The role of cost modeling in competitive bid procurement

Yan Yin, Hyun-Soo Ahn, Damian R. Beil
Ross School of Business at the University of Michigan, 701 Tappan Street, Ann Arbor, MI, 48109

Industrial buyers often create cost models to estimate the costs of potential suppliers and use the information to negotiate an attractive price with suppliers. However, in settings where the buyer wishes to negotiate a contract price by soliciting competitive bids from multiple suppliers, the benefit of cost modeling is not clear since the bidding competition among suppliers can itself reveal cost information. In this paper, we examine this interplay, and examine if and when cost modeling should be used prior to competitive bidding. We show that although bid competition sometimes duplicates the information gleaned by cost modeling, the latter can still be beneficial when it helps the buyer set an effective reserve price. We then analyze how the buyer can gain the most benefit through cost modeling. Specifically, we characterize which supplier(s) to learn about, which portion(s) of the costs to learn, and how deeply the buyer should learn. Interestingly, learning about the supplier whose cost is the most uncertain is not necessarily optimal, nor is learning about the cost portion that contributes most to the total cost. We also show that conventional intuition that the benefit of additional information has a diminishing rate of return does not always apply.

Key words: Cost Modeling, Procurement Design, Optimal Learning Strategy

1. Introduction

The average US manufacturer spends roughly 57% of its revenue on purchases from external suppliers (U.S. Department of Commerce 2011). As manufactures come to increasingly rely on suppliers for intermediate and even finished goods, many have entirely eliminated their in-house production capabilities. A large tier-one automotive firm we recently worked with closed down their last internal US plant for an automotive part, and now must purchase the part externally. In other industries such as electronics, Cisco relies on outside suppliers for manufacturing virtually all its products.

For firms who rely on suppliers for production, it is natural to experience atrophying capabilities in manufacturing, as process engineers, production technology experts, material experts, etc. are shed (Pisano and Shih 2009). However, this eventually leaves the firm with little insight into current production trends, technologies, and conditions that are important in determining production costs. A firm who lacks such information is at a disadvantage when buying from suppliers, since suppliers
can earn outsized profits by exploiting the buyer’s lack of knowledge about the true production cost. For example, at the automotive firm we worked with, several years after closing down their internal US plant they found themselves with little insight into what the true production cost was for the auto part. They were concerned that the price they had paid for external supply might be far above the true cost.

One way to address this concern is through cost modeling, which enables the buyer to estimate a supplier’s true cost. The firm mentioned above did just this: It allocated two employees for more than one month to painstakingly create a cost model for the auto part. The team tracked down a production engineer formerly employed at the firm’s now shuttered internal plant to learn how the part had been made. They also interviewed industry experts to learn about current manufacturing practices, scoured reports on prevailing labor and utility costs, gathered, cleaned and analyzed data, etc.

As this example illustrates, cost modeling is an expensive and time-consuming endeavor. It should only be used when the firm is confident that it will get its money’s worth out of the exercise. Alternatively, the firm can discover cost information by having suppliers submit competitive bids, since suppliers have to lower their bids - potentially close to their true costs — in order to win the business. In the aforementioned example, the firm actually knew of multiple suppliers who were capable of producing the needed auto part, and planned to put the production contract up for competitive bidding.

In this paper we ask how these two common ways of discovering cost information — cost modeling and competitive bidding — should be used by a buyer who seeks to manage its procurement costs. After deciding to source (e.g., after winning a new production contract from a downstream customer), the buyer needs to select an upstream supplier in a timely fashion to allow adequate ramp-up time to avoid production delays. Since cost models take many weeks or possibly months to develop, we take the perspective of a buyer who, at the beginning of a bidding cycle, must decide which supplier(s) to learn about, which portion(s) of their costs to learn, and how deeply to learn. More specifically, we investigate the following research questions:

**Research Question 1:** How does one characterize the benefit of cost modeling? Is cost modeling beneficial to the buyer even when suppliers compete?

**Research Question 2:** If the buyer chooses to develop a cost model, which supplier(s) should the buyer learn about?

**Research Question 3:** If the buyer can choose to learn about various portions of the cost, e.g., labor cost, input cost, packaging cost, etc., which portion(s) should the buyer learn about?

**Research Question 4:** If the buyer chooses to develop a cost model, how deeply into supplier(s)’ costs should the buyer learn? Would the buyer ever want to learn about multiple competing suppliers?
Answers to these research questions are summarized in Section 7. We review related literature in Section 2, and introduce the model in Section 3. Sections 4 and 5 study the two-supplier and N-supplier cases, respectively. We extend our analysis to include the cost of learning and provide an optimal learning strategy incorporating it in Section 6. Proofs of results are furnished in the Appendix.

2. Literature Review

There is a vast economics literature on competitive bidding (auctions). Krishna (2002) provides an excellent introduction to the early literature. There is a growing literature on procurement auctions, examining various operational aspects such as the need to consider quality as well as price (Che 1993, Beil and Wein 2003, Kostamis et al. 2009), the need to perform costly supplier qualification screening (Wan and Beil 2009, Wan et al. 2010), auctions with quantity-flexible pricing (Chen 2007, Duenyas et al. 2011, Tunca and Wu 2009), post-award audits and profit-sharing (Chen at al. 2008), supplier capacity decisions (Li 2011), unreliable suppliers (Chaturvedi et al. 2010), double auctions (Chu 2009), and large-scale combinatorial auctions (Olivares et al. 2011). However, none of these papers study cost modeling prior to competitive bidding, which is the key element of study in our paper.

A number of papers apply econometric analyses on historical bid data to estimate bidders’ private information. For example, Paarsch (1997) uses historical auction field data to derive an estimate of the optimal reserve price in English auctions for government timber sales (see the recent text by Paarsch and Hong (2006) for more examples). In contrast to these papers, we take the buyer’s prior beliefs about suppliers’ costs as our point of departure. We examine how the buyer can deploy a totally different approach to refine her prior information, namely creating a cost model to derive suppliers’ costs using a bottom-up approach. The main aim of our paper is to examine if the buyer should engage in the time-consuming cost modeling prior to competitive bidding, and how and to what extent the buyer should deploy it. To our knowledge, ours is the first paper in the literature to examine cost modeling in this context.

The practical importance of cost modeling has grown in tandem with the increased use of outsourcing. Chapter 3 in Laseter (1998) provides an introduction to cost modeling, and describes various ways that a supplier’s cost can be decomposed into a number of cost drivers, each of which can then be estimated through various means of data collection. One way of gathering data about suppliers firsthand is the so-called Rapid Plant Assessment technique. Goodson (2002) describes how this careful approach to factory tours allows a trained team of procurement specialists to estimate a plant’s cost of sales based on key observations made while touring the plant.
3. Model Description and Preliminaries

We consider a buyer seeking to allocate a single, indivisible contract to a supplier for a good that the buyer cannot produce herself. To build intuition, we first examine a setting where the buyer has a single supplier.

3.1. One supplier

Let $c$ be the supplier’s cost, which is the supplier’s private information. The value $c$ is a realization of random variable $C$ drawn from a continuous distribution $H$. We assume that $C$ is bounded above by $\bar{c}$. The buyer can employ a reserve price, $r$, as a take-it-or-leave-it offer to the supplier, and a supplier will accept this offer if this price can cover his production cost, i.e., $r \geq c$.

The process of locating and qualifying a supplier who is capable of producing the needed good is typically very time-consuming. For example, at Fortune 500 manufacturers we interacted with, the supplier qualification processes can take months, even for “commodity-type” goods like cables, connectors, and screw machine parts. At these firms, functional areas such as engineering invest time and resources in the supplier qualification to verify supplier capabilities and approve the supplier. Because the buyer has no internal production capabilities, the supplier is already approved, and it would take months and be politically costly to find and qualify a new supplier, the procurement manager’s job is to strike a deal with this supplier. Of course, the buyer would like to strike a deal at the lowest possible contract payment. However, to avoid the risk of the supplier walking away, the buyer will set the reserve price, $r$, high enough to ensure that the supplier will be interested in the contract. The best price the buyer can offer is $r^o = \bar{c}$. This reserve price may leave the supplier a windfall profit if the upper bound $\bar{c}$ is high above the supplier’s true cost $c$.

To get a better estimate of the supplier’s cost, the buyer can exert effort to build a cost model and collect data on the supplier before deciding her reserve price. For instance, for a manufactured good, a buyer might decompose the supplier’s production cost into the sum of several attributes — such as raw material cost, direct labor cost, electricity cost, technology cost, and overhead cost — and learn some of these attribute costs through data collection. For example, a buyer can contact raw material providers to learn the cost of material that a supplier uses, estimate labor cost using local labor cost reports, or send specialists to visit a supplier’s plant to learn the type of machines used and hence the technology cost. Of course, not all attributes can be learned; for example, exact overhead cost is usually difficult to learn.

To model this, we assume that the supplier’s cost $C$ can be decomposed into two portions: $A$, the portion of cost attributes that can be learned via cost modeling, and $X$, the portion that cannot be learned. We assume that $A$ follows a continuous distribution $F$ and is bounded above by $\bar{a}$, $X$ follows a continuous distribution $G$ and is bounded above by $\bar{x}$. As such, the supplier’s cost $C$’s
distribution $H$ is the convolution of $F$ and $G$ ($H = F \otimes G$) and the upper bound of $C$ satisfies $\bar{c} = \bar{a} + \bar{x}$. Let $a$ and $x$ denote the realizations of $A$ and $X$, respectively.

After learning the supplier’s portion, $A = a$, the buyer can update her knowledge about the supplier’s cost: $[C| (A = a)] = a + X$. This random variable $C|(A = a)$ has a new upper bound $[\bar{c}^\prime|(A = a)] = a + \bar{x}$, so the buyer can set a new reserve price $[r^\prime|(A = a)] = [\bar{c}^\prime|(A = a)] = a + \bar{x}$, which is the new contract payment. The benefit of learning is the difference in the contract payment without learning and with learning, denoted by $\psi$:

$$
\psi = r^o - r^l = \bar{c} - (A + \bar{x}) = \bar{a} - A.
$$

Obviously, learning brings the buyer a positive benefit, as the following shows:

**Lemma 1.** If the buyer’s supply base has only one supplier, $P(\psi > 0) = 1$ and $E[\psi] = \bar{a} - E[A]$.

### 3.2. Two suppliers

Now we consider the case when the buyer has two suppliers. It is less clear whether learning about the supplier(s)’ cost information can bring the buyer a positive benefit, because competition between suppliers itself acts as a cost-discovery tool.

To explore this interplay, we assume the two suppliers’ costs, $C_1$ and $C_2$, are independent, and follow distributions $H_1$ and $H_2$ with upper bounds $\bar{c}_1$ and $\bar{c}_2$, respectively. The buyer first sets a reserve price $r$ as a starting point and then asks the suppliers to give price offers at or below price $r$. The buyer can play the suppliers’ offers against one another by treating the current lowest offer as the winner unless this offer is unseated by a lower offer from the other supplier. As we have observed in industries ranging from electronics to automotive parts, such competitive bid processes are often administered via an online platform where the bidding takes place over a short period of time (e.g., 30 minutes). The ensuing process drives down the contract price until it becomes so low that only one supplier (the contract winner) remains. This competitive process can be modeled as a reverse open-descending auction for the contract, for which the dominant strategy for a bidder is to continue bidding until reaching their true cost or winning the auction. The predicted equilibrium outcome is as follows: The lower-cost supplier wins the contract at a price equalling the cost of the other supplier. The reserve price set by the buyer, $r$, acts as a price ceiling; hence, the buyer’s contract payment with two suppliers is

$$
\pi = \min(r, \max(C_1, C_2)).
$$

The buyer must transact with one of the suppliers qualified to produce the good, but seeks to set a reserve price that minimizes her expected contracting payment:

$$
\min \quad E[\pi]
$$

s.t. $P(\min(C_1, C_2) \leq r) = 1$. 

The solution of the above problem is:

\[ r^o = \min(\bar{c}_1, \bar{c}_2), \]

so the buyer’s contracting payment is:

\[ \pi^o = \min(r^o, \max(C_1, C_2)). \]

Notice that \( r^o \) is the lowest reserve price that can avoid the risk of non-transaction since the supplier with the lower upper bound can generate non-negative surplus and will surely be interested in the contract.

Similar to the one-supplier case, the buyer can create cost models to better estimate the suppliers’ costs. For supplier \( i = 1, 2 \), we let random variable \( A_i \) denote the learnable portion, with distribution \( F_i \) whose domain is bounded above by \( \bar{a}_i \). Likewise, let \( X_i \sim G_i \) denote the unlearnable portion, which is bounded above by \( \bar{x}_i \). After learning supplier \( i \)'s portion \( A_i = a_i \), the buyer can update her knowledge of supplier \( i \)'s cost \[ C_i | (A_i = a_i) = a_i + X_i \] which has a new upper bound \[ \bar{c}_i | (A_i = a_i) = a_i + \bar{x}_i \]. Hence, if the buyer learns about a subset of suppliers \( S \) (\( S \) can be \( \{1\} \), \( \{2\} \), or \( \{1, 2\} \)), the buyer can set a new reserve price:

\[ r^l = \min_{i \in S, j \notin S} (\bar{c}_i, \bar{c}_j) = \min_{i \in S, j \notin S} (A_i + \bar{x}_i, \bar{c}_j). \]

Accordingly, the buyer’s contract payment after learning is:

\[ \pi^l = \min(r^l, \max(C_1, C_2)). \]

The benefit of learning is:

\[ \psi = \pi^o - \pi^l. \]  \( (1) \)

Unlike the one-supplier case, when there exists competition between suppliers learning does not always bring a positive benefit for the buyer.

**Lemma 2.** If there are two competing suppliers and one supplier’s cost is not dominated by the other supplier’s cost (i.e., \( 0 < P(C_1 > C_2) < 1 \)), then with strictly positive probability learning brings no benefit at all (i.e., \( P(\psi = 0) > 0 \)).

To understand the reason why competition between the two suppliers can make learning redundant, suppose the buyer learns about just one of the two suppliers. Competition will automatically identify the lower-cost supplier as the contract winner and can push the price down to the cost of the other supplier, so a good reserve price needs to be set close to the winner’s true cost so that the buyer can squeeze the winner’s surplus. Simply setting a lower reserve price does not guarantee it
will be effective. Indeed, the buyer may be able to set a new, lower the reserve price even when the supplier she learns about eventually loses. However, in this case, the new reserve price is never low enough to affect the final payment. Even if the buyer learns about the contract winner, learning can again be useless since if the new reserve price is not low enough it will not bind the payment. It should be noted that learning results in a positive benefit to the buyer only when the buyer learns about the contract winner and can set a new reserve price low enough to be the new contract payment.

However, this makes the buyer’s problem difficult since suppliers’ true costs are their private information and the buyer can not \textit{ex ante} predict with certainty who will be the contract winner or how close the suppliers’ costs will be. In the next section, we will discuss when to learn and how to learn so as to get the maximum expected benefit. We define the expected benefit of learning as \( \Psi = E[\psi] \). In this paper, the sample path value, \( \psi \), and its expectation, \( \Psi \), will be subscripted to denote which supplier(s) are learned about.

Thus far we have introduced the expected benefit of learning, which we analyze in Sections 4 and 5. The models in these sections do not consider the cost of learning. Of course, cost modeling is time-consuming and expensive as the buyer may need to visit the supplier’s site, interview industry experts, analyze industry reports and forecasts, etc. Practitioners need to consider the tradeoff between the expense and the benefit of cost modeling. Building on our base model analyses, we incorporate the cost of learning in Section 6.

4. Analysis – Two suppliers

We now answer the Introduction’s four research questions when the buyer has two competing suppliers. We allow the possibility that the two suppliers are heterogeneous. For example, the cost structure of a supplier using labor-intensive production processes will differ from that of a supplier whose production is highly automated. Suppliers located in different regions may have different energy and shipping cost structures. In these cases, cost models are supplier-specific and a single cost model is developed to estimate the cost information of one supplier. Without loss of generality, we consider learning about supplier 1’s learnable portion \( A_1 \), and let \( \Psi_1 \) denote the expected benefit of learning.

We also allow the other possibility, namely the two suppliers are homogeneous. In such cases, the suppliers can have different cost realizations but both share the same cost distributions, i.e., \( C_1, C_2 \) follow \( H \) with upper bound \( \bar{c} \), \( A_1, A_2 \) follow \( F \) with upper bound \( \bar{a} \) and \( X_1, X_2 \) follow \( G \) with upper bound \( \bar{x} \). For example, the suppliers could have similar, labor-intensive production processes, so one cost model can be used for both suppliers. Of course, even if they share the same cost drivers (e.g., labor) the suppliers’ costs can be different (due to differences in exact minutes
of direct labor per unit, suppliers’ labor wage rates, or fringe benefits for employees). Let $\Psi_{1&2}$ denote the expected benefit of learning about portion $A$ of both suppliers.

The following result characterizes the benefit of cost modeling prior to competitive bidding. This result and Lemma 2 characterize the benefit of learning and therefore answer research question 1.

**Proposition 1.** [Relation between learning and cost distributions]

(i) When the two suppliers are heterogeneous, the expected benefit of learning $A_1$ is

$$\Psi_1 = E[\min(\bar{a}_1 - A_1, (C_2 - (A_1 + \bar{x}_1))^+)].$$

As long as the domains of $F_1$ and $G_1$ do not change,

(a) $\Psi_1$ decreases as $F_1$ becomes stochastically larger;

(b) $\Psi_1$ is independent of $G_1$;

(c) $\Psi_1$ increases as $H_2$ becomes stochastically larger.

(ii) When the two suppliers are homogeneous, the expected benefit of learning $A_1$ and $A_2$ is

$$\Psi_{1&2} = E[(\max(C_1, C_2) - (\min(A_1, A_2) + \bar{x}))^+].$$

As long as the domains of $F$ and $G$ do not change,

(a) $\Psi_{1&2}$ is not necessarily monotonic as $F$ becomes stochastically larger;

(b) $\Psi_{1&2}$ increases as $G$ becomes stochastically larger.

Proposition 1 shows that the benefit of learning depends on the suppliers’ cost distribution in interesting ways. Consider the case in which the buyer builds a cost model for a single supplier in a heterogeneous supply base. As the distribution of the learnable portion, $F_1$, becomes stochastically larger, the benefit of learning decreases, as shown in part (i.a). This is intuitive because the new reserve price (which will stochastically increase in $F_1$) is less likely to bind the contract payment. Likewise, the benefit grows as the cost distribution of supplier 1’s opponent whom we do not learn about increases, explaining part (i.c). However, part (i.b) implies that the expected benefit of learning is independent of the distribution that governs supplier 1’s unlearnable cost, $G_1$. This is because after learning, the new reserve price, $a_1 + \bar{x}_1$, depends on $G_1$ only through its upper bound.

Next, consider the case where the buyer creates a cost model that applies to both suppliers in the supply base. Interestingly, the results for this case differ from what we saw above. As the learnable cost decreases, the reserve price, $\min(A_1, A_2) + \bar{x}$, decreases as well. But at the same time, the cost the buyer would have paid without learning, $\max(C_1, C_2)$, also decreases. As a result, the benefit of learning — the difference between these two — does not necessarily grow or shrink, as stated in part (ii.a). Moreover, part (ii.b) shows that the benefit of learning increases as the unlearnable portion becomes stochastically larger. This is because the cost the buyer would have paid without learning increases, but the reserve price remains unchanged.
4.1. Which supplier(s) should the buyer learn about?

For the heterogeneous suppliers case in the previous section, we assumed that the buyer learned about supplier 1. However, would the buyer have been better off learning instead about supplier 2? To help answer this, we need to compare the expected benefit of learning either supplier. We now examine this issue, addressing research question 2.

Intuitively, one may expect that the buyer wishes to resolve the maximal amount of cost uncertainty and so she will optimally choose whichever supplier has the most uncertain learnable cost. This logic is true when the buyer has only one supplier in her supply base. Suppose supplier 1’s learnable cost $A_1 \sim U[5, 10]$ and unlearnable cost $X_1 \sim U[0, 5]$, while supplier 2’s cost $A_2 \sim U[0, 10]$ and $X_2 \sim U[10, 11]$. Were the buyer facing a single-supplier situation with supplier 1, learning would bring an expected benefit of $\bar{a}_1 - E[A_1] = 2.5$ (Lemma 1). This is smaller than $\bar{a}_2 - E[A_2] = 5$, which would be her expected benefit from learning if supplier 2 was her only supplier. The intuition in the single-supplier case is that the buyer would gain more benefit from learning when the learnable portion has more uncertainty.

However, this result is no longer true when the buyer has two competing suppliers. The benefit of learning does not solely depend on the $A_i$’s. In fact, in the example above, learning about supplier 1 brings a much larger benefit (per equation (1), $\Psi_1 = 1.79 > 0.04 = \Psi_2$). Competition is what differentiates the one-supplier and two-supplier cases. Competition itself works as a cost-discovery tool, so learning brings a positive benefit only when it collects information that cannot be duplicated by competition. More precisely, the buyer needs information about the contract winner’s cost that helps the buyer to set an effective new reserve price. In this example, supplier 1’s cost $C_1$ is stochastically smaller than supplier 2’s cost $C_2$; this means that supplier 1 has a better chance of winning the contract. In addition, the new reserve price set by learning about supplier 1 (i.e., $A_1 + \bar{x}_1$) is more likely to be smaller than that for supplier 2 (i.e., $A_2 + \bar{x}_2$). As a result, learning about supplier 1 is more beneficial. Our result below formalizes.

**Proposition 2.** [Preferred supplier to learn about] For any distributions $F_i, G_i$ and $H_i$, $i = 1, 2$, the buyer prefers to learn about supplier 1 rather than supplier 2 if

$$A_1 + \bar{x}_1 \leq_{st} A_2 + \bar{x}_2 \text{ and } C_1 \leq_{st} C_2.$$ 

It should be noted that both conditions in Proposition 2 are necessary to draw the conclusion. Even if learning is likely to dramatically reduce the reserve price, it will be useless if it does not bind the contract payment to the contract winner.

Instead of learning about just one supplier, the buyer could instead choose to undertake learning about both suppliers. If suppliers are homogeneous, the buyer can create a single cost model that
applies to both suppliers. If suppliers are heterogeneous, the buyer needs to develop two different cost models concurrently. Since cost models take several months to develop, in most cases, it is not feasible to conduct multiple rounds of learning in a single bidding cycle. Thus, decisions on whom to learn about should be made judiciously at the beginning.

Utilizing the same machinery, our next result characterizes the benefit of learning about multiple suppliers. One may intuitively reason that the marginal benefit of learning will diminish in the number of suppliers that the buyer learns about — after all, the buyer ultimately only sets one new reserve price which applies to all suppliers, thus more learning will likely just generate duplicative information. In other words, the benefit of learning about two suppliers should be smaller than the sum of the benefits of learning about either supplier individually. Surprisingly, the next result shows that this is not the case: The benefit is actually additive.

**Proposition 3.** [Benefit of learning about multiple suppliers]

(i) For two heterogeneous suppliers, $\psi_{1,2} = \psi_1 + \psi_2$ almost surely; hence $\Psi_{1,2} = \Psi_1 + \Psi_2$.

(ii) For two homogeneous suppliers, $\Psi_{1,2} = 2 \cdot \Psi_1$.

To understand the intuition behind Proposition 3, recall that the buyer realizes a benefit from learning only if she learns about the contract winner. For a given sample path, at most one supplier can yield a positive benefit, i.e., $\psi_{1,2} = \psi_1, \psi_2 = 0$ or $\psi_{1,2} = \psi_2, \psi_1 = 0$, which explains part (i). Part (ii) follows from part (i) and the fact that the two suppliers have the same cost distribution.

4.2. Which portion(s) of cost should the buyer learn?

Another interesting challenge for the buyer is to determine which portion(s) of the suppliers’ costs she should learn. To address this question, in this section we assume that portions $A_i$ and $X_i$ are both learnable. Note that this setup is without loss of generality: our results still hold when supplier $i$’s cost is $C_i = A_i + X_i + \epsilon_i$, where $\epsilon_i$ is unlearnable. To address research question 3, we first ask which portion, if learned, would bring more benefit to the buyer.

Suppose two cost portions, labor and utility, comprise roughly 50% and 20% of the total cost, respectively. Intuitively, one might expect that the buyer would prefer to learn the labor cost, since it represents a larger share of the total cost. However, this intuition does not tell the whole story. What is more important for the buyer to consider is the amount of cost uncertainty which will be resolved by learning, not the absolute magnitude of the cost. This is the intuition behind the following proposition.

**Proposition 4.** [Preferred portion to learn]

(i) When the two suppliers are heterogeneous, for any distributions $F_i, G_i$ and $H_i$, $i = 1, 2$, the buyer prefers to learn portion $A_i$ rather than portion $X_i$ if

$$a_1 - A_1 \geq a_1 - X_1.$$
(ii) When the two suppliers are homogeneous and the buyer learns about one portion of both suppliers, then for any distributions $F, G$ and $H$, the buyer prefers to learn portion $A$ rather than portion $X$ if

$$\bar{a} - A \geq \bar{x} - X.$$  

Thus, it is not the case that the buyer only wishes to learn whichever portion contributes most to the supplier’s overall total cost. Instead, Proposition 4 reveals that the buyer prefers to learn whichever portion will reduce the reserve price the most.

Having discussed which portion ($A$ or $X$) the buyer should learn, we now address research question 4 by considering how deep the buyer should learn, that is, the buyer’s preference between learning one portion ($A$ or $X$) or both portions ($A$ and $X$). To examine the buyer’s preference, suppose again that the two cost portions are labor and utility. Suppose if the buyer learns only the labor cost of supplier 1, her expected benefit is $10,000, and if the buyer learns only the utility cost of supplier 1, her expected benefit is $17,000. Intuitively, one might expect that “doubling down” on the supplier by learning the supplier’s labor and utility costs would yield an expected benefit less than the sum of the individual benefits (i.e., the expected benefit should be smaller than $27,000)—after all, there is a chance that supplier 1 will not win the contract and the learning will be useless. However, the following result shows that, when it comes to the depth of learning, the benefit of learning the whole is greater than the sum of the benefits of learning its parts (superadditivity).

**Proposition 5.** [Benefit of learning multiple portions] Consider three cost-learning models: Model $A$ in which the buyer learns about portion $A$, Model $X$ in which the buyer learns about portion $X$, and Model $AX$ in which the buyer learns about both portions ($A$ and $X$) simultaneously.  

(i) When the two suppliers are heterogeneous and cost modeling is applied only to supplier 1, we have $\psi_{1AX} \geq \psi_{1A} + \psi_{1X}$ almost surely; hence $\Psi_{1AX} \geq \Psi_{1A} + \Psi_{1X}$.

(ii) When the two suppliers are homogeneous and the same cost model can be applied to both suppliers, we have $\psi_{AX} \geq \psi_A + \psi_X$ almost surely; hence $\Psi_{AX} \geq \Psi_A + \Psi_X$.

Both competition and cost modeling are tools that the buyer can use to reduce the winning supplier’s surplus. Our earlier results establish that cost modeling prior to competition is only beneficial when the new reserve price squeezes the winning supplier’s surplus more than competition alone would have. In other words, to yield a positive benefit, the new reserve price must be lower than the second-lowest supplier’s cost. The new reserve price is much more likely to accomplish this if two cost portions are learned instead of just one, because the reserve price reduction is cumulative in cost portions learned. To use a golf analogy, when the buyer only learns one portion, she has to “hit a hole-in-one” but when she learns two portions she only needs to hit a hole in two strokes; the proposition means that hitting a hole-in-one in two attempts is much less likely
to occur than hitting a hole in two strokes. Combining Propositions 3 and 5, the takeaway for procurement managers is that, all else equal, depth in cost modeling is likely to be more beneficial than breadth.

5. More than two suppliers

We now generalize our results to the case of \( N > 2 \) suppliers. Some of the \( N \) suppliers may share the the same cost drivers (i.e., are homogeneous), in which case the buyer can learn about those suppliers simultaneously by developing a cost model that applies to all of them. We also allow supplier heterogeneity, meaning that not all \( N \) suppliers have the same cost structure. Formally, we divide the \( N \) suppliers into \( S \) groups, where group \( s \) \((s = 1, 2, ..., S)\) contains \( N_s \) homogeneous suppliers whose cost is represented by random variable \( C_s = A_s + X_s \). As in the two-supplier case, random variables \( C_s, A_s, \) and \( X_s \) respectively follow distributions \( H_s, F_s, \) and \( G_s \), whose domains are bounded above by \( \bar{c}_s, \bar{a}_s \) and \( \bar{x}_s \). When referring to the cost realizations for a specific supplier \( i \) within group \( s \) we will use notation \( c_{is}, a_{is} \) and \( x_{is} \). Suppliers in the same group are homogeneous and suppliers in different groups are heterogeneous. For example, suppose the buyer has a total of eight suppliers: five are located in northeastern China and use very similar labor-intensive production techniques; three are located in the midwestern US and all these suppliers utilize very similar automated production lines. We can divide the eight suppliers into two separate groups: group 1 has five suppliers in China and group 2 has three suppliers in the US.

We find that all of Section 4’s propositions for the two-supplier case can be generalized: Suppliers across different groups are heterogeneous (analogous to part (i) of the two-supplier propositions) and suppliers within a group are homogeneous (analogous to part (ii) of the two-supplier propositions). Building on the two-supplier case results, we find that the buyer follows similar strategies when deciding which group(s) of suppliers to learn about and which portion(s) of cost to learn. To denote the \( N \)-supplier generalizations of earlier propositions, we append their label with “G” for general.

Characterizing the benefit of cost modeling. Addressing research question 1, Lemma 2 directly applies to the \( N \)-supplier case, and shows that learning might not be beneficial due to the fact that learning and competition can act as substitutes. Moreover, we extend Proposition 1:

**Proposition 1-G.** [Relation between learning and cost distributions] Let \( \Psi_1 \) denote the expected benefit of learning all suppliers in group 1. As long as the domains of \( F_1 \) and \( G_1 \) do not change,

(i) when group 1 has only one supplier, i.e., \( N_1 = 1 \),

(a) \( \Psi_1 \) decreases as \( F_1 \) becomes stochastically larger;
The intuition is similar to that for Proposition 1. Learning about group 1 brings a positive benefit only when group 1 contains the winner. Furthermore, when learning does bring a positive benefit, the benefit’s magnitude is governed by the gap between the reserve price (which is determined by \( F_1 \) and the domain of \( X_1 \)) and the second-lowest supplier’s cost (determined by all the non-winning suppliers’ costs). For example, the benefit of learning is not affected by \( G_1 \) when there is one supplier in group 1 (part (i.b)) but is affected when there are two suppliers in group 1 (part (ii.b)). This is because when group 1 contains the winner, the second-lowest supplier’s cost does not change in the former case but depends on \( G_1 \) in the latter case.

Choosing which group(s) of suppliers to learn about. Addressing research question 2, we have:

**Proposition 2-G.** [Preferred group to learn about] For any distributions \( F_s, G_s \) and \( H_s \), \( s = 1, 2, \ldots, S \), the buyer prefers to learn about group 1 rather than group 2 if

\[
A_1 + \bar{x}_1 \leq_{st} A_2 + \bar{x}_2, \quad C_1 \leq_{st} C_2 \quad \text{and} \quad N_1 \geq N_2.
\]

Compared to Proposition 2, in Proposition 2-G we have an additional condition, \( N_1 \geq N_2 \). The intuition is similar to that for Proposition 2. As before, the buyer’s preference about which group to learn about depends on how low the new reserve price can be and how likely it is that a supplier in the group will win the contract. Both factors improve as the group’s size increases: Larger groups offer a bigger chance to set a small reserve price and are more likely to contain the contract winner.

The next result characterizes how the breadth of learning across groups affects the buyer’s benefit.

**Proposition 3-G.** [Benefit of learning about multiple groups] Let \( \Psi_{1,2}, \Psi_s, \text{ and } \Psi_i^s \) denote, respectively, the expected benefits of learning about all suppliers in groups 1 and 2, all suppliers in group \( s \), and supplier \( i \) in group \( s \). We have:

(i) \( \psi_{1,2} = \psi_1 + \psi_2 \) almost surely; hence \( \Psi_{1,2} = \Psi_1 + \Psi_2 \).

(ii) \( \Psi_s = N_s \cdot \Psi_i^s \).
One may reason that, particularly with N suppliers, the expected marginal benefit of learning will diminish in the number of groups that the buyer learns about — after all, the buyer ultimately only sets one new reserve price which applies to all N suppliers. However, the expected benefit of learning multiple groups is in fact additive across groups and linear in the number of learned suppliers within each group. The intuition behind this is the same as for Proposition 3: Along any sample path only one supplier wins the contract, and the buyer only benefits when she learns about the winner.

Choosing which portion(s) of cost to learn and how deeply to learn. Addressing research questions 3 and 4, we have:

**Proposition 4-G.** [Preferred portion to learn] Suppose the buyer learns about all suppliers in group 1. For any distributions $F_s$, $G_s$ and $H_s$, $s = 1, 2, ..., S$, the buyer prefers to learn portion $A_1$ rather than portion $X_1$ if

$$\bar{a}_1 - \bar{A}_1 \geq st \bar{x}_1 - \bar{X}_1.$$ 

The proposition reveals that what matters to the buyer is the amount of cost uncertainty that learning will resolve: She prefers to learn about whichever portion will (stochastically) reduce the reserve price the most. The takeaway is similar to what we found with two suppliers, namely it is not the case that the buyer only wishes to learn about whichever cost portion contributes most to the suppliers’ overall total costs. Next we examine how the benefit changes in depth of learning.

**Proposition 5-G.** [Benefit of learning multiple portions] Let $\Psi^A_1$, $\Psi^X_1$, and $\Psi^{AX}_1$ denote, respectively, the expected benefits when cost modeling is applied to portion $A$, portion $X$, and both portions $A$ and $X$ for all suppliers in group 1. We have

$$\psi^{AX}_1 \geq \psi^A_1 + \psi^X_1 \text{ almost surely; hence } \Psi^{AX}_1 \geq \Psi^A_1 + \Psi^X_1.$$ 

The intuition is akin to that for Proposition 5: To be effective the reserve price must bind the winner, and this is particularly encouraged by deeper learning because reserve price reductions are cumulative in cost portions learned. Combining this result with Proposition 3-G, which shows that the benefit across groups is linear, we can see that, all else equal, the buyer gets greater benefit from learning a single group in more depth than she does by learning two groups more superficially.

6. Optimal learning strategy

The previous sections characterized the buyer’s benefit of learning without explicitly considering the cost of learning. However, as established earlier, cost modeling is time-consuming and expensive. If we consider the cost of learning, should the buyer learn at all? If so, which suppliers should the
buyer learn about, and how deep into suppliers’ costs should the buyer learn when deeper learning incurs higher cost? The buyer’s cost of learning can depend on which suppliers she learns about, how many suppliers she learns about, and how deeply she learns about them; we will show how the previous sections’ analytical results can directly be applied in all these cases to provide insight on the structure of the buyer’s optimal learning.

**Learning cost is fixed for all suppliers in the same group.** To begin, suppose there is simply a fixed cost $K_s$ to create a cost model for each supplier group $s$. Recall that $\Psi_s$ is the expected benefit of learning about group $s$ without considering the cost of learning. Thus, the net benefit of learning about group $s$ is $\Psi_s - K_s$. Label the groups such that $\Psi_1 - K_1 \geq \Psi_2 - K_2 \geq \cdots \geq \Psi_S - K_S$. If the buyer can learn about just one group, it is optimal to learn about group 1 (provided that $\Psi_1 - K_1 > 0$). Surprisingly, if the buyer can learn about multiple groups, it is optimal for her to learn about the first $s$ groups, where $s$ is largest value such that $\Psi_s - K_s > 0$. In fact, although it may seem intractible to consider the buyer’s cost versus benefit of learning while taking into account the possibility that she learns from multiple cost models at once, the problem in fact decomposes and can be solved by considering each supplier group in isolation. This is a direct consequence of Proposition 3-G(i), which proves that the benefit of learning is additive across groups.

**Learning cost increases in the number of suppliers learned.** Next suppose that there is a cost associated with the number of suppliers learned in any group $s$, $K_s(n)$. Let $\Psi_s(n)$ be the expected benefit of learning about $n$ suppliers in group $s$. Putting these two together, the net benefit of learning is $\Psi_s(n) - K_s(n)$.

If the cost function $K_s(n)$ is concave in $n$, Proposition 3-G(ii) directly applies to prove that the optimal number of suppliers to learn about must be $n^*_s = 0$ or $N_s$. This comes from the fact that the benefit of learning is linear in the number of suppliers learned, which makes the net benefit function convex in $n$. On the other hand, if $K_s(n)$ is convex, Proposition 3-G(ii) directly applies to prove that the net benefit function, $\Psi_s(n) - K_s(n)$, is concave in $n$. Hence, one can easily determine the optimal number of suppliers to learn about. For general increasing functions, the optimal number of suppliers to learn about depends on the particular shape of the cost function $K_s(n)$.

Once we determine the optimal number of suppliers to learn about for each group, label the groups such that $\Psi_1(n^*_1) - K_1(n^*_1) \geq \Psi_2(n^*_2) - K_2(n^*_2) \geq \cdots \geq \Psi_S(n^*_S) - K_S(n^*_S)$. If the buyer can learn about just one group, it is optimal to learn about group 1 if $\Psi_1(n^*_1) - K_1(n^*_1) > 0$. If the buyer can learn about multiple groups, it is optimal to learn about the first $s$ groups, where $s$ is largest value such that $\Psi_s(n^*_s) - K_s(n^*_s) > 0$. Once again, what initially appears to be an intractible problem — considering the buyer’s cost versus benefit of learning while taking into account the
possibility that she learns from multiple cost models at once and the cost of learning depends on how many suppliers she chooses to learn — the problem in fact decomposes and again we can solve it by examining each supplier group separately.

**Learning cost increases in the depth of learning.** Suppose the buyer can learn portions $A$ and $X$, but the cost of learning increases with the portions learned. (For simplicity, we focus on just two cost portions, but the results easily extend to multiple portions.) Our previous sections’ analytical results also apply in this case to determine the optimal depth of learning.

If $K_s(A) + K_s(X) \geq K_s(AX)$ (i.e., the cost of learning is subadditive in depth), it immediately follows from Proposition 5-G that the optimal depth to learn in group $s$ is $d^*_s = 0$ or $AX$. This is because, as we proved, the expected benefit of learning is superadditive in the portions learned, so subadditive learning cost makes the net benefit superadditive in the portions learned. On the other hand, if $K_s(A) + K_s(X) < K_s(AX)$ (superadditive), then it is possible that an intermediate level of learning becomes optimal, i.e., $d^*_s \in \{A, X\}$. In this case, the optimal depth to learn depends on the particular shape of the learning cost function.

Regardless of the learning cost function’s shape, once we determine the optimal depth of learning for each group, label the groups such that $\Psi_1(d^*_1) - K_1(d^*_1) \geq \Psi_2(d^*_2) - K_2(d^*_2) \geq \cdots \geq \Psi_S(d^*_S) - K_S(d^*_S)$. If the buyer can learn about just one group, she prefers to learn about group 1 if $\Psi_1(d^*_1) - K_1(d^*_1) > 0$. If the buyer can learn about multiple groups, she prefers to learn about the first $s$ groups, where $s$ is largest value such that $\Psi_s(d^*_s) - K_s(d^*_s) > 0$.

Propositions 3-G and 5-G can be used to render similar results when the cost of learning is a function of both the number of suppliers and portions learned. For example, for any given group $s$, Proposition 3-G implies that the optimal number of suppliers to learn about will be $0$ or $N_s$ if the cost is concave in the number of suppliers for any given depth $d_s$. Similarly, from Proposition 5-G, the optimal depth to learn about each group will be $0$ or $AX$ if the cost of learning is subadditive in depth for any number of suppliers learned. Finally, after ordering the groups by their optimal benefits of learning, if the buyer can only learn about one group she prefers to learn about the highest benefit group, and if she can learn about multiple groups she learns about all the groups with positive individual benefit.

**Example:** In this example, all costs are in thousands of dollars. Suppose there are two groups of suppliers: group 1 has two suppliers, where $A_1 \sim U[200, 575], X_1 \sim U[300, 360]$, only portion $A_1$ is learnable and the cost of learning is a function of the number of suppliers learned, $K_1(n) =$
10 + 2√π (n = 1, 2); group 2 has only one supplier, where \( A_2 \sim U[360, 530] \), \( X_2 \sim U[270, 600] \), both portions \( A_2 \) and \( X_2 \) are learnable and the cost of learning is \( K_2(A) = 12 \), \( K_2(X) = 13 \), \( K_2(AX) = 15 \).

We first determine the optimal learning strategy for group 1. Since \( K_1(n) \) is concave in \( n \), the optimal number of suppliers to learn about is either 0 or 2. The benefit of learning 2 suppliers in group 1 is \( \Psi_1(2) = 83.80 \) which is greater than the cost of learning \( K_1(2) = 12.828 \), so \( n^*_1 = 2 \) and \( \Psi_1(n^*_1) - K_1(n^*_1) = 70.992 \). As for group 2, since \( K_2(A) + K_2(X) \geq K_2(AX) \), the optimal depth to learn is either 0 or \( AX \). The benefit of learning \( AX \) is \( \Psi_2(AX) = 3.31 \) which can not cover the cost of learning \( K_2(AX) = 15 \), so the optimal depth to learn is \( d^*_2 = 0 \). Thus, the optimal learning strategy that the buyer should follow is to learn about the two suppliers in group 1 and not to learn about group 2. In this example, the expected payment without learning is 764.73. Through learning, the buyer reduces the expected payment by \( 83.80/764.73 = 10.96\% \) and the net saving is \( 70.992/764.73 = 9.28\% \).

7. Conclusions

In this paper, we ask whether a buyer should deploy cost modeling to learn about suppliers who will compete for a contract. Intuitively, because competitive bidding and cost modeling are both ways that a buyer can discover information about suppliers’ costs, and moreover since cost modeling is itself an expensive undertaking, careful analysis is needed to understand if and when cost modeling should be used in conjunction with supplier competition. To our knowledge, this paper is the first study of cost modeling and its interactions with competitive bidding.

Addressing research question 1 (How does one characterize the benefit of cost modeling?), we show that learning prior to competitive bidding enables the buyer to set a tighter reserve price. In doing so the buyer seeks to truncate suppliers’ surplus. However, when there are competing suppliers, the buyer does not always receive a positive benefit from learning. In fact, the benefit of learning is positive only when the new reserve price actively truncates the surplus of the contract winner, and learning about suppliers who wind up losing the contract provides no useful information to the buyer, since these suppliers’ costs will be revealed during the competitive bid process anyway. Of course, this makes the buyer’s problem difficult because she cannot \textit{ex ante} predict which supplier will win — after all, the suppliers’ costs (which the buyer does not know) determine who wins the contract. Summarizing the above, learning is more valuable to the extent that the reserve price is likely to truncate the surplus of the contract winner.

Turning to research question 2 (If the buyer chooses to develop a cost model, which supplier(s) should she learn about?), we find that the key criterion is not to learn about the supplier whose learnable cost is more uncertain, but to learn about the supplier who is more likely to win and whose learnable cost, once learned, will enable the buyer to lower the reserve price the most. Moreover, one
might intuitively expect it to be redundant to learn about several competing suppliers. However, we find that the benefit of learning about multiple suppliers is actually additive across suppliers. Hence, managers should strongly consider learning about multiple suppliers at once, when the cost of learning is linear or concave in the number of suppliers learned.

Addressing research question 3 (Which cost portion(s) should the buyer learn?), we find that the buyer prefers to learn about the portion which can resolve more uncertainty. Surprisingly, we also find that the benefit of learning multiple portions at once is even greater than the sum of the benefits of learning each portion in isolation. Coupled with our insights regarding research question 2, this suggests that — all else equal — learning fewer suppliers in depth is preferable to learning more suppliers superficially. Thus, staffing a procurement department with specialists having deep but narrow domain expertise (e.g., of a particular type of production method used by some suppliers) may be preferable to having generalists with broader but more limited knowledge about the industry’s general cost drivers.

Finally, research question 4 considers the optimal learning strategy. We find that the buyer can first divide all suppliers into several groups based on their cost structures, next figure out the optimal level of learning for each group, and then the optimal learning strategy is to learn about the groups with a positive net benefit of learning. Depending on the cost to learn various groups to varying depths, we show that the buyer can adopt a mix-and-match strategy of learning, whereby she learns some groups deeply, others superficially, and some not at all.

For simplicity our discussion focused on two cost portions, $A$ and $X$. However, our results extend to multiple cost portions. Suppose that supplier $i$’s cost is given by $C_i = Z^1_i + Z^2_i + \cdots + Z^m_i + \epsilon_i$, where the $Z^j_i$’s are learnable cost portions and $\epsilon_i$ is unlearnable. For any subsets $R, T \subseteq \{1, 2, \ldots, m\}$, $R \cap T = \emptyset$, we can apply all the paper’s propositions by defining $A_i = \sum_{j \in R} Z^j_i$, $X_i = \sum_{j \in T} Z^j_i$, and $\epsilon_i = C_i - A_i - X_i$.

In this paper, we focus our analysis on the reverse open-descending auction where the buyer sets a reserve price as a starting point and then the suppliers compete by lowering their bids until only one bidder remains. We study this mechanism because it is analytically tractable and extremely prevalent in practice. However, one may wonder what the buyer’s learning preferences will be under other mechanisms. The buyer’s bargaining power determines which mechanism she can choose. Following the reasoning of Bulow and Klemperer (1996), we can classify a buyer’s bargaining power into three levels: (1) Zero bargaining power. A buyer with zero bargaining power can not set a reserve price and the contract payment is completely determined by suppliers’ bid competition in a reverse open-descending auction without a reserve price. In this case, one can show that learning provides no benefit since the buyer has no power to apply any cost information. (2) Moderate bargaining power. A buyer with moderate bargaining power can set a reserve price. Through
learning, the buyer can apply the cost information to set a more aggressive reserve price and reduce the contract payment. This is the case that we analyze in this paper. (3) Complete bargaining power. A buyer with complete bargaining power can design and commit to any mechanism. This allows her to design, for example, an optimal mechanism. One can show that a buyer with complete bargaining power does not have to exert effort to learn about suppliers’ costs before the auction. Instead, she can induce the suppliers to voluntarily divulge any learnable cost information practically for free before the auction, simply by employing random post-auction audits on learnable portions and levying large penalties on suppliers found to have lied about their costs. In reality, these penalties could take the form of disallowing future bids from the supplier, which might be significant if the buyer comprises a substantial portion of market demand.

Thus, although one might expect that buyers with more bargaining power will have more incentive to utilize cost modeling since a powerful buyer can utilize the information better than a weak buyer, we conclude from the above paragraph that a buyer’s inclination to undertake cost modeling on her suppliers is not monotonic in her bargaining power. Since buyers with zero bargaining power are too weak to apply any information and buyers with complete bargaining power essentially obtain learnable information for free without cost modeling, learning via cost modeling is only interesting for buyers with moderate bargaining power. This is why our paper focused on this latter case.

Our paper is an important first step in understanding how two very common procurement tools — cost modeling and competitive bidding — interact. It has the potential to help procurement managers make better decisions about how to apply cost modeling in practice. We hope that it spurs further research into the role of cost modeling in supply chain and sourcing strategies.

Appendix. Proofs of Propositions

In our proofs we let \( A_i, X_i, \) and \( C_i \) denote supplier \( i \)'s cost random variables, and let \( a_i, x_i, \) and \( c_i \) denote the cost realizations. This notation suppresses the group index \( s \), but it is understood that if suppliers \( i \) and \( j \) are in the same group (i.e., are homogeneous) then \( A_i \) and \( A_j \) have the same distribution, \( X_i \) and \( X_j \) have the same distribution, and \( C_i \) and \( C_j \) have the same distribution. Also, we let \( \psi^1, \psi^2, \) and \( \psi^{1k2} \), respectively, denote the benefits of learning about supplier 1, supplier 2, and both supplier 1 and supplier 2, where \( \Psi^1 = E[\psi^1], \Psi^2 = E[\psi^2], \) and \( \Psi^{1k2} = E[\psi^{1k2}] \) denote the expected benefits of learning.

A. Proof of Proposition 3-G

The proof of Proposition 1-G and 2-G utilizes Proposition 3-G; therefore we prove Proposition 3-G first.

It suffices to prove that when there are \( N \) suppliers, the benefit of learning any two of them is additive. Consider any suppliers 1 and 2, who may or may not be in the same group. We want to show \( \psi^{1k2} = \psi^1 + \psi^2 \). We consider the sample paths such that \( C_i \neq C_j (\forall i \neq j) \). Since \( C_i (i = 1, 2, ..., N) \) are continuous random
variables, the probability that \( C_i \neq C_j (\forall i \neq j) \) is 1. Let \( C_{[k]} \) denote the \( k^{th} \) order statistic of \( C_1, C_2, \ldots, C_N \), and define \( C_{-i} \triangleq \min_{j \neq i} \{ C_j \} \). We have

\[
\psi^{1k/2} = \min(r^o, C_{[2]}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)^+,
\]

\[
= I_{(C_1 = C_{[1]})} [\min(r^o, C_{[2]}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)]^+ + I_{(C_2 = C_{[1]})} [\min(r^o, C_{[2]}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)]^+ + \sum_{i=3}^{N} I_{(C_i = C_{[1]})} [\min(r^o, C_{[2]}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)]^+,
\]

\[
= I_{(C_1 = C_{[1]})} [\min(r^o, C_{[2]}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)]^+ + I_{(C_2 = C_{[1]})} [\min(r^o, C_{[2]}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)]^+ + I_{(C_1 = C_{[1]})} [\min(r^o, C_{-1}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)]^+ + I_{(C_2 = C_{[1]})} [\min(r^o, C_{-2}) - \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2)]^+ + I_{(C_1 = C_{[1]})} [\min(r^o, C_{-1}) - (A_1 + \bar{x}_1)]^+ + I_{(C_2 = C_{[1]})} [\min(r^o, C_{-2}) - (A_2 + \bar{x}_2)]^+.
\]

The third equality follows since \( C_i = C_{[1]}, i \geq 3 \) implies \( C_1 \geq C_{[2]}, C_2 \geq C_{[2]} \), so \( \min(A_1 + \bar{x}_1, A_2 + \bar{x}_2) \geq \min(C_1, C_2) \geq C_{[2]} \geq \min(r^o, C_{[2]}) \). The final equality follows since \( C_{-1} \leq C_2 \leq A_2 + \bar{x}_2 \Rightarrow \min(r^o, C_{-1}) - (A_2 + \bar{x}_2) \leq 0 \).

By similar arguments,

\[
\psi^1 = [\min(r^o, C_{[2]}) - (A_1 + \bar{x}_1)]^+ = I_{(C_1 = C_{[1]})} [\min(r^o, C_{-1}) - (A_1 + \bar{x}_1)]^+,
\]

\[
\psi^2 = [\min(r^o, C_{[2]}) - (A_2 + \bar{x}_2)]^+ = I_{(C_2 = C_{[1]})} [\min(r^o, C_{-2}) - (A_2 + \bar{x}_2)]^+.
\]

Hence \( \psi^{1k/2} = \psi^1 + \psi^2 \) almost surely. \( \square \)

**B. Proof of Proposition 1-G**

Suppose supplier 1 is in group 1. We have the benefit of learning about supplier 1

\[
\psi^1 = [\min(r^o, C_{[2]}) - (A_1 + \bar{x}_1)]^+ = [\min(r^o, C_{-1}) - (A_1 + \bar{x}_1)]^+.
\]

This is because: when \( C_1 = C_{[1]}, C_{[2]} = C_{-1} \); when \( C_1 \geq C_{[2]}, A_1 + \bar{x}_1 \geq C_1 \geq C_{[2]} \geq C_{-1} \Rightarrow \min(r^o, C_{-1}) \leq \min(r^o, C_{[2]}) \leq A_1 + \bar{x}_1 \).

Since \( \psi^1 \) increases in \( C_{-1} \) and decreases in \( A_1 \), the expected benefit of learning group 1, \( \Psi_1 = N_1 \cdot E[\psi^1] \), increases as the distribution of \( C_{-1} \) becomes larger and decreases as the distribution of \( A_1 \) becomes larger.

This explains part (i.a), (i.b), (i.c), (ii.b) and (ii.c). Part (ii.a) is because when group 1 has multiple suppliers, \( F_1 \) affects not only \( A_1 \) but also \( C_{-1} \). \( \square \)

**C. Proof of Proposition 2-G**

Suppose supplier 1 is in group 1 and supplier 2 is in group 2. We have

\[
\psi^1 = [\min(r^o, C_{[2]}) - (A_1 + \bar{x}_1)]^+ = [\min(r^o, C_{-1}) - (A_1 + \bar{x}_1)]^+, \quad \text{and}
\]

\[
\psi^2 = [\min(r^o, C_{[2]}) - (A_2 + \bar{x}_2)]^+ = [\min(r^o, C_{-2}) - (A_2 + \bar{x}_2)]^+.
\]

Since \( r^o \) is constant and \( C_1 \leq_{st} C_2 \), we have \( \min(r^o, C_{-1}) \geq_{st} \min(r^o, C_{-2}) \) (see page 7 in Muller and Stoyan (2002)). Also, since \( A_1 + \bar{x}_1 \leq_{st} A_2 + \bar{x}_2 \), we have \( -(A_1 + \bar{x}_1) \geq_{st} -(A_2 + \bar{x}_2) \). Because \( (C_{-1}, A_1) \)
are independent random variables and \((C_{-2}, A_2)\) are independent random variables, from Theorem 1.2.17 in Muller and Stoyan (2002), we can conclude that \(\min(r^o, C_{-1}) - (A_1 + \bar{x}_1) \geq_{st} \min(r^o, C_{-1}) - (A_2 + \bar{x}_2)\) and \(\psi^1 \geq_{st} \psi^2\). It then follows that \(\Psi^1 \geq \Psi^2\). Finally, by applying Proposition 3-G, the expected benefit of learning group 1 is \(\Psi^1 = N_1 \cdot \Psi^1\) and the expected benefit of learning group 2 is \(\Psi^2 = N_2 \cdot \Psi^2\). Since \(N_1 \geq N_2\) and \(\Psi^1 \geq \Psi^2\), we have \(\Psi_1 \geq \Psi_2\). □

D. Proof of Proposition 4-G

Suppose supplier 1 is in group 1. The benefit of learning supplier 1’s portion \(A_1\), or \(X_1\), are

\[
\psi^{1,A} = [\min(r^o, C_{[2]}) - (A_1 + \bar{x}_1)]^+ = [\min(r^o, C_{-1}) - (A_1 + \bar{x}_1)]^+, \\
\psi^{1,X} = [\min(r^o, C_{[2]}) - (X_1 + \bar{a}_1)]^+ = [\min(r^o, C_{-1}) - (X_1 + \bar{a}_1)]^+.
\]

Since \(C_{-1}, A_1, X_1\) are independent,

\[\bar{a}_1 - A_1 \geq_{st} \bar{x}_1 - X_1 \Rightarrow A_1 + \bar{x}_1 \leq_{st} X_1 + \bar{a}_1 \Rightarrow \psi^{1,A} \geq_{st} \psi^{1,X},\]

which implies \(\Psi^{1,A} = E[\psi^{1,A}] \geq \Psi^{1,X} = E[\psi^{1,X}]\).

The result extends to the case of any number of suppliers in the same group. By Proposition 3-G, the expected benefit of learning portion \(A\) of group 1 is \(\Psi^1 = N_1 \cdot \Psi^{1,A}\) and the expected benefit of learning portion \(X\) of group 1 is \(\Psi^X = N_1 \cdot \Psi^{1,X}\). Thus we have \(\Psi^1 \geq \Psi^X\). □

E. Proof of Proposition 5-G

The proof uses the following technical lemma, which is stated and proved below.

**Lemma 3.** If \(Y, Z, T\) are positive, then \((Y + Z - T)^+ \geq (Y - T)^+ + (Z - T)^+\).

**Proof of Lemma 3.** If \(Y \geq T, Z \geq T\), then \((Y + Z - T)^+ = Y + Z - T \geq (Y - T)^+ + (Z - T)^+\);

if \(Y \geq T, Z < T\), then \((Y + Z - T)^+ = Y + Z - T \geq Y - T = (Y - T)^+ + (Z - T)^+\);

if \(Y < T, Z \geq T\), then \((Y + Z - T)^+ = Y + Z - T \geq Z - T = (Y - T)^+ + (Z - T)^+\);

if \(Y < T, Z < T\), then \((Y + Z - T)^+ \geq 0 = (Y - T)^+ + (Z - T)^+\).

Hence, \((Y + Z - T)^+ \geq (Y - T)^+ + (Z - T)^+\). □

Suppose supplier 1 is in group 1. The benefit of learning supplier 1’s portion \(A_1\) is

\[\psi^{1,A} = [\min(r^o, C_{-1}) - (A_1 + \bar{x}_1)]^+ = [(\bar{a}_1 - A_1) - (\bar{c}_1 - \min(r^o, C_{-1}))]^+.\]

Similarly, the benefit of learning supplier 1’s portion \(X_1\) is

\[\psi^{1,X} = [(\bar{x}_1 - X_1) - (\bar{c}_1 - \min(r^o, C_{-1}))]^+\]

and the benefit of learning supplier 1’s both portions \(A_1\) and \(X_1\) is

\[\psi^{1,AX} = [(\bar{c}_1 - C_1) - (\bar{c}_1 - \min(r^o, C_{-1}))]^+\]

\[= [(\bar{a}_1 - A_1) + (\bar{x}_1 - X_1) - (\bar{c}_1 - \min(r^o, C_{-1}))]^+.\]

Since the reserve price without learning is \(r^o = \min_{i=1}^N (\bar{c}_i) \leq \bar{c}_1\), we have \(\bar{c}_1 - \min(r^o, C_{-1}) \geq 0\). Also, because \(\bar{a}_1 - A_1 \geq 0\) and \(\bar{x}_1 - X_1 \geq 0\), Lemma 3 implies \(\psi^{1,AX} \geq \psi^{1,A} + \psi^{1,X}\). By Proposition 3-G, we have \(\psi^{1,AX} \geq \psi^1 + \psi^X\). □
F. Proof of Propositions 1-5

Propositions 1, 2, 3, 4 and 5 are direct consequences of the corresponding generalized propositions.

References


