Using Cost Modeling to Set a More Informed Procurement Auction Reserve Price

Yan Yin
HEC Paris, 78350 Jouy-en-Josas, France, yin@hec.fr

Hyun-Soo Ahn, Damian R. Beil
Ross School of Business, University of Michigan, Ann Arbor, MI 48109, hsahn@umich.edu, dbeil@umich.edu

Reserve prices are commonly used in procurement auctions, and the logic is simple: They help the buyer avoid paying too much for the contract. The reserve price is set by the buyer before the auction based on the buyer’s knowledge about suppliers’ costs. In practice, buyers can refine their knowledge about suppliers’ costs through activities known in industry as cost modeling. In this paper we explore the value of such activities to help the buyer set a more informed reserve price prior to an auction. Specifically, we analyze which supplier(s) to learn about, which portion(s) of the costs to learn, the value of increasing the number of suppliers learned about, and the depth to which the buyer should learn about each supplier. Our analysis accommodates a rich set of supply base characteristics, such as ex ante asymmetric suppliers, supplier cost correlations, and possible collusion between suppliers. 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 value of additional information has a diminishing rate of return does not apply. Finally, we compare the benefit of inviting additional suppliers versus setting a more informed reserve price and find that the latter can outweigh the former, which is in stark contrast with received wisdom from past work (e.g., the classic paper Bulow and Klemperer (1996)) that ignored the possibility of cost modeling.

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). Seeking to manage their input costs, many firms select external suppliers via competitive bidding (reverse auctions). In an ideal setting, competitive bidding between potential suppliers drives the price to nearly the minimum the winning supplier
is willing to accept, including of course a necessary operating profit for the supplier. However, in practice this does not necessarily happen: As a buyer at Pfizer put it, “In our online auctions, suppliers can see what each other bids. If they were willing to go much lower, but none of the other participants forced them to do so, [the buyer] can lose out” (Atkinson 2008). In practice, there are several reasons why this can happen.

First, buyers often auction off contracts for non-standardized items (e.g., Beall et al. (2003), pages 10, 31-32) for which supplier costs can be heterogeneous (e.g., driven by the supplier’s production technology), and, as a result, the first- and second-lowest-cost suppliers may have dissimilar costs. For this reason, buyers in practice utilize a reserve price. A reserve price is the price at which the buyer opens the bidding — it is the maximum amount the buyer is willing to pay for the contract. Setting the reserve price is important because the closer it is to the eventual winner’s minimum price, the less money the buyer leaves on the table.

Second, quality and reliability are necessities. Thus, it is standard practice for buyers to invite only qualified suppliers to a competitive bidding event and announce that the lowest bidding supplier will win the contract. This is because, even when a contract for a relatively standard item is auctioned off, the process of locating and qualifying a supplier who is capable of producing the needed good is typically very time-consuming. For example, at automotive and electronics Fortune 500 manufacturers we interacted with, the supplier qualification processes can take months, even for relatively uncomplicated goods like cables, connectors, and screw machine parts. At these firms, functional areas such as engineering invest considerable time and resources to verify supplier capabilities and approve the supplier(s). The costly nature of supplier qualification (in terms of time, monetary costs of testing, and political costs of imposing on the engineering department’s time) often makes a large pool of qualified suppliers an either impossible or impractical option. For example, one manufacturer we worked with had just two suppliers who would be asked to bid for many parts. Unfortunately, a small supply pool can result in a large gap between the first- and second-lowest-cost suppliers, which again is another reason why buyers in practice utilize a reserve price.
The above discusses why buyers in practice take the setting of a reserve price very seriously in their auctions. As the practitioner survey Beall et al. (2003) points out (page 44), a critical issue when procuring an item is “understanding the unique characteristics of the item’s supply market structure, degree of competitiveness, key cost drivers, and current open capacity (i.e., is it a buyer’s, suppliers’, or neutral market?).” This information is essential for setting an appropriate reserve price.

To set a more informed reserve price, buyers collect information pertinent to suppliers’ costs through activities, in industry, known as cost modeling. At an automotive firm we worked with, two employees were assigned to painstakingly collect cost information about an auto part for more than one month. They 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. At an electronics firm we worked with, a procurement specialist with a background in corporate finance spent a considerable amount of time combing through data such as financial statements to gather information related to suppliers’ financial status and operating profits. At the same time, other procurement managers at the firm performed reconnaissance on the ground, by visiting suppliers and gauging the supplier’s capacity utilization (for example, seeing if they could spot a recently installed new production line that was under-utilized). These cost modeling activities help the buyer to estimate the minimum price at which the supplier is willing to accept the contract, and hence allows the buyer to set a more informed reserve price when auctioning off the contract.

Setting an informed reserve price becomes even more important when supply base characteristics such as collusion or correlated costs can weaken the auction’s ability to lower the price. In fact, at the automotive firm we worked with, the buyer was concerned about collusion among suppliers (who were co-located and in close contact with one another). In this case, the buyer had another reason to invest time and energy to engage cost modeling activities.
Cost modeling is expensive and time-consuming. Therefore it is important to know what value it brings. In this paper we explore the value of cost modeling activities that the buyers engage in in an effort to set a more informed reserve price prior to an auction. Since these activities are often time consuming, 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:

**Question 1:** Which supplier(s) should the buyer learn about?

**Question 2:** 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?

**Question 3:** How deeply into supplier(s)’ costs should she learn? Would the buyer ever want to learn about multiple competing suppliers?

**Question 4:** How do answers to the above questions depend on underlying business characteristics, such as correlations across suppliers’ costs, and supplier collusion?

To study the above questions, we consider a buyer who uses an open-descending auction to award a contract to one of \( N \) qualified suppliers. Before the auction the buyer announces the starting bid, or reserve price. This approach is common in practice, because it is intuitive and easy to explain to bidders. Formally such a reserve price is called “non-discriminatory” because it applies equally to all bidders in the auction. To answer our research questions analytically, one needs to characterize the optimal, non-discriminatory reserve price of an auction where the bidders have different cost distributions (i.e., are \( ex \ ante \) asymmetric). Given the progress made by Myerson (1981) almost 40 years ago on optimal reserve prices in the context of optimal mechanism design where reserve prices and biasing rules can be bidder-specific, i.e., discriminatory, one would think the non-discriminatory case would be relatively easy. However, compared to the optimal mechanism setting, the more implementable setting of a non-discriminatory reserve price is much more difficult to analyze (Gunay et al. 2013) and no general solution exists in the literature. This fact presents an obstacle to addressing our research questions.
At first glance, it would seem impossible to analytically study the research questions we have posed in this paper, let alone arrive at meaningful insights from the analysis. We circumvent this roadblock by leveraging the fact that the optimal non-discriminatory reserve price can be readily derived in a natural setting that occurs frequently in practice: The buyer wishes to set a reserve price, but does not wish to risk non-transaction (we furnish details about why this setting frequently arises in practice in the Model section). Incorporating this practical consideration facilitates solving for the reserve price even when bidder costs are \textit{ex ante} asymmetric. Furthermore, focusing on this setting enables us to model and analyze a number of practically relevant characteristics that are often ignored in the literature. We explicitly allow for bidder costs to follow possibly different distributions. This facilitates a model that can capture richness and complexity in the supply base topology, e.g., certain subgroups of suppliers having similar production technologies and/or cost drivers, which we have not seen in other auction models. We capture various levels of correlation across suppliers such as a common cost factor shared across all suppliers, a regional cost factor (like labor) shared only among suppliers within a given group, as well as a supplier-specific cost factor. In contrast, much auction research assumes just independent supplier costs. Additionally we consider collusion across suppliers, whereas the vast majority of auction work ignores collusion. The cost of learning can also have a very general structure — the functions could be convex or concave in the number of suppliers, sub-additive or super-additive in the depth of learning. Despite all these general assumptions, we are able to derive analytical results. Our paper contributes to the literature by identifying a novel and important problem and modeling it in a way that permits analysis and insight for a realistic and practical setting.

We review related literature in §2, and introduce the model in §3. Section 4 analyzes the expected value of learning, ignoring the learning cost. We extend our analysis to include the cost of learning and provide an optimal learning strategy incorporating it in §5. Section 6 studies correlations across suppliers’ costs and collusion among suppliers. Section 7 compares the value of learning to the value of adding suppliers, and §8 concludes. Proofs of results are furnished in the Appendix.
2. Literature Review

There is a vast economics literature on competitive bidding (auctions), for which Krishna (2009) provides an excellent introduction. The central idea in this literature is that competition can be used to reveal private information held by bidders. Our paper combines this central feature of auctions with another practical way in which cost information is discovered in procurement settings: cost modeling. To our knowledge, ours is the first paper to combine these two features and study how a buyer can use cost modeling to help reveal supplier cost information prior to competitive bidding. Wan and Beil (2009) and Wan et al. (2012) study settings where the buyer exerts effort prior to the auction to learn whether or not suppliers are qualified. However, unlike cost modeling which is meant to discover supplier cost information, qualification screening simply informs the buyer of whether or not the supplier is capable of fulfilling the contract. Aral et al. (2014) study a setting where a buyer exerts effort prior to a total-cost procurement auction in order to refine her estimates of suppliers’ non-price attributes, the goal being a better total-cost supplier selection decision rather than a pre-auction discovery of some supplier cost information, as is the goal of cost modeling. Finally, Chen et al. (2008) study a buyer who awards a contract to a supplier and, to the extent that the buyer can subsequently learn the supplier’s cost, the buyer can then extract profit and coordinate the supply chain. By contrast, in our paper the supplier keeps what profit they earn in the auction, but the buyer can exert effort (conduct cost modeling) to set a more informed reserve price for the auction.

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 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 how the buyer should best
deploy cost modeling prior to competitive bidding. To our knowledge, ours is the first paper in the literature to examine cost modeling in this context.

A practitioner literature exists on how to construct cost models. For example, chapter 3 in Laseter (1998) describes various ways that buyers can decompose a supplier’s cost into a number of cost drivers, each of which can then be estimated through various means of data collection, as well as the challenges of doing so. One way of gathering data about suppliers firsthand is Rapid Plant Assessment; 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. There is also a literature on using cost modeling to inform new product development decisions where the designer must balance functionality against cost (see Locascio (2000) and references therein).

We study how the buyer should best deploy cost modeling to manage her expected payment in the auction. It is well known that deriving a closed-form expression for the buyer’s expected payment in an auction with *ex ante* asymmetric bidders is extremely difficult. This challenge arises because there do not exist closed-form solutions for order statistics when the underlying random draws are not independent and identically distributed. The lack of closed-form solutions thus makes it difficult to analytically describe how the expected auction price would change if, say, the cost distributions of some subset of bidders were changed. Our paper provides analytical results in this setting. We do so by exploiting separability across bidders in our analyses, which enables us to analyze the auction outcome with a sample path approach. In addition, we study the cases where supplier costs are not independent due to cost correlation or collusion, reflecting real challenges faced by the firm. Thus our paper makes both theoretical and practical contributions by gaining structural results in a complex bidder environment.

3. Model Description and Preliminaries

3.1. Basic setup

We consider a buyer seeking to allocate a single, indivisible production contract to one of $N$ potential suppliers. Each supplier has a minimum price that he would accept for the contract.
Following auction theoretic convention we simply refer to this as the supplier’s “cost”, which of course is the supplier’s private information.

We divide the $N$ suppliers into $S$ groups, where each group $s$, $s = 1, 2, ..., S$, contains $N_s$ suppliers: $N = N_1 + N_2 + ... + N_S$. The cost of supplier $i$ in group $s$ is represented by random variable $C_i^s$ (with realization $c_i^s$), which follows distribution $H_s$ and is bounded above by $\bar{c}_s$. This setup allows for the fact that suppliers may be heterogeneous and have different cost distributions. 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. If all suppliers are heterogeneous, then the supplier groups are all singletons. The model also allows for the case where a subset of suppliers have costs that are drawn from the same distribution; these could represent suppliers located within the same region possessing similar production technologies. These cases represent groups of suppliers where $N_s > 1$. For example, suppose the buyer has a total of eight suppliers ($N = 8$): five are located in northeastern China and use similar production processes and three are located in the US and all of them utilize 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, that is, $N_1 = 5$ and $N_2 = 3$. For now we assume supplier costs are statistically independent and they do not collude; correlation and collusion are studied in §6.

To award her contract, 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 each other by treating the current lowest offer as the winner unless this offer is unseated by a lower offer from some other supplier. As we have observed in a variety of industries 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 his true cost or winning the auction. Unless there is no winner (no bidder meets the reserve price), the predicted equilibrium outcome is as follows: The lowest-cost supplier wins the contract at a price equaling the second-lowest supplier cost, or the reserve price, whichever is lower. Hence, the buyer’s contract payment is $\pi = \min(r, C_{[2]})$ where $C_{[j]}$ denotes the $j$th lowest of random variables $\{C_i^s\}$, where $s = 1, \ldots, S$ and $i = 1, \ldots, N_s$.

Procurement managers in practice are often loathe to set a reserve price that no bidder can meet. To understand why, note that if the reserve price is not met, the buyer has to exclude the original bidders from further consideration. (If suppliers anticipate that the buyer will increase the reserve price and re-invite them to bid if the original reserve price is not met, then no supplier would ever let their payment be capped by a reserve price.) The primary job of the procurement department is to procure supply necessary for the firm to function, and in many cases the firm has no internal production of necessary inputs; for example, the electronics and automotive industry firms we dealt with did not have internal production facilities for many of the parts they needed. Moreover, it can take a long time to qualify suppliers (up to six months at many firms we interacted with, even for simple parts), meaning it may not be feasible to qualify totally new suppliers within a given procurement cycle. With this in mind, one can see why the practitioner survey of Fortune 500 firms Beall et al. (2003) points out (page 9), “misreading the market and setting a reserve price that is too far below the market price, resulting in no bidder responses” is considered a “dysfunctional outcome” of an auction that the buyer wants to avoid.

Reflecting the above aspects of reality and also to facilitate tractability, in our model we capture the realistic situation in which the buyer must transact with one of the $N$ suppliers in the auction, but at the same time seeks to set a reserve price that minimizes her expected contracting payment:

$$\min_r E[\pi] \quad s.t. \quad P(C_{[1]} \leq r) = 1.$$  

The solution of the above problem, and the buyer’s contracting payment are, respectively:

$$r^o = \min_{s=1,2,\ldots,S} (\bar{c}_s), \text{ and } \pi^o = \min(r^o, C_{[2]}).$$
In other words, the buyer’s optimal reserve price is the lowest upper bound on the suppliers’ costs (see also Wan and Beil (2014)). The important thing to note here is that the better the buyer can estimate the suppliers’ costs, the better she can set her reserve price.

3.2. Learning about suppliers (“cost modeling”)

As mentioned in §1, to get a better estimate of suppliers’ costs the buyer can build a cost model and collect data on the suppliers before deciding her reserve price. For instance, for a manufactured good, a buyer might decompose a supplier’s production cost into the sum of several cost elements — such as raw material cost, direct labor cost, electricity cost, technology cost, overhead cost, and minimum acceptable profit margin — and learn some of these 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, overhead cost is usually difficult to learn.

To model this, for a supplier \( i \) in group \( s \), we assume cost \( C_i^s \) can be decomposed into two portions: \( A_i^s \), the portion of cost attributes that can be learned via cost modeling, and \( X_i^s \), the portion that cannot be learned. We assume that \( A_i^s \) follows a continuous distribution \( F_s \) and is bounded above by \( \bar{a}_s \), \( X_i^s \) follows a continuous distribution \( G_s \) and is bounded above by \( \bar{x}_s \). As such, the supplier’s cost \( C_i^s \)'s distribution \( H_s \) is the convolution of \( F_s \) and \( G_s \) (\( H_s = F_s \otimes G_s \)) and the upper bound of \( C_i^s \) satisfies \( \bar{c}_s = \bar{a}_s + \bar{x}_s \). Let \( a_i^s \) and \( x_i^s \) denote the realizations of \( A_i^s \) and \( X_i^s \), respectively.

In practice, a cost model analysis can be used to learn about an entire group of suppliers. For instance, if suppliers within group \( s \) have similar, labor-intensive production processes, the buyer can learn about those suppliers simultaneously by developing a cost model that applies to all of them. 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). To capture this, we will describe learning on a group level. That
is, we say that when the buyer “learns group\(s\)” (i.e., creates a cost model for group\(s\)), she learns about all \(N_s\) suppliers in the group, namely, she learns \(A^i_s = a^i_s\), for \(i = 1, \ldots, N_s\). However, all our results can be easily applied to learning a subset of \(n_s (< N_s)\) suppliers in the group.

After learning about group \(s\), for suppliers \(i = 1, \ldots, N_s\) the buyer can update her knowledge of supplier \(i\)’s cost, \([C^i_s(A^i_s = a^i_s)] = a^i_s + X^i_s\) (note that this is still a random variable, where \(X^i_s\) captures the unlearned portion of the costs), which has a new upper bound \([\tilde{c}^i_s(A^i_s = a^i_s)] = a^i_s + \bar{x}_s\).

Hence, if the buyer learns about a subset of groups \(L(\subseteq \{1, 2, \ldots, S\})\), the buyer can set a new reserve price:

\[
\hat{r} = \min_{s \in L, t \notin L} \left\{ \min_{i = 1, \ldots, N_s} \tilde{c}^i_s, \tilde{c}_t \right\} = \min_{s \in L, t \notin L} \left\{ \min_{i = 1, \ldots, N_s} (A^i_s + \bar{x}_s), \tilde{c}_t \right\}.
\]

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

\[
\pi^l = \min(\hat{r}, C_{[2]}).
\]

The value of learning is:

\[
\psi = \pi^o - \pi^l.
\]  

(1)

The crux of our paper is to understand this value of learning. The auction competition will automatically identify the lowest-cost supplier as the contract winner, and pushes the price down to the second-lowest supplier cost. Consequently, one can see that learning results in a positive value 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.

We define the expected value of learning as

\[
\Psi = E[\psi].
\]

In §4, we will discuss when to learn and how to learn so as to get the maximum expected value. There can be significant benefit to setting a more informed reserve price. To provide a simple numerical example, if the buyer has two suppliers in a single group \((N_1 = N = 2)\) with uniform cost
distributions \( (A^i \sim U[0,1], X^i \sim U[0,1], i = 1, 2) \), by learning both suppliers’ learnable portion \( A \), the buyer’s expected reduction in contract price (i.e., \( \Psi \)) is around 7%. Note that given the large amount of money many firms spend on procurement, even a small percentage savings on price can be significant.

Thus far we have introduced the expected value of learning. Of course, learning does not come for free. For ease of analysis, we will initially examine the value of learning in the next section without factoring in the cost of learning. We then formally incorporate the cost of learning in §5.

4. Value of Learning (Without Considering Cost of Learning)

We consider learning about group \( s \)'s learnable portion \( A^i_s \) for \( i = 1, 2, ..., N_s \), and let \( \Psi_s \) denote the expected value of learning about suppliers in group \( s \). The following result characterizes the value of cost modeling prior to competitive bidding.

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

(i) When group \( s \) has only one supplier \( (N_s = 1) \), \( \Psi_s = E\{[\min(C_2, r_o) - (A^1_s + \bar{x}_s)]^+\} \), and

(a) \( \Psi_s \) decreases as \( F_s \) becomes stochastically larger;

(b) \( \Psi_s \) is independent of \( G_s \);

(c) \( \Psi_s \) increases as \( H_t \) \( (t \neq s) \) becomes stochastically larger;

(ii) When \( N_s \geq 2 \), \( \Psi_s = E\{[\min(C_2, r_o) - (\min(A^1_s, A^2_s, ..., A^{N_s}_s) + \bar{x}_s)]^+\} \), and

(a) \( \Psi_s \) is not necessarily monotonic as \( F_s \) becomes stochastically larger;

(b) \( \Psi_s \) increases as \( G_s \) becomes stochastically larger;

(c) \( \Psi_s \) increases as \( H_t \) \( (t \neq s) \) becomes stochastically larger.

Proposition 1 shows that the value of learning depends on the suppliers’ cost distributions as well as the number of suppliers in a group which the buyer learns about. Suppose that the buyer learns about the cost of a single heterogeneous supplier whose cost distribution is different from all others. Part (i.a) says that as the distribution of the learnable portion, \( F_s \), becomes stochastically larger (while the support remains the same), the value of learning decreases. This is intuitive because the new reserve price, \( \min(A^1_s + \bar{x}_s, r_o) \) is stochastically increasing in \( F_s \), thus it is less likely to bind
the contract payment. Likewise, the value of learning grows as the cost distribution of the other suppliers — whom we do not learn about — increases, explaining part (i.c). However, part (i.b) implies that the expected value of learning is independent of the distribution $G_s$, which represents the portion of cost that cannot be discovered even after learning about the supplier in group $s$. This is because after learning, the new reserve price, $\min(A_{1s} + \bar{x}_s, r^o)$, depends on $G_s$ only through its upper bound.

On the other hand, consider the case where the buyer learns about multiple suppliers in group $s$ whose costs are drawn from the same distribution. Interestingly, the results for this case differ from the result when the buyer learns about a uniquely heterogeneous supplier. As the learnable cost stochastically decreases, the reserve price decreases as well. But at the same time, the cost the buyer would have paid without learning, $\min(C_{[2]}, r^o)$, also decreases. As a result, the value 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 value 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?**

We assumed above that the buyer learned about suppliers in group $s$. However, would the buyer have been better off learning instead about another group $t \neq s$ instead? To help answer this, we need to compare the expected value of learning either group. We now examine this issue, addressing research question 1.

Intuitively, one may expect that the buyer wishes to resolve the maximal amount of cost uncertainty and so she will optimally choose whichever group of suppliers has the most uncertain learnable cost. This logic works when the buyer has only one supplier in her supply base. Suppose group 1 has a single supplier whose learnable cost $A_{11} \sim U[5, 10]$ and unlearnable cost $X_{11} \sim U[0, 5]$, while group 2 has a single supplier where $A_{12} \sim U[0, 10]$ and $X_{12} \sim U[10, 11]$. Were the buyer facing a single-supplier situation with the single supplier in group 1, learning would bring an expected value
of \( \bar{a}_1 - E[A_1] = 2.5 \). This is smaller than \( \bar{a}_2 - E[A_2] = 5 \), which would be her expected value from learning if the supplier in group 2 was her only supplier. Note that with one supplier, the buyer’s reserve price is a take-it-or-leave-it offer, and the best the buyer can do while ensuring transaction is to offer the upper bound of the supplier’s cost distribution. Hence, in the single-supplier case, the buyer would gain more value from learning when the learnable portion has more uncertainty.

However, this result is no longer true when the buyer has multiple competing suppliers. The value of learning does not solely depend on the learnable portions. In fact, in the example above if the buyer has both suppliers (one in group 1 and the other in group 2), learning about the supplier in group 1 brings a much larger value (per Equation (1), \( \Psi_1 = 1.79 > 0.04 = \Psi_2 \)). Competition is what differentiates the one-supplier and multiple-supplier cases. Competition itself works as a cost-discovery tool, so learning brings a positive value 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, the group 1-supplier’s cost \( C_1 \) is stochastically smaller than the group 2-supplier’s cost \( C_2 \); this means that the group 1-supplier has a better chance of winning the contract. In addition, the new reserve price set by learning about the group 1-supplier (i.e., \( \min\{A_1 + \bar{x}_1, 15\} \), where \( r^o = \min\{10 + 5, 10 + 11\} = 15 \)) is more likely to be smaller than that for the group 2-supplier (i.e., \( \min\{A_2 + \bar{x}_2, 15\} \)). As a result, learning about the group 1-supplier is more valuable. Our result below formalizes this.

**Proposition 2.** [Preferred group to learn about] The buyer prefers to learn about group \( s \) rather than group \( t \) if

\[
A_s + \bar{x}_s \leq_{st} A_t + \bar{x}_t, \quad C_s \leq_{st} C_t \quad \text{and} \quad N_s \geq N_t.
\]

For convenience, in the proposition and throughout the remainder of the paper, we suppress superscript \( i \) when not referring to a specific supplier, i.e., \( C_s, A_s, \) and \( X_s \) (\( C_t, A_t, \) and \( X_t \)) are simply random variables following distributions \( H_s, F_s, \) and \( G_s \) (\( H_t, F_t, \) and \( G_t \)), respectively.

The conditions given in Proposition 2 reflect the fact that the buyer prefers to learn about a supplier who is more likely to win and whose learnable cost, once learned, will enable the buyer to
lower the reserve price the most (which is captured by the first two conditions). The third condition, \( N_s \geq N_t \), indicates that the size of the group also matters. 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.

Instead of learning about just one group of suppliers, the buyer could instead choose to undertake learning about multiple groups 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. As a result, decisions on whom to learn about must be made at the beginning of a procurement cycle. Our next result characterizes the value of learning about multiple suppliers. One may intuitively reason that the marginal value 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, thus more learning will likely just generate duplicative information. In other words, the value of learning about two groups should be smaller than the sum of the values of learning about either group individually. Surprisingly, the next result shows that this is not the case: The value of learning is in fact additive across groups and linear in the number of learned suppliers within each group.

**Proposition 3.** [Value of learning about multiple groups] Let \( \Psi_{s,t}, \Psi_s, \Psi_t, \) and \( \Psi(s_i) \) denote, respectively, the expected values of learning about all suppliers in groups \( s \) and \( t \), all suppliers in group \( s \), all suppliers in group \( t \), and supplier \( i \) in group \( s \). We have:

(i) \( \psi_{s,t} = \psi_s + \psi_t \) almost surely; hence \( \Psi_{s,t} = \Psi_s + \Psi_t \).

(ii) \( \Psi_s = N_s \cdot \Psi(s_i) \).

To understand the intuition behind Proposition 3, recall that the buyer realizes a positive value from learning only if she learns about the contract winner. For a given sample path, at most one group can yield a positive value, i.e., \( \psi_{s,t} = \psi_s, \psi_t = 0 \) or \( \psi_{s,t} = \psi_t, \psi_s = 0 \), which explains part (i). Part (ii) follows from part (i) and the fact that the suppliers in the same group have the same cost distributions.
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 $A$ and $X$ are both learnable. (Note that this setup is without loss of generality: our results still hold when the cost of a supplier in group $s$ is $C_s = A_s + X_s + \epsilon_s$, where $\epsilon_s$ is unlearnable.)

To address research question 2, we first ask which portion, if learned, would bring more value 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.

**Proposition 4.** [Preferred portion to learn] When learning about the suppliers in group $s$, the buyer prefers to learn portion $A$ rather than portion $X$ if

$$\bar{a}_s - A_s \geq \bar{x}_s - X_s.$$ 

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, the condition in Proposition 4 reveals that the buyer would always prefer to learn about the portion that will reduce the reserve price the most. Interestingly, this differs from what we saw when determining which group to learn (Proposition 2), where the overall cost magnitude was important because it related to the chance of the group containing the winning supplier.

Having discussed which portion ($A$ or $X$) the buyer should learn, we now address the question about how deep the buyer should learn (research question 3). We start this by examining the buyer’s preference between learning one portion ($A$ or $X$) or both portions ($A$ and $X$). To illustrate this, suppose again that the two cost portions are labor and utilities. Suppose the buyer’s expected value is $10,000 from learning only the labor cost, and her expected value is $17,000 from learning only the utility cost. Intuitively, one might expect that “doubling down” on the supplier by learning the
supplier’s labor and utility costs would yield an expected value less than the sum of the individual values (i.e., the expected value should be smaller than $27,000)—after all, there is a chance that no supplier in the learned group will win the contract and then learning will be useless. However, the following result shows that, when it comes to the depth of learning, the value of learning the whole is greater than the sum of the values of learning its parts (superadditivity).

**Proposition 5.** [Value of learning multiple portions] Let $\Psi^A_s$, $\Psi^X_s$, and $\Psi^{AX}_s$ denote, respectively, the expected value of learning when cost modeling is applied to portion $A$, portion $X$, and both portions $A$ and $X$ for all suppliers in group $s$. We have

$$\psi^{AX}_s \geq \psi^A_s + \psi^X_s \quad \text{almost surely; hence} \quad \Psi^{AX}_s \geq \Psi^A_s + \Psi^X_s.$$

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 valuable when the new reserve price squeezes the winning supplier’s surplus more than competition alone would have. In other words, to yield a positive value, 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 valuable than breadth.

**5. Optimal Learning Strategy**

The previous section characterized the buyer’s value of learning without explicitly considering the cost of learning. However, as established in the Introduction, 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 section’s analytical results can directly be applied in all these cases to
provide insight on the structure of the buyer’s optimal learning.

We first discuss the optimal learning strategy for each group. Let $\Psi_s(n,Y)$ be the expected value
of learning portion $Y$ from $n$ suppliers in group $s$, and $K_s(n,Y)$ be the cost of learning portion
$Y$ from $n$ suppliers in group $s$, where $n = 0, 1, ..., N_s$ and $Y = A, X$ or $AX$. We have the following
structural result by comparing the expected value of learning with the associated cost of learning.

**Proposition 6.** [Optimal learning strategy for the suppliers in group $s$]

(i) If $K_s(n,\cdot)$ is concave in $n$, there exists a threshold $n'$ above which learning is profitable, and
the optimal number of suppliers to learn about is either 0 or $N_s$;

(ii) If $K_s(n,\cdot)$ is convex in $n$, the net profit is concave in $n$ and hence it is possible that the
optimal number of suppliers to learn about is strictly between 0 and $N_s$;

(iii) If $K_s(\cdot,Y)$ is subadditive in $Y$, i.e., $K_s(n,A) + K_s(n,X) \geq K_s(n,AX)$, the optimal depth of
learning is 0 or $AX$;

(iv) If $K_s(\cdot,Y)$ is superadditive in $Y$, i.e., $K_s(n,A) + K_s(n,X) < K_s(n,AX)$, it is possible that
an intermediate level of learning becomes optimal (i.e., $A$ or $X$).

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

If the cost of learning is subadditive in the depth of learning, it immediately follows from Propo-
sition 5 that the optimal depth to learn to in group $s$ is 0 or $AX$. This is because, as we proved,
the expected value 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 the cost of learning is superadditive in depth of learning, then it is possible that an intermediate level of learning becomes optimal, i.e., \( A \) or \( X \). In this case, the optimal depth to learn depends on the particular shape of the learning cost function.

Once we determine the optimal number of suppliers to learn and the optimal depth of learning for each group, we can construct the optimal learning strategy for all suppliers as follows: If the buyer can learn about just one group, it is optimal to learn about the group with largest positive net benefit; if the buyer can learn about multiple groups, it is optimal for her to learn about all the groups with positive net benefit. In fact, what initially appears to be an intractable problem — considering the buyer’s cost versus value of learning while taking into account the possibility that she learns about multiple heterogeneous supplier groups at once and the cost of learning depends on how many suppliers she chooses to learn and how deep to learn — the problem in fact decomposes and can be solved by considering each supplier group separately. This is a direct consequence of Proposition 3(i), which proves that the value of learning is additive across groups.

**Illustrative example (all costs are in thousands of dollars).** Suppose there are two groups of suppliers. Group 1 has two suppliers, where \( A_i \sim U[200, 575], X_i \sim U[300, 360] \) for \( i = 1, 2 \), only portion \( A \) can be learned, and the cost of learning is a function of the number of suppliers learned, \( K_1(n) = 10 + 2\sqrt{n} \) (\( n = 1, 2 \)). Group 2 has only one supplier, where \( A_2 \sim U[360, 530], X_2 \sim U[270, 600] \), both portions \( A \) and \( X \) 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 \( n \), either learning two or zero suppliers is optimal. The value of learning two suppliers is \( \Psi_1(2) = 83.80 \) which is greater than the learning cost, \( K_1(2) = 12.828 \), so \( n^*_1 = 2 \) and \( \Psi_1(n^*_1) - K_1(n^*_1) = 70.972 \). 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 value of learning \( AX \) is \( \Psi_2(AX) = 3.31 \) which can not cover the cost of learning \( K_2(AX) = 15 \).
so it is optimal depth to learn is 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. The expected payment without learning is 764.73. Through learning, the buyer reduces the expected payment by $\frac{83.80}{764.73} = 10.96\%$, resulting in a net saving of $\frac{70.972}{764.73} = 9.28\%$.

**Remark:** We assume that the total cost of learning is the sum of the cost of learning in each group. This is appropriate as learning about the suppliers in another group requires the buyer to engage in separate learning activities. If we relax this assumption, the buyer needs to solve a combinatorial optimization problem that jointly determines which group(s) to learn, how many suppliers (within each group) to learn, and how deep to learn for each group. Even in this case, our insight that the value of learning (i.e., excluding the cost of learning) is additive and linear across groups does not change.

6. Relaxing Independence

6.1. Correlated Costs

Thus far our analysis has assumed that suppliers’ costs are statistically independent from each other. However in practice it is quite common that some portion of a supplier’s cost may be correlated to costs of other suppliers. For example, all suppliers might rely on a certain key material or component from the same upstream source, or perhaps suppliers who are located in the same industrial zone (e.g., southern China) incur similar labor rates. Unlike the independent cost case, learning about one supplier may reveal cost information about his competitors when costs are correlated, which is relevant to the buyer’s decision about cost modeling (e.g., whether to learn or not, how much to learn, etc.). In this section we show how our results extend to the correlated cost setting.

To extend our model, we assume that the cost of a supplier in group $s$ is represented by a random variable $C_s = A_s + X_s$, where $A_s = \tilde{A} + \tilde{A}_s + A'_s$, and $X_s = \tilde{X} + \tilde{X}_s + X'_s$. Random variables $\tilde{A}$ and $\tilde{X}$ represent portions of $A$ and $X$ that are common for all suppliers. Random variables $\tilde{A}_s$ and $\tilde{X}_s$ represent portions that are common for all the suppliers in group $s$, but independent from all other...
suppliers in other groups (if a group \( s \) consists of just one supplier, we take \( \tilde{A}_s = \tilde{X}_s = 0 \)). Finally, \( A'_s \) and \( X'_s \) represent portions that are independent across all suppliers. Let \( \tilde{F}, \tilde{F}_s \) and \( F'_s \) be the distributions of \( \tilde{A}, \tilde{A}_s \) and \( A'_s \), respectively. Likewise, \( \tilde{X} \sim \tilde{G}, \tilde{X}_s \sim \tilde{G}_s \) and \( X'_s \sim G'_s \). As before, \( H_s = F_s \otimes G_s \), where \( A_s \sim F_s, X_s \sim G_s \). Additionally, we define \( H'_s = F'_s \otimes G'_s \).

With this setup, supplier cost realizations are arrived at as follows: one draw from \( \tilde{F} \) and one draw from \( \tilde{G} \) are made and applied to all suppliers; for each group \( s \), one draw from \( \tilde{F}_s \) and one draw from \( \tilde{G}_s \) are made and applied to just suppliers in group \( s \); and for each supplier in group \( s \), one draw from \( F'_s \) and one draw from \( G'_s \) are made and applied to just that individual supplier.

As a check, note that the case with \( \tilde{A} \equiv \tilde{X} \equiv \tilde{A}_s \equiv \tilde{X}_s \equiv 0 \) corresponds to our model of independent supplier costs. Recall that \( N_s \) is the number of suppliers in group \( s \), and \( N_1 + \cdots + N_S = N \). We let random variable \( C_{-s} \) denote the minimum cost among \( N_s - 1 \) draws from \( H_s \) and \( N_t \) draws from \( H_t \) for each \( t \neq s \) (these \( N - 1 \) draws may be correlated). Let \( r^o \) denote the original reserve price set without learning, \( r^o = \min(\bar{c}_1, \bar{c}_2, \ldots, \bar{c}_s) \). We have the following result.

**Proposition 7.** All previous results hold with the following changes to accommodate the correlation model:

(i) Adapting Proposition 1 [Relation between learning and cost distributions]: \( F_s, G_s \) and \( H_t \) are replaced with \( F'_s, G'_s \) and \( H'_t \), respectively. Moreover, the value of learning \( \Psi_s \) decreases in \( \tilde{F} \) and \( \tilde{F}_s \), but increases in \( \tilde{G} \) and \( \tilde{G}_s \).

(ii) Adapting Proposition 2 [Preferred group to learn about]: Condition \( A_s + \bar{x}_s \leq_{st} A_t + \bar{x}_t \) is replaced by \( A'_s + \bar{x}_s \leq_{st} A'_t + \bar{x}_t \), and condition \( C_s \leq_{st} C_t \) is replaced by \( \min(r^o, C_{-s}) - \tilde{A} - \tilde{A}_s \geq_{st} \min(r^o, C_{-t}) - \tilde{A} - \tilde{A}_t \);

(iii) Adapting Proposition 4 [Preferred portion to learn]: Condition \( \bar{a}_s - A_s \geq_{st} \bar{x}_s - X_s \) is replaced by \( \bar{a}_s - A'_s \geq_{st} \bar{x}_s - X'_s, \min(r^o, C_{-s}) - \tilde{A} - \tilde{A}_s \geq_{st} \min(r^o, C_{-t}) - \tilde{X} - \tilde{X}_s \);

(iv) Proposition 3 [Value of learning about multiple groups] and Proposition 5 [Value of learning multiple portions] hold without any changes.
Under the cost correlation model, some supplier cost portions are independent across suppliers while others are correlated. It turns out that the effect these costs’ distributions have on the buyer’s value of learning depends on whether the cost distribution governs an independent or correlated portion. For the independent portions, the effects are exactly as we saw before for the case with completely independent supplier costs (Proposition 1). For the correlated portions, the value of learning decreases with $\tilde{F}$ and $\tilde{F}_s$, the distributions governing the portions of learned cost that are common across multiple suppliers. The intuition is that the new reserve price stochastically increases in $\tilde{F}$ and $\tilde{F}_s$ and therefore recovers less profit for the buyer when it binds. Moreover, increasing $\tilde{F}$ and $\tilde{F}_s$ may increase the payment without learning, but unlike the independent cost case, this change “washes out” due to correlation because the new reserve price would increase by exactly the same amount. Analogously, the value of learning increases with $\tilde{G}$ and $\tilde{G}_s$, the distributions governing the portions of the unlearned cost that are common across multiple suppliers, since the new reserve price is unchanged by this but the contract price in the absence of learning becomes higher. To summarize, we find that learning is less valuable when the learned cost portions exhibit more substantial correlation ($\tilde{F}$ and $\tilde{F}_s$ become larger), but learning becomes more valuable when the unlearned cost portions exhibit more substantial correlation ($\tilde{G}$ and $\tilde{G}_s$ become larger). The managerial takeaway is that, although one may intuitively expect that more correlation will always reduce the value of learning since supplier costs will be more closely matched and competition will be fiercer, increasing the magnitude of correlated cost drivers can actually increase the value of learning.

In choosing which group to learn about, we need to account for correlation when determining how effective the revised reserve price will be in lowering the buyer’s payment. In adapting Proposition 2 for correlation, the first condition focuses on the supplier-specific uncertainty ($A$ is replaced by $A'$). This condition is just like the condition we had in the independent case, capturing the fact that — all else equal — the buyer prefers learning that will lead to a smaller new reserve price. In the second condition we subtract out the correlated portions ($\tilde{A} + \tilde{A}_s$, $\tilde{A} + \tilde{A}_t$) from the original price
(\min(r^o, C_{-s}), \min(r^o, C_{-t}))$. The reason is that in reaching the original price (without learning), the correlated portions will be revealed “for free” through competition, so intuitively the buyer prefers to learn about groups that have less correlation in the costs that will be learned; as a check, notice that if group $t$ has a large within-group correlation governed by $\tilde{A}_t$, the second condition is more likely to hold and the buyer would prefer to learn about group $s$ instead of group $t$. We make a similar modification when determining which cost portion to learn (adapting Proposition 4). The managerial insight here is that, all else equal, the buyer prefers learning about groups and/or cost portions that exhibit weaker cost correlations in the learned portion because strongly correlated costs are more apt to be revealed via competition regardless of learning.

For the fully independent costs case we saw that the value of learning about multiple suppliers was additive, and we next ask whether this insight changes in the presence of supplier cost correlation. One might expect it would, since we just saw how sensitivity of the value of learning and the decision of which group or portion to learn should account for whether or not the cost portions are correlated across suppliers. However, Proposition 7 shows that the value of cost modeling remains additive across suppliers and groups, even with correlation. The reason is that, regardless of cost correlation, learning is valuable provided that the buyer learns about the contract winner. Similar reasoning gives our final result, namely that the value of cost modeling is superadditive in cost portions that are learned about, even under correlation.

### 6.2. Supplier Collusion

Cost correlation is not the only reason that supplier bids are not independent. At firms we have interacted with — in industries ranging from electronics to automotive parts — buyers have related suspicions of supplier collusion. Collusion dampens bidding competition, making it less effective. We now extend our analysis to study the effect of cost modeling in cases with supplier collusion.

In this subsection we allow the possibility that a group $t \in \{1, \ldots, S\}$ can be a bidding ring. Literature on bidding rings dates back to the seminal paper Graham and Marshall (1987). Krishna (2009) offers a nice distillation of the canonical theory as follows: (1) the bidders in a ring identify
which bidder among them has the lowest cost; (2) only this bidder submits meaningful bids for the contract (other bidders submit bids that do not affect the final winning price, e.g., drop out at the reserve price); (3) after the bidding, the ring shares any surplus gained from collusion. This 3-step distillation describes so-called efficient collusion, meaning only the “best” bidder from a colluding ring submits a serious bid for the contract. Graham and Marshall (1987) identified a simple way that the ring can coordinate its members to achieve efficient collusion.

Suppose that $T$ of the $S$ supplier groups are bidding rings, and label the indices such that the bidding rings are groups $\{1, \ldots, T\}$. Since a ring can maximize its gain by colluding efficiently, we assume that each ring colludes efficiently and so effectively collapses into a single “best bidder” who represents that entire ring in the bidding process. Not surprisingly, it is a dominant strategy of each ring’s representative bidder to lower his bid until winning or reaching his true production cost. It is also a dominant strategy of bidders in groups $\{T + 1, \ldots, S\}$ to lower their bid until winning the contract or reaching their true production cost. All other bidders (members of rings who are not ring representatives) submit non-serious bids (meaning they intentionally drop out early, e.g., at the reserve price). For each ring $t$, define cost distributions $F_{t}^{\text{ring}} = 1 - (1 - F_{t})^{N_{t}}$, $G_{t}^{\text{ring}} = 1 - (1 - G_{t})^{N_{t}}$, and $H_{t}^{\text{ring}} = 1 - (1 - H_{t})^{N_{t}}$. We have the following result.

**Proposition 8.** Propositions 1, 2, 3 and 4 hold if each bidding ring $t$ is replaced with a synthetic group, $t^{\text{ring}}$, containing a single supplier whose learnable, unlearnable, and total costs are governed by distributions $F_{t}^{\text{ring}}$, $G_{t}^{\text{ring}}$, and $H_{t}^{\text{ring}}$, respectively.

In words, synthetic group $t^{\text{ring}}$ consists of a single supplier whose costs equal the first order statistic of costs from all suppliers in the bidding ring $t$. The implication behind Proposition 8 is that buyers can understand the role of cost modeling in the presence of supplier collusion simply by considering the colluding groups to be collapsed to representative suppliers.

However, there are some interesting differences that arise. First, note that we no longer get superadditivity of values of learning multiple portions for a colluding group (namely, Proposition 5 does not hold under collusion, see Appendix G(ii)). This is because in Proposition 8 we have
collapsed the ring to a single representative supplier. In fact, if we look at individual suppliers within a ring, we recover the superadditivity result (see Appendix G(iii)). Thus, unlike the non-collusion case where superadditivity across portions applied at the individual supplier level and at the group level, under collusion it only applies at the individual level. The opposite situation arises for the additivity of values of learning multiple groups (Proposition 3): Under collusion additivity of values holds at the group level (per Proposition 8), but it turns out that it does not hold at an individual supplier level (see Appendix G(i); recall that Proposition 8 had collapsed the ring into a single representative bidder). Thus, under collusion there is an interesting difference between group and individual-level learning values.

The reason behind this is that collusion changes the way learning value accrues across suppliers. Recall that when suppliers do not collude, losing suppliers’ cost information would always be revealed through the bidding competition, so the buyer only benefitted when she learned about whoever eventually won the contract. By contrast, if a losing supplier is in a winning bidding ring, their cost information will not be revealed through the competition. In fact, under collusion the buyer can gain a positive value from learning even if the supplier she learns about is not the eventual contract winner. The upshot is that under collusion the successive values of learning about additional suppliers in the same bidding ring need not be additive. However, more importantly, the value of any amount of learning in the presence of collusion is always at least as large as it would have been were the collusion absent.

**Proposition 9.** Compared to the non-colluding case, the presence of collusion can only increase the buyer’s value of learning about supplier costs.

### 7. Learning versus More Bidders

Thus far we have shown that learning about suppliers can decrease the buyer’s procurement cost. Of course, the buyer could also lower her cost by locating new suppliers who will bid for the contract. In practice, this option is very expensive and time-consuming; at Fortune 500 companies we interacted with it typically takes months to locate and qualify new suppliers to bid in an
auction. This raises an important question: When designing a supply base, should the buyer focus on fostering her knowledge about her existing suppliers, or instead focus on locating and qualifying new suppliers? We address this question by comparing the value of adding new suppliers to the value of learning about existing suppliers.

We consider a buyer who has a supply base consisting of a group of \( N \) non-colluding suppliers whose costs are \textit{ex ante} symmetric but possibly correlated. We use the correlation model from §6.1, but drop the group index \( s \) since there effectively is only one group; thus, for example, \( H = F \otimes G \) represents the suppliers’ cost distribution. (With multiple groups (\textit{ex ante} asymmetric suppliers) and/or collusion, the key insights do not change but the expressions are more complicated.)

As in §5, suppose that there is a cost to learn about each supplier. We compare the following two options: (a) The buyer learns about \( m \) out of \( N \) suppliers (and learns the realization of portion \( A \) for each of these \( m \) suppliers); (b) The buyer locates \( y \) more suppliers, for a total of \( N + y \) suppliers. The value of option (a), learning \( m \) suppliers, is given by \( m \Psi \), where \( \Psi \) is the value of learning about a single supplier (see Proposition 7 part (iv)). Let \( \Phi(y) \equiv E[C_{[2:N]} - C_{[2:N+y]}] \) denote the value of adding \( y \) suppliers, where \( C_{[2:n]} \) denotes the second-lowest order statistic from \( n \) draws from distribution \( H \).

We compare the value of learning, \( m \Psi \), and the value of locating, \( \Phi(y) \), in the following proposition. We let \( \lceil m \rceil \) denote the smallest integer that is greater than or equal to \( m \).

**Proposition 10.** The buyer prefers learning about existing suppliers to locating new suppliers if and only if \( m \geq m^* = \lceil \Phi(y) / \Psi \rceil \).

As one can see, if the value of learning about each individual supplier, \( \Psi \), is very small then \( m^* > N \) and learning will not be better than locating even if the buyer learns about all \( N \) existing suppliers. Likewise, the same is true if the number of additional suppliers, \( y \), is very large. To make the comparison between learning and locating more meaningful, we consider a case where learning is effective and \( y \) is not extremely large. Namely, suppose that the buyer can learn a supplier’s exact cost through learning, and compare this to the value of adding one new supplier.
Proposition 11. Let \( \gamma_N = \frac{\int H'(c)H'(d)N^{-1}dc}{\int H'(c)H'(d)N^{-1}dc} \), and \( m^* = \lceil \gamma_N \cdot N \rceil \). Learning about \( m \) suppliers completely is better than locating one more supplier if and only if \( m \geq m^* \). Moreover, for any cost distribution \( H' \), \( \gamma_N \in (0, 1) \), and thus \( m^* \leq N \).

Costs that are common to all suppliers “wash out” in competition and therefore do not affect the value of learning. This is why the threshold \( m^* \) in Proposition 11 only depends on the cost portion that is independent across suppliers, which is governed by \( H' \) (whose tail is denoted by \( \tilde{H}' \)). The proposition implies that, when cost modeling is precise enough, learning can provide larger cost saving than locating a new supplier.

It is interesting to note that the minimum number of existing suppliers that the buyer needs to learn about in order to prefer it to locating one new supplier can be quite small. See Table 1, which gives the minimum (rightmost column) for several well-known distributions. For all the cases in the table, the value from learning the costs of just two (out of \( N \)) suppliers exceeds the value from locating a new supplier. In fact, when supplier costs follow a uniform or power function distribution, our result shows that for a buyer with small existing supply base (\( N = 2 \) for uniform and \( N \leq \nu + 1 \) for power distribution), learning the cost of just one supplier is better than locating a new supplier. Our result contrasts with the famous finding of Bulow and Klemperer (1996), who showed it is optimal for a buyer to locate one more supplier (and intensify competition) rather than set an optimal reserve price with the existing supply base. This is because a buyer in our paper has the ability to learn about suppliers’ costs (which is certainly true in procurement settings), and this capability enables the buyer to set a more effective reserve price than a buyer in Bulow and Klemperer (1996).

8. Conclusions

Buyers often use procurement auctions to award production contracts. To mitigate the risk of paying too much for the contract, the buyer sets the reserve price before the auction based on the buyer’s knowledge about suppliers’ costs. In this paper we consider a practical setting in which buyers can refine their knowledge about suppliers’ costs through activities known in industry as cost
modeling. We model this important practical fact and provide guidance about how cost modeling should be deployed to help the buyer set a more informed reserve price prior to an auction.

Addressing research question 1 (Which supplier(s) the buyer should learn about), we find that it is not sufficient to simply focus on learning about the supplier whose learnable cost is more uncertain. Rather, the buyer should learn about the supplier who is more likely to win and whose learnable cost 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 value of learning about multiple suppliers is actually additive across suppliers. Hence, procurement managers should strongly consider learning about multiple suppliers at once, especially when the cost of learning is linear or concave in the number of suppliers learned.

Addressing research question 2 (Which cost portion(s) should the buyer learn?), we find that the buyer can simply focus on learning about the portion which can resolve more uncertainty. This result is different from what we saw above for choosing which supplier to learn about. Moreover, we also find that the value of learning multiple portions at once is even greater than the sum of the values of learning each portion in isolation. Coupled with our insights regarding research question 1, 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.

Addressing research question 3 (What is 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.

Finally, research question 4 considers the effect of the underlying business context. Our insights are robust to correlations across supplier costs, and surprisingly — even though correlation tends to make competition fiercer by increasing the parity between suppliers’ costs — larger correlated costs across suppliers can actually increase the value of learning. We show that supplier collusion short circuits the price discovery of the competitive bidding process and makes cost modeling aimed at learning about supplier costs even more valuable for the buyer. Lastly, we show that learning just a few suppliers deeply can be preferable to adding another bidder, which contrasts with the well-known finding of Bulow and Klemperer (1996).

We focus our analysis on a reverse open-descending auction. However, one may wonder what the buyer’s learning preferences will be under other mechanisms. Following the reasoning of Bulow and Klemperer (1996), we can classify a buyer’s power (relative to her suppliers) into three levels: (i) A buyer with zero power can not set a reserve price and the contract payment is completely determined by suppliers’ bid competition, in which case learning provides no value. (ii) A buyer with absolute power can design and commit to any mechanism. 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. (iii) A buyer with moderate power can set a reserve price. This was the case documented in Beall et al. (2003), which surveyed large (Fortune 500) firms, many of whom spend billions of dollars on procurement each year. Through learning, the buyer can apply the cost information to set a better reserve price and reduce the contract payment. To summarize, buyers with zero power are too weak to apply any information and buyers with absolute power essentially obtain learnable information for free without cost modeling. But learning
via cost modeling matters for buyers with moderate power, which is why our paper focuses on that case.

For simplicity our discussion focused on two cost portions, $A$ and $X$. However, our results extend to multiple cost portions. Suppose that the cost of supplier $i$ in group $s$ is given by $C^i_s = Z^{i,1}_s + Z^{i,2}_s + \cdots + Z^{i,m}_s + \epsilon^i_s$, where the $Z^{i,l}_s$'s are learnable cost portions and $\epsilon^i_s$ 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 $A^i_s = \sum_{l \in R} Z^{i,l}_s$, $X^i_s = \sum_{l \in T} Z^{i,l}_s$, and $\epsilon^i_s = C^i_s - A^i_s - X^i_s$. The buyer could also consider non-price costs of doing business with supplier $i$, such as shipping costs if the buyer is responsible for transportation. Accordingly, she would run a total-cost open-descending auction, where she takes into account such non-price costs (see, for example, Kostamis et al. (2009) and Engelbrecht-Wiggans et al. (2007)).

We could model this by introducing an additional cost, $Z^{i,m+1}_s$, representing the buyer’s non-price costs associated with doing business with supplier $i$. The only difference is that this cost is already known by the buyer at the outset of the auction, and this cost is borne by the buyer should supplier $i$ win the contract.

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 inform reserve prices 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

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

A. Proof of Proposition 3

It suffices to prove that the value of learning any two different suppliers is additive. Consider supplier $i$ in group $s$ and supplier $j$ in group $t$, where $s$ and $t$ may or may not be the same group. Let $\psi(s_i)$, $\psi(t_j)$, and $\psi(s_i, t_j)$, respectively, denote the values of learning about supplier $i$ in group $s$, supplier $j$ in group $t$, and both supplier $i$ in group $s$ and supplier $j$ in group $t$. We want to show $\psi(s_i, t_j) = \psi(s_i) + \psi(t_j)$. We consider the sample paths such that any two suppliers’ costs
are different. Since suppliers’ costs are are all continuous random variables, the probability that any two suppliers’ costs are different is 1. Define $C_{-s_i} \triangleq \min_{(w,k) \neq (s,i)} \{C_w^k\}$. We have

$$
\psi(s_i, t_j) = \min(r^o, C_{[2]}(s_i) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+.
$$

$$
= \min(r^o, C_{[2]}(s_i) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+ + \sum_{(w,k) \neq (s,i),(t,j)} I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{[2]}(s_i) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+ + I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{[2]}(s_i) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+ + I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+ + I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+ + I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+ + I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s, A_i^j + \bar{x}_t))^+.
$$

The third equality follows since $C_w^k = C_{[1]}$, $(w,k) \neq (s,i),(t,j)$ implies $C_i \geq C_{[2]}$, $C_i \geq C_{[2]}$, so $\min(A_i^j + \bar{x}_s, A_i^j + \bar{x}_t) \geq \min(C_i, C_{[2]}) \geq C_{[2]} \geq \min(r^o, C_{[2]})$. The final equality follows since $C_i = C_{[1]} \Rightarrow C_{-s_i} \leq C_i \leq A_i^j + \bar{x}_t \Rightarrow \min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_t) \leq 0$.

By similar arguments, $\psi(s_i) = \min(r^o, C_{[2]}(s_i) - (A_i^j + \bar{x}_s))^+ + I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s))^+ + I_{\{C_w^k = C_{[1]}\}} [\min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s))^+.

Hence $\psi(s_i, t_j) = \psi(s_i) + \psi(t_j)$ almost surely.

**B. Proof of Proposition 1**

We have the value of learning about supplier $i$ in group $s$

$$
\psi(s_i) = \min(r^o, C_{[2]}(s_i) - (A_i^j + \bar{x}_s))^+ = \min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s))^+.
$$

This is because: when $C_i = C_{[1]}$, $C_{[2]} = C_{-s_i}$; when $C_i \geq C_{[2]}$, $A_i^j + \bar{x}_s \geq C_i \geq C_{[2]} \geq C_{-s_i} \Rightarrow \min(r^o, C_{-s_i}) \leq \min(r^o, C_{[2]}) \leq A_i^j + \bar{x}_s$. Since $\psi(s_i)$ increases in $C_{-s_i}$ and decreases in $A_i^j$, the expected value of learning group $s$, $\Psi_s = N_s \cdot E[\psi(s_i)]$, increases as the distribution of $C_{-s_i}$ becomes larger and decreases as the distribution of $A_i^j$ becomes larger. This explains part (i.a), (i.b), (i.c), (ii.b) and (ii.c). Part (ii.a) is because when group $s$ has multiple suppliers, $F_s$ affects not only $A_i^j$ but also $C_{-s_i}$.

**C. Proof of Proposition 2**

The values of learning about supplier $i$ in group $s$, and learning about supplier $j$ in group $t$, are

$$
\psi(s_i) = \min(r^o, C_{[2]}(s_i) - (A_i^j + \bar{x}_s))^+ = \min(r^o, C_{-s_i}) - (A_i^j + \bar{x}_s))^+,
$$

$$
\psi(t_j) = \min(r^o, C_{[2]}(t_j) - (A_i^j + \bar{x}_t))^+ = \min(r^o, C_{-t_j}) - (A_i^j + \bar{x}_t))^+.
$$
Since \( r^o \) is constant and \( C_s \leq C \), we have \( \min(r^o, C_{-s}) \geq \min(r^o, C_{-s}) \) (see page 7 in Müller and Stoyan (2002)). Also, since \( A^i + \bar{x}_s \leq A^i + \bar{x}_s \), we have \( -(A^i + \bar{x}_s) \geq -(A^i + \bar{x}_t) \). Because \((C_{-s}, A^i)\) are independent random variables and \((C_{-s}, A^i)\) are independent random variables, from Theorem 1.2.17 in Müller and Stoyan (2002), we can conclude that \( \min(r^o, C_{-s}) - (A^i + \bar{x}_s) \geq \min(r^o, C_{-s}) - (A^i + \bar{x}_t) \). Because \((C_{-s}, A^i)\) are independent random variables and \((C_{-s}, A^i)\) are independent random variables, from Theorem 1.2.17 in Müller and Stoyan (2002), we can conclude that \( \min(r^o, C_{-s}) - (A^i + \bar{x}_s) \geq \min(r^o, C_{-s}) - (A^i + \bar{x}_t) \) and \( \psi(s_i) \geq \psi(t_j) \). It then follows that \( \Psi(s_i) \geq \Psi(t_j) \). Finally, by applying Proposition 3, the expected value of learning group \( s \) is \( \Psi_s = N_s \cdot \Psi(s_i) \) and the expected value of learning group \( t \) is \( \Psi_t = N_t \cdot \Psi(t_j) \). Since \( N_s \geq N_t \) and \( \Psi(s_i) \geq \Psi(t_j) \), we have \( \Psi_s \geq \Psi_t \).

D. Proof of Proposition 4

For supplier \( i \) in group \( s \), the values of learning his portion \( A \), or \( X \), are

\[
\psi^A(s_i) = [\min(r^o, C_{[q]}) - (A^i + \bar{x}_s)]^+ = [\min(r^o, C_{-s}) - (A^i + \bar{x}_s)]^+, \\
\psi^X(s_i) = [\min(r^o, C_{[q]}) - (X^i + \bar{a}_s)]^+ = [\min(r^o, C_{-s}) - (X^i + \bar{a}_s)]^+.
\]

Since \((C_{-s}, A^i, X^i)\) are independent, \( \bar{a}_s - A^i \geq \bar{x}_s - X^i \Rightarrow A^i + \bar{x}_s \leq X^i + \bar{a}_s \Rightarrow \psi^A(s_i) \geq s \psi^X(s_i) \), which implies that the expected value of learning \( \Psi^A(s_i) = E[\psi^A(s_i)] \geq \Psi^X(s_i) = E[\psi^X(s_i)] \).

The result extends to the case of any number of suppliers in the same group. By Proposition 3, the expected value of learning portion \( A \) of group \( s \) is \( \Psi^A_s = N_s \cdot \Psi^A(s_i) \) and the expected value of learning portion \( X \) of group \( s \) is \( \Psi^X_s = N_s \cdot \Psi^X(s_i) \). Thus we have \( \Psi^A_s \geq \Psi^X_s \).

E. Proof of Proposition 5

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

**Lemma 1.** If \( Y, Z, T \) are positive, then \( (Y + Z - T)^+ \geq (Y - T)^+ + (Z - T)^+ \).

**Proof:** If \( Y \geq T, Z \geq T \), then \( (Y + Z - T)^+ = Y + Z - T \geq (Y - T) + (Z - T) = (Y - T)^+ + (Z - T)^+ \); 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, \( (Y + Z - T)^+ \geq (Y - T)^+ + (Z - T)^+ \). □

For supplier \( i \) in group \( s \), the value of learning his portion \( A \) is

\[
\psi^A(s_i) = [\min(r^o, C_{-s}) - (A^i + \bar{x}_s)]^+ = [(\bar{a}_s - A^i) - (\bar{c}_s - \min(r^o, C_{-s}))]^+.
\]

Similarly, the value of learning his \( X \) is

\[
\psi^X(s_i) = [(\bar{x}_s - X^i) - (\bar{c}_s - \min(r^o, C_{-s}))]^+,
\]

and the value of learning his both portions \( A \) and \( X \) is

\[
\psi^{AX}(s_i) = [(\bar{c}_s - C^i) - (\bar{c}_s - \min(r^o, C_{-s}))]^+ = [(\bar{a}_s - A^i) + (\bar{x}_s - X^i) - (\bar{c}_s - \min(r^o, C_{-s}))]^+.
\]

Since \( r^o = \min_{t=1,2,...,S}(\bar{c}_t) \leq \bar{c}_s \), we have \( \bar{c}_s - \min(r^o, C_{-s}) \geq 0 \). Also, because \( \bar{a}_s - A^i_s \geq 0 \) and \( \bar{x}_s - X^i_s \geq 0 \), Lemma 1 implies \( \psi^{AX}(s_i) \geq \psi^A(s_i) + \psi^X(s_i) \). By Proposition 3, we have \( \psi_{s_s}^{AX} \geq \psi_{s_s}^A + \psi_{s_s}^X \).
F. Proof of Proposition 6

By Proposition 3, \( \Psi_s(n, \cdot) \) is linear in \( n \), so the net benefit \( \Psi_s - K_s \) is convex (concave) in \( n \) when \( K_s \) is concave (convex) in \( n \), which explains part (i) and (ii). By Proposition 5, \( \Psi_s(\cdot, Y) \) is superadditive in \( Y \), so the net benefit \( \Psi_s - K_s \) is superadditive in \( Y \) when \( K_s \) is subadditive in \( Y \), which explains part (iii). When \( K_s \) is superadditive in \( Y \), the net benefit \( \Psi_s - K_s \) is not necessarily superadditive in \( Y \), so the optimal learning level could be intermediate, which explains part (iv).

G. Proof of Proposition 7

(i) Adapting Proposition 1: We have the value of learning about one supplier \( i \) in group \( s \)

\[
\psi(s_i) = [\min(r^o, C_{s,i}) - (A_s + \bar{x}_s)]^+ = [\min(r^o - \bar{A} - A_s, C_{s,i} - \bar{A} - 2A_s) - (A'_s + \bar{x}_s)]^+.
\]

Since \( \min(r^o - \bar{A} - A_s, C_{s,i} - \bar{A} - 2A_s) \) decreases in \( \bar{A}, A_s \) and increases in \( \bar{X}, \bar{X}_s \), the expected value of learning group \( s \), \( \Psi_s = N_s \cdot E[\psi(s_i)] \), decreases in \( F, \bar{F}_s \) and increases in \( \bar{G}, \bar{G}_s \).

(ii) Adapting Proposition 2: For supplier \( i \) in group \( s \) and supplier \( j \) in group \( t \), we have

\[
\psi(s_i) = [\min(r^o, C_{s,j}) - (A_s + \bar{x}_s)]^+ = [\min(r^o, C_{s,i}) - \bar{A} - A_s - (A'_s + \bar{x}_s)]^+ \quad \text{and} \quad \psi(t_j) = [\min(r^o, C_{s,j}) - (A_t + \bar{X}_t)]^+ = [\min(r^o, C_{s,i}) - \bar{A} - A_s - (A'_t + \bar{x}_t)]^+.
\]

Because \( \min(r^o, C_{s,w}) - \bar{A} - \bar{X}_w \) and \( A'_w + \bar{x}_w (w = s, t) \) are independent random variables, from Theorem 1.2.17 in Müller and Stoyan (2002), we can conclude that \( \min(r^o, C_{s,i}) - \bar{A} - A_s - (A'_s + \bar{x}_s) \geq \min(r^o, C_{s,t}) - \bar{A} - A_s - (A'_t + \bar{x}_t) \) and \( \psi(s_i) \geq \psi(t_j) \).

(iii) Adapting Proposition 4: For supplier \( i \) in group \( s \), the value of learning his portion \( A \), or \( X \), are

\[
\psi^A(s_i) = [\min(r^o, C_{s,j}) - (A_s + \bar{x}_s)]^+ = [\min(r^o, C_{s,i}) - \bar{A} - A_s - (A'_s + \bar{x}_s)]^+,
\]

\[
\psi^X(s_i) = [\min(r^o, C_{s,j}) - (X_s + \bar{a}_s)]^+ = [\min(r^o, C_{s,i}) - \bar{X} - X_s - (X'_s + \bar{a}_s)]^+.
\]

Since \( \bar{a}_s - A'_s \geq \bar{x}_s - X'_s, A'_s + \bar{x}_s \leq X'_s + \bar{a}_s \). Because \( \min(r^o, C_{s,i}) - \bar{A} - A_s \) and \( A'_s + \bar{x}_s \) are independent, and \( \min(r^o, C_{s,i}) - \bar{X} - X_s \) and \( X'_s + \bar{a}_s \) are independent, from Theorem 1.2.17 in Müller and Stoyan (2002), we can conclude that \( \min(r^o, C_{s,i}) - \bar{A} - A_s - (A'_s + \bar{x}_s) \geq \min(r^o, C_{s,i}) - \bar{X} - X_s - (X'_s + \bar{a}_s) \) and \( \psi^A(s_i) \geq \psi^X(s_i) \).

(iv) Proofs of Proposition 3 and 5 in Appendix A and E can be directly applied to the correlation case.

H. Proof of Proposition 8

For a bidding ring \( t \), the representative bid is the lowest of all suppliers’ costs, which is governed by distribution \( H_t^{\text{ring}} \). If the buyer learns about portion \( A \) of all suppliers in bidding ring \( t \), then the new reserve price is \( \min(r^o, A_{i|N_t} + \bar{x}_t) \), in which \( A_{i|N_t} \) is governed by distribution \( F_t^{\text{ring}} \). Similarly for portion \( X \), the new reserve price set by learning about portion \( X \) is \( \min(r^o, X_{i|N_t} + \bar{a}_t) \), in which \( X_{i|N_t} \) is governed by distribution \( G_t^{\text{ring}} \). This explains why Propositions 1, 2, 3 and 4 hold if
We utilize the following recursive relationship from David and Joshi (1968):

\[ C_\text{Suppliers} = \psi(x, a, \bar{a}) \]

Let us consider the following scenario:

(i) Proposition 3 does not hold at individual level. We provide a counter-example. Suppose there are two colluding suppliers who have the same cost structure with upper bounds \( \bar{a} = 1, \bar{x} = 1, \bar{c} = 2 \). The cost realizations are \( a^1 = 0.2, x^1 = 0.5, c^1 = 0.7; a^2 = 0.5, x^2 = 0.3, c^2 = 0.8 \). Since the two suppliers are colluding, supplier 2 will not submit a meaningful bid and the contract price without learning is \( \bar{c} = 2 \), so the values of learning supplier 1, learning supplier 2, and learning both suppliers, are \( \psi(1) = [\bar{c} - (a^1 + \bar{x})^+] = 0.8, \psi(2) = [\bar{c} - (a^2 + \bar{x})^+] = 0.5 \) and \( \psi(1, 2) = [\bar{c} - (\min(a^1, a^2) + \bar{x})^+] = 0.8 \). The additivity does not hold since \( \psi(1, 2) < \psi(1) + \psi(2) \).

(ii) Proposition 5 does not hold at group level. Consider the same example as in (i), the values of learning both suppliers are \( \psi^A(1, 2) = [\bar{c} - (\min(a^1, a^2) + \bar{x})^+] = 0.8, \psi^X(1, 2) = [\bar{c} - (\min(x^1, x^2) + \bar{a})^+] = 0.7 \) and \( \psi^{AX}(1, 2) = [\bar{c} - (\min(c^1, c^2))^+] = 1.3 \). The superadditivity does not hold since \( \psi^{AX}(1, 2) < \psi^A(1, 2) + \psi^X(1, 2) \).

(iii) Proposition 5 applies at individual level. We can use the proof in Appendix E, while \( C_{-s} \) in that proof stands for the minimum cost of all other suppliers not in bidding ring \( s \).

I. Proof of Proposition 9

When suppliers collude, fewer bids are submitted and thus the contract price without learning is higher. Since the new reserve price is the same, the value of learning (which equals to the gap between the contract price without learning and the new reserve price) is higher in the presence of collusion.

J. Proof of Proposition 10

The value of learning \( m \) suppliers is \( m^n \), and the value of locating \( y \) suppliers is \( \Phi(y) \). Hence, the buyer prefers learning to locating if and only if \( m^n \geq \Phi(y) \iff m \geq \frac{\Phi(y)}{m} \iff m \geq m^* = \left[ \frac{\Phi(y)}{m} \right] \).

K. Proof of Proposition 11

We only need to prove that \( \gamma_N = \frac{\Phi(1)}{N^n} = \int h'(c)^2 H'(c) N^{-1} dc \in (0, 1) \). Let \( \mu_{k:M} \) denote \( E[C'_{[k:M]}] \). Since \( \Phi(1) = E[C_{[2:N]} - C_{[2:N+1]}] = E[C'_{[2,N]} - C'_{[2,N+1]}] = \mu_{2:N} - \mu_{2:N+1} \) and \( \Psi = E[C_{[2,N]} - C_{[1:N]}] = P(C_1 = C_{[1,N]}) \cdot E[C_{[2,N]} - C_{[1,N]}] = \frac{1}{\Psi} E[C'_{[2,N]} - C'_{[1,N]}] = \frac{1}{\Psi} (\mu_{2:N} - \mu_{1:N}) \), we have \( \gamma_N = \frac{\Phi(1)}{N^n} = \frac{\mu_{2:N} - \mu_{2:N+1}}{\mu_{2:N} - \mu_{1:N}} \).

Suppose \( C' \) is distributed over interval \([c', \bar{c}] \) with density function \( h' \) and CDF \( H' \), then the density function of \( C'_{[1,M]} \) is \( p(c) = M h'(c) H'(c)^{M-1} \) and the expected value of the first-order statistic is

\[ \mu_{1:M} = \int_{\underline{c}}^{\bar{c}} c M h'(c) H'(c)^{M-1} dc = - \int_{\underline{c}}^{\bar{c}} c d(H'(c)^M) = -c\bar{H}'(c)^M \bigg|_{\underline{c}}^{\bar{c}} + \int_{\underline{c}}^{\bar{c}} \bar{H}'(c)^M dc = \bar{H}'(c)^M \bigg|_{\underline{c}}^{\bar{c}} + \int_{\underline{c}}^{\bar{c}} \bar{H}'(c)^M dc. \]

We utilize the following recursive relationship from David and Joshi (1968):

\[ \mu_{2:N} = N \mu_{1:N-1} - (N-1) \mu_{1:N}, \text{ and } \mu_{2:N+1} = (N+1) \mu_{1:N} - N \mu_{1:N+1}. \]
Then, we have

\[ \gamma_{N} = \frac{\mu_{2,N} - \mu_{2,N+1}}{\mu_{1,N} - \mu_{1,N+1}} = \frac{\mu_{1,N-1} - 2\mu_{1,N} + \mu_{1,N+1}}{\mu_{1,N-1} - \mu_{1,N}} \]

\[ = \frac{\int H'(c)^{N-1}dc - 2 \int H'(c)^{N}dc + \int H'(c)^{N+1}dc}{\int H'(c)^{N-1}dc} = \frac{\int H'(c)^{2}H'(c)^{N-1}dc}{\int H'(c)^{N}H'(c)^{N-1}dc} < 1. \]

References


