well. So the backorder level \((s - \varepsilon^{L_2})^+\) in each period is verifiable. The supplier’s expected profit can be now written as

\[
\pi_2(c, s) = \mu(c)w(c) - \omega E[(s - \varepsilon^{L_2})^+] - c\mu(c) - H_2(c, s),
\]

When choosing \(\omega\), we require the following condition to hold at \(s = \tilde{s}\):

\[
\frac{d\pi_2(c, s)}{ds} = -\frac{dE[\omega(s - \varepsilon^{L_2})^+]}{ds} - \frac{dH_2(c, s)}{ds} = 0,
\]

which yields

\[
\omega = (h_0 + rc) \frac{\Phi^{L_2}(\tilde{s})}{1 - \Phi^{L_2}(\tilde{s})}.
\]

This condition guarantees that it is in the supplier’s best interest to choose \(s = \tilde{s}\) and can be easily solved. In addition, as a compensation, the buyer pays the supplier a lump-sum \(\hat{T} = E[\omega(\tilde{s} - \varepsilon^{L_2})^-]\). This lump-sum payment ensures that for any cost \(c\), the supplier’s and the buyer’s expected profits remain unchanged. Thus the inclusion of \(\omega\) and \(\hat{T}\) does not affect the original contract \(\{\tilde{s}, \mu(\cdot), w(\cdot)\}\).\(^5\) The notion of penalizing unsatisfactory suppliers is intuitive. It seems to be an effective tool to manage supplier performance given its prevalence in practice (see Kay, 2005 and Thomas, 2005).

Next we consider the buyer’s action. The buyer’s order quantity decision \(\mu\) is equivalent to her market price decision \(p\). For a procurement price \(w\), the buyer’s optimal market price \(\hat{p}(w)\) is given by

\[
\hat{p}(w) = \arg \max_p \pi_1 = (p - w)\mu(p) - H_1(s, s_1(s))
\]

\(^5\)This analysis depends on the assumption of stationary base stock policy in Section 3. If non-stationary inventory policies are allowed, then the firms’ ordering behavior may change. For instance, when the penalty cost \(\omega\) is large enough relative to the inventory holding cost \(h_1\), the buyer may over-order in certain periods to take advantage of the backorder penalty charged to the supplier. The buyer’s dynamic ordering behavior will trigger a non-stationary inventory policy from the supplier as well. A full analysis of such a dynamic inventory game between the firms is quite complex and beyond the scope of this paper.
and must satisfy the first-order-condition

$$\frac{d\mu(p)}{dp}(p - w) + \mu(p) = 0. \quad (33)$$

Note that $H_1(s, s_1(s))$ is independent of $p$ based on the transformation explained in Section 3. Thus, given the optimal contract $\{\mu^o(\cdot), s^o(\cdot), w^o(\cdot)\}$ derived in Section 4.1, the buyer may wish to deviate from $p^o(\cdot)$ to $\hat{p}(w^o(\cdot))$. Let $\hat{w}(\cdot)$ be a price function satisfying $\hat{p}(\hat{w}(\cdot)) = p^o(\cdot)$ (i.e., under $\hat{w}(\cdot)$, it is in the buyer’s best interest to choose $p^o(\cdot)$). By comparing (33) and (10), we know

$$\hat{w}(x) = x + (h_0 + rx)L_1 + (1 + rL_1)\frac{F(x)}{f(x)}.$$

Let $\hat{T}(\cdot)$ be a transfer payment from the buyer to the supplier such that $\hat{T}(\cdot) + \hat{w}(\cdot)\mu(p^o(\cdot)) = w^o(\cdot)\mu^o(\cdot)$ (i.e., the total transfer payment remains unchanged if $p^o(\cdot)$ is chosen). Now consider the following new contract $\{\mu^o(\cdot), s^o(\cdot), \hat{w}(\cdot), \hat{T}(\cdot)\}$. In this new contract, the buyer will choose $p^o(\cdot)$ by the definition of $\hat{w}(\cdot)$. Moreover, both firms’ payoffs are the same as before, which implies that the outcome of the new contract will be the same as that of the optimal contract $\{\mu^o(\cdot), s^o(\cdot), w^o(\cdot)\}$. Therefore, with a carefully designed incentive scheme, the supplier can be assured that the buyer will adopt the contract-specified market price $p^o(\cdot)$ even if her action is not verifiable. A numerical example that visualizes the above incentive scheme is given in the Supplemental Appendix.